Journal of Public Economics 118 (2014) 128–142

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

The promise of positive optimal taxation: normative diversity and a role for equal sacrifice☆

Matthew Weinzierl

277 Morgan Hall, Harvard Business School, Boston MA 02163, United States
NBER, Cambridge MA 02138, United States

ARTICLE INFO

Article history:
Received 10 September 2013
Received in revised form 17 June 2014
Accepted 18 June 2014
Available online 1 July 2014

Keywords:
Optimal taxation
Tagging
Equal sacrifice

ABSTRACT

A prominent assumption in modern optimal tax research is that the objective of taxation is Utilitarian. I present new survey evidence that most people reject this assumption's implications for several prominent features of tax policy, instead preferring tax policies based at least in part on a classic alternative objective: the principle of Equal Sacrifice. I generalize the standard model to accommodate this preference for a mixed objective, proposing a method by which to make disparate criteria commensurable while respecting Pareto efficiency. Then, I show that optimal policy in this generalized model, calibrated to the survey evidence and U.S. microdata, is capable of quantitatively matching several features of existing tax policy that are incompatible in the conventional model but widely endorsed in the survey and reality, including the coexistence of substantial redistribution and limited tagging. Together, these findings demonstrate the potential of a positive theory of optimal taxation.

1. Introduction

Modern tax theorists have a workhorse model. Created by Mirrlees (1971) more than four decades ago, that model has been used to study countless aspects of tax policy. It provides the benchmark guidelines against which policy proposals are often judged, and its recommendations form the basis of prominent policy advice.

When this standard model has been used to generate quantitative lessons for policy, theorists commonly have imposed a strong assumption: the objective of tax policy is Utilitarian, either in its simplest form as a sum of individual utilities or in a more general form as the sum of a concave transformation of individual utilities. Mirrlees (1971) himself introduced this assumption with little explanation, but virtually all optimal tax research in the last four decades has adopted it [1]. To the extent that this assumption has been relaxed, it has usually been to allow for a more redistributive normative criterion, such as the Rawlsian priority on the least advantaged. The conventional case for Utilitarianism is usually traced to Harsanyi (1953, 1955).

Some theorists have taken a more agnostic approach by examining only whether policies are optimal given some set of weights on individuals' welfares; that is, Pareto efficient. An open question in that approach is what weights to use when choosing between a wide range of Pareto-efficient policy options; in practice, Utilitarian (or Rawlsian) weights are typically the default assumption.[2] The relatively little attention paid to the Utilitarian assumption and its alternatives, as opposed to its policy implications, is especially surprising given that optimal tax theory is one of few forthrightly normative fields in economic research.

The first contribution of this paper is to present evidence of wide disagreement with this core assumption, at least in the United States. I design and implement a novel survey in which respondents are asked to choose between sets of feasible and incentive compatible tax policies for a society with the income distribution of the current United States. First, I ask them to choose between two policies: one based on the standard (simple sum) Utilitarian criterion and the other based on the principle of Equal Sacrifice, a less redistributive and historically prominent alternative criterion for optimal tax design. In that case, nearly 60% of respondents prefer the Equal Sacrifice alternative over the conventional Utilitarian objective. Disagreement with the conventional Utilitarian

☆ A number of technical details, footnotes, contextual information, and clarifications omitted from this published version can be found in the working paper version of this paper and its Appendix, Weinzierl (2014), at the author's website. Portions of this paper incorporate and build on material from Weinzierl (2012), which was entitled "Why do we Redistribute so Much but Tag so Little?" I am grateful to Alan Auerbach, Felix Bierbrauer, Kim Clausing, Raj Chetty, Mihir Desai, Rafael di Tella, Amy Finkelstein, Victor Fleischer, John Friedman, Alex Gelber, Mikhail Golosov, Caroline Hoxby, Bas Jacobs, Louis Kaplow, Wojciech Kopczuk, Camille Landais, Benjamin B. Lockwood, Greg Mankiw, Yoram Margalioth, Joe Mazor, Jean-Baptiste Michaud, Jeff Miron, Dina Pomeranz, Alex Raskolnikov, Meg Rithmire, Julio Rotemberg, Emmanuel Saez, Bernard Salanie, Larry Samuelson, Florian Scheuer, Eytan Sheshinski, Ali Shourideh, Stefanie Stantcheva, Alain Trannoy, Aleh Tsyvinski, David Weisbach, Glen Weyl, John Weymark, Danny Yagan, and several anonymous referees for their helpful discussions.

E-mail address: mweinzierl@hbs.edu.

[1] Harsanyi (1953) provides the classic discussion of the applicability of Utilitarianism in this context.

[2] See, e.g., Stiglitz, 1987; Werning, 2007; Rothschild and Scheuer, 2012, and Saez and Stantcheva, 2014. In addition, specific normative limitations of the conventional model have been addressed directly (see Section 1.4 and Weinzierl, 2014).

http://dx.doi.org/10.1016/j.jpubeco.2014.06.012
0047-2727/

assumption is even more striking when I give respondents a range of choices, including options that are based in part on Utilitarianism and in part on Equal Sacrifice. I find that 81% of individuals prefer policies other than the pure Utilitarian or Rawlsian policies, and nearly half most prefer policies based on a combination of Utilitarianism and Equal Sacrifice.[3] Of course, these responses may be due to a variety of factors other than an affinity for Equal Sacrifice, so I use additional questions in the survey to test for more direct evidence on the relevance of Equal Sacrifice. When asked explicitly how "sacrifice" from paying taxes should be distributed, respondents prefer a distribution between that implied by Utilitarianism and Equal Sacrifice. And the more enthusiastic a respondent is about Equal Sacrifice, the more likely he or she is to reject tagging, the taxation of personal characteristics that is a feature of Utilitarian-optimal tax policy but that is rejected by Equal Sacrifice.

This evidence is admittedly far from definitive. The survey respondents are not a random representative sample of Americans, and many variations in the survey's design, framing, and implementation are possible and could have large effects (see McCaffery and Baron, 2004, for example). Nevertheless, the results are robust across subsamples, and the survey is designed to guard against a number of potential design concerns. In the end, the survey evidence suggests that a number of features of tax policy implied by the Utilitarian objective of conventional theory may not be, in the terminology of Diamond and Saez (2011), "socially acceptable." That is, a large majority of individuals appear to place substantial value on an alternative normative principle – Equal Sacrifice – that rejects some of the conventional objective's policy implications.

While my finding of a preference for a mixed objective is foreign to the optimal tax literature, it is consistent with a large body of existing research showing that most individuals are not normative purists. In that research, whether individuals are asked to evaluate income distributions, answer conceptual questions, or participate in allocation games, few appear to use a single normative criterion. As Scott et al. (2001) write: "Experimental research reveals that distributive justice judgments usually involve several distinct allocation principles."

How should we respond to this evidence? One possible response is to ignore it. We may decide that a normative theory ought to choose its objective based on philosophical reasoning regardless of popular opinion. An alternative approach is to incorporate as much evidence as possible on the way the agents included in these models think about these very same issues. In their important synthesis of "empirical social choice" research, Gaertner and Schokkaert (2012) make a strong argument for the value of eliciting public attitudes toward such issues. In optimal tax research, incorporating key aspects of reality into the conventional model has been a hallmark of major contributions such as Diamond (1998), Saez (2001, 2002), Golosov and Tsyvinski (2006), and Farhi and Werning (2010), and often these efforts have improved the match between the theory's recommendations and real-world policy. Diamond and Saez (2011) suggest a similar effort with regard to the normative aspects of the model, advocating a requirement of "social acceptability" under which real-world normative beliefs would constrain the set of relevant policy results.[4] My paper falls in this tradition and proposes that we go one step further: it gathers formal evidence about people's views and interprets that evidence as motivation for constructing a *positive* optimal tax theory. The broad aim of this positive optimal taxation project is, then, to pursue empirically-supported generalizations of the standard optimal tax model to better match the way in which real societies appear to evaluate tax policy. Specifically, this paper's survey evidence, and a large body of prior work, suggests that we generalize the standard model to include a mixed policy objective.

The second main contribution of this paper is to formally develop a generalized model that can be used for positive optimal tax analysis. The generalized model combines multiple normative criteria into a single policy objective while retaining both Pareto efficiency and the remainder of the familiar formal apparatus of conventional optimal tax theory. In this way, I am following up on a suggestion made more than three decades ago by Martin Feldstein (1976), that "optimal tax design involves a balancing of conflicting criteria." This generalization of the standard theory requires addressing long-standing concerns about commensurability of different normative criteria. In keeping with the survey evidence, I develop in depth the specific case of an objective that combines Utilitarianism and Equal Sacrifice.

A complementary approach to generalizing the conventional model's objective – part of the Pareto-efficient optimal tax approach mentioned above – can be found in contemporaneous research by Saez and Stantcheva (2014). They focus on the role of marginal social welfare weights in the aggregation of a given tax reform's effects on individuals. By allowing these weights to take any non-negative values, they include the possibility that they may be based in part on normative criteria other than Utilitarianism.[5] Their approach and this paper's can be seen as two sides of the same coin: one might translate a mixed objective function into a profile of marginal social welfare weights or vice versa. Each approach has applications for which it is more naturally suited, and both contribute toward the broader goal of constructing a positive theory of optimal taxation.

One attractive feature of this paper's approach is that it requires a clear statement of each component of the set of criteria by which policy is judged. This requirement acts as a second test of the theory (in addition to its ability to match observed policy features), in that criteria lacking intellectual coherence can be rejected and we can avoid the risk that fully-flexible welfare weights lose any explanatory power. More generally, under the Pareto-efficient optimal tax approach, assumptions on the welfare weights are often made in the interests of deriving more powerful results. One way to interpret my contribution in the context of that approach is that I look for evidence on the normative criteria that seem to hold in reality and that, therefore, might inform the values of those weights that society would endorse. Specifically, I am able to use the principle of Equal Sacrifice as a disciplined way to give weight to a point on the Pareto frontier that appears to matter to the public but has been largely ignored by modern tax theory.

This paper's approach has a number of limitations. Positive optimal tax theory as developed here is not a positive tax theory, i.e., I have not modeled the political economy that translates the public's preferences into policy. While recognizing that establishing such a link is essential for a full understanding of how any normative principles affect real-world allocations, that task is outside the main objective of this paper. At the same time, positive optimal tax theory is not normative optimal tax theory, and we may reject the implications of the former if we believe the public is – at any given time – subject to biases or mistaken beliefs. It is because of this very real and important risk that I emphasize the search for recognizable, and at least arguably defensible, philosophical principles in the development of the model. Moreover, it is important to clarify that positive optimal tax theory is not a substitute for traditional normative optimal tax theory based on considered judgments of what society's objective function ought to be.

The third contribution of this paper is to show that this generalized model, when calibrated to this survey evidence, can reconcile a number of features of tax policy that are incompatible in conventional theory but endorsed in the survey evidence as well as in reality. In particular, I simulate optimal policy using the survey respondents' most-preferred

[3] A note on terminology: from this point on I will use "Utilitarian" to refer to the simple sum of individual utilities, not the more general version in which transformations of those utilities are made prior to aggregation.

[4] Gaertner and Schokkaert (2012) provide a lucid and insightful discussion of the relationship between normative and positive analyses of social preferences.

[5] Saez and Stantcheva also note that welfare weights could be derived from existing policies or survey evidence. Bourguignon and Spadaro (2012) take the former approach to calibrating the welfare weights in a standard model, as do Spadaro et al. (2012); Bargain et al. (2011, 2013); Zoutman et al. (2013a, 2013b); Hendren (2014); and Lockwood and Weinzierl (2014).

normative objective, which combines Utilitarianism and Equal Sacrifice, and U.S. microdata. That policy simultaneously rejects the use of height, gender, and race as tags; accepts the use of blindness as a tag, endorsing a quantitatively realistic blindness benefit; and provides redistribution through a progressive schedule of average income tax rates that closely resembles actual policy. I also briefly note two other examples: it substantially reduces the extent of utility rank reversals in the first-best policy, and it implies top marginal tax rates lower than what conventional theory would recommend and closer to reality.[6]

The results on tagging are important because they provide evidence that Equal Sacrifice's explanatory power is due to more than simply being a "less redistributive" criterion than Utilitarianism. While the optimal extent of both tagging and redistribution decrease when Equal Sacrifice is given more weight, tagging is reduced much more dramatically. Intuitively, Equal Sacrifice strongly rejects horizontal inequity in taxes (conditional on income-earning ability) because a taxpayer's sacrifice is determined by his or her income-earning ability only, while Equal Sacrifice can accommodate a range of tax progressivity across ability levels. Consistent with this indirect evidence, I find more direct evidence for Equal Sacrifice's role in explaining limited tagging in the survey, where a greater share of survey respondents who oppose height and blindness tags prefers policies based in part on Equal Sacrifice.

Taken together, the survey results, theoretical analysis, and calibrated simulations of this paper demonstrate the potential of a positive optimal taxation research agenda. They show that we can rigorously capture empirical evidence on what tax policies individuals find acceptable and, as one might hope, use the resulting model to better understand how actual tax policy is and (arguably) ought to be designed.

The support that I find for policies based in part on Equal Sacrifice may seem surprising, but in fact it ought not to be. Though Equal Sacrifice has played only a minor role in tax research since 1971, it was originally proposed by no less a Utilitarian than John Stuart Mill, and it avoids a prominent critique of Utilitarianism put forward by John Rawls (1971), among others. In the early years of modern optimal tax theory, Martin Feldstein (1976) saw a connection between Equal Sacrifice and Robert Nozick's (1974) Libertarianism, arguing that "…tax schedules that impose equal utility sacrifice have an appeal that is clearly lacking in the utilitarian framework." The pioneering work of H. Peyton Young (1987, 1988, 1990, 1994) and Berliant and Gouveia (1993) showed that existing income tax rate schedules were consistent with the Equal Sacrifice principle by itself. In a sense, it would be surprising if Equal Sacrifice did *not* feature at least somewhat in the views of many, especially in the United States.

The paper proceeds as follows. Section 1 reports the new survey evidence on normative preferences and discusses similar findings in prior work. Section 2 generalizes the standard model to allow for a mixed objective, discusses Equal Sacrifice as an alternative to Utilitarianism, and applies the model to the case of these two criteria. Section 3 shows that the parameterizations of that model most preferred by survey respondents imply policies that resolve several disparities between conventional theory and real-world policy, especially the puzzle of limited tagging. Section 4 concludes, and Weinzierl (2014) contains supporting material.

2. New results on empirical normative preferences

In this section, I describe the design and results of a novel survey eliciting normative preferences over realistic tax policies. I also provide a range of robustness checks, all of which confirm the main findings: few individuals prefer the pure Utilitarian criterion standard in conventional optimal tax theory or the commonly-used Rawlsian alternative, and a plurality of individuals prefers tax policies reflecting a mixed normative objective including both Utilitarianism and the classic alternative criterion of Equal Sacrifice. Despite those checks, of course these results are far from definitive. Future research could explore many variations on the surveys I perform, including changes to the way the data are presented, the design of the survey itself, and the choice of the respondent sample.

One important ambiguity in a subset of the survey results is that the influence of the Equal Sacrifice criterion on the policy choices facing survey respondents is by no means unique to Equal Sacrifice. Respondents endorsing the less redistributive policies in the survey may be motivated by a variety of factors other than an affinity for Equal Sacrifice. To clarify, this paper is not intended to show that Equal Sacrifice influences preferences to the exclusion of other, similar factors. Nevertheless, to address this ambiguity and bolster the case that Equal Sacrifice does matter for preferences, later in the paper I show the results of survey questions that explicitly gauge support for Equal Sacrifice, and I show that Equal Sacrifice can have explanatory power along a dimension of policy – tagging – not directly tied to the extent of redistribution (in Section 3).

This paper's survey makes a methodological contribution to empirical research on tax preferences by having respondents face a task that mimics the conventional social planner's optimal tax problem: that is, policies are constrained by both feasibility (in the context of government spending) and incentive compatibility. This innovation over most prior work[7] allows me to use the evidence on participant preferences to calibrate a fully-specified optimal policy model.

2.1. Survey design

The survey, shown in full in Appendix to Weinzierl (2014), has three parts. The first part tests whether respondents understand and can perform simple calculations related to the concepts of before-tax income, after-tax income, and average tax rates. It defines each of these terms, shows a graphical illustration of them that parallels the figures used in the remainder of the survey, and then asks four multiple-choice questions to test comprehension. The third part of the survey asks respondents about their opinions on aspects of tax policy, political views, and personal traits, including economic status. To address any concerns that these characteristics matter for the results, I examine my findings' robustness across all subgroups.

The second part is the centerpiece of the survey. Respondents are shown a graphical gross income distribution divided into eight types of households (based on CBO data as discussed below). These types represent the four lower quintiles and a division of the top quintile into the next 10, 5, 4, and 1 percentiles. We might worry that respondents attribute some of the variation in incomes to differences in preferences for which, in the influential terminology of Fleurbaey and Maniquet (2006), individuals ought to be held "responsible" and not taxed or subsidized. I try to minimize that risk by clarifying in the survey text that differences in earnings are not due to effort as follows[8]: "If there were no taxes, these households would all work equally hard. But, type 2 would earn more than type 1, type 3 would earn more than type 2, and so on."

Respondents are then put in the position of objective policymakers facing a constrained optimal tax problem. They are told "You are given the chance to choose taxes for this society. Please think of yourself as a policymaker for this society." They are given information about the constraints affecting their choices, as the survey states the required level of exogenous government spending (i.e., feasibility) and emphasizes that households' labor supplies include responses to tax policy

[6] Weinzierl (2014) contains much more detail on these applications than is shown here.

[7] The recent paper by Kuziemko et al. (2013) also presents respondents with realistic policy choices.

[8] See Lockwood and Weinzierl (2014) on how greater preference heterogeneity lowers optimal redistribution in the standard model. To the extent that heterogeneity in preferences exists in reality, the survey text to the contrary would be expected to push respondents toward a more redistributive policy than they would endorse in reality.

(i.e., incentive compatibility). Respondents are reminded that taxes may serve a variety of purposes, from funding public goods to redistributing before-tax income. [9]

In a series of choices, respondents rank sets of tax policies. For each policy option, the survey displays two overlapping income distributions (see Fig. 1 for an example). The pretax distribution is shown as empty outlined columns while the after tax distribution is shown as filled-in columns. The average tax rate for each household type is shown in a text box above their columns. Respondents are asked to rank the policies from "best" to "worst" by clicking on numbered radio buttons. By using the general terms "best" and "worst" without further defining the criteria by which tax policies ought to be judged, the survey leaves the respondent free to use his or her own definition of optimality.

The survey was listed in November, 2012 as an available task to up to 400 members of the Amazon Mechanical Turk worker population from the United States who demonstrated good past performance on tasks.[10] The title of the task was "We want your opinions on tax policy", the description was "Rank possible tax policies and give us your opinions on taxes," and the survey requestor was identified as "TaxSurvey." Respondents had up to 30 min to complete the survey, and they were asked to enter their MTurk identification number as well as a completion code at the end of the survey for verification purposes. The respondents completed the survey in an average of 13 min and 6 s. They were paid $2.00 for the task, implying an average hourly rate of $9.16.

2.2. Results

Respondents' first rankings provide straightforward evidence that the Utilitarian criterion is less popular than the conventional model implies. Fig. 1 shows the two policies respondents rank, labeled A and B. In Section 2, I provide the details of how I calculated these (and all other) policy options in the survey. For now, note that option A reflects a conventional, pure Utilitarian objective for tax policy, while option B reflects an objective based entirely on the principle of Equal Sacrifice (which sets the utility cost of taxation equal for all individuals). Policy A is redistributive, while B is not.

The results of this first choice are strikingly at odds with the conventional model's assumed objective. The share of respondents preferring the Utilitarian policy A is 42%, with a standard error of 2%. In other words, nearly three-fifths of respondents prefer the pure Equal Sacrifice policy B to policy A.

Respondents are then asked, over the course of two questions, to rank a wider range of seven policy options. Table 1 summarizes these choices, decreasing in redistributiveness from left to right. For each option, it shows the average tax rates levied on each household. In the second column of the table, before-tax incomes in the no-tax scenario are shown.

The middle five policies in Table 1 combine Utilitarianism and Equal Sacrifice, using a range of values for the weight on Equal Sacrifice, α_{ES}, discussed below. The value $\alpha_{ES} = 0.00$ yields the conventional Utilitarian policy, while $\alpha_{ES} = 1.00$ yields the Equal Sacrifice policy. In between these polar values, three values generate intermediate policies: $\alpha_{ES} = \left\{0.0\overline{3},\ 0.10,\ 0.20\right\}$.

I also generate the two "endpoint" policies shown in Table 1. The left-most policy, C, is a "Rawlsian" policy that maximizes the utility of the lowest-ability household. The right-most policy is a "poll tax" that splits the financial cost of government spending G evenly across households.[11]. Adding these endpoint policies yields two benefits. First, offering the Rawlsian and poll tax options addresses a potential framing problem with presenting respondents with only the set of policies along the Utilitarian-Equal Sacrifice spectrum. To the extent that individuals shy away from options that seem "extreme," having policies A and B as endpoints could bias us toward finding support for a mixed objective. Adding the Rawlsian and poll tax options as the endpoints on the redistributive spectrum may alleviate this concern. Second, the most common deviation from simple-sum Utilitarianism in conventional optimal tax theory is a generalized Utilitarianism under which the planner takes a concave transformation of utilities before summing them. The Rawlsian option is often included as an extreme version of this generalization. By including a Rawlsian policy as a choice, we can gauge the empirical support for this prominent criterion.

Respondents first compare option A to three additional options, two of which are less redistributive than A, namely D and E, while C is more redistributive. Respondents then compare option B from the first choice to three additional options, two of which are more redistributive than B, namely D and G, while F is less redistributive. Both choices can be seen in the survey as reproduced in Weinzierl (2014). Policy D is included in both sets of four-option rankings so that we can infer respondents' preferences across the full range of seven policies.[12] Note that the redistributive spectrum is masked in both the alphabetical policy labels and the physical placement of policies within the four-policy rankings (visible in Weinzierl 2014).

Fig. 2 shows each of these seven policy options and the share of respondents who placed them in their top-ranked or second-ranked group of policies.

Fig. 2 reveals two main results from these rankings: one, support for the conventional Utilitarian assumption and the Rawlsian alternative is low; two, a plurality of respondents prefer a mixed objective.[13] The purely Utilitarian policy (A) makes up only 10% of the top-ranked choices and 11% of the second-ranked choices.[14] For the Rawlsian alternative, these figures are 9% and 10%. Together, then, policies at least as

[9] I ran a follow-up survey, requested by a referee, that explicitly explained that transfers were made in kind. Though a greater percentage of the survey respondents self-identified as politically left-leaning, the main results highlighted below were unchanged. In particular, a majority of respondents preferred policy B to policy A, and in the seven-policy choice three-quarters preferred a policy other than Utilitarian or Rawlsian, with a plurality preferring one of the three intermediate policies.

[10] Specifically, only respondents registered as in the United States whose work had been accepted on 95% of previous tasks could take the survey. Horton et al. (2011) study the use of online labor markets, and specifically of Mechanical Turk, and find: "Online experiments, we show, can be just as valid—both internally and externally—as laboratory and field experiments, while often requiring far less money and time to design and conduct."

[11] Formally, using notation specified below, these policies are as follows. To generate the Rawlsian policy option, the planner solves the problem

$$\max_{\{c_*^i, y_*^i\}_{i=1}^I \in \{\mathbb{F}\cap\mathbb{IC}\}} U\left(c_*^1, y_*^1/w^1\right),$$

where $\mathbb{F}$ is defined in Eq. (2); $\mathbb{IC}$ is defined in Eq. (3); and $U(c^1, y^1 / w^1)$ is defined in Eq. (12). The Rawlsian policy is the most redistributive policy option. The poll tax policy is defined as follows:

$$c_*^i = y_*^i - \frac{1}{I}G \text{ for all } i \in \{1, \ldots, I\},$$

where households maximize utility subject to this constraint. The poll tax is the least redistributive policy option (it is, in fact, regressive).

[12] An alternative approach would be to give the respondents control over a continuous policy lever that would trace out the entire range of redistribution (for instance, α_{ES}). That alternative has two drawbacks, however. First, communicating the meaning of that policy lever would be difficult without influencing the respondents' answers. Second, we are likely interested not merely in the respondents' ideal points but in their attitudes toward options along the entire range. Those would be difficult to elicit with this alternative approach.

[13] Note that the survey was designed to minimize the risk that aversion to choosing endpoints is generating the observed preference for the intermediate policies D, E, and G. Each of the two scenarios in which respondents chose between four policies had policy D as an endpoint option. In contrast, policies A and B were intermediate policies in each of their four-option scenarios.

[14] These classifications are made as follows. Using option D, which was included in both of the four-option policy choices, we can create a weak ranking of all seven policy options for each respondent. The top-ranked group includes any policy strictly dominated by no other policy. Less than 6% of respondents had more than one top-ranked policy. The second-ranked group includes any policy strictly dominated by only policies in the top-ranked group. For example, if a respondent ranks option D as their first choice in both four-option choices, they will rank two policies second, each of which is therefore placed in the "second-ranked" group for that respondent.

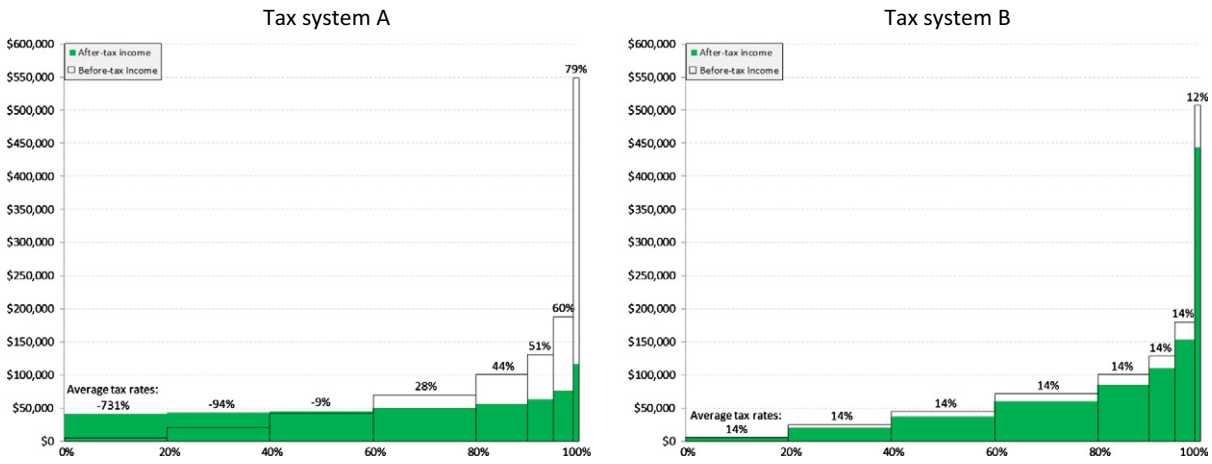


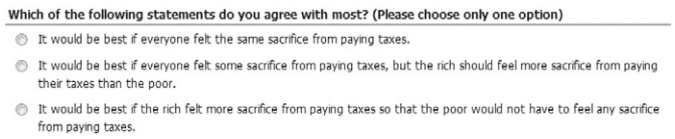
Fig. 1. The choice between the utilitarian policy A and the equal sacrifice policy B.

redistributive as the conventional Utilitarianism make up less than 20% of the most-preferred policies in this survey, the same share claimed by the pure Equal Sacrifice policy B. In contrast, nearly half – 48% – of the top-ranked policies were one of the three (E, D, and G) that correspond to a mixed normative criterion. These mixed policies also dominate the second-ranked preferences of respondents, making up 56% of those choices.

Together, these results sharply contradict the normative assumptions that dominate modern optimal tax research. Respondents give little support to using the conventional Utilitarian criterion as the optimal policy benchmark. They also appear to disagree with the most commonly used alternative to pure Utilitarianism – a more concave social welfare function – as respondents are less enthusiastic about the Rawlsian policy than any other option except (perhaps) the poll tax. Instead, empirical normative preferences appear to favor the use of a mixed objective with some weight on a less redistributive criterion such as Equal Sacrifice.

More direct evidence of a preference for an objective that combines Utilitarianism with Equal Sacrifice in particular is revealed when respondents are asked explicitly about the optimal distribution of sacrifice in a tax system:

The responses to this question are as follows: 33% choose the first



option; 48% choose the second; and 19% choose the third option.[15] As these responses demonstrate, the preference for mixed objectives that was apparent in respondents' choices over tax systems is echoed by their stated preferences over the distribution of sacrifice from the tax system. The conventional Utilitarian policy is most consistent with the third option in this question, though it would in fact recommend a more redistributive option in which the poor received a net benefit from the tax system. The pure Equal Sacrifice policy is most consistent with the first option. Fewer survey respondents show enthusiasm for the conventional Utilitarian outcome than for the Equal Sacrifice alternative, while the largest percentage of respondents prefers the intermediate option. In other words, more than three-quarters of respondents choose policies reflecting some weight on Equal Sacrifice, the same share as in the choices over tax policies as shown in Fig. 2.

[15] In the follow-up survey, the order of possible answers was randomized at the request of a referee. The results were 37, 39, and 24%. Though slightly less enthusiasm for the middle answer was apparent, that option still was chosen by the most respondents, and the lessons are the same as those drawn from the initial results.

Finally, Fig. 3 shows a degree of consistency that suggests the survey is accurately eliciting respondents' policy preferences. It shows these preferences according to individuals' views on Equal Sacrifice. Policy option B, based purely on Equal Sacrifice, claims more than twice the share of the top rankings among those who state a preference for equal sacrifice than among those who prefer distributing sacrifice less equally. This pattern holds despite that connection never being made apparent in the survey. Similarly, the Utilitarian and Rawlsian policies are supported more by those who prefer to have the poor bear no sacrifice, and intermediate policies are supported more by those who prefer the intermediate distribution of sacrifice.

2.3. Robustness

Here, I analyze the data along several dimensions to check the robustness of these results.

2.3.1. Respondents' understanding

If respondents fail to understand the questions being asked, we might worry that their answers poorly reflect their true preferences. Two sets of observations offer reassurance on this point.

First, respondents appear to understand the economic concepts used in the survey. The survey begins with definitions of the concepts of before-tax income, after-tax income and the average tax rate. It then asks respondents to: 1) use before-tax income and taxes paid to calculate after-tax income; 2) use before-tax income and taxes paid to calculate a (positive) average tax rate; 3) use before-tax income and taxes paid to calculate a (negative) average tax rate; 4) calculate the average of three before-tax incomes. These questions test comprehension

Table 1
Features of the tax systems among which respondents choose.

HH type	No-tax earnings	Tax system:						
		C	A	E	D	G	B	F
		Average tax rates (in percent)						
1	$6,205	−895	−731	−504	−345	−260	14	97
2	$24,314	−119	−94	−43	−8	11	14	33
3	$43,961	−14	−9	16	16	16	14	18
4	$70,254	30	28	22	19	17	14	12
5	$99,114	48	44	28	22	19	14	8
6	$127,252	56	51	33	25	21	14	6
7	$177,199	68	60	39	29	23	14	5
8	$476,167	81	79	59	45	35	12	2
		Weight on equal sacrifice: α_{ES}s						
		Rawls	0.00	$0.0\overline{3}$	0.10	0.20	1.00	*Poll tax*

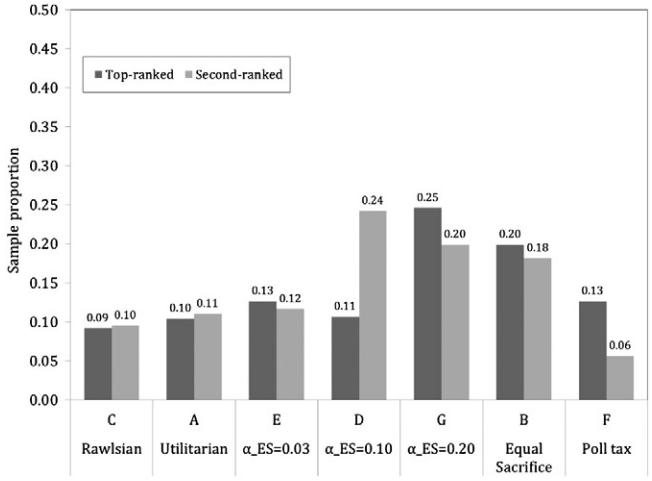


Fig. 2. Respondent preferences across a range of policy options.

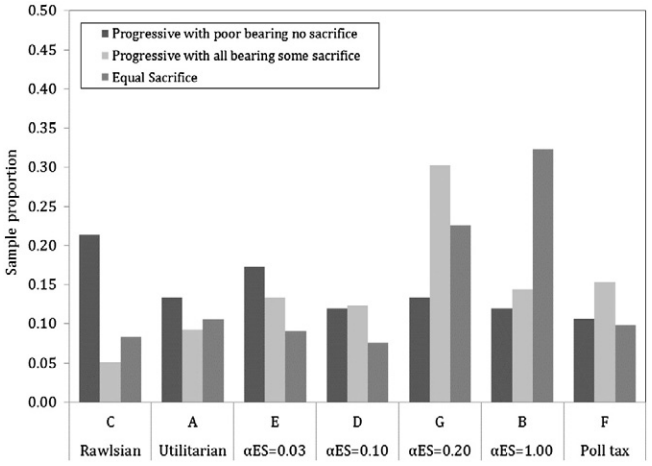


Fig. 3. Policy preferences by view on equal sacrifice (top choices only).

and the ability to work with the concepts, as well as numeracy. The results show that respondent understanding was very high, with 75% of respondents correctly answering all four questions and 87% answering at least three correctly.[16]

Second, the pattern of rankings by most respondents suggests that they understood the choices they were making. If a respondent reports single-peaked preferences across the five policy options along the spectrum between pure Utilitarianism and Equal Sacrifice, we might be confident in his or her understanding of the relationship among policies, not to mention the respondent's rationality. In fact, 68% of respondents exhibited single-peaked preferences across these five policy choices. Importantly, that result does not imply that 32% of respondents were making irrational choices—someone may prefer policies that commit fully to one normative criterion or another, or multiple-peaked preferences may suggest other influences on respondents' normative reasoning that are not captured in this two-component objective function.

2.3.2. Robustness across demographic groups

Natural concerns in any survey of this kind are whether the results are driven by particular demographic groups and whether economic status is systematically related to respondents' preferences. To examine these concerns, I ask the respondents to report their gender, age, education, and economic status when young and when an adult. Summary demographic data for the 381 respondents who successfully completed the survey is provided in Table 2.

Though not meant to be a representative sample, this group exhibits substantial variation in (self-reported) personal characteristics and backgrounds. The self-reported distribution of respondents across household types when they were children matches the overall U.S. income distribution remarkably well, with 36% reporting being from the bottom two quintiles, 45% from the next two quintiles, and 19% from the top quintile. The respondent population also appears to be (or to expect to be) upwardly mobile, with only about one-quarter of those who report their childhood household most resembled one of the two lowest-earning household types reporting that their household at age 40 was also one of those types. Consistent with that fact, the respondents were generally more well-educated than the population aged 18–65 in the United States, where approximately 30% of adults are college graduates.[17]

Table 2 also shows that there are few large differences in preferred policies across demographic groups. In the table, I group policies into three groups: Utilitarian or Rawlsian (C or A); Mixed (E, D, or G); and Equal Sacrifice (B).[18] Respondents with more education and higher economic status tend to be more supportive of Mixed policies. However, a large majority of each demographic group prefers policies other than the Utilitarian or Rawlsian options, and a plurality of all but one group prefers policies that result from a mixed normative criterion that combines Utilitarianism and Equal Sacrifice. The exception is those with lower status as adults, where a slight plurality (though less than one-third) prefers the more redistributive policies.[19] With this exception, both main findings from the full survey apply across demographic groups.

2.3.3. Robustness across political views

A major conceptual question raised by this paper's results is how individuals' preferences are aggregated in a political system. Though I largely set that question aside, we can analyze the survey results to test whether the paper's main conclusions are likely to be sensitive to the details of that aggregation. For example, if we found that individuals of only a particular political perspective were driving the results, we might discount their relevance. To address these concerns, I ask the respondents to self-classify at three points on the (U.S.) political spectrum with regard to economic issues: 1) Left-leaning, or Liberal; 2) Centrist, or Moderate; 3) Right-leaning, or Conservative. I also ask them to classify themselves as (strongly or somewhat) supportive of or opposed to Libertarianism (which is left undefined in the survey). Table 3 shows the distribution of responses. A plurality of the respondents, 44%, self-classifies as left-leaning.[20] Support for Libertarianism in this sample is consistent with the magnitudes for the U.S. population cited by Boaz and Kirby (2007).

Table 3 also shows that both main findings from the full survey characterize respondents across a wide range of political opinions.

[16] Respondents with fewer correct answers were more supportive of both extremes of the policy distribution, especially the poll tax. As with those who answered at least three correctly, a substantial majority (more than two-thirds) of top choices for this group were for a policy other than the Utilitarian or Rawlsian options.

[17] See Table 229 of the Statistical Abstract of the United States: 2012, published by the U.S. Census Bureau.

[18] I do not show the results for the Poll Tax because it received little support and is too distinct from the Equal Sacrifice policy to be grouped with it.

[19] In a follow-up replication of the survey, requested by a referee, I also ask respondents to self-report race, using the General Social Survey categories: white, black, other. The results for "white" and "other" respondents conform to the two main lessons highlighted in the paper. For "black" respondents, less than 10% of the sample, a majority prefers policies other than the Utilitarian or Rawlsian, consistent with the overall respondent group. In contrast with the overall group, however, relatively few black respondents prefer the combination policies, with more than 40% preferring the Equal Sacrifice or Poll Tax policies and just under 40% preferring the Utilitarian or Rawlsian policies.

[20] We might expect this group to be more supportive of redistributive policy than a sample centered on the "centrist" position.

Table 2
Preferences across policy groups by demographic trait.

Tax policy group:	Rawls or utilitarian (C or A)	Mixed (E, D, or G)	Equal sacrifice (B)	Share of respondents
Gender				
Male	0.16	0.48	0.23	0.56
Female	0.24	0.47	0.16	0.44
Age				
18–25	0.16	0.52	0.17	0.30
26–40	0.23	0.49	0.19	0.48
41–65	0.18	0.40	0.25	0.21
Education				
High school grad	0.21	0.36	0.29	0.13
Some college	0.20	0.48	0.18	0.35
College grad	0.19	0.52	0.18	0.51
Status when child				
Types 1–2 (lower)	0.22	0.41	0.22	0.36
Types 3–4 (middle)	0.19	0.50	0.19	0.45
Types 5–8 (higher)	0.16	0.55	0.17	0.19
Status when adult				
Types 1–2 (lower)	0.31	0.28	0.24	0.16
Types 3–4 (middle)	0.19	0.50	0.19	0.51
Types 5–8 (higher)	0.15	0.55	0.18	0.33

Table 3
Preferences across policy groups by political views.

Tax policy group:	Rawls or utilitarian (C or A)	Mixed (E, D, or G)	Equal sacrifice (B)	Share of respondents
Political position				
Left-leaning	0.19	0.57	0.11	0.42
Centrist	0.21	0.52	0.18	0.30
Right-leaning	0.18	0.29	0.37	0.22
View on libertarianism				
Support	0.17	0.42	0.25	0.46
Oppose	0.19	0.57	0.15	0.27

As might be expected, right-leaning and Libertarian respondents are more likely to favor less redistributive policies.[21] However, across all groups, and even among those who self-classify as left-leaning or liberal, a large majority of respondents prefer policies other than those reflecting conventional objectives, and a plurality prefer a mixed normative framework with some (or all) weight on Equal Sacrifice.

2.4. Relation to existing evidence on normative preferences

This paper's survey evidence and the large body of prior empirical work on normative preferences share a common main conclusion: individuals use and prefer a mixed normative criterion. Elsewhere, I discuss the related research in detail, but summary statements from studies representing three research designs in that literature illustrate the main point.[22] Frohlich et al. (1987) use surveys in which participants are asked to rank different distributions of resources, much as in this paper, and find that "…subjects preferred a compromise. This implies that *individuals treat choice between principles as involving marginal decisions. Principles are much like economic goods inasmuch as individuals are willing to trade off between them* [italics in the original]." Feldman and Zaller (1992) ask a large group of Americans open-ended questions on distributive justice and write: "Most people are internally conflicted about exactly what kind of welfare system they want…Ambivalence with respect to social welfare policy is more pronounced among welfare liberals…They end up acknowledging the values of economic individualism even as they try to justify their liberal preferences." Engelmann and Strobel (2004) use allocation games among individuals to elicit values and conclude: "a combination of efficiency concerns, maximin preferences, and selfishness can rationalize most of the data."

This prior work is only indirectly, not directly, supportive of the role of Equal Sacrifice as a factor in normative preferences. As far as I am aware, this paper is the first attempt to elicit (or infer) attitudes toward that principle, though support for the Libertarian viewpoint that some have linked to Equal Sacrifice has been found by a number of researchers (e.g., Cappelen et al., 2011; Boaz and Kirby, 2007; Frohlich et al., 2004; Konow, 2003).[23]

An alternative normative perspective that has received substantial support in empirical work, such as in Fong (2001) and Gaertner and Schokkaert (2012), emphasizes individual responsibility for differences in effort as well as compensation for differences in ability. Fleurbaey and Maniquet (2006) provide an influential theory of optimal taxation sensitive to this distinction, and Lockwood and Weinzierl (2012) discuss how incorporating a simple version of that distinction into the standard model affects optimal policy. While the framework of this paper is not designed to include that alternative perspective directly, our conclusions here are not in conflict with those of that literature, and a synthesis of them (along with others) may prove fruitful.

3. Generalizing the optimal tax model for multiple objectives

The survey results and related literature presented in the previous section suggest two lessons for a positive theory of optimal taxation: first, the conventional optimal tax model's assumption of a Utilitarian objective is counterfactually narrow; second, an accurate positive optimal tax theory must be able to accommodate multiple normative objectives simultaneously. In this section, I generalize the conventional model to allow for this normative diversity, retaining much of the standard theory's (familiar) formal apparatus. I then develop the details of that model for the case of the two main normative criteria used in the survey: Utilitarianism and Equal Sacrifice. Finally, I show the parameterizations of the model that correspond to the policies offered to survey respondents.

3.1. The general model with multiple criteria

Appealing as it may be to generalize the normative objective in the optimal tax model, there is a methodological obstacle: many plausible normative criteria evaluate outcomes in ways that are not directly commensurable. For example, Utilitarianism ranks all possible allocations, but Equal Sacrifice yields only a most-preferred outcome and fails to rank alternative allocations. To obtain a ranking of allocations that reflects the judgments of both criteria therefore requires a translation of Equal Sacrifice into a more complete form. This case is an example of a more general problem with capturing unconventional principles in a framework amenable to economic analysis.[24]

This paper ensures commensurability by representing the priorities of each normative criterion with a loss function that depends on

[21] Support for the poll tax (policy F, not shown in the table) is generally low, but as might be expected it is higher among those who identify as on the "right" (i.e., 16% of top choices vs. 11% for the rest of the sample) and among those who support Libertarianism (i.e., 16% vs. 9%).

[22] See Weinzierl (2014).

[23] Gaertner and Schokkaert (2012) discuss some work related to equal losses in "claims" problems, where they find support for proportional monetary losses: see their Section 4.2.

[24] For example, Utilitarianism has a consequentialist (i.e., welfarist) criterion, namely maximal aggregate utility, that ranks all possible allocations based exclusively on the utility levels of the individuals in society. In contrast, some normative frameworks stress the moral relevance of concerns such as freedom, rights, and rules, rather than the ends emphasized by Utilitarianism. These frameworks are often referred to as deontological, and a long-standing concern in moral philosophy is whether the judgments of consequentialist and deontological frameworks can be compared. See Sen (1982).

deviations of the actual allocation of resources from each criterion's optimal allocation. Of course, specifying these loss functions is a matter of judgment, and some may object to their use altogether. In the end, the appeal of my analysis will depend on how closely the optimal allocations and loss functions I use align with the priorities of the normative criteria. An important feature of this approach is that these loss functions can be specified in a way that respects Pareto efficiency, as the examples below illustrate, avoiding the problem with non-welfarist criteria noted by Kaplow and Shavell (2001).

In other words, one interpretation of this paper's contribution is as providing a basis, in the form of a specific alternative normative criterion, for including in the objective for policy some affinity for a point along the Pareto-efficient frontier that is far from the conventional Utilitarian or Rawlsian points. This interpretation relates to the distinction made earlier between this paper's approach and one in which we allow for any point on the Pareto frontier to be chosen, regardless of whether it is connected to any normative principle. It may be useful to draw an analogy to how Rawls' theory has been reduced, in conventional optimal tax analyses, to a simple maximin objective. While this reductivism no doubt betrays many fundamental aspects of Rawls' framework, focusing on the maximin point along the Pareto-efficient frontier is commonly justified by appeals to his work. Similarly, while Equal Sacrifice as a principle may not correspond perfectly to any point on that frontier, it provides an intuition for considering one point that conventional criteria do not.

Thus, the key formal innovation in this paper's generalization of the standard model is that the social planner minimizes a "social loss function" that is the weighted sum of these criterion-specific losses. The weight on a given criterion's loss represents the force that criterion exerts on society's moral evaluations. The social planner is therefore interpreted as an authority using a diverse normative criterion that is the product of an (unspecified) political process.

This loss-minimization approach to combining disparate normative criteria appears to be consistent with the "consequential evaluation" of Amartya Sen (2000). Sen does not specify how these criteria ought to be combined, but a suggestive passage indicates that my approach of social loss minimization may not be far off the mark: "…rights-inclusive objectives in a system of consequential evaluation can accommodate certain rights the fulfillment of which would be excellent but not guaranteed, and we can still try to minimize the shortfall."

In most other respects, the model economy in this paper is identical to that considered in standard modern optimal tax models. Individuals differ in their innate ability to earn income, denoted w^i for types $i \in \{1, 2, \ldots, I\}$, with the proportion of the population with ability i denoted p^i such that $\sum_{i=1}^{I} p^i = 1$. Individuals derive utility from consumption c and disutility from exerting labor effort y/w to earn income y. Denote the interpersonally-comparable utility function $U(c, y/w)$.

A planner chooses allocations $\{c^i_*, y^i_*\}_{i=1}^I$ to minimize social loss subject to feasibility and incentive compatibility constraints. Formally, the planner's problem is:

Problem 1. Social planner's problem (general case)

$$\min_{\left\{c^i_*, y^i_*\right\}_{i=1}^I \in \{\mathbb{F}\cap\mathbb{IC}\}} \mathcal{L} = \sum_{\phi\in\Phi} \alpha_\phi \mathcal{L}_\phi\left(\left\{c^i_\phi, y^i_\phi\right\}_{i=1}^I, \left\{c^i_*, y^i_*\right\}_{i=1}^I\right), \tag{1}$$

where the criterion-specific loss functions $\mathcal{L}_\phi$ for each criterion ϕ in the set Φ are defined below; and $\mathbb{F}$ denotes the set of feasible allocations for the economy:

$$\mathbb{F} = \left\{\left\{c^i, y^i\right\}_{i=1}^I : \sum_{i=1}^{I} p^i\left(y^i - c^i\right) \geq G\right\}, \tag{2}$$

where G is exogenous, required government spending on public goods; and $\mathbb{IC}$ denotes the set of incentive compatible allocations:

$$\mathbb{IC} = \left\{\left\{c^i, y^i\right\}_{i=1}^I : U\left(c^i, y^i/w^i\right) \geq U\left(c^j, y^j/w^i\right) \text{ for all } i, j \in \{1, 2, \ldots, I\}\right\}. \tag{3}$$

The weights $\{\alpha_\phi\}_{\phi\in\Phi}$ applied to each loss function represent the importance of each normative criterion in society's evaluations of policy. A number of models of the policymaking process could be used to generate such weights, but incorporating a convincing model of the political economy of policymaking is beyond this paper's scope.

The losses to which these weights apply are calculated using two components that, together, capture the priorities of each normative criterion.

First, each criterion generates a preferred, economically-feasible allocation of consumption and income across types, which I label the "ϕ-optimal feasible allocation." To identify these allocations, start by assuming that each normative criterion $\phi \in \Phi$ implies a (possibly incomplete) preference relation $\succeq_\phi$ on the set $\mathbb{F}$, so that we say allocation $\left\{c^i_1, y^i_1\right\}_{i=1}^I \in \mathbb{F}$ is weakly preferred under criterion ϕ to allocation $\left\{c^i_2, y^i_2\right\}_{i=1}^I \in \mathbb{F}$ if

$$\left\{c^i_1, y^i_1\right\}_{i=1}^I \succeq_\phi \left\{c^i_2, y^i_2\right\}_{i=1}^I.$$

Given $\succeq_\phi$, the strict preference relation $\succ_\phi$ is defined as usual. For any $\left\{c^i_1, y^i_1\right\}_{i=1}^I, \left\{c^i_2, y^i_2\right\}_{i=1}^I \in \mathbb{F}$,

$$\left\{c^i_1, y^i_1\right\}_{i=1}^I \succ_\phi \left\{c^i_2, y^i_2\right\}_{i=1}^I \Leftrightarrow \left\{c^i_1, y^i_1\right\}_{i=1}^I \succeq_\phi \left\{c^i_2, y^i_2\right\}_{i=1}^I \text{ but not } \left\{c^i_2, y^i_2\right\}_{i=1}^I \succeq_\phi \left\{c^i_1, y^i_1\right\}_{i=1}^I.$$

These preference relations allow the identification of the ϕ-optimal feasible allocations, which I denote $\{c^i_\phi, y^i_\phi\}_{i=1}^I$, and formally define as follows.

Definition 1. An ϕ-optimal feasible allocation $\{c^i_\phi, y^i_\phi\}_{i=1}^I$ is any allocation in the set $\mathbb{F}$ for which there is no other allocation $\{c^i, y^i\}_{i=1}^I$ in the set $\mathbb{F}$ such that: $\{c^i, y^i\}_{i=1}^I \succ_\phi \{c^i_\phi, y^i_\phi\}_{i=1}^I$.

These ϕ-optimal feasible allocations provide a key link across normative criteria. Note that no incentive compatibility constraints are imposed when defining the ϕ-optimal feasible allocations, so that they equal each criterion's "first-best" allocation in this context (i.e., when ability is observable).

Second, each criterion's priorities are represented by a loss function that measures the costs of deviations from the criterion's most preferred allocation. I denote these loss functions $\mathcal{L}_\phi(\{c^i_\phi, y^i_\phi\}_{i=1}^I, \{c^i_*, y^i_*\}_{i=1}^I)$

The loss functions $\{\mathcal{L}_\phi\}_{\phi\in\Phi}$ that I use in this paper satisfy the following three conditions. The first two are straightforward. The third, Pareto Efficiency, may be more controversial among political philosophers but is generally viewed as a reasonable requirement in the optimal taxation literature.

Remark 1. For all $\phi \in \Phi$, the loss function $\mathcal{L}_\phi(x, y)$ satisfies:

1. Ordinality: For any $\left\{c^i_1, y^i_1\right\}_{i=1}^I, \left\{c^i_2, y^i_2\right\}_{i=1}^I \in \mathbb{F}$,

$$\mathcal{L}_\phi\left(\left\{c^i_\phi, y^i_\phi\right\}_{i=1}^I, \left\{c^i_1, y^i_1\right\}_{i=1}^I\right) \leq \mathcal{L}_\phi\left(\left\{c^i_\phi, y^i_\phi\right\}_{i=1}^I, \left\{c^i_2, y^i_2\right\}_{i=1}^I\right) \Leftrightarrow \left\{c^i_1, y^i_1\right\}_{i=1}^I \succeq_\phi \left\{c^i_2, y^i_2\right\}_{i=1}^I,$$

so that the loss from one allocation is no greater than that from another to which it is weakly preferred under criterion ϕ;

2. Normalization: $\mathcal{L}_\phi(\{c^i_\phi, y^i_\phi\}^I_{i=1}, \{c^i_\phi, y^i_\phi\}^I_{i=1}) = 0$, so that the loss is zero when the equilibrium allocation equals the ϕ-optimal feasible allocation.
3. Weak Pareto efficiency:

$$U\left(c^i_1, y^i_1/w^i\right) \geq U\left(c^i_2, y^i_2/w^i\right) \text{ for all } i \in \{1,2,\ldots I\} \Rightarrow \mathcal{L}_\phi\left(\left\{c^i_\phi, y^i_\phi\right\}^I_{i=1}, \left\{c^i_1, y^i_1\right\}^I_{i=1}\right) \leq \mathcal{L}_\phi\left(\left\{c^i_\phi, y^i_\phi\right\}^I_{i=1}, \left\{c^i_2, y^i_2\right\}^I_{i=1}\right), \tag{4}$$

which can be converted into Strong Pareto Efficiency if desired.[25]

In words, Weak Pareto Efficiency as defined here says that if all individuals do at least as well under allocation 1 as they do under allocation 2, the loss from allocation 1 cannot be greater than the loss from allocation 2. This condition will prevent the planner from rejecting Pareto-improving allocations. It is too weak, however, to guarantee that the planner will avoid Pareto-inefficient allocations—for that, Strong Pareto Efficiency is required.

Below, I apply this general approach to the case of the two main criteria between which I have respondents to the survey choose: the conventional Utilitarian criterion and the principle of Equal Sacrifice.

3.2. Equal Sacrifice as an alternative to Utilitarianism

First, I provide a brief discussion of why Equal Sacrifice is a natural choice as an alternative to Utilitarianism.[26] John Stuart Mill (1871) was the most famous proponent of Equal Sacrifice, and his argument for it is worth quoting at length.

> "For what reason ought equality to be the rule in matters of taxation? For the reason, that it ought to be so in all affairs of government…Equality of taxation, therefore, as a maxim of politics, means equality of sacrifice. It means apportioning the contribution of each person towards the expenses of government so that he shall feel neither more nor less inconvenience from his share of the payment than every other person experiences from his."

To Mill (1871), the appeal of Equal Sacrifice was simple: it treats all individuals equally. This argument for Equal Sacrifice was endorsed by other influential thinkers, including Alfred Marshall and Henry Sidgwick, the latter of whom claimed it was the obviously equitable principle—assuming that the existing distribution of wealth is accepted as just or not unjust.[27]

Utilitarianism, in contrast, is willing to trade the losses of some for greater gains of others, a willingness that thinkers as diverse as John Rawls and Robert Nozick have seen as a serious failing. The specific context in which this concern has been seen as most forceful is "endowment" taxation, where individuals would be taxed on their potential to earn income rather than their actual earned income. Of course, endowment taxation is exactly the preferred policy of the conventional Utilitarian optimal tax model.[28] While Rawls and Nozick take from their critiques very different lessons, they share a similar target: Utilitarianism's potential to violate individual liberty due to its acceptance of unequal treatment. This critique of Utilitarianism makes clear why Mill's Equal Sacrifice, with its emphasis on equal treatment of all individuals, is a natural alternative normative criterion.

While a priority on equal treatment may be of paramount concern to only a small minority of individuals, this paper's evidence and that cited in Section 2.4 suggests that it has at least some appeal to most. For many, that appeal is linked to an affinity for a non-welfarist normative perspective such as Libertarianism.[29] As Feldstein (1976) noted prior to linking Nozick's logic to Equal Sacrifice: "Those who are fully persuaded by Nozick will thus completely redefine the problem of optimal taxation. Others will reject Nozick completely…Many will be persuaded that the entitlement principle limits the desirable degree of redistribution."

Mill himself provides a telling example of exactly this form of mixed normative reasoning, writing approvingly of both Equal Sacrifice and minimal total sacrifice (which is similar to the Utilitarian criterion):

> "As a government ought to make no distinction of persons or classes in the strength of their claims on it, whatever sacrifices it requires from them should be made to bear as nearly as possible with the same pressure upon all, which, it must be observed, is the mode by which least sacrifice is occasioned on the whole."

Mill is incorrect, as many others have noted, in the assertion that Equal Sacrifice implies minimized total sacrifice. But this mistake reveals that, for Mill, both equal and minimized total sacrifice were principles he believed appealing and likely to be accepted by his readers. This paper is built on the idea that Mill's split normative intuition is more the rule than the exception.[30]

3.3. A two-criterion case: Utilitarianism and Equal Sacrifice

In this section, I apply the general approach from above to the case of the two main criteria used in the survey of Section 1.

3.3.1. ϕ-optimal feasible allocations

The first step in this application is to define the preference relations that determine the ϕ-optimal feasible allocations. The preference relation for Utilitarianism is familiar from the conventional optimal tax literature: allocations are preferred that generate a greater sum of individual utilities. Formally, $\succeq_{Util}$ is defined by:

$$\left\{c^i_1, y^i_1\right\}^I_{i=1} \succeq_{Util} \left\{c^i_2, y^i_2\right\}^I_{i=1} \Leftrightarrow \sum_{i=1}^I p^i U\left(c^i_1, y^i_1/w^i\right) \geq \sum_{i=1}^I p^i U\left(c^i_2, y^i_2/w^i\right). \tag{5}$$

The Utilitarian-optimal feasible allocation is therefore:

$$\left\{c^i_{Util}, y^i_{Util}\right\}^I_{i=1} \in \mathbb{F} : \sum_{i=1}^I p^i U\left(c^i_{Util}, y^i_{Util}/w^i\right) \geq \sum_{i=1}^I p^i U\left(c^i, y^i/w^i\right),$$

for all possible $\left\{c^i, y^i\right\}^I_{i=1} \in \mathbb{F}$.

The preference relation for the principle of Equal Sacrifice requires more discussion. The key question is from what starting point is each individual's sacrifice to be calculated? Though one could defend a number of choices for that starting point, one natural option is the allocation that would obtain absent any government intervention, i.e., the no-tax

[25] Namely, $U(c^i_1, y^i_1/w^i) \geq U(c^i_2, y^i_2/w^i)$ for all $i \in \{1, 2, \ldots, I\}$ and $U\left(c^{i'}_1, y^{i'}_1/w^{i'}\right) > U\left(c^{i'}_2, y^{i'}_2/w^{i'}\right)$ for some $i' \in \{1, 2, \ldots, I\} \Rightarrow \mathcal{L}_\phi(\{c^i_\phi, y^i_\phi\}^I_{i=1}, \{c^i_1, y^i_1\}^I_{i=1}) < \mathcal{L}_\phi(\{c^i_\phi, y^i_\phi\}^I_{i=1}, \{c^i_2, y^i_2\}^I_{i=1})$.

[26] Weinzierl (2014) includes an expanded version of this discussion.

[27] In addition to the work of H. Peyton Young and Berliant and Gouveia mentioned earlier, Yaari (1988), Ok (1995), Mitra and Ok (1996), D'Antoni (1999), and Moyes (2003) helped establish conditions on the progressivity of taxes designed in accordance with Equal Sacrifice and argue for the centrality of that principle. Lambert and Naughton (2009) is a recent contribution that reviews much of this literature.

[28] Legal scholars have extensively analyzed this issue with endowment (ability) taxation under the heading of "talent slavery," the heavy taxation of those with high ability that forces them to work exceptionally hard or at an occupation they dislike. See, for instance, Hasen (2007), Markovits (2003), Rakowski (2000), Shaviro (2002), Stark (2005), Sugin (2011), and Zelenak (2006).

[29] Murphy and Nagel (2002) have argued: "If (and only if) [libertarianism] is the theory of distributive justice we accept, the principle of equal sacrifice does make sense."

[30] Mill (1871) also wrote: "An income not exceeding 50 l. should not be taxed at all, either directly or by taxes on necessaries," again illustrating how his affinity for Equal Sacrifice was tempered by Utilitarian (or even Rawlsian) intuitions as well. I thank a referee for bringing this example to my attention.

allocation. In particular, the allocation with no taxation is the preferred allocation of the Libertarian framework with which the principle of equal sacrifice has been linked. As Murphy and Nagel (2002) have argued: "The implication for tax policy of rights-based libertarianism in its pure or absolute form is that no compulsory taxation is legitimate…" For clarity, I will refer to the allocation with no taxation as the *laissez-faire allocation* and formally define it as follows.

Definition 2. The laissez-faire allocation, $\left\{c_{lf}^i, y_{lf}^i\right\}_{i=1}^I \in \mathbb{F}$, where $G = 0$, satisfies the following conditions (where $U_x(c, y/w)$ denotes the partial derivative of individual utility with respect to x):

1. $U_{c_{lf}^i}\left(c_{lf}^i, y_{lf}^i/w^i\right) = U_{y_{lf}^i}\left(c_{lf}^i, y_{lf}^i/w^i\right)/w^i$
2. $c_{lf}^i = y_{lf}^i$.

These conditions are simply that each individual maximizes utility and there are no interpersonal transfers. In the statement of the definition, I clarify that $G = 0$, as this is the allocation with no government.

A well-known conceptual issue with the idea of the laissez-faire allocation is that any economy is, in reality, inseparable from the government and state institutions that taxes fund. The laissez-faire allocation is, therefore, not well-defined, because $G = 0$ implies a very different economy than that the status quo. Without a well-defined starting point, calculating "sacrifice" is impossible. In formal terms, if $G > 0$ is required for the status quo economy to function, the laissez-faire allocation is not in the feasible set $\mathbb{F}$.

Fortunately, though I am not aware of this being recognized before, the Equal Sacrifice principle provides a natural way to convert the infeasible hypothetical laissez-faire allocation into a feasible one. Consider the following thought experiment. Suppose that the public goods necessary to support the current economy are sustained without any cost to the economy, so that $G = 0$ but the status quo economic system is feasible. According to Equal Sacrifice, the (no tax) laissez-faire outcome in this scenario is surely optimal, as it satisfies Equal Sacrifice with the smallest possible uniform sacrifice – that is, zero – for all individuals. Now, suppose that sustaining those public goods is costly, so that $G > 0$. The Equal Sacrifice principle implies that the cost of the public goods will be distributed across individuals such that the utility loss is identical (and as small as possible) for all.

Formally, define $\mathbb{ES}$ as the set of all feasible allocations that satisfy the principle of Equal Sacrifice relative to the laissez-faire allocation:

$$\mathbb{ES} = \left\{ \left\{c^i, y^i\right\}_{i=1}^I \in \mathbb{F} : U\left(c_{lf}^i, y_{lf}^i/w^i\right) - U\left(c^i, y^i/w^i\right) = U\left(c_{lf}^j, y_{lf}^j/w^j\right) - U\left(c^j, y^j/w^j\right) \text{ for all } i, j \in \{1, 2, \ldots, I\} \right\}. \quad (6)$$

The Equal Sacrifice preference relation, denoted $\succeq_{ES}$, indicates that one allocation in $\mathbb{ES}$ is preferred to another if it generates a smaller uniform sacrifice:

$$\left\{c_1^i, y_1^i\right\}_{i=1}^I \succeq_{ES} \left\{c_2^i, y_2^i\right\}_{i=1}^I \Leftrightarrow U\left(c_{lf}^i, y_{lf}^i/w^i\right) - U\left(c_1^i, y_1^i/w^i\right) \leq U\left(c_{lf}^i, y_{lf}^i/w^i\right) - U\left(c_2^i, y_2^i/w^i\right), \quad (7)$$

for $\left\{c_1^i, y_1^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I \in \mathbb{ES}$ and for any $i \in \{1, 2, \ldots, I\}$.

Consequently, the Equal Sacrifice-optimal feasible allocation is that which achieves the smallest equal sacrifice while funding G. Formally, we define $\{c_{ES}^i, y_{ES}^i\}_{i=1}^I$ as follows:

$$\left\{c_{ES}^i, y_{ES}^i\right\}_{i=1}^I \in \mathbb{ES} : U\left(c_{lf}^i, y_{lf}^i/w^i\right) - U\left(c_{ES}^i, y_{ES}^i/w^i\right) \leq U\left(c_{lf}^i, y_{lf}^i/w^i\right) - U\left(c^i, y^i/w^i\right),$$

for any $i \in \{1, 2, \ldots, I\}$ and for all possible $\left\{c^i, y^i\right\}_{i=1}^I \in \mathbb{ES}$.

Once we have specified the ϕ-optimal feasible allocations, the next step is to specify the loss functions for the planner.

3.3.2. Loss functions

The Utilitarian loss function $\mathcal{L}_{Util}$ is:

$$\mathcal{L}_{Util}\left(\left\{c_{Util}^i, y_{Util}^i\right\}_i, \left\{c_*^i, y_*^i\right\}_i\right) = \sum_{i=1}^I p^i \left[U\left(c_{Util}^i, y_{Util}^i/w^i\right) - U\left(c_*^i, y_*^i/w^i\right)\right]. \quad (8)$$

In words, it is the sum of individuals' utility losses from having the equilibrium allocation $\{c_*^i, y_*^i\}_i$ deviate from the Utilitarian-optimal feasible allocation. This loss function has the appealing property that it directly adopts the cardinal welfare comparisons underlying the Utilitarian preference relation and, thus, the conventional optimal tax model.[31] Note that it converts the familiar goal of aggregate utility maximization into aggregate sacrifice minimization.

Unlike Utilitarianism, the Equal Sacrifice criterion does not rank allocations that deviate from its preferred allocation. As far as I am aware, no previous work has studied how to obtain a complete ranking of allocations based on Equal Sacrifice. While my approach is, therefore, by necessity somewhat speculative, I design the Equal Sacrifice loss function to reflect the priorities of that principle. In words, these priorities are simple: deviations from equal sacrifice are costly, even if they reduce the aggregate level of sacrifice, and outcomes with less sacrifice for some and no more for all are preferred (i.e., Pareto efficiency). Of course, future research may discover alternative specifications that prove more useful. The goal of this paper is to propose one reasonable way, not the definitive way, to capture the priorities of the Equal Sacrifice principle.

I will assume an Equal Sacrifice loss function $\mathcal{L}_{ES}$ with three features: first, deviations of individual utility below the Equal Sacrifice-optimal feasible allocation are costly but deviations above the Equal Sacrifice-optimal feasible allocation yield little or no offsetting benefits[32]; second, losses increase more than proportionally with the size of the deviation of individual utility below the Equal Sacrifice-optimal feasible allocation; third, gains are concave in the size of the deviation of individual utility above the Equal Sacrifice-optimal feasible allocation.

I formalize these properties as follows:

$$\mathcal{L}_{ES}\left(\left\{c_{ES}^i, y_{ES}^i\right\}_i, \left\{c_*^i, y_*^i\right\}_i\right) = \sum_{i=1}^I p^i V\left(U\left(c_{ES}^i, y_{ES}^i/w^i\right), U\left(c_*^i, y_*^i/w^i\right)\right), \quad (9)$$

[31] An alternative approach would be to use a common loss function for all criteria. While this has the seeming advantage of consistency, it in fact would lead to pathologies. For example, if full weight was put on the Utilitarian criterion, but the loss function used was not the same as expression (8), the model would yield a different ranking of policies than the conventional Utilitarian model.

[32] This property is consistent with the classic "loss aversion" of Kahneman and Tversky (1979). However, equal sacrifice is not consistent with the diminishing sensitivity to losses that is part of classic prospect theory.

where

$$V\left(U^i_{ES}, U^i_*\right) = \begin{cases} -\left[\delta\left(U^i_* - U^i_{ES}\right)\right]^\theta & \text{if } U^i_{ES} < U^i_* \\ \left[\lambda\left(U^i_{ES} - U^i_*\right)\right]^\rho & \text{if } U^i_{ES} \geq U^i_* \end{cases}, \tag{10}$$

for scalars $\{\delta \geq 0,\ \lambda > \delta,\ \theta \in (0,1],\ \rho > 1\}$.

Consistent with the first property, the loss function in expressions (9) and (10) applies weights δ and λ, where $0 \leq \delta < \lambda$, to deviations of individual utility above and below the Equal Sacrifice-optimal feasible allocation.

The kink at the Equal Sacrifice-optimal feasible allocation implied by $\delta < \lambda$, and thus the asymmetric punishment of downward deviations from that allocation, rejects the Utilitarian idea that the distribution of utility across individuals is irrelevant. Though nondifferentiability is technically inconvenient, it is conceptually important to capture the Equal Sacrifice criterion's priorities. The reason is that the alternative – a smooth loss function at the Equal Sacrifice-optimal feasible allocation – implies local indifference to symmetric deviations from that allocation. That indifference is in direct conflict with the principle's priority on equal sacrifice.

The assumption that $\delta \geq 0$ respects Weak Pareto Efficiency as discussed above ($\delta > 0$ would respect Strong Pareto Efficiency). Consistent with the second and third properties, the parameters $\rho > 1$ and $\theta \in (0, 1]$ imply losses that increase more than proportionally with deviations below and gains that increase (weakly) less than proportionally for deviations above the Equal Sacrifice-optimal feasible allocation. As noted above, I do not mean to claim that these functional forms or these assumptions on its parameter values are the only possible choices, but rather to construct a plausible representation of the Equal Sacrifice principle for analysis.

3.3.3. Planner's problem

With the loss functions defined by expressions (8), (9) and (10), the planner in this case chooses $\{c^i_*, y^i_*\}^I_{i=1}$ to solve the following problem.

Problem 2. Social planner's problem (specific case)

$$\min_{\left\{c^i_*, y^i_*\right\}^I_{i=1} \in \{\mathbb{F} \cap \mathbb{IC}\}} \left\{ \begin{array}{l} \alpha_{Util} \sum_{i=1}^{I} p^i \left[U\left(c^i_{Util}, y^i_{Util}/w^i\right) - U\left(c^i_*, y^i_*/w^i\right)\right] \\ \quad + \alpha_{ES} \sum_{i=1}^{I} p^i V\left(U\left(c^i_{ES}, y^i_{ES}/w^i\right), U\left(c^i_*, y^i_*/w^i\right)\right) \end{array} \right\}, \tag{11}$$

where

$$\alpha_{Util} + \alpha_{ES} = 1,$$

$V(\cdot)$ is defined in (10), $\mathbb{F}$ is defined in (2), and $\mathbb{IC}$ is defined in (3).

This planner's problem is equivalent to the conventional approach if $\alpha_{ES} = 0$.

To illustrate the effect of positive α_{ES} on optimal policy, I simulate a simple model with two types of workers and show how this form of normative diversity affects the well-being of individuals in the economy.

3.3.4. Example with two types

Individual income-earning ability is either $w^1 = 10$ or $w^2 = 50$, each of which makes up half the population, so $p^1 = p^2 = 0.5$. The individual utility function is

$$U\left(c^i, y^i/w^i\right) = \frac{\left(c^i\right)^{1-\gamma} - 1}{1-\gamma} - \frac{1}{\sigma}\left(\frac{y^i}{w^i}\right)^\sigma,$$

where $\gamma = 1.5, \sigma = 3$. The Equal Sacrifice loss function's parameters are $\delta = 0.5, \lambda = 20, \rho = 2.0, \theta = 1.0$, and the social loss function's weight on the Equal Sacrifice loss function is $\alpha_{ES} = 0.20$. Government spending G is set to zero.

This simple example is most useful for showing the effect of such a mixed objective on the allocation of utility across individuals. Fig. 4 plots the utility of the high-ability individual against that of the low-ability individual. The bold solid line shows the utility possibilities frontier (UPF): that is, the highest incentive-compatible, feasible utility for the low-ability individual given a utility level for the high-ability individual. The thin solid and dotted lines are the indifference curves passing through the ϕ-optimal feasible (but not necessarily incentive compatible) allocations for the Utilitarian and Equal Sacrifice criteria. The dashed line is the indifference curve for the planner that chooses (by tangency with the UPF) the optimal allocation for the economy. Also shown are the optimal feasible and incentive-compatible allocations chosen by each criterion.

Fig. 4 shows how the Equal Sacrifice loss function, $\mathcal{L}_{ES}$, differs from the Utilitarian, $\mathcal{L}_{Util}$. To remain indifferent while moving away from its optimal allocation, $\mathcal{L}_{ES}$ requires a greater gain for the low-ability individual in exchange for a given loss for the high-ability individual. Moreover, $\mathcal{L}_{ES}$ increases more than proportionally with these deviations, while $\mathcal{L}_{Util}$ is linear. The impact of incorporating this loss function in the planner's decisions is as expected: the planner compromises between the competing normative criteria, implementing some redistribution but stopping well short of what a Utilitarian would choose. By varying α_{Util}, we can shift the planner's chosen allocation along the UPF.

3.4. Generating the survey's policy options from the model

The set of policy options presented to respondents in the survey of Section 1 was generated using this section's generalized optimal tax model. Here, I describe the calibration of the model to data on the U.S. income distribution and the parameterizations of the model that generate those policies.

For each policy objective I simulate a constrained planner's problem as in expression (11), calibrated to data on the U.S. income distribution from 2006 as calculated by the Congressional Budget Office. In particular, I take the gross labor income distribution as calculated (by the CBO) into eight bins: the bottom four quintiles and the next 10, 5, 4, and 1 percentiles. The CBO also provides taxes paid

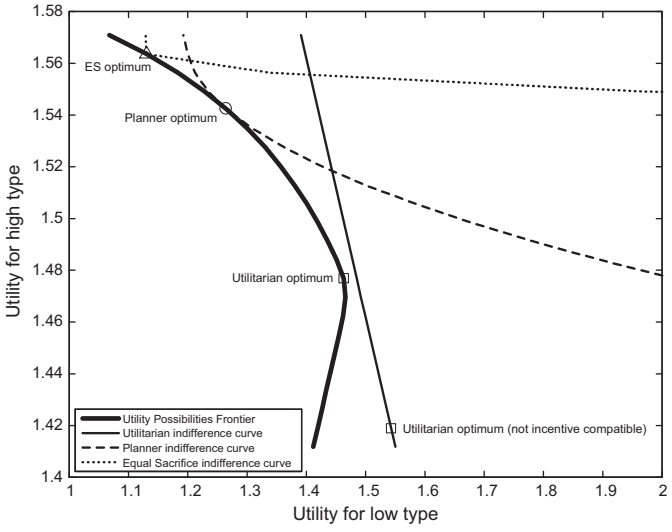


Fig. 4. The utility possibilities frontier and indifference curves in the two-type example.

for these households,[33] so I use the utility function specified below to back out the earnings ability implied by the households' pre-tax earnings and tax payments. Then, I calculate the earnings each household would choose if there were no taxation, again using the individual utility function defined below. This calculation yields the distribution presented to respondents as the baseline "no tax" income distribution. The distribution of ability for the model, where w^i denotes the ability and p^i denotes the population proportion of type i, is as follows.

Ability distribution								
w^i	3.10	12.16	21.98	35.13	49.56	63.63	88.59	238.08
p^i	0.20	0.20	0.20	0.20	0.10	0.05	0.04	0.01

All model parameters other than α_{ES} and α_{Util} retain the same values across simulations. These parameters, and the underlying formal structure of the problem, are never disclosed to respondents. I assume the following parameter values.

Parameter values						
ρ	θ	δ	λ	$\frac{1}{\sigma-1}$	φ	G
2.0	1.0	0	10	0.10	$12^{-\sigma}$	8.26

The utility function for all households is

$$U\left(c^i, y^i/w^i\right) = \ln\left(c^i\right) - \frac{\varphi}{\sigma}\left(\frac{y^i}{w^i}\right)^{\sigma}. \tag{12}$$

The values of ρ, θ, δ, and λ determine the shape of the Equal Sacrifice loss function. The parameter σ controls the elasticity of labor supply, while φ is a taste shifter used only to normalize labor effort. The value of G is chosen so that government expenditure as a share of equilibrium total output roughly matches that in the United States. Several of these parameter values deserve additional comment.

First, assuming $\delta = 0$ implies that deviations of individual utility above the Equal Sacrifice-optimal feasible allocation generate no gains according to the Equal Sacrifice criterion. This is the strictest version of the Equal Sacrifice loss function, in that it rejects redistribution even if it generates enormous gains for some as long as it generates any losses for others. To the extent that respondents are, in reality, sympathetic to a more moderate version of Equal Sacrifice, this assumption biases the survey toward support for more Utilitarian policies.

Second, the implied Frisch labor supply elasticity $\frac{1}{\sigma-1}$ is low in these parameterizations, at 0.10. That is below most mainstream estimates, though not for prime-aged heads of households. Lower labor supply elasticities will reduce the efficiency costs of redistributive policies, increasing their appeal. Therefore, our survey results are likely to be biased toward Utilitarianism due to this choice.

Third, assuming logarithmic utility of consumption has two implications. If it underestimates the concavity of that subutility function, the simulations generate policies with less income redistribution than what a more realistic calibration would produce. As with the other assumptions above, this bias would tend to increase the reported support for Utilitarianism in the survey, as the survey results show that most people prefer less redistributive policies than the purely Utilitarian one. Log utility of consumption also means that average tax rates are flat under the Equal Sacrifice criterion, and a flat tax may have some appeal to respondents due to its simplicity. Working against this, if we were to use a more concave form of utility, Equal Sacrifice would yield progressive average taxes, making it appear closer to the most-preferred policies in the survey.[34]

Related to the choices of both $\frac{1}{\sigma-1}$ and the concavity of the utility of consumption (which I will later parameterize with the coefficient of relative risk aversion γ) is the question of whether the Utilitarian and Equal Sacrifice criteria become quantitatively indistinguishable as redistribution becomes more costly ($\frac{1}{\sigma-1}$ increases) and less rewarding (γ decreases). It turns out that the answer is "no." The clearest example is that Utilitarianism endorses negative average tax rates at low incomes, while Equal Sacrifice does not because, by definition, Equal Sacrifice requires all individuals to bear some sacrifice. More generally, the two policies remain quite distinct even for unconventionally high values of $\frac{1}{\sigma-1}$ and low values of γ.[35]

Finally, the values of ρ, θ, and λ are necessarily chosen without any direct empirical guidance. I do not mean to suggest, therefore, that these are necessarily the correct values, or that the results of this paper are robust to their choice. They are chosen to represent a plausible specification of the Equal Sacrifice criterion, and the (indirect) test of their appropriateness is that they contribute to the explanatory power of this paper's model for real-world policy.

4. Descriptive power of the positive optimal tax model

In this brief section I illustrate how the optimal tax model, as proposed and empirically estimated in this paper, is able to explain aspects of existing policy that are difficult to reconcile in conventional theory but widely endorsed in reality. I focus especially on the puzzle of limited tagging—the taxation of personal characteristics in addition to income introduced to the literature in Akerlof (1978). I also briefly discuss the implications of Equal Sacrifice for the topic of rank-reversals in the case of optimal policy with full information.[36]

In the modern theory of optimal taxation, tagging is a free lunch, and a wide variety of candidate tags exist. Any observable and largely inelastic characteristic across which the distribution of abilities differs ought to affect tax schedules.[37] In comparison, the role for tagging in modern tax policy is highly constrained. Some sizeable tagging does occur, but only for tags that are virtually guaranteed to indicate that a taxpayer has low income-earning ability.[38] In this section, I use numerical simulations calibrated to the U.S. microdata to show that the combinations of Utilitarian and Equal Sacrifice principles preferred by the survey respondents can quantitatively match the simultaneous rejection of most, but not all, forms of tagging and the acceptance of substantial income redistribution in U.S. policy.

The importance of the tagging application to this paper is that it suggests Equal Sacrifice has explanatory power beyond simply being a "less redistributive" criterion than Utilitarianism. I show that, while the optimal extent of both tagging and redistribution decrease when Equal Sacrifice is given more weight, tagging is disproportionately discouraged. Intuitively, Equal Sacrifice rejects tagging because it causes differential sacrifice across people of the same underlying ability, not because it causes too much sacrifice by those with high ability. In an

[33] The CBO provides income taxes, which are due to both labor and capital income tax payments. For simplicity, I multiply total personal income taxes paid by the labor share of income for each type of household to generate labor income taxes. Variations on this approach yield very similar results in terms of the distribution of abilities.

[34] As noted by Berliant and Gouveia (1993), among others, Equal Sacrifice endorses progressivity if, in the notation of this paper, utility is separable across consumption and leisure and $\gamma > 1$. Mill's (1871) writings suggest that he thought logarithmic utility was a natural specification.

[35] See Weinzierl (2014) for numerical results on this point.

[36] Detail on the applications in this section, including analytical results on the optimal extent of tagging under Equal Sacrifice, can be found in Weinzierl (2014).

[37] This statement assumes that there are no differences in preferences or elasticities across tagged groups, which may reinforce or weaken the case for a given tag.

[38] Aside from blindness (analyzed below), disability and the presence of children are the dominant tags used (age is not properly thought of as a tag, as all individuals pass through all ages, subject to mortality differences—see Weinzierl, 2011). To the extent that disability status implies zero earning ability, it by definition merits tagging. Future work could usefully focus on showing whether the model can explain the substantial tagging on dependent children in existing policy. That task will require making judgments on the proper modeling and normative treatment of households.

optimal policy that gives weight to both Equal Sacrifice and Utilitarianism, only those tags that provide sufficiently strong information about ability, and therefore Utilitarian welfare gains, will be optimal.

4.1. Numerical results on optimal tagging and redistribution

First, I consider three prominent potential tags – height, gender, and race – that are not used in reality.[39] Then, I consider blindness, one of the few personal characteristics explicitly tagged in the U.S. tax code.

For the optimal tax policy simulations in this section, I use the following parameter values:

Parameter values							
α_{ES}	ρ	θ	δ	λ	γ	$\frac{1}{\sigma-1}$	G
$\{0.00,\ 0.0\overline{3},\ 0.10,\ 0.20,\ 1.00\}$	2.0	1.0	0.01	10	1.5	0.5	20

A few of these values differ from those used to generate the policy options for the survey. While I chose values for the survey of δ, γ, and σ to increase the appeal of the conventional Utilitarian policy, here I choose values to maximize realism.[40]

Data on ability distributions by height, gender, and race come from the National Longitudinal Survey of Youth. Mean wages for each tagged group are shown in Table 4. For each of the five values of α_{ES}, Table 4 also reports measures of the optimal extent of tagging: the "extra" average tax paid by or transfer made to the members of each tagged group as a share of their income when the planner can use tagging as compared to when it cannot.

To gauge the progressivity of the optimal income tax under each value of α_{ES}, Table 5 reports the average tax rate paid by the members of each wage range.

The results in these tables show that the support expressed in the survey of Section 1 for objectives that include Equal Sacrifice can yield an optimal tax policy that rejects most forms of tagging but pursues substantial income redistribution, as in reality. Note, in particular, the case of $\alpha_{ES} = 0.20$, the most-preferred policy in the survey. In that case, tagging is negligible but average tax rates (and top-income marginal tax rates, not shown) closely mimic those in current U.S. policy.[41]

The U.S. tax code includes a special deduction or exemption for individuals with substantially impaired vision. To analyze optimal blindness tagging, I use data from the Statistics of Income microdata of the U.S. Internal Revenue Service. Those claiming the blindness exemption make up 0.3% of the population, 79% of whom report zero income.

Table 6 shows the optimal extent of tagging in the conventional calibration with $\alpha_{ES} = 0$ and in the most-favored calibration in the survey of Section 1, with $\alpha_{ES} = 0.20$. All other parameters are as before (though G is adjusted to be a similar share of total income).

As with height, gender, and race, Table 6 shows that adding this weight on Equal Sacrifice to the objective function substantially reduces the optimal extent of tagging on blindness. Unlike those other tags, however, the optimal extent of tagging on blindness in the Utilitarian benchmark is so great that even the dramatically reduced extent of optimal tagging is sizeable—namely, a 20 percent transfer to the blind on average. Using the data from the IRS, we can calculate mean income for the blind (including those with zero income) to be approximately $2350 per year. A 16 percent transfer to the blind on average is therefore equivalent to approximately $376, not far from the value of actual blindness deductions and exemptions in the mid-1980s.

4.2. Equal sacrifice, horizontal equity, and vertical equity

The popularity of the principle of horizontal equity – that "equals ought to be treated equally" – has long been used to explain the limited use of tagging. Boadway and Pestieau (2006) write: "Of course, such a system may be resisted because, if the tagging characteristic has no direct utility consequences, a differentiated tax system violates the principle of horizontal equity". Similar statements are made by, e.g., Atkinson and Stiglitz (1980), Auerbach and Hassett (1999), and King (1983).

Importantly, the analysis of this paper is consistent with these arguments stressing horizontal equity. Horizontal equity is implied by Equal Sacrifice, as Equal Sacrifice values both it and "vertical" equity as part of its overall priority on equal treatment. Thus, one possibility raised by this paper is that the popular enthusiasm for horizontal equity comes out of an underlying affinity for Equal Sacrifice. This possibility is especially interesting in light of the critiques, for example by Musgrave (1959) and Kaplow (2008), of the intellectual coherence of horizontal equity.

In fact, when respondents to the survey of Section 1 were given the logic for height and blindness tags and asked whether they supported or opposed them, the individuals who oppose tagging disproportionately support Equal Sacrifice. In Weinzierl (2014), I show that this result is statistically significant and holds when we control for gender, age, education, race, and income status when a child and adult.

Horizontal equity, however, has no implications for the "vertical" distribution of income, while the ability of Equal Sacrifice to imply empirically-realistic levels of redistribution along with limited tagging bolsters the case for its use. A particularly stark version of this argument relates to the issue of so-called rank reversals. It has been known since the analyses in Mirrlees (1971) and Mirrlees (1974) that an optimal Utilitarian tax policy in the case of full information generally induces a negative relationship between innate ability and the allocation of utility across individuals. This reversal of pre-tax and post-tax utility orderings has generated considerable discomfort among optimal tax theorists (Saez and Stantcheva, 2014) and tax law scholars (see, for example, King, 1983; Zelenak, 2006). Horizontal equity raises no concerns about rank reversals. In contrast, simulations following the same process as above for the U.S. income distribution show that mixed objective most preferred in the survey generates a nearly uniform utility distribution in the first-best, substantially limiting rank reversals. Intuitively, Equal Sacrifice leaves the utility ordering of agents unchanged, even in the case of full information, because its reluctance to redistribute does not depend on the distortionary costs of taxation.

5. Conclusion

The optimal tax literature occupies a rare place in economic research in which the normative assumptions of economists are given priority. The conventional use of Utilitarianism as the criterion for quantitative analyses of optimal policy is expedient, as it narrows the range of models to consider. It may also be compelling, if we believe the Utilitarian criterion is the right one.

An alternative to the conventional approach is to use empirical evidence on normative preferences to develop a *positive* optimal tax theory in which economists' normative intuitions are replaced by those that hold sway among voters and taxpayers in reality. Of course, a number of classic questions arise about such an approach, such as: whose preferences matter for policymaking, how are individual preferences aggregated, and what are the admissable normative criteria. This paper has not focused on these questions, which are important topics for future work. The conventional approach sidesteps these questions, but at the potential cost of relevance.

[39] See Mankiw and Weinzierl (2009); Alesina, Ichino, and Karabarbounis (2011) and Cremer, Gahvari, and Lozachmeur (2010); and Blumkin, Margalioth, and Sadka (2009)

[40] I set $\delta > 0$ (rather than $\delta = 0$) to capture a less strict version of Equal Sacrifice. Utility from consumption is $\frac{1}{1-\gamma}(c^{1-\gamma}-1)$, where I set $\gamma > 1$ (rather than $\gamma = 1$, log utility). I set $\sigma = 3$ (rather than $\sigma = 11$) to be closer to mainstream estimates. I set G to the share of the U.S. government expenditure in total income.

[41] See Weinzierl (2014) for more detail.

Table 4
Extent of Tagging: extra tax (+) or transfer (−) rate, in percent.

α_{ES}	*1* Tall M White	2 Med. M White	3 Short M White	4 Tall M NW	5 Tall F White	6 Short M NW	7 Med. M NW	8 Med. F White	9 Short F White	10 Tall F NW	11 Med. F NW	12 Short F NW
0	10.5	8.1	6.3	1.6	−4.3	−5.5	−3.5	−11.7	−13.4	−17.7	−22.0	−23.4
$0.0\overline{3}$	4.5	3.6	3.1	1.5	−1.1	−1.5	−1.6	−5.0	−6.5	−8.4	−11.8	−12.7
0.10	1.8	1.3	1.3	0.9	−0.4	−0.5	−1.3	−2.0	−2.6	−2.8	−4.9	−5.2
0.20	0.8	0.5	0.5	0.7	−0.1	−0.1	−1.0	−0.9	−1.2	−0.8	−2.0	−2.7
1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mean wage	17.7	16.9	16.3	15.3	14.3	13.6	13.5	12.8	12.3	11.2	10.7	10.5

Table 5
Extent of progressivity (average tax rates, in percent).

	Average wage rate in range									
α_{ES}	2.81	6.50	10.03	13.82	17.80	21.70	27.28	43.25	62.06	95.96
0	−396	−64	−5	17	27	32	38	50	52	53
$0.0\overline{3}$	−346	−51	1	18	25	30	35	47	50	52
0.10	−300	−39	3	18	23	27	31	43	47	50
0.20	−258	−29	7	17	22	24	28	40	44	47
1.00	−5	11	13	14	16	17	19	22	25	29

In this paper, I make three contributions toward demonstrating the promise of such a positive optimal tax theory.

First, I present novel survey evidence on the empirical normative preferences of individuals in the United States. Using a fully-specified planner's problem, I generate feasible and incentive-compatible tax policies that are optimal according to a range of social objective functions, and I have respondents rank these policies. I find striking and robust results: few respondents prefer the conventional Utilitarian policy or the Rawlsian alternative, and a plurality (nearly half) prefers policies that reflect a mixed objective that gives weight to both Utilitarianism and Equal Sacrifice. Additional questions in the survey provide more direct evidence in support of these results. When asked explicitly how "sacrifice" from paying taxes should be distributed, respondents prefer a distribution between that implied by Utilitarianism and Equal Sacrifice. And the more enthusiastic a respondent is about Equal Sacrifice, the more likely he or she is to reject tagging, the taxation of personal characteristics that is a feature of Utilitarian-optimal tax policy but that is rejected by Equal Sacrifice. This evidence is consistent with a substantial body of previous work showing that the normative reasoning of most individuals draws on a diverse set of criteria.

Second, I generalize the conventional optimal tax model to accommodate this evidence of a mixed objective for taxation. This generalization requires overcoming the challenge of combining disparate, sometimes incommensurable, criteria for optimality. I develop a method by which any set of criteria can be integrated into a unified objective that respects Pareto efficiency, and I apply that method to the specific case of two criteria at the heart of this paper: Utilitarianism and Equal Sacrifice. More generally, this method provides a way to inform the choice of welfare weights in the generalized Pareto-efficient approach to optimal taxation.

Third, I show that the empirically-preferred calibration of the generalized theory can explain some conventionally puzzling features of real-world tax policy. I focus on the model's ability to explain the limited role of tagging in policy that is otherwise quite redistributive. I simulate optimal policy with the objective functions favored by survey respondents, calibrated to microdata from the United States. That policy rejects the use of height, gender, and race as tags; it accepts the use of blindness as a tag, endorsing a quantitatively realistic blindness benefit; and it provides redistribution through a progressive schedule of average income tax rates that closely resembles actual policy.

Table 6
Extent of tagging on blindness (extra tax or transfer rate, in percent).

α_{ES}	Not blind	Blind
0	0.07	−130
0.20	0.01	−16

References

Akerlof, George, 1978. The economics of 'Tagging' as applied to the optimal income tax, welfare programs, and manpower planning. Am. Econ. Rev. 68 (1), 8–19 (March).
Alesina, Alberto, Ichino, Andrea, Karabarbounis, Loukas, 2011. Gender-based taxation and the division of family chores, Am. Econ. J. Econ. Policy 3 (2), 1–40 May.
Atkinson, Anthony B., Stiglitz, Joseph, 1980. Lectures on Public Economics. McGraw-Hill.
Auerbach, Alan, Hassett, Kevin, 1999. A new measure of horizontal equity. Working Paper.
Bargain, Olivier, Dolls, Mathias, Neumann, Dirk, Peichl, Andreas, Siegloch, Sebastian, 2011. "Tax-benefit Systems in Europe and the US: Between Equity and Efficiency", IZA DP No 5440.
Bargain, Olivier, Dolls, Mathias, Neumann, Dirk, Peichl, Andreas, Siegloch, Sebastian, 2013. Partisan Tax Policy and Income Inequality in the US, 1979–2007, IZA DP No 7190.
Berliant, Gouveia, 1993. Equal sacrifice and incentive compatible income taxation. J. Public Econ. 51.
Blumkin, Tomer, Margalioth, Yoram, Sadka, Efraim, 2009. Incorporating affirmative action into the welfare state. J. Pub. Econ. 93.
Boadway, Robin, Pestieau, Pierre, 2006. Tagging and redistributive taxation. Queen's University Working Paper 5-2006.
Boaz, David, Kirby, Robin, 2007. Libertarian voters in 2004 and 2006. CATO Policy Report (January/February).
Bourguignon, Francois, Spadaro, Amedeo, 2012. Tax-benefit revealed social preferences. J. Econ. Inequal. 10.
Cappelen, Alexander W., Hole, Astri Drange, Sorensen, Erik O., Tungodden, Bertil, 2011. The pluralism of fairness ideals: an experimental approach. Am. Econ. Rev. 97 (3).
Cremer, Helmuth, Gahvari, Firouz, Lozachmeur, Jean-Marie, 2010. Tagging and income taxation: theory and an application. Am. Econ. J. Econ. Policy 2 (1).
D'Antoni, Massimo, 1999. Piecewise linear tax functions, progressivity, and the principle of equal sacrifice. Econ. Lett. 65, 191–197.
Diamond, Peter, 1998. Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. Am. Econ. Rev. 88 (1).
Diamond, Peter, Saez, Emmanuel, 2011. The case for a progressive tax: from basic research to policy recommendations. J. Econ. Perspect. 25 (4).
Engelmann, Dirk, Strobel, Martin, 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. Am. Econ. Rev. 94 (4).
Farhi, Emmanuel, Werning, Ivan, 2010. Progressive estate taxation. Q. J. Econ. 125 (2).
Feldman, Stanley, Zaller, John, 1992. The political culture of ambivalence: ideological responses to the welfare state. Am. J. Polit. Sci. 36 (1).
Feldstein, Martin, 1976. On the theory of tax reform. J. Public Econ. 6, 77–104.
Fleurbaey, Marc, Maniquet, Francois, 2006. Fair income tax. Rev. Econ. Stud. 73, 55–83.
Fong, Christina, 2001. Social preferences, self-interest, and the demand for redistribution. J. Public Econ. 82.
Frohlich, Norman, Oppenheimer, Joe, Kurki, Anja, 2004. Modeling other-regarding preferences and an experimental test. Public Choice 119 (1/2) (April).
Frolich, Norman, Oppenheimer, Joe A., Eavey, Cheryl L., 1987. Laboratory results on Rawls's distributive justice. Br. J. Polit. Sci. 17.
Gaertner, Wulf, Schokkaert, Erik, 2012. Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice. Cambridge University Press.
Golosov, Mikhail, Tsyvinski, Aleh, 2006. Designing optimal disability insurance: a case for asset-testing. J. Polit. Econ. 114 (2).
Harsanyi, John C., 1953. Cardinal utility in welfare economics and in the theory of risk-taking. J. Polit. Econ. 61 (5), 434–435 (October).
Hasen, David M., 2007. Liberalism and ability taxation. Texas Law Rev. 85 (5) (April).

Hendren, Nathaniel, 2014. The inequality deflator: interpersonal comparisons without a social welfare function. Working Paper.
Horton, John J., Rand, David G., Zeckhauser, Richard J., 2011. The Online Laboratory: Conducting Experiments in a Real Labor Market. Exp. Econ. 14, 399–425.
Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–292.
Kaplow, Louis, 2008. The Theory of Taxation and Public Economics. Princeton.
Kaplow, Louis, Shavell, Steven, 2001. Any non-welfarist method of policy assessment violates the Pareto principle. J. Polit. Econ. 109 (2).
King, Mervyn A., 1983. An index of inequality: with applications to horizontal equity and social mobility. Econometrica 51 (1), 99–115.
Konow, James, 2003. A positive analysis of justice theories. J. Econ. Lit. XLI (December).
Kuziemko, Ilyana, Norton, Michael, Saez, Emmanuel, Stantcheva, Stefanie, 2013. How elastic are preferences for redistribution: evidence from randomized survey experiments. Working Paper.
Lambert, Peter J., Naughton, Helen T., 2009. The equal absolute sacrifice principle revisited. J. Econ. Surv. 23 (2).
Lockwood, Benjamin B., Weinzierl, Matthew, 2012. De gustibus non est taxandum: theory and evidence on preference heterogeneity and redistribution. NBER Working Paper 17784.
Mankiw, N. Gregory, Weinzierl, Matthew, 2009. The optimal taxation of height: a case study of utilitarian income redistribution. Am. Econ. J. Econ. Policy 51 (1), 155–176.
Markovits, Daniel, 2003. How much redistribution should there be? Yale Faculty Scholarship Series. Paper 412.
McCaffery, Edward J., Baron, Jonathan, 2004. Framing and taxation: evaluation of tax policies involving household composition. J. Econ. Psychol. 25.
Mill, John Stuart, 1871. Principles of Political Economy. Oxford University Press, p. 169 (1994, Book V).
Mirrlees, J.A., 1971. An exploration in the theory of optimal income taxation. Rev. Econ. Stud. 38, 175–208.
Mirrlees, J.A., 1974. Notes on welfare economics, information, and uncertainty. In: McFadden, Balch, Wu (Eds.), Essays on Equilibrium Behavior under Uncertainty. North-Holland.
Mitra, Tapan, Ok, Efe A., 1996. Personal income taxation and the principle of equal sacrifice revisited. Int. Econ. Rev. 37 (4), 925–948.
Moyes, Patrick, 2003. Redistributive effects of minimal equal sacrifice taxation. J. Econ. Theory 108, 111–140.
Murphy, Liam, Nagel, Thomas, 2002. The Myth of Ownership. Oxford University Press.
Musgrave, Richard, 1959. The Theory of Public Finance. McGraw-Hill.
Nozick, Robert, 1974. Anarchy, State, and Utopia. Basic Books, New York.
Ok, Efe A., 1995. On the principle of equal sacrifice in income taxation. J. Public Econ. 58, 453–467.
Rakowski, Eric, 2000. Can wealth taxes be justified. NYU Tax Law Rev. 53, 263–375.
Rawls, John, 1971. A Theory of Justice. Harvard University Press, Cambridge.
Rothschild, Casey, Scheuer, Florian, 2012. Optimal taxation with rent-seeking. NBER Working Paper, 17035.
Saez, Emmanuel, 2001. Using elasticities to derive optimal income tax rates. Rev. Econ. Stud. 68.
Saez, Emmanuel, 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. J. Public Econ. 83, 217–230.
Saez, Emmanuel, Stantcheva, Stefanie, 2014. Generalized social marginal welfare weights for optimal tax theory. Available as NBER working paper 18835.
Scott, John T., Matland, Richard E., Michelbach, Philip A., Bornstein, Brian H., 2001. Just deserts: an experimental study of distributive justice norms. Am. J. Polit. Sci. 45 (4).
Sen, Amartya, 1982. Rights and agency. Philos. Public Aff. 11 (1), 3–39 (Winter).
Sen, Amartya, 2000. Consequential evaluation and practical reason. J. Philos. 97 (9), 477–502 (September).
Shaviro, Daniel, 2002. Endowment and inequality. In: Thorndike, Ventry Jr. (Eds.), Tax Justice: The Ongoing Debate. Urban Institute.
Spadaro, Amedeo, Mangiavacchi, Lucia, Piccoli, Luca, 2012. Optimal taxation, social contract, and the four worlds of welfare capitalism. Working Paper.
Stark, Kirk J., 2005. Enslaving the beachcomber: some thoughts on the liberty objections to endowment taxation. UCLA Law and Economics Research Series.
Stiglitz, Joseph E., 1987. Pareto efficient and optimal taxation and the new welfare economics. In: Auerbach, Alan, Feldstein, Martin (Eds.), Handbook on Public Economics. Elsevier. Science Publishers, North Holland, pp. 991–1042.
Sugin, Linda, 2011. A philosophical objection to the optimal tax model. NYU Tax Law Rev. 64, 229–281.
Weinzierl, Matthew, 2011. The surprising power of age-dependent taxes. Rev. Econ. Stud. 78 (4).
Weinzierl, Matthew, 2012. Why do we redistribute so much but tag so little? Normative diversity, equal sacrifice, and optimal taxation. NBER Working Paper 18045.
Weinzierl, Matthew, 2014. The promise of positive optimal taxation. In: NBER Working Paper 18599 (Ed.), originally 2012. (version, at www.matthewweinzierl.com).
Werning, Ivan, 2007. Pareto efficient income taxation. Working Paper. MIT (April).
Yaari, Menahem, 1988. A controversial proposal concerning inequality measurement. J. Econ. Theory 44, 381–397.
Young, H. Peyton, 1987. Progressive taxation and the equal sacrifice principle. J. Public Econ. 32 (2).
Young, H. Peyton, 1988. Distributive justice in taxation. J. Econ. Theory 44.
Young, H. Peyton, 1990. Taxation and equal sacrifice. Am. Econ. Rev. 80 (1).
Young, H. Peyton, 1994. Equity: Theory and Practice. Princeton University Press, Princeton.
Zelenak, Lawrence, 2006. Taxing endowment. Duke Law J. 55 (6), 1145–1181.
Zoutman, Floris, Jacobs, Bas, Jongen, Egbert L.W., 2013a. Optimal redistributive taxes and redistributive preferences in the Netherlands. Working Paper.
Zoutman, Floris, Jacobs, Bas, Jongen, Egbert L.W., 2013b. Revealed social preferences of Dutch political parties. Working Paper.