NBER WORKING PAPER SERIES

THE PROMISE OF POSITIVE OPTIMAL TAXATION

Matthew C. Weinzierl

Working Paper 18599
http://www.nber.org/papers/w18599

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2012

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 helpful discussions. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Promise of Positive Optimal Taxation
Matthew C. Weinzierl
NBER Working Paper No. 18599
December 2012, Revised April 2014
JEL No. D6,D63,H2,H21

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 disagree with this assumption, 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, quantitatively matches several features of existing tax policy that are incompatible in the conventional model but widely endorsed in reality, including the coexistence of substantial redistribution and limited tagging. Together, these findings demonstrate the potential of a positive theory of optimal taxation.

Matthew C. Weinzierl
Harvard Business School
277 Morgan
Soldiers Field
Boston, MA 02163
and NBER
mweinzierl@hbs.edu

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. To cite only one example, the recent authoritative summary of modern tax theory and its policy implications was chaired by James Mirrlees and entitled *The Mirrlees Review* (2010).

When this standard model has been used to generate quantitative lessons for policy, theorists have generally imposed a strong assumption: the objective of tax policy is Utilitarian. Mirrlees 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. 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 that most people appear to disagree 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 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 percent of respondents prefer the Equal Sacrifice alternative over the conventional Utilitarian objective. Disagreement with the conventional Utilitarian 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 percent 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.

This evidence is far from definitive, as many variations in its design, framing, and implementation are possible. Nevertheless, it suggests that key aspects of the policy implied by the Utilitarian objective of conventional theory may not be, in the terminology of Diamond and Saez (2011), "socially acceptable."[3] That is, a large majority of individuals appear to place substantial value on an alternative normative principle–Equal Sacrifice–that rejects 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

[1] Economists in general and optimal tax theorists in particular have largely embraced the defense of Utilitarianism given by Harsanyi (1953), that expected utility maximization applies just as well to uncertainty across one's place in society as it does to one's risky economic choices. Mirrlees (1971) used a generalized Utiltarianism in which the planner may take concave transformations of individual utilities before summing across them. In the limit, this generalized Utilitarianism resembles the Rawlsian priority on the least well-off.

[2] See, e.g., Stiglitz 1987, Werning 2007, Rothschild and Scheuer 2012, and Saez and Stantcheva 2012 (also discussed later in the Introduction). In addition, specific normative limitations of the conventional model have been addressed directly. Fleurbaey and Maniquet (2006) allow for considerations of fairness and responsibility with respect to preference heterogeneity. Besley and Coate (1992) allow for society to place particular emphasis on poverty alleviation. This paper's framework can accommodate these concerns.

[3] In a recent, influential overview Diamond and Saez (2011) argue that to be "fruitfully used as part of forming a policy recommendation," a result from theory "needs to be socially acceptable," by which they mean "there should not be very widely held normative views that make such policies seem implausible and inappropriate at pretty much all times."

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, Matland, Michelbach, and Bornstein (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, not popular opinion. If the moral case for the conventional objective is strong enough, we may continue with the optimal tax project as is, disregarding the apparent preferences of most people. 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. 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. 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.[4] 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. This generalization of the standard theory requires addressing long-standing concerns about commensurability of different normative criteria, and to my knowledge it represents the first effort of its kind.[5] 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 (2012). 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.[6] Their approach and this

[4] Of course, some real-world policies are no doubt far from what would be desired by most people, and instead are due to accident, political economy, and other factors. Identifying which aspects of policy truly reflect society's preferences is a key component of the development of positive optimal tax theory.

[5] 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." An alternative approach to accommodating multiple objectives is to solve for each objective's recommendations separately and then consider how these recommendations interact. The approach in this paper has two advantages over this alternative. First, perhaps the central challenge in combining objectives is to evaluate allocations that are optimal according to none of the component criteria. The alternative approach fails to address this challenge, leaving it unclear how different criteria's priorities would be integrated. Second, the interaction of criteria generates results that would be difficult to obtain through an interaction of their results. For example, this paper's finding that tagging may be limited while substantial redistribution persists would not necessarily be obtained by combining the results of Utilitarianism and Equal Sacrifice, as the former endorses both tagging and redistribution while the latter prohibits both. I am grateful to a referee for a comment prompting this discussion.

[6] 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.

paper's can be seen as two sides of the same coin: one might translate a mixed objective function from my approach into a profile of marginal social welfare weights in theirs, or one might discipline the choice of welfare weights in their approach by seeing whether the mixed objective function from mine can produce them using plausible normative criteria. 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 the full set of criteria by which policy is judged, avoiding the risk acknowledged by Saez and Stantcheva that "the endogenous welfare weights might appear ad-hoc and specified exactly so as to explain the puzzle at hand."[7] 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.

The third contribution of this paper is to show that this generalized model, when calibrated to this survey evidence, has remarkable explanatory power for real-world tax policy. In particular, I simulate optimal policy using the survey respondents' most-preferred normative objective, which combines Utilitarianism and Equal Sacrifice, and U.S. microdata. That policy quantitatively matches three aspects of the U.S. tax code that are incompatible in conventional theory but widely endorsed in reality: it 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. Intuitively, a mix of Utilitarianism and Equal Sacrifice is able to match these features because Equal Sacrifice alone cannot match observed redistribution and Utilitarianism alone cannot match limited tagging. Moreover, I find direct evidence for Equal Sacrifice's role in explaining limited tagging, as a greater share of survey respondents who oppose height and blindness tags prefer policies based in part on Equal Sacrifice. Beyond resolving this puzzle, this paper's model has at least two additional appealing implications: 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.

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 strong 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 the main critique of Utilitarianism put forward by John Rawls, 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 Gouviea (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.

[7] They go on explain that "However, we see this flexibility of our approach as a virtue that opens two new avenues of investigation. First, endogenous welfare weights can be derived from social justice principles, leading to a normative theory of taxation...Second, endogenous welfare weights could also be derived empirically, by estimating actual social preferences of the public, leading to a positive theory of taxation." These avenues of implementing their approach are exactly those pursued in this paper, making the deep connection between the approaches clear.

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 an Appendix contains supporting material.

1 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 prefer 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.

This paper's survey makes a methodological contribution to empirical research on tax preferences by, for the first time as far as I am aware, having respondents face a task that mimics the conventional social planner's optimal tax problem.[8] Especially novel is that these policies are constrained by both feasibility (in the context of government spending) and incentive compatability. This innovation over prior work allow me to use the evidence on participant preferences to calibrate a fully-specified optimal policy model.

1.1 Survey design

The survey, shown in full in the Appendix, 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.[9] These types represent the four lower quintiles and a division of the top quintile into the next 10, 5, 4, and 1 percentiles. To establish that differences in earnings are not due to effort,[10] the text states: "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

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

[9] This distribution is calculated based on data from the Congressional Budget Office for the United States in 2006. See the discussion in Section 2.4.1 for details.

[10] If respondents attribute some of income variation to effort, rather than ability, they are likely to be less likely to endorse redistributive criteria (see Lockwood and Weinzierl, 2012). Therefore, this clarification makes it more likely that we will find strong support for Utilitarianism.

as a policymaker for this society." They are given information about the constraints facing their choice, as the survey states the required level of exogenous government spending (i.e., feasibility) and emphasizes that households' labor supplies may respond to tax policy (i.e., incentive compatibility). Both of these constraints are reflected in the figures. Respondents are reminded that taxes may serve a variety of purposes, from funding public goods to redistributing before-tax income.

In a series of choices, respondents rank sets of tax policies. For each policy option, the survey displays two overlapping income distributions (see Figure 1 below for an example). The pretax distribution is shown as empty outlined columns while the aftertax distribution is shown as filled-in green 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.[11] 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 minutes 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 minutes and 6 seconds. They were paid $2.00 for the task, implying an average hourly rate of $9.16.

1.2 Results

Respondents' first rankings provide straightforward evidence that the Utilitarian criterion is less popular than the conventional model implies. Figure 1 shows the two policies respondents rank, labeled A and B.

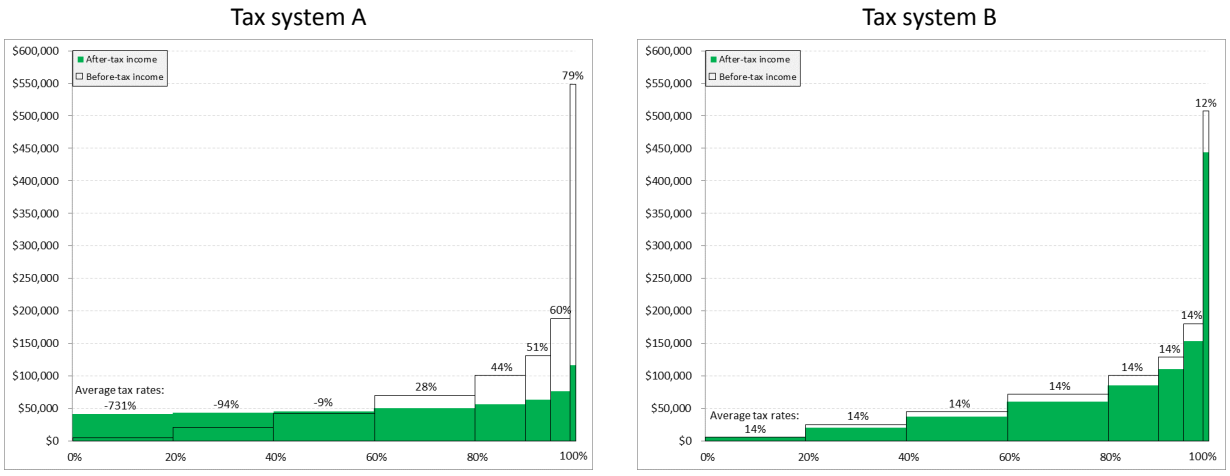


Figure 1: The choice between the Utilitarian policy A and the Equal Sacrifice policy B

[11] Specifically, only respondents registered as in the United States whose work had been accepted on 95 percent of previous tasks could take the survey. Horton, Rand, and Zeckhauser (2010) 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."

In Section 2, I provide the details of how I calculated these (and all other) policy options in the survey. For now, note simply 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 percent, with a standard error of 2 percent. 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. These seven include the two policies from the first choice (the Utilitarian option A and the Equal Sacrifice option B); three intermediate policies that are weighted combinations of A and B (namely E, D, and G); and two more extreme policies, a Rawlsian policy C and a lump-sum "poll tax" policy F. 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 the Appendix. 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 the Appendix).

Figure 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.

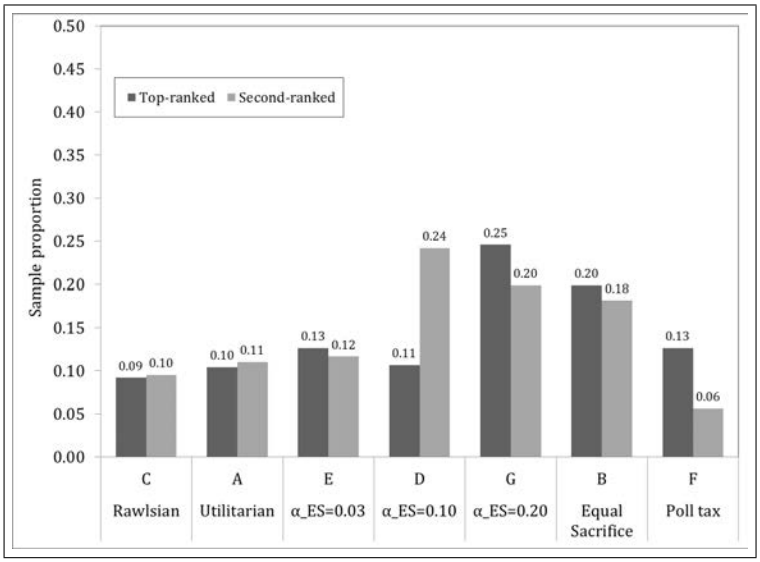


Figure 2: Respondent preferences across a range of policy options

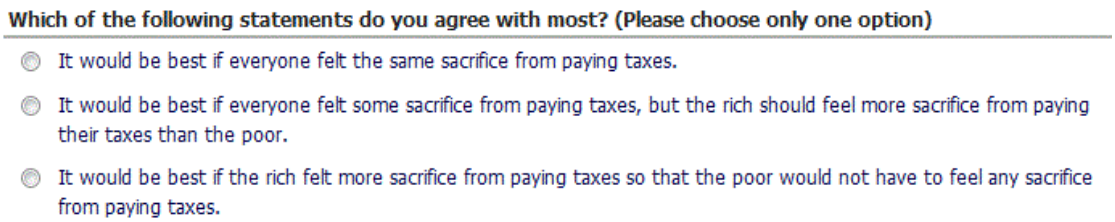
Figure 2 reveals two main results from these rankings: one, support for the conventional Utiltarianism

[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.

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 percent of the top-ranked choices and 11 percent of the second-ranked choices.[14] For the Rawlsian alternative, these figures are 9 percent and 10 percent. Together, then, policies at least as redistributive as the conventional Utilitarianism make up less than 20 percent of the most-preferred policies in this survey, the same share claimed by the pure Equal Sacrifice policy B. In contrast, nearly half–48 percent–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 percent 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.

Further evidence of a preference for an objective that combines Utilitarianism with Equal Sacrifice 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 percent choose the first option; 48 percent choose the second; and 19 percent choose the third option. 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. One-fifth of the survey respondents show enthusiasm for the conventional Utilitarian outcome, one-third support the Equal Sacrifice alternative, and nearly one-half choose the intermediate option. In other words, approximately four out of five respondents choose policies reflecting some weight on Equal Sacrifice, the same share as in the choices over tax policies as shown in Figure 2.

Finally, Figure 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

[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 six percent 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.

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.

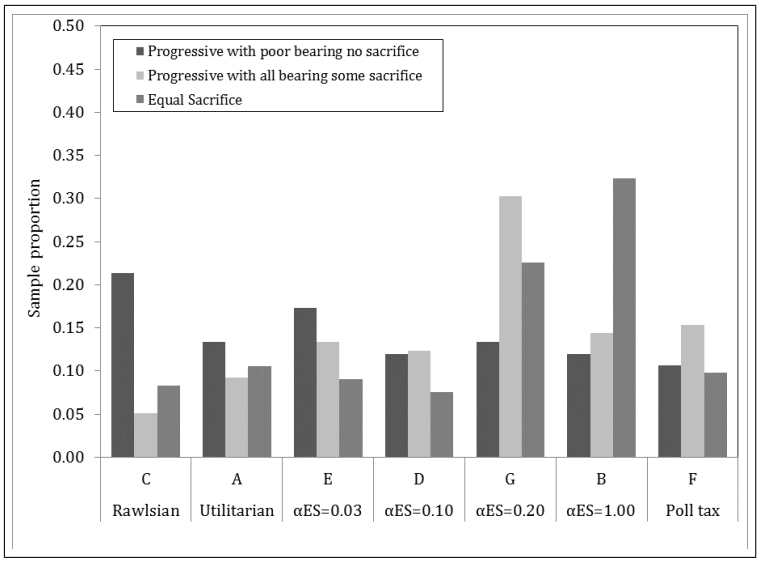


Figure 3: Policy preferences by view on Equal Sacrifice (top choices only)

1.3 Robustness

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

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

Second, the pattern of rankings by most respondents suggests 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 percent of respondents exhibited single-peaked preferences across these five policy choices. Importantly, that result does not imply

that 32 percent of respondents were making irrational choices–someone may prefer policies that commit fully to one normative criterion or another, for example.

1.3.2 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 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 1.

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 percent reporting being from the bottom two quintiles, 45 percent from the next two quintiles, and 19 percent 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 percent of adults are college graduates.[15]

Table 1: 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

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

Table 1 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).[16] Respondents with less education tend to be more supportive of redistributive policies, and respondents' economic status as adults tends to be negatively related to support for redistribution. 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 (which supports Equal Sacrifice) prefers policies that result from a mixed normative criterion that combines Utilitarianism and Equal Sacrifice. Thus, both main findings from the full survey apply across demographic groups.

1.3.3 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 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 2 shows the distribution of responses. A plurality of the respondents, 44 percent, self-classifies as left-leaning.[17] Support for Libertarianism in this sample is consistent with the magnitudes for the U.S. population cited by Boaz and Kirby (2007).

Table 2 also shows that both main findings from the full survey characterize respondents across a wide range of political opinions.

Table 2: 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. 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 weight on Equal Sacrifice.

[16] 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.

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

1.4 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 criteria. In the Appendix, I discuss the related research in detail, but summary statements from studies representing the three research designs in that literature illustrate the main point. Frohlich, Oppenheimer, and Eavey (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." Englemann 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." As these statements demonstrate, the findings of this paper's survey are consistent with and build on a substantial base of prior work.

2 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.

2.1 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.[18]

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

[18] 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 consequentalist and deontological frameworks can be compared. See Sen (1982).

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).

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).[19] 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 $i \in \{1, 2, ..., I\}$, 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 utility function $U(c, y/w)$.[20]

A planner chooses allocations $\left\{c_*^i, y_*^i\right\}_{i=1}^{I}$ 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)*

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

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

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

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

[19] In Sen (1982) he writes: "...both welfarist consequentialism (such as utilitarianism) and constraint-based deontology are fundamentally inadequate because of their failure to deal with certain important types of interdependences present in moral problems. This leads to an alternative approach... which incorporates, among other things, some types of rights in the evaluation of states of affairs, and which gives these rights influence on the choice of actions through the evaluation of consequent states of affairs."

[20] As in most optimal tax analyses, I assume utilities are interpersonally comparable.

such weights. The most straightforward is that the pivotal voter has his or her own weights on each normative criterion, adopted by policymakers as a result of electoral competition.[21] One implication of this paper's analysis is that future research estimating the values of these weights and how they are generated by the political process would be valuable.

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 $\left\{c_1^i, y_1^i\right\}_{i=1}^I \in \mathbb{F}$ is weakly preferred under criterion ϕ to allocation $\left\{c_2^i, y_2^i\right\}_{i=1}^I \in \mathbb{F}$ if

$$\left\{c_1^i, y_1^i\right\}_{i=1}^I \succeq_\phi \left\{c_2^i, y_2^i\right\}_{i=1}^I.$$

Given $\succeq_\phi$, the strict preference relation $\succ_\phi$ is defined as usual. For any $\left\{c_1^i, y_1^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I \in \mathbb{F}$,

$$\left\{c_1^i, y_1^i\right\}_{i=1}^I \succ_\phi \left\{c_2^i, y_2^i\right\}_{i=1}^I \Leftrightarrow \left\{c_1^i, y_1^i\right\}_{i=1}^I \succeq_\phi \left\{c_2^i, y_2^i\right\}_{i=1}^I \text{ but not } \left\{c_2^i, y_2^i\right\}_{i=1}^I \succeq_\phi \left\{c_1^i, y_1^i\right\}_{i=1}^I.$$

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

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

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). Omitting incentive compatibility constraints enables the ϕ-optimal feasible allocations to provide a stable target for each criterion against which to measure the appeal of different policy proposals. The alternative, namely to use a "second-best" allocation, requires assumptions on the planner's information set that can lead to path-dependencies in policy evaluations and, thus, unstable judgments.[22] The problem is that no obvious principle exists by which to determine which information is to be included or excluded from the planner's information set when defining ϕ-optimal feasible allocations.[23] No such inconsistency exists if the information set includes all possible information.

[21] If one wished to consider, instead, different groups engaged in a policy-setting game, alternative approaches could be used. For example, the Nash bargaining solution would optimize a weighted combination of their interests. "Veto" models such as that in Moulin (1981) would allow a coalition of voters to block some alternatives. Such formulations are conceptually similar to this paper's, as the key to this paper's results is not the specific formalization of the tradeoff between normative criteria but rather that the tradeoff is included at all.

[22] An example will illustrate this best. Recall that for each criterion ϕ, loss is calculated by comparing the distributions of individual utility under the ϕ-optimal feasible allocation and the actual, constrained allocation. Consider two cases for how to define the ϕ-optimal feasible allocations. In case 1, allow the planner to condition taxes on ability. In case 2, assume that the planner cannot condition taxes on ability but can condition taxes on gender, an observable characteristic related to ability in the data. The distributions of utility in the ϕ-optimal feasible allocations for these two cases will differ. Now suppose we want to gauge the appeal of conditioning taxes on height or race (as in Section 3 below). Assume that height and gender are correlated but race and gender are not. Then, the use of gender in defining the ϕ-optimal feasible allocations in case 2 will differentially affect the loss calculations for the allocations that use a height tax or a race tax. By the same reasoning, if we allowed a different trait to be used in case 2, the relative appeal of height and race taxes would be affected yet differently.

[23] In particular, note that limiting the information set to the standard observables of optimal tax theory, namely income and consumption, does not solve the problem. Even in that case, evaluating a height-gender-race combination tag may yield different results than evaluating each tag separately, whether one-by-one or cumulatively.

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 $\mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_*^i, y_*^i\right\}_{i=1}^I\right)$

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.[24]

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

1. *Ordinality: For any $\left\{c_1^i, y_1^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I \in \mathbb{F}$,*

$$\mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_1^i, y_1^i\right\}_{i=1}^I\right) \leq \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I\right) \Leftrightarrow \left\{c_1^i, y_1^i\right\}_{i=1}^I \succeq_\phi \left\{c_2^i, y_2^i\right\}_{i=1}^I,$$

   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: $\mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I\right) = 0$, so that the loss is zero[25] when the equilibrium allocation equals the ϕ-optimal feasible allocation.*

3. *Weak Pareto Efficiency:*

$$\begin{aligned} U\left(c_1^i, y_1^i/w^i\right) \geq U\left(c_2^i, y_2^i/w^i\right) \text{ for all } i \in \{1, 2, ..., I\} \\ \Rightarrow \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_1^i, y_1^i\right\}_{i=1}^I\right) \leq \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I\right), \end{aligned} \tag{4}$$

   which can be converted into Strong Pareto Efficiency if desired.[26]

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.[27]

Together, ϕ-optimal feasible allocations and loss functions allow us to make commensurable a diversity of normative frameworks that, then, can jointly influence the determination of optimal policy. 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. First, I provide some background on why Equal Sacrifice is a natural choice as the alternative to Utilitarianism.

2.2 Equal Sacrifice as an alternative to Utilitarianism

John Stuart Mill (1871) was the most famous proponent of Equal Sacrifice, and his argument for it is worth quoting at length.

[24] See, for examples of contrasting views, Sen and Williams (1982, introductory chapter) and Kaplow and Shavell (2001).

[25] Any constant would accomplish the same normalization, though zero is the natural choice.

[26] Namely, $U\left(c_1^i, y_1^i/w^i\right) \geq U\left(c_2^i, y_2^i/w^i\right)$ for all $i \in \{1, 2, ..., I\}$ and $U\left(c_1^{i'}, y_1^{i'}/w^{i'}\right) > U\left(c_2^{i'}, y_2^{i'}/w^{i'}\right)$ for some $i' \in \{1, 2, ..., I\} \Rightarrow \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_1^i, y_1^i\right\}_{i=1}^I\right) < \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I\right)$

[27] The Strong Pareto Efficiency condition states, in words, that if all individuals do at least as well under allocation 1 as under allocation 2, and at least one individual does better, then the loss from allocation 1 must be strictly less than the loss from allocation 2.

> "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, 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." [28]

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.[29] Of course, endowment taxation is exactly the preferred policy of the conventional Utilitarian optimal tax model. Rawls (1971) wrote that Utilitarianism "does not take seriously the distinction between persons," and that an endowment tax "would force the more able into those occupations in which earnings were high enough for them to pay off the tax in the required period of time; it would interfere with their liberty to conduct their life within the scope of the principles of justice."[30] The broad force of this critique is made clear when it is coupled with Robert Nozick's (1974) claim that "taxation of earnings from labor is on a par with forced labor" because "it is like forcing the person to work n hours for another's purpose." 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.[31]

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. Relatedly, some have suggested that there is an explicit connection between Equal Sacrifice and Libertarianism. As noted earlier, Feldstein (1976) writes: "Nozick (1974) has recently presented an extensive criticism of the use of utilitarian principles to justify the redistribution of income and wealth...In this context, the principle of benefit taxation or of tax schedules that impose equal utility sacrifice have an appeal that is clearly lacking in the utilitarian framework." Similarly, Liam Murphy and Thomas 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." Sidgwick's statement above, with its caveat that speaks to the core of Libertarianism, suggests the same link.

A priority on equal treatment may be of paramount concern to only a small minority of individuals, but evidence strongly suggests that it has at least some appeal to most. Public opinion surveys, including this

[28] In addition to the work of H. Peyton Young and Berliant and Gouveia mentioned earlier, Yaari (1988), Moyes (1989), Ok (1995), Mitra and Ok (1996), and D'Antoni (1999) helped established 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.

[29] 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).

[30] This latter quote is from Rawls (2001). Political philosophers and legal scholars have developed this critique in depth. As an example of the former, see Mazor (2012) and Richard Arneson (2000), who writes: "It is better to regard Rawls as making the point that ...it is a flaw that utilitarianism would have the decision about what should be done vary only with the utility total that different acts could achieve."

[31] Stark (2005) offers a detailed argument that the concerns of Rawls and Nozick are closely connected. A related perspective is captured in Immanuel Kant's (1785) dictum "to treat himself and all others never merely as means but always at the same time as ends in themselves."

paper's, estimate that the proportion of individuals with largely Libertarian views is 10 to 20 percent in the United States (Boaz and Kirby 2007).[32] But, research has shown that even those predisposed toward redistribution feel some pull toward normative principles that prioritize the individual.[33] Feldman and Zaller's (1992) statement cited in Section 1.4 makes this point, and this paper's survey evidence clearly supports that conclusion. As Feldstein (1976) noted: "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. Mill's split normative intuition is more the rule than the exception, and I explore the implications of it in this paper.

2.3 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: the Utilitarian criterion of maximal aggregate utility and the principle of Equal Sacrifice.

2.3.1 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:

$$\left\{c_1^i, y_1^i\right\}_{i=1}^I \succeq_{Util} \left\{c_2^i, y_2^i\right\}_{i=1}^I \Leftrightarrow \sum_{i=1}^I p^i U\left(c_1^i, y_1^i/w^i\right) \geq \sum_{i=1}^I p^i U\left(c_2^i, y_2^i/w^i\right). \tag{5}$$

The Utilitarian-optimal feasible allocation is therefore:

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

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

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

[32] Cappelen et al. (2011) conduct experiments in which participants' choices imply a preference among competing "fairness ideals," and in their preferred specification 18.7 percent of participants are classified as "libertarians." Konow (2003) reports results consistent with these magnitudes.

[33] Though the connection to problems of taxation is imperfect, Frohlich, Oppenheimer, and Kurki (2004) show that "just deserts" or "entitlements" exert an influence on allocations for most dictators in allocation games with production.

choices for that starting point, one natural option is the allocation that would obtain absent any government intervention, i.e., the no-tax 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 Liam Murphy and Thomas 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* $U_x\left(c, y/w\right)$ *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:

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

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:

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

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 $i \in \{1, 2, ..., I\}$.

Consequently, the ES-optimal feasible allocation is that which achieves the smallest equal sacrifice while funding G. Formally, we define $\left\{c_{ES}^i, y_{ES}^i\right\}_{i=1}^I$ 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 $i \in \{1, 2, ..., I\}$ and for all possible $\left\{c^i, y^i\right\}_{i=1}^{I} \in \mathbb{ES}$.

Even with this technical solution, a prominent normative concern remains: is the ES-optimal feasible allocation a just starting point?[34] Many tax specialists and political philosophers argue that the laissez-faire allocation, and by implication the ES-optimal feasible allocation, is not. But this paper takes a positive perspective on the objective for taxation, and the evidence suggests that most individuals support Equal Sacrifice defined in this way. In particular, as will be made clear in Section 2.4, the survey evidence of Section 1 shows that most respondents prefer a tax policy reflecting some weight on Equal Sacrifice with sacrifice calculated from the ES-optimal feasible allocation. Moreover, when they are asked about the proper distribution of "sacrifice" from taxes, respondents endorse a more equal distribution than the conventional Utiltarian model would suggest. Though the interpretation of this result is debatable in many ways, it strongly suggests that most respondents do not believe the existing economic distribution is unjust in a way that merits substantial redistribution.

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

2.3.2 Loss functions

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

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

In words, it is the sum of individuals' utility losses from having the equilibrium allocation $\left\{c^i_*, y^i_*\right\}_i$ 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.[35] 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.

I will assume an Equal Sacrifice loss function $\mathcal{L}_{ES}$ with three features: first, deviations of individual utility below the ES-optimal feasible allocation are costly but deviations above the ES-optimal feasible allocation yield little or no offsetting benefits;[36] second, losses increase more than proportionally with the size of the deviation of individual utility below the ES-optimal feasible allocation; third, gains are concave in the size of the deviation of individual utility above the ES-optimal feasible allocation.

[34] The caveat in Sidgwick's earlier statement in support of Equal Sacrifice directly raises this issue.

[35] 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 were 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.

[36] 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.

I formalize these properties as follows:

$$\mathcal{L}_{ES}\left(\left\{c^i_{ES}, y^i_{ES}\right\}_i, \left\{c^i_*, y^i_*\right\}_i\right) = \sum_{i=1}^{I} p^i V\left(U\left(c^i_{ES}, y^i_{ES}/w^i\right), U\left(c^i_*, y^i_*/w^i\right)\right), \tag{9}$$

where

$$V\left(U^i_{ES}, U^i_*\right) = \begin{cases} -\left(\delta\left[U^i_* - U^i_{ES}\right]\right)^\theta \text{ if } U^i_{ES} < U^i_* \\ \left[\lambda\left(U^i_{ES} - U^i_*\right)\right]^\rho \text{ if } U^i_{ES} \geq U^i_* \end{cases}, \tag{10}$$
$$\text{for scalars } \{\delta \geq 0,\ \lambda > \delta,\ \theta \in (0,1],\ \rho > 1\}.$$

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 ES-optimal feasible allocation.

The asymmetric punishment of downward deviations from the ES-optimal feasible allocation implied by $\delta < \lambda$ rejects the Utilitarian idea that the distribution of utility across individuals is irrelevant. 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 ES-optimal feasible allocation.

2.3.3 Planner's problem

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

Problem 2 *Social Planner's Problem (specific case)*

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

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.

2.3.4 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. Figure 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.

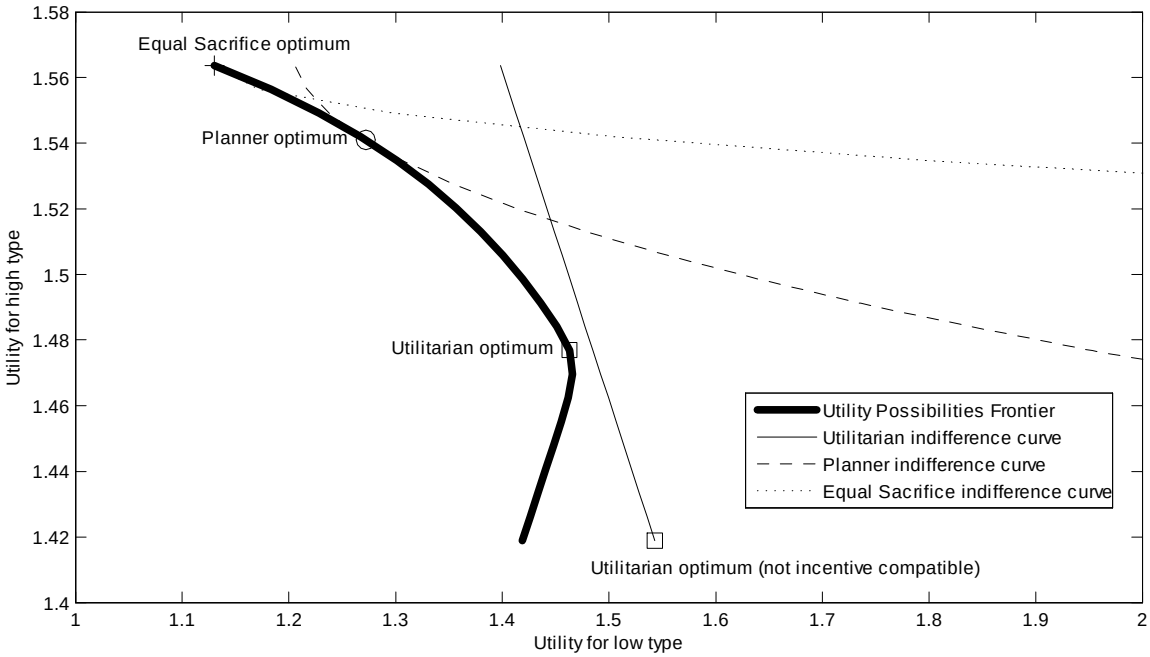


Figure 4: The Utility Possibilities Frontier and Indifference Curves in the Two-Type Example

Figure 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.

2.4 The set of policy options offered to survey respondents

The set of policy options presented to respondents in the survey of Section 1 were 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. I also describe the two additional policies, based on other criteria, that were offered as options in the survey.

2.4.1 Calibration of the model

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 for these households,[37] 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 the population proportion of type i, is as follows.

Ability distribution								
w^i	3.01	12.65	23.21	37.42	53.05	68.43	95.66	258.64
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	$2^{-\sigma}$	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. Three parameter values deserve additional comment.

First, assuming $\delta = 0$ implies that deviations of individual utility above the ES-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

[37] 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.

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 may underestimate the concavity of that subutility function. If so, this assumption generates 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.

2.4.2 Generating the survey's policy options

Table 3 shows the seven policy options in the survey of Section 1, decreasing in redistributiveness from left to right. For each option, it shows the average tax rates they levy on each household. In the second column of the table, before-tax incomes in the no-tax scenario are shown.

Table 3: Features of the tax systems among which respondents choose

		Tax system:						
		C	A	E	D	G	B	F
HH type	No-tax earnings	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 ES: α_{ES}						
	Rawls	0.00	$0.0\bar{3}$	0.10	0.20	1.00	*Poll tax*

The middle five policies in Table 3 combine Utilitarianism and Equal Sacrifice, using a range of values for α_{ES} in the planner's problem from expression (11). 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: $\alpha_{ES} = \{0.0\bar{3},\ 0.10,\ 0.20\}$.

I also generate the two "endpoint" policies shown in Table 3. The left-most policy, C, is a "Rawlsian" policy that maximizes the utility of the lowest-ability household, $i = 1$. Formally, 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 (2); $\mathbb{IC}$ is defined in (3); and $U\left(c^1, y^1/w^1\right)$ is defined in (12). The Rawlsian policy is

the most redistributive policy option. The right-most policy is a "poll tax" that splits the financial cost of government spending G evenly across households. Formally, under a poll tax,

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

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

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 pure Utilitarianism in conventional optimal tax theory is a generalized Utiltarianism under which the planner takes a concave transformation of utilities before summing them. The Rawlsian option is often included as a simple way to capture this generalization. By including a Rawlsian policy as a choice, we can gauge the empirical support for this prominent criterion.

3 Descriptive power of the positive optimal tax model

In this section I show that the optimal tax model, as proposed and empirically estimated in this paper, is remarkably successful at matching 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. I prove analytically that the model with an objective for taxation that puts some weight on Equal Sacrifice will do less tagging than the conventional model. I then use numerical simulations to show, quantitatively, that the calibrations of that model favored in the survey evidence of Section 1 can explain specific patterns of tagging and redistribution in the United States. I go on to show that optimal policy according to the same calibration can help resolve two additional puzzling gaps between conventional theory and actual policy.

3.1 Why do we redistribute so much but tag so little?

Tagging has an illustrious theoretical pedigree. James Mirrlees (1971) noted the potential of tagging in only the fifth sentence of his Nobel Prize-winning analysis of optimal taxation. George Akerlof (1978), also a recipient of the Nobel Prize, worked out the basic theory of tagging in a seminal paper just seven years later. Forty years into the modern optimal tax literature, recent analyses have shown the substantial potential gains from tagging according to three specific personal characteristics: height, gender, and race (see Mankiw and Weinzierl 2010; Alesina, Ichino, and Karabarbounis 2011 and Cremer, Gahvari, and Lozachmeur 2010; and Blumkin, Margalioth, and Sadka 2009). Though the most general version of the standard optimal tax model does not necessarily imply tagging, the specifications of that model that dominate research strongly recommend it.[38]

[38] I am grateful to a referee for suggesting this clarification. A fully general model in which the social welfare function simply uses Pareto weights to value individual utility may not endorse tagging, depending on the values assumed for those Pareto weights. For example, if "needs" vary with a tag that is positive correlated with ability, optimal policy may avoid using the tag. Or, if Pareto weights put high value on those with high income-earning ability, optimal policy may not wish to redistribute toward low-ability individuals, making tags less valuable. Such interpretations of the model are relatively rare, however, and the claim that tagging is optimal under a Utilitarian criterion is largely uncontroversial. To cite just one example, Piketty and Saez (2012) write in their new chapter for the Handbook of Public Economics that "We have assumed that T(z) depends

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. For example, groups with higher mean ability ought to be taxed to support other groups, while groups with a higher variance of ability ought to face a more progressive within-group tax policy. As Mirrlees writes: "One might obtain information about a man's income-earning potential from his apparent I.Q., the number of his degrees, his address, age or colour..."[39] There are many other potential tags–height, gender, facial symmetry, place in birth order, native language, parental traits, macroeconomic conditions at age 18, and so on–all of which relate systematically to income-earning ability and are largely exogenous to the individual. Genetic information may someday provide particularly powerful tags.[40]

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. For example, disability benefits are common among developed countries, as are programs aimed at alleviating poverty among the elderly. Indeed, nearly two-thirds of U.S. federal entitlement spending goes to programs generally limited to the elderly and disabled (Viard, 2001). These groups are the prototypical examples of those with systematically low income-earning ability.[41] The other large example of tagging is payments to families with young children, where the *per capita* ability to earn income is mechanically low when compared to childless households. Other, isolated programs such as benefits for the blind follow a similar pattern, so that existing tagging bears little resemblance to the broad and nuanced application recommended by modern optimal tax theory.

The generalized model proposed in this paper can resolve this puzzle. The Equal Sacrifice principle says that all taxpayers should bear the same sacrifice (in terms of reduced well-being) from paying taxes. Tagging violates Equal Sacrifice because it causes, for example, a tall person to pay more tax–and therefore bear a greater sacrifice–than a short person who has the same ability to earn income. A revised optimal tax theory that values Equal Sacrifice, as do the preferred policies in the survey of Section 1, will determine whether to use a given tag by weighing the costs of such violations against the gains it generates according to Utilitarianism. Only tags providing sufficiently strong information about ability, and therefore Utilitarian welfare gains, will be optimal according to this mixed objective.

3.1.1 Analytical results on optimal tagging

To analyze optimal tagging, I modify the social planner's problem so that individuals differ in two characteristics: unobservable ability w indexed by i, and an observable, tagged variable indexed by $m = \{1, 2, ..., M\}$. Therefore, allocations are denoted $\left\{c^{i,m}, y^{i,m}\right\}_{i=1,m=1}^{I,M}$ and the population proportion of the individual with

only on earnings z. In reality, the government can observe many other characteristics (denoted by vector X) also correlated with ability (and hence social welfare weights) such as gender, race, age, disability, family structure, height, etc. Hence, the government could set T(z,X) and use the characteristic X as a "tag" in the tax system. There are two noteworthy theoretical results. First, if characteristic X is immutable then there should be full redistribution across groups with different X..."

[39] Despite this quotation, age should not be considered a tag. Unlike these other characteristics, age is shared by all individuals (abstracting from mortality variation), so that age-dependent taxes do not achieve support for a disadvantaged group by taxing another. In particular, age-dependent taxes do not violate equal sacrifice once the full lifecycle of each taxpayer is considered. See Weinzierl (2011) for a study of this and other aspects of age-dependent taxes.

[40] Note that privacy concerns may be relevant for some potential tags, such as genetic information. A concern for privacy is one example of a value that could be incorporated into the optimal tax model using the approach of this paper, provided that it can be translated convincingly into a preference over final allocations.

[41] The economic prospects for people over the age of 65 have improved in the decades since the programs designed to support the elderly were created. The current debate over raising the retirement age in these programs may reflect, in part, skepticism that age 65 is still a reliable indicator of lower income-earning ability. Also, see the earlier note in this section on age not being a proper tag.

ability i and tagged variable value m is denoted $p^{i,m}$ where $\sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m} = 1$. The modified planner's problem is as follows.

Problem 3 *Social Planner's Problem with Tagging*

$$\min_{\left\{c_*^{i,m},y_*^{i,m}\right\}_{i=1,m=1}^{I,M}\in\{\mathbb{F}\cap\mathbb{IC}\}} \left\{ \begin{array}{c} \alpha_{Util}\sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m}\left[U\left(c_{Util}^{i,m},y_{Util}^{i,m}/w^i\right)-U\left(c_*^{i,m},y_*^{i,m}/w^i\right)\right] \\ +\ \alpha_{ES}\sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m}V\left(U\left(c_{ES}^{i,m},y_{ES}^{i,m}/w^i\right),U\left(c_*^{i,m},y_*^{i,m}/w^i\right)\right) \end{array} \right\}, \tag{13}$$

where

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

$V(\cdot)$ *is a modified version of* (10),

$$V\left(U_{ES}^{i,m},U_*^{i,m}\right)=\left\{ \begin{array}{c} -\left(\delta\left[U_*^{i,m}-U_{ES}^{i,m}\right]\right)^{\theta} \text{ if } U_{ES}^{i,m}<U_*^{i,m} \\ \left[\lambda\left(U_{ES}^{i,m}-U_*^{i,m}\right)\right]^{\rho} \text{ if } U_{ES}^{i,m}\geq U_*^{i,m} \end{array} \right., \tag{14}$$

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

the feasibility set is a natural modification of expression (2),

$$\mathbb{F}=\left\{\left\{c^{i,m},y^{i,m}\right\}_{i=1,m=1}^{I,M}:\sum_{i=1}^{I}\sum_{m=1}^{M}p^{i,m}\left(y^{i,m}-c^{i,m}\right)\geq G\right\}, \tag{15}$$

and the set of incentive compatible allocations $\mathbb{IC}$ *is:*

$$\mathbb{IC}=\left\{\left\{c^{i,m},y^{i,m}\right\}_{i=1,1}^{I,M}:U\left(c^{i,m},y^{i,m}/w^i\right)\geq U\left(c^{j,m},y^{j,m}/w^i\right)\ \text{for all } i,j\in\{1,2,...,I\}\ \text{ and } m\in\{1,2,...,M\}\right\}. \tag{16}$$

In this problem the incentive constraints (16) are m-specific. That is, the planner can restrict each individual to the allocations within his or her tagged group, whereas if tagging were excluded the planner would be required to ensure that each individual preferred his or her allocation to that of any individual in any tagged group.

The following proposition is implied by the first-order conditions of this planner's problem, assuming separable utility between consumption and labor effort. The proof can be found in the Appendix.

Proposition 1 *If* $U_{c,y/w}(c,y/w)=0$*, the solution to the Social Planner's Problem with Tagging satisfies:*

$$\frac{E_i\left[\left(U_{c_*^{i,m}}\right)^{-1}\right]}{E_i\left[\left(U_{c_*^{i,n}}\right)^{-1}\right]}=\frac{E_i\left[\alpha_{Util}-\alpha_{ES}\frac{\partial V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\partial U_*^{i,m}}\right]}{E_i\left[\alpha_{Util}-\alpha_{ES}\frac{\partial V\left(U_{ES}^{i,n},U_*^{i,n}\right)}{\partial U_*^{i,n}}\right]}, \tag{17}$$

where $U_*^{i,m}$ *denotes* $U\left(c_*^{i,m},y_*^{i,m}/w^i\right)$ *and* $U_{c_*^{i,m}}$ *denotes* $\partial U\left(c_*^{i,m},y_*^{i,m}/w^i\right)/\partial c_*^{i,m}$.

The left-hand side of (17) is the ratio of the expected inverse marginal utilities of consumption across tagged types.[42] This equals the ratio of the cost in consumption units of an incentive-compatible marginal

[42] Note that the terms in brackets on the right-hand side of expression (17) are related to marginal social welfare weights, as in Saez and Stantcheva (2012). In the conventional model, these terms are equal to one. Here, they differ from one because of

increase in utility across all individuals with tagged value m versus n. The following corollary makes plain why this ratio is of interest.

Corollary 1 *If $\alpha_{ES} = 0$, equation (17) simplifies to:*

$$\frac{E_i\left[\left(U_{c_*^{i,m}}\right)^{-1}\right]}{E_i\left[\left(U_{c_*^{i,n}}\right)^{-1}\right]} = 1. \tag{18}$$

This result, also shown in Weinzierl (2011) for age-dependent taxes and labeled the Symmetric Inverse Euler equation in that context, shows that the Utilitarian planner with access to tagging will equalize the cost of providing utility to tagged groups.[43] Intuitively, the planner has full information about the tag, so any opportunity to raise overall welfare by transfers across tag values will be exploited.

Next, I derive a condition analogous to (18) for positive α_{ES}. I make two mild assumptions to provide a clean benchmark case.[44] Importantly, both of these assumptions hold in the numerical simulations below.

Assumption 1: *At least one pair of tagged groups $(m, n) \in \{1, 2, ..., M\}$ can be ordered such that $m < n$ implies that the solution to the Social Planner's Problem with Tagging when $\alpha_{ES} < 1$ satisfies*

$$U_*^{i,m} \geq U_*^{i,n} \text{ for all } i = \{1, 2, ..., I\}, \tag{19}$$

and

$$U_*^{j,m} > U_*^{j,n} \text{ for at least one } j = \{1, 2, ..., I\}. \tag{20}$$

In words, Assumption 1 holds that tagged groups can be "ranked", for instance by some function of the mean and variance of wages within each group, so that individuals in at least one higher-ranked group fare no better, and in some cases worse, than individuals of the same abilities in a lower-ranked group when the planner is at least in part Utilitarian. That is, individuals of any given ability obtain allocations that generate greater losses or smaller gains when they are members of a higher-ranked group.

Assumption 1 is closely related to a well-known result from previous optimal tax analyses that an "advantaged" tagged group is taxed heavily by a conventional Utilitarian-optimal tax policy. Mankiw and Weinzierl (2010) show this numerically for the optimal height tax in the United States, under which a tall taxpayer ends up with lower utility than a short taxpayer of the same ability. Intuitively, the planner treats those with the advantaged tag as higher-skilled workers on average, requiring them to produce more income than others. Mirrlees (1971, 1974) showed much the same result for higher ability individuals in the full information case (which is the relevant analogue) of his optimal tax problem, a result discussed in a different context (rank reversals) below.

Assumption 2: In *the solution to the Social Planner's Problem with Tagging when $\alpha_{ES} < 1$,*

$$U_{ES}^{i,m} \neq U_*^{i,m} \text{ for all } i = \{1, 2, ..., I\} \text{ and } m \in \{1, 2, ..., M\}. \tag{21}$$

Assumption 2 is a technical assumption that rules out the scenario in which the utility allocated to any individual under the optimal policy exactly equals the utility that individual obtains under the ES-

the planner's aversion to unequal sacrifice. I am grateful to Bernard Salanie for suggesting this note.

[43] A referee brought to my attention that this condition also appears in Parsons (1982).

[44] These assumptions are sufficient, but not necessary, for the result in Corollary 2.

optimal feasible allocation.[45] This assumption is unlikely to bind because the optimal allocations with $\alpha_{ES} < 1$ reflect not only the Equal Sacrifice priorities but also the Utilitarian ones, and because incentive compatibility is imposed on the optimal allocations but not on the ES-optimal feasible allocations. Again, note that Assumption 2 is satisfied in all cases in the numerical simulations below.

With these assumptions, the following corollary to Proposition 1 can be derived and compared with Corollary 1 above. The proof is in the Appendix.

Corollary 2 *If Assumptions 1 and 2 hold, then the solution to Social Planner's Problem with Tagging satisfies, for some pair of tagged groups* $(m, n) \in \{1, 2, ..., M\}$ *such that* $m < n$,

$$\frac{E_i\left[\left(U_{c_*^{i,m}}\right)^{-1}\right]}{E_i\left[\left(U_{c_*^{i,n}}\right)^{-1}\right]} < 1. \tag{22}$$

Corollary 2 states that the planner who puts positive weight on Equal Sacrifice allocates consumption in a way that leaves the cost of raising utility for the disadvantaged group (i.e., m in this example) lower than that for the advantaged group. As shown in result (18), a purely Utilitarian planner would transfer additional resources to the disadvantaged group, but the planner with this more diverse objective stops short, redistributing less. The numerical simulations below reinforce this lesson.[46]

Intuitively, taxing the advantaged tagged group to aid the disadvantaged group generates costs in unequal sacrifice to this planner. A Utilitarian planner ignores the distribution of sacrifice, caring only about total sacrifice (which tagging helps to minimize). This disparity in the treatment of transfers across tagged groups causes an optimal policy based in part on Equal Sacrifice to use tagging less than in conventional theory.

3.1.2 Numerical results on optimal tagging

Next, I use numerical simulations calibrated to micro-level data for the United States to show that the positive optimal tax model developed and estimated in this paper can quantitatively explain the puzzle of tagging. First, I consider three prominent potential tags–height, gender, and race–and show that the parameterizations of the model preferred in the survey of Section 1 yield an optimal policy that rejects the use of these tags but accepts redistributive income taxes driven by differences in income-earning ability. Second, I show that the most preferred parameterization yields a policy that endorses a sizeable and empirically reasonable tag on blindness, one of the few personal characteristics explicitly tagged in the U.S. tax code.[47]

I use the following parameter values in the planning problem of expression (13):

Parameter values							
α_{ES}	ρ	θ	δ	λ	γ	$\frac{1}{\sigma-1}$	G
$\{0.00,\ 0.0\bar{3},\ 0.10,\ 0.20,\ 1.00\}$	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

[45] In particular, the scenario it rules out, where these utility levels coincide, generates complications due to the nondifferentiability of the Equal Sacrifice loss function at the point. An alternative assumption to Assumption 2 that yields the same technical simplification is that $\delta = 0$.

[46] Corollaries 1 and 2 hold in the simulations of Section 6.

[47] For simplicity, I do not consider differences in preferences or elasticities across these groups, though such differences provide an alternative justification for tagging.

values to maximize realism. I set $\delta > 0$ (rather than $\delta = 0$) to capture a less strict version of Equal Sacrifice, in particular one that satisfies the Strong Pareto Efficiency property defined in Section 2.[48] I set $\gamma > 1$ (rather than $\gamma = 1$) to reflect many estimates of the concavity of the utility from consumption that suggest logarithmic utility is too conservative. I set $\sigma = 3$ (rather than $\sigma = 11$) to be closer to mainstream estimates of the labor supply elasticity for a broad population. I set G to approximate the current value of government expenditure as a share of total income in the United States.

Rejecting tagging on height, gender, and race but retaining redistribution

The first data I use allows me to simulate optimal height, gender, and race taxes. To obtain ability distributions by tagged type, I classify respondents to the National Longitudinal Survey of Youth into three height categories, two gender categories, and two race categories.[49] For height, I use gender-dependent ranges, as the height distributions of males and females are substantially different: for men the thresholds are 70 and 72 inches; for women the thresholds are 63 and 66 inches. Table 4 lists the twelve tagged groups that these divisions generate in descending order of their mean wage, where the wage is reported earnings divided by reported hours in 1996.[50] The table shows the mean and standard deviation of each group's reported wages and the population proportion of each group, all adjusted for the NLSY sample weights, as well as each group's raw sample size in the NLSY.

Table 4: Tagged groups

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
	Tall	Med.	Short	Tall	Tall	Short	Med.	Med.	Short	Tall	Med.	Short
	M	M	M	M	F	M	M	F	F	F	F	F
	White	White	White	NW	White	NW	NW	White	White	NW	NW	NW
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
SD wage	11.3	11.0	10.4	12.3	11.6	9.9	10.4	11.6	10.3	5.9	6.2	5.7
Pop. share	0.11	0.13	0.21	0.02	0.09	0.08	0.02	0.14	0.10	0.01	0.03	0.05
Obs.	411	507	785	226	340	994	314	557	405	223	469	653

The differences in wages among these twelve tagged groups are substantial. The highest-earning group in Table 4 earns a mean wage nearly 70 percent greater than the lowest-earning group. Overall, average wages are higher for those who are tall, male, and white. Appendix Table 1 provides more detail than Table 4, reporting the (sample weights-adjusted) distributions of the members of the tagged groups across ten wage bins. These wage distributions are the second key input to the numerical simulations (in addition to the assumed parameters described above).

For each of the five values of α_{ES}, I report measures of the optimal extent of tagging and income tax progressivity in Table 5 and Table 6. To measure the extent of tagging, Table 5 reports 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. This is the ratio of total tax payments to total income for each group under the optimal policy less the same ratio under the constrained-optimal policy

[48] Simulations with the special case of $\delta = 0$ show that the results are virtually identical to those reported in the paper.

[49] I omit individuals who report negative wages or earnings or who report less than 1,000 or more than 4,000 hours of annual work. The results are not sensitive to these restrictions, which are likely to remove misreported data.

[50] Using all three tags in concert maximizes the power of tagging in the conventional model.

with no tagging. If that difference is positive, the group is paying taxes in addition to what it would pay without tagging. If that difference is negative, it is receiving an extra transfer.

Table 5: Extent of Tagging (Extra tax or transfer rate, in percent)

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
	Tall	Med.	Short	Tall	Tall	Short	Med.	Med.	Short	Tall	Med.	Short
	M	M	M	M	F	M	M	F	F	F	F	F
α_{ES}	White	White	White	NW	White	NW	NW	White	White	NW	NW	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

To gauge the progressivity of the optimal income tax, Table 6 reports the average tax rate paid by the members of each wage range under each parameterization.

Table 6: 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

Finally, Table 7 shows the welfare gain obtainable from tagging in each case. To compute this welfare gain, I calculate the increase in consumption for all individuals that would lower the total social loss under the policy without tagging to the level of total social loss obtained by the optimal policy.

Table 7: Welfare Gain from Tagging

α_{ES}	Percent of aggregate consumption
0	0.96
$0.0\overline{3}$	0.45
0.10	0.20
0.20	0.10
1.00	0.00

The results in these three tables show that the support expressed in the survey of Section 1 for objectives that include Equal Sacrifice can explain the coexistence of limited tagging and substantial income redistribution through progressive taxes observed in policy. Table 5 shows that Equal Sacrifice dramatically reduces the appeal of tagging according to height, gender, and race, despite the substantial information that these

three tags carry about income-earning ability. While large group-specific taxes and transfers are optimal when $\alpha_{ES} = 0$ and none are optimal once $\alpha_{ES} = 1$, the values for α_{ES} behind the most favored policies in this paper's survey (namely, $\alpha_{ES} = 0.10$ and $\alpha_{ES} = 0.20$) generate a steep decline in the use of tags. At the same time, for these values of α_{ES}, Table 6 shows that in all cases the extent of redistribution and progressivity remains quite high when measured by either the maximal average tax rate or the gap between the maximal and minimal average tax rates. Table 7 shows that the welfare gains one might achieve through tagging are estimated to be large in the conventional case of $\alpha_{ES} = 0$ but are small in the cases preferred in the survey of Section 1.

As a specific example, consider the single most popular policy in the survey of Section 1, in which $\alpha_{ES} = 0.20$. The optimal tag-based tax is 0.8 percent of the highest-earning group's total income in this parameterization, whereas the conventional model suggests a tax of 10.5 percent. Consistent with this reduced role for tagging, the welfare gain from tagging in this parameterization is negligible: translated into the magnitudes of the current U.S. economy, it is equivalent to approximately $15 billion. Assuming some costs from false tagging and administration (Akerlof 1978), these tags would likely be welfare-reducing, on net, in this parameterization. In contrast, the conventional model implies a gain worth nearly $150 billion. Nevertheless, in this parameterization top earners pay an average tax rate of 47 percent, close to the 53 percent recommended by the conventional model, and a substantial transfer is made to the poor. Moreover, this most-preferred policy recommends a schedule of average tax rates that is quite similar to actual U.S. policy. To see this resemblence, consider Table 8, which gives the schedule of average tax rates for this favored parameterization (policy G) and the schedule implied by the same CBO data used to calibrate the model.[51] The close match between these schedules provides further evidence for the power and relevance of this paper's calibrated positive optimal tax model.

Table 8: Comparing most-preferred policy to actual U.S. average tax rates

	Average tax rate in income percentile range (in percent)							
Status	1st quintile	20-39	40-59	60-79	80-89	90-94	95-98	99
Most-preferred policy ($\alpha_{ES} = 0.20$)	-260	11	16	17	19	21	23	35
Actual U.S. tax schedule (CBO)	-298	-22	10	21	26	28	31	35

The intuition for these results is as follows. The principle of Equal Sacrifice is consistent with the use of progressive taxes to pay for public goods if a given rate of taxation causes a smaller utility loss for a higher-income individual than a lower-income one.[52] But, that principle places little to no value on redistribution.[53] Similarly, while both Utilitarianism and Equal Sacrifice value the efficiency gains from tagging, tagging violates Equal Sacrifice because such personal characteristics have no bearing on individual utility. Altogether, the introduction of Equal Sacrifice considerations into the evaluation of outcomes causes optimal policy to move away from redistribution and, especially, tagging. For the range of parameters

[51] Individual income taxes include taxes on capital income, while the simulations are based on labor income only (to match the model). To calculate tax liability, I multiply each household type's labor share of total market income by its total federal taxes and use the product as the measure of total taxes paid. While variations on this approach affect the estimated rates for top earners, the bottom four quintile's average tax rates are robust to using reasonable alterantive approaches.

[52] 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 $\gamma > 1$.

[53] Note that the average tax rate on the lowest earner when $\alpha_{ES} = 1.0$ is slightly negative in this figure. If $\delta = 0$, the otherwise same simulation sets that average tax rate to a positive value. To see why, note that $\delta = 0$ represents the most severe adherence to Equal Sacrifice, which rejects redistribution. I use $\delta = 0.01$ in the baseline simulation to avoid the concern that $\delta = 0$ is a special case, and $\delta > 0$ causes the purely Equal Sacrifice policy to admit some, although quite limited, redistribution despite the inequality of sacrifice it entails.

considered here, those effects are enough to make the optimal extent of tagging on height, gender, and race negligible but leave substantial redistribution and progressivity intact.

As this intuitive explanation suggests, the key forces determining the optimal extent of tagging in this model will apply to different degrees for different tags. Most important, the costs that tagging generates from the perspective of the Equal Sacrifice principle will be smaller when a tag is closely correlated with ability. If a tag were a perfect indicator of ability, it would generate no costs according to Equal Sacrifice. Given that such a tag would continue to generate efficiency gains by being inelastic to taxation, it would be more valuable to the social planner. In other words, the model suggests that personal characteristics are more likely to be used as tags when they provide stronger and more reliable signals of income-earning ability.[54] I now turn to demonstrating this effect for blindness.

Tagging blindness To demonstrate the model's potential not only to reject most tags but to accept those few tags that predict ability sufficiently well, I consider blindness, one of the few characteristics used as a tag in existing (i.e., U.S.) tax policy.[55] Since 1943, the U.S. tax code has included a special deduction or exemption for individuals with substantially impaired vision. To claim the exemption, individuals simply check a box on their tax forms.

The data source used for the previous tagging analysis has too few observations on the blind, so I combine three years (1985, 1986, and 1987) of the Statistics of Income (SOI) microdata from the U.S. Internal Revenue Service to obtain an earnings distribution of those who claim the blindness exemption. Lacking any information on hours worked, I assume all individuals work the same number of hours (2,000 per year) and calculate hourly wages using individuals' reported wage and salary incomes. I limit the sample to individuals filing as singles, to avoid complications with the proper treatment of couples that are abstracted from in the model above. The distributions of calculated wages, adjusted for sampling weights provided in the SOI, are shown in Table 9. The share of the population in each category also can be estimated from the SOI sample, adjusting for sampling weights. Those claiming the blindness exemption make up 0.3 percent of the population, with 99.7 percent not claiming the exemption.

Table 9: Wage distributions for blind and non-blind

	Average wage rate in range										
Status	0.00	1.73	4.44	7.12	9.60	12.61	15.08	19.17	27.56	44.51	264.19
Blind	0.79	0.08	0.05	0.06	0.02	0.00	0.01	0.005	0.001	0.000	0.000
Not blind	0.17	0.31	0.17	0.12	0.09	0.05	0.03	0.030	0.010	0.002	0.001

As Table 9 makes clear, a large majority of those claiming the blindness exemption earned no wage and salary income and are therefore assigned a zero wage by this calculation. Of course, these individuals would be likely to earn positive wages in the labor market, but we cannot observe those wages, and a zero wage may serve as a rough proxy for a combination of high fixed costs of work and low true wages. Moreover, I will assign all of those who do not claim the blindness exemption but earn zero income a zero wage as well,

[54] The conventional, Utilitarian model also recommends more fully utilizing tags that more accurately signal ability, whether because of fixed costs of tagging or because concavity of individual utility means that tagging errors have some cost. As the simulation results show, however, a plausible calibration of the conventional model recommends dramatically different levels of tagging than that seen in U.S. policy. Explaining these levels as optimal in a conventional model is likely to be difficult.

[55] 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.

so both groups are treated the same.[56]

Table 10 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).

Table 10: Extent of Tagging on Blindness
(Extra tax or transfer rate, in percent)

α_{ES}	Not blind	Blind
0	0.07	-102.11
0.20	0.01	-19.82

As with height, gender, and race, Table 10 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 Table 9, we can calculate mean income for the blind (including those with zero income) to be approximately $2,350 per year. A 20 percent transfer to the blind on average is therefore equivalent to approximately $470, not far from the value of actual blindness deductions and exemptions in the mid-1980s.

3.1.3 Direct evidence that those who oppose tagging support Equal Sacrifice

Here, I present survey results showing that individuals who oppose tagging disproportionately support Equal Sacrifice, augmenting the indirect evidence for this relationship presented above.

In the survey, the respondents see the following text (see the Appendix for the full survey screens):

> Suppose that reliable studies show the following fact: on average, short men earn lower incomes than tall men in the United States. Of course, many short men earn high incomes, and many tall men earn low incomes. But, these studies show that the average income among short men is lower than the average income among tall men. The reasons for this difference are uncertain.
>
> Now, please consider the following proposal: decrease taxes slightly on short men and increase taxes slightly on tall men. Because short men, on average, earn less than tall men, this proposal would, on average, decrease taxes on lower-income men and increase taxes on higher-income men.

A similar question is asked with regard to blindness, for which the respondents are told that the income difference between groups is substantially larger than height. Respondents are asked to say whether they (strongly or somewhat) support or oppose these proposals.

In their responses to this question, those who support a height tag are almost evenly split between the Utilitarian policy A and the Equal Sacrifice policy B, while nearly two-thirds of those who oppose the height tax prefer the Equal Sacrifice policy to the Utilitarian policy. The gaps are yet wider for the proposed blindness tag.[57] These patterns provide direct support to the mechanism proposed in this paper, namely that a concern for Equal Sacrifice limits popular support for tagging.

[56] I excluded those who earn no income from the main analysis of tagging because they are so rare in the NLSY data. However, simulations including these individuals leave the results on height, gender, and race taxation unchanged.

[57] One can look at these data in the "other direction," as well. Namely, 75 percent of those who choose policy A oppose a height tag, compared to 82 percent of those who choose policy B. Only about half of those who choose policy A oppose a blindness tag, compared to more than two-thirds of those who choose policy B.

3.1.4 Relationship between this analysis and the principle of horizontal equity

The way in which this paper's positive optimal tax model resolves the puzzle of tagging reveals that it provides a normatively rigorous foundation for the concern over horizontal equity long intuited as the obstacle to greater tagging.

Concerns over horizontal equity and the other main reasons why tagging may be unappealing in practice have been discussed from the beginning.[58] Akerlof (1978) himself writes: "the disadvantages of tagging... are the perverse incentives to people to be identified as needy (to be tagged), the inequity of such a system, and its cost of administration." Akerlof's first and third disadvantages of tagging are straightforward but of limited effect.[59] In contrast, the concern over horizontal equity–the notoriously difficult-to-define principle that "equals ought to be treated equally"–has been prominent. 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).

As has been widely recognized, the intuitive appeal of horizontal equity is not supported by a solid normative foundation, at least in the context of the optimal tax problem. As Kaplow (2008) writes, the principle of horizontal equity "lacks affirmative justification." The core of the problem for horizontal equity is that its intuitive appeal is based on its guarantee of equal treatment, but in fact it provides only a limited and arbitrary form of equal treatment, namely across those characteristics deemed "horizontal." The principle offers no reliable guidance as to how that classification ought to be made. Musgrave (1959) puts it best: "If there is no specified reason for discriminating among unequals, how can there be a reason for avoiding discrimination among equals?"

Note that the flaws in horizontal equity as it applies to optimal taxation may not matter in other contexts, such as in the use of statistical discrimination or in determining legal liability. If we are not intending to treat individuals differentially across an exogenous dimension of heterogeneity, then horizontal equity may be thought of as equivalent to equal treatment more generically. The Utilitarian optimal tax policy, however, treats individuals of different abilities differently, making horizontal equity's incomplete version of equal treatment apparent.[60]

In contrast, the principle of Equal Sacrifice is a comprehensive criterion of optimal taxation with a solid normative foundation of equal treatment for all individuals that, as one of its outcomes, discourages tagging. In other words, rather than a requirement of horizontal equity acting as an ad hoc explanation for limited tagging, in this paper a concern for horizontal equity arises endogenously out of the classic principle of Equal Sacrifice.[61]

[58] Additional concerns about tagging exist. First, tagging could induce stigma. Stigma in this context is plausibly related to the normative appeal of equal sacrifice, as those receiving tag-based transfers would be sacrificing less. Second, tagging could slow the resolution of underlying distortions. If those distortions are due to irrational behavior by employers, it is unclear why tagging would exacerbate their mistakes. If not, the distortions are likely to be persistent. Third, tagging may be against the laws or constitutions of various nations. Any such prohibitions on tagging beg the question of why they are accepted by voters.

[59] Tags are undoubtedly less appealing if they are easily mimicked–as they would then distort behavior while failing to redistribute–or costly to monitor and administer. Most of the candidate tags mentioned above and considered in modern tax theory, however, are inelastic and cheap to enforce. Even a statistic such as "apparent I.Q.", which may seem both elastic and costly to monitor, has such large implications outside the tax system for individuals that we might argue it would be largely immune to these concerns. Mirrlees (1971) makes the same point on I.Q. See page 208. Certainly a characteristic such as gender is highly inelastic and could be cheaply incorporated into the tax system.

[60] I am grateful to Yoram Margalioth and to a referee for comments that prompted this discussion. An early analysis by Balcer and Sadka (1982) examined the conditions under which a Utilitarian tax policy would violate or respect horizontal equity.

[61] This appealing property of Equal Sacrifice is related to what Berliant and Gouveia (1993) label "an ethically undesirable property of the solutions implied by the [conventional] optimal income tax formulations: the marginal tax faced by a taxpayer depends, other things equal, on the density of the population on the domain of taxpayer characteristics."

Finally, while the distinction between Equal Sacrifice and horizontal equity may seem subtle, the paper presents substantial evidence on prevailing normative preferences that cannot be explained as a concern for horizontal equity. The survey evidence of Section 1 shows implicit support for Equal Sacrifice (or a similar criterion) through respondents' choices over income tax policies for which horizontal equity is irrelevant. That survey also showed that most individuals directly support, when asked, a more equal distribution of sacrifice than the Utilitarian criterion would recommend, even if it were augmented with a concern for horizontal equity. Moreover, neither of the additional puzzles addressed in Section 3.2 below can be addressed by horizontal equity. Together with the puzzle of tagging, these additional findings suggest that a criterion broader than horizontal equity, but one that also yields a similar intuition as does horizontal equity in the context of tagging, is a required component of a viable positive optimal tax model.

3.2 Using the generalized model to address two additional puzzles

Next, I show that the same favored calibration of this paper's positive optimal tax model can help address two additional puzzles in optimal tax research.

3.2.1 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, especially, tax law scholars (see, for example, King 1983 and Zelenak 2006).

In this section, I use a detailed calibration of the U.S. ability distribution[62] to simulate first-best (i.e., full information) feasible income tax policies for a range of model parameterizations. Figure 5 shows the results by plotting utility as a function of ability in four scenarios.

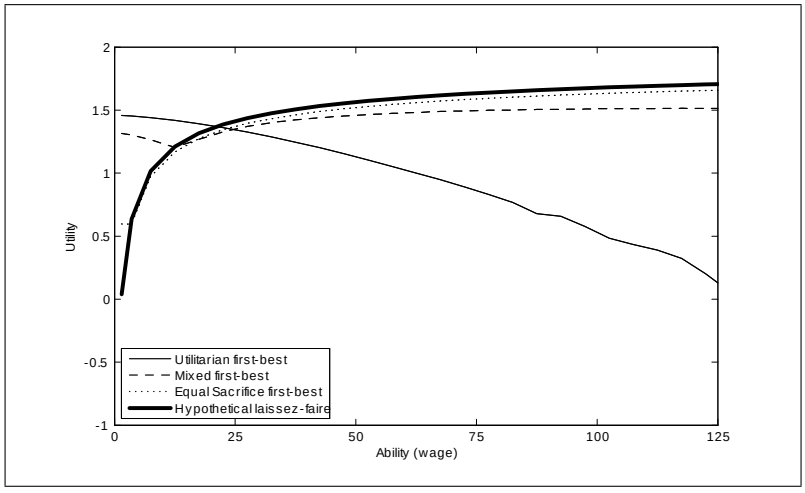


Figure 5: Utility levels by ability type under different objective functions in full-information first-best allocations.

[62] The previous section's simulation used a calibration of the U.S. ability (i.e., wage) distribution that was limited by the availability of tagging data. Here, I use a lognormal-Pareto calibration of the U.S. wage distribution originally calculated by Mankiw, Weinzierl, and Yagan (2009).

The thick solid line is the hypothetical laissez-faire allocation in which no taxes are collected. It shows how utility increases monotonically with ability absent government intervention. The other three lines show utilities under three parameterizations: the thin solid line is for the Utilitarian case of $\alpha_{ES} = 0.00$, the dotted line is for the Equal Sacrifice case of $\alpha_{ES} = 1.00$, and the dashed line is for the mixed case preferred in the survey of Section 1 in which $\alpha_{ES} = 0.20$. All other parameters are as in the tagging simulations above (though G is adjusted to be a similar share of total income).

Figure 5 shows the rank reversals when going from the laissez-faire or Equal Sacrifice allocations to the Utilitarian allocation, as the upward sloping thick solid and dotted lines contrast sharply with the downward sloping thin solid line. The empirically-preferred, mixed objective (shown as the dashed line) generates a far more uniform utility distribution than either of the more pure objective functions. More important, the empirically-preferred objective chooses a first-best allocation that substantially limits rank reversals. The reason for this result is that Equal Sacrifice's optimal allocation reduces each individual's utility by the same quantity and thus leaves the utility ordering of agents unchanged. In the mixed objective functions used here, the Utilitarian preference for rank reversals is tempered.

3.2.2 Optimal top marginal income tax rates

Finally, I use the more detailed ability distribution from the previous simulation to explore in depth the effects of a role for Equal Sacrifice on optimal marginal income tax rates. I use the same preferred calibration ($\alpha_{ES} = 0.20$) in which the optimal policy rejected height, gender, and race tags, accepted tagging on blindness and substantial redistribution, and largely avoided rank reversals in utility in the first-best. All other parameters are as before.[63]

Figure 6 shows the optimal schedule of marginal tax rates for this calibration, calculated as the distortions to individuals consumption-leisure margins (as defined formally in the Appendix). For comparison, the figure also shows the optimal results under a pure Utilitarian criterion, that is when $\alpha_{ES} = 0$ as in the conventional model.

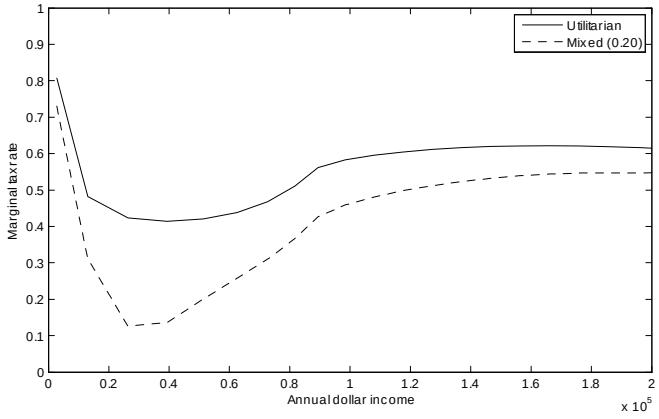


Figure 6: Optimal Marginal Tax Rates

Figure 6 shows that the optimal marginal income tax rate at high incomes falls substantially, by about seven percentage points, with the empirically-preferred role for the principle of Equal Sacrifice. The marginal

[63] The appendix contains analytical results on optimal marginal rates in this model.

tax rate schedule has the U-shape introduced by Diamond (1998) and Saez (2001) whether $\alpha_{ES} = 0$ or 0.20, though positive α_{ES} does lead to lower rates for all workers. The explanation for this pattern is that the planner with $\alpha_{ES} > 0$ redistributes less. This reduces higher earners' temptation to mimic lower income earners and thus the required distortions throughout the income distribution. At the same time, substantial redistribution persists despite this role for the principle of Equal Sacrifice. Though not shown here, the high-skilled continue to pay sizeable average tax rates of 45 percent, not far from the 54 percent rate under the Utilitarian policy. A related result is that the lowest-ability type enjoys a level of consumption worth 52 percent of average consumption in the economy under the policy with $\alpha_{ES} = 0.20$ compared to 63 percent under the Utilitarian policy with $\alpha_{ES} = 0$.

Therefore, this paper's calibrated positive optimal tax model may help address a gap between conventional theory and existing policy noted by Diamond and Saez (2011). Using the conventional model, they conclude that the optimal top rate is "73 percent, substantially higher than the current 42.5 percent top US marginal tax rate (combining all taxes)." The top rate in the mixed policy shown in Figure 6 is 55 percent, compared to 62 percent under the conventional Utilitarian criterion. Of course, a number of other potential explanations exist for why top marginal tax rates are not higher, such as a higher elasticity of taxable incomes at high income levels or the existence of preference heterogeneity. But these simulation results suggest that a difference between the objective assumed in the conventional model and the prevailing normative preferences in society may play a role.

4 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.

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) prefer policies that reflect a mixed objective that gives weight to both Utilitarianism and 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 has remarkable explanatory power as a positive optimal tax model. I focus on the model's ability to explain the limited role of tagging in policy that is otherwise quite redistributive. I show, analytically, that a concern for Equal Sacrifice limits the optimal extent of tagging. I then 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. Moreover, 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. In sum, optimal policy in this calibrated model matches remarkably well several prominent characteristics of existing policy that are puzzling from the perspective of conventional theory but widely endorsed in reality.

The traditional goal of optimal tax research has been to choose the "right" normative objective for policy and characterize the tax system that best attains it. Public opinion on the appropriate normative criterion has been seen as beside the point. An alternative goal, pursued in this paper, is to characterize the tax system that best attains the normative objective that prevails in reality. Recognizing that moral authority in reality rests not with tax theorists but with voters and policymakers, this paper heeds the advice of Samuelson (1980) that "Basic questions concerning right and wrong goals to be pursued cannot be settled by mere science as such....The citizenry must ultimately decide such issues."

References

[1] Akerlof, George, (1978). "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," American Economic Review, 68(1), March, pp. 8-19.

[2] Atkinson, Anthony B. and Joseph Stiglitz, (1980). Lectures on Public Economics. McGraw-Hill.

[3] Auerbach, Alan, and Kevin Hassett, (1999). "A New Measure of Horizontal Equity." Working Paper.

[4] Boadway, Robin and Pierre Pestieau, (2006). "Tagging and Redistributive Taxation," Queen's University Working Paper 5-2006.

[5] Balcer, Yves and Efraim Sadka (1982), "Horizontal Equity, Income Taxation and Self-Selection with an Application to Income Tax Credits," Journal of Public Economics, 19.

[6] Berliant and Gouveia (1993). "Equal sacrifice and incentive compatible income taxation," Journal of Public Economics, 51.

[7] Besley, Timothy, and Stephen Coate (1992). "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," American Economic Review 82(1).

[8] Boaz, David and Robin Kirby (2007)."Libertarian Voters in 2004 and 2006", CATO Policy Report, January/February.

[9] Bolton, Gary and Axel Ockenfels (2000). A Theory of Equity, Reciprocity, and Competition," American Economic Review, 90(1)

[10] Bolton, Gary and Axel Ockenfels (2006). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiements: Comment" American Economic Review, 96(5).

[11] Bourguignon, Francois and Amedeo Spadaro (2012), "Tax-benefit revealed social preferences," Journal of Economic Inequality, 10.

[12] Cappelen, Alexander W., Astri Drange Hole, Erik O. Sorensen, and Bertil Tungodden, (2011). "The Pluralism of Fairness Ideals: An Experimental Approach," American Economic Review 97(3).

[13] Charness, Gary and Matthew Rabin (2002). "Understanding Social Preferences with Simple Tests," Quarterly Journal of Economics, 117(3).

[14] Cremer, Helmuth, Firouz Gahvari, and Jean-Marie Lozachmeur (2010). "Tagging and Income Taxation: Theory and an Application," American Economic Journal: Economic Policy 2(1).

[15] da Costa, Carlos and Thiago Pereira (2012). "Efficiency of Equal Sacrifice Income Tax Schedules," Working paper.

[16] Deutsch, Morton (1985). Distributive Justice: A Social-Psychological Perspective. Binghamton: NY: Vail-Ballou Press.

[17] Diamond, Peter, (1998). "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," American Economic Review, 88(1).

[18] Diamond, Peter and Emmanuel Saez (2011). "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," Journal of Economic Perspectives 25(4).

[19] Engelmann, Dirk and Martin Strobel (2004). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiements," American Economic Review, 94(4).

[20] Engelmann, Dirk and Martin Strobel (2006). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiements: Reply," American Economic Review, 96(5).

[21] Feldman, Stanley and John Zaller (1992), "The Political Culture of Ambivalence: Ideological Responses to the Welfare State," American Journal of Political Science 36(1).

[22] Farhi, Emmanuel and Ivan Werning (2010). "Progressive Estate Taxation", Quarterly Journal of Economics, 125(2).

[23] Feldstein, Martin (1976). "On the Theory of Tax Reform," Journal of Public Economics, 6, pp. 77-104.

[24] Fehr, Ernst and Klaus M. Schmidt (1998). "A Theory of Fairness, Competition, and Cooperation," Quarterly Journal of Economics, August.

[25] Fehr, Ernst, Michael Naef, and Klaus M. Schmidt (2006). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiements: Comment" American Economic Review, 96(5).

[26] Fleurbaey, Marc and Francois Maniquet (2006). "Fair Income Tax," Review of Economic Studies, 73, pp. 55-83.

[27] Fong, Christina. (2001). "Social preferences, self-interest, and the demand for redistribution," Journal of Public Economics, 82.

[28] Free, Lloyd A. and Hadley Cantril, (1968). The Political Beliefs of Americans. New York: Simon and Schuster.

[29] Friedman, Milton (1962). Capitalism and Freedom. Chicago: University of Chicago Press.

[30] Frolich, Norman, Joe A. Oppenheimer, and Cheryl L. Eavey, (1987). "Laboratory Results on Rawls's Distributive Justice," British Journal of Political Science, 17.,

[31] Frolich, Norman and Joe A. Oppenheimer (1992). Choosing Justice: An Experimental Approach to Ethical Theory. Berkeley, CA: UC Press.

[32] Frohlich, Norman, Joe Oppenheimer, and Anja Kurki, (2004). "Modeling Other-Regarding Preferences and an Experimental Test," Public Choice 119(1/2), April.

[33] Gainous, Jason, and Michael D. Martinez (2005). "What Happens When We Simultaneously Want Opposite Things? Ambivalence about Social Welfare." Ambivalence, Politics, and Public Policy, edited by Stephen C. Craig and Michael D. Martinez, Palgrave.

[34] Golosov, Mikhail and Aleh Tsyvinski (2006). "Designing Optimal Disability Insurance: A Case for Asset-Testing," Journal of Political Economy, 114(2).

[35] Harsanyi, John C. (1953) "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," Journal of Political Economy, 61(5), (October), pp. 434-435.

[36] Harsanyi, John C. (1955). "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," Journal of Political Economy, 63(4), (August), pp. 309-321.

[37] Hochschild, Jennifer (1981. What's Fair? American Beliefs about Distributive Justice. Cambridge, MA: Harvard University Press.

[38] Horton, John J., David G. Rand, and Richard J. Zeckhauser, "The Online Laboratory: Conducting Experiments in a Real Labor Market," Experimental Economics.

[39] Kant, Immanuel (1785). Groundwork of the Metaphysics of Morals. See section 4:434 of the Cambridge University Press edition, 1997, Mary Gregor, ed.

[40] Kaplow, Louis and Steven Shavell (2001). "Any non-welfarist Method of Policy Assessment Violates the Pareto Principle," Journal of Political Economy 109(2).

[41] Kaplow, Louis (2008). *The Theory of Taxation and Public Economics*, Princeton.

[42] King, Mervyn A. (1983). "An Index of Inequality: With Applications to Horizontal Equity and Social Mobility," Econometrica.

[43] Konow, James (2001). "Fair and square: the four sides of distributive justice," Journal of Economic Behavior and Organization 46.

[44] Konow, James (2003). "A Positive Analysis of Justice Theories," Journal of Economic Literature XLI, December.

[45] Kuziemko, Ilyana, Michael Norton, Emmanuel Saez, and Stefanie Stantcheva, (2013). "How Elastic are Preferences for Redistribution: Evidence from Randomized Survey Experiments," Working Paper.

[46] Lockwood, Benjamin B. and Matthew Weinzierl (2012). "De Gustibus non est Taxandum: Theory and Evidence on Preference Heterogeneity and Redistribution," NBER Working Paper 17784.

[47] Mankiw, N. Gregory and Matthew Weinzierl (2009). "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," American Economic Journals: Economic Policy.

[48] Mill, John Stuart (1871). Principles of Political Economy. Oxford University Press, 1994. Book V, page 169.

[49] Miller, David (1976). Social Justice. London: Oxford University Press.

[50] Mirrlees, J.A., (1971). "An Exploration in the Theory of Optimal Income Taxation," Review of Economic Studies 38, 175-208.

[51] Mirrlees, J.A., S. Adam, T. Besley, R. Blundell, S. Bon, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba (eds.) (2010) Dimensions of Tax Design: The Mirrlees Review. Oxford University Press.

[52] Mitchell, Gregory, Philip E. Tetlock, Barbara A. Mellers, and Lisa D. Ordonez (1993). "Judgments of Social Justice: Compromises between Equality and Efficiency," Journal of Personality and Social Psychology 65(4).

[53] Musgrave, Richard (1959). The Theory of Public Finance. McGraw-Hill.

[54] Nozick, Robert (1974). Anarchy, State, and Utopia. New York: Basic Books.

[55] Parsons, Donald O. (1982). "Imperfect 'tagging' in Social Insurance Programs. Journal of Public Economics.

[56] Rawls, John (1971). A Theory of Justice. Cambridge: Harvard University Press.

[57] Rothschild, Casey and Florian Scheuer (2012). "Optimal Taxation with Rent-Seeking," NBER Working Paper, 17035.

[58] Saez, Emmanuel, (2001). "Using Elasticities to Derive Optimal Income Tax Rates," Review of Economic Studies, 68.

[59] Saez, Emmanuel (2002). "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," Journal of Public Economics, 83: 217-230.

[60] Saez, Emmanuel and Stefanie Stantcheva (2012). "Optimal Tax Theory with Endogenous Social Marginal Welfare Weights." Working paper.

[61] Samuelson, Paul (1980). Economics 11th edition. McGraw Hill.

[62] Scott, John T., Richard E. Matland, Philip A. Michelbach, and Brian H. Bornstein (2001). "Just Deserts: An Experimental Study of Distributive Justice Norms," American Journal of Political Science 45(4).

[63] Sen, Amartya, (1982). "Rights and Agency," Philosophy and Public Affairs, 11(1), Winter, pp. 3-39.

[64] Sen, Amartya, (2000). "Consequential Evaluation and Practical Reason." Journal of Philosophy, 97(9), September, pp. 477-502.

[65] Stiglitz, Joseph E., (1987). "Pareto efficient and optimal taxation and the new new welfare economics," in: Auerbach, Alan and Martin Feldstein (eds.), Handbook on Public Economics. North Holland: Elsevier.Science Publishers, pp. 991-1042.

[66] Weinzierl, Matthew (2011). "The Surprising Power of Age-Dependent Taxes," Review of Economic Studies, 78(4).

[67] Weinzierl, Matthew (2012). "Why do we Redistribute so Much but Tag so Little? Normative Diversity, Equal Sacrifice, and Optimal Taxation," NBER Working Paper 18045.

[68] Werning, Ivan, (2007). "Pareto Efficient Income Taxation," Working Paper, MIT, April.

[69] Young, H. Peyton (1987). "Progressive taxation and the equal sacrifice principle," Journal of Public Economics, 32(2).

[70] Young, H. Peyton (1988). "Distributive Justice in Taxation," Journal of Economic Theory, 44.

[71] Young, H. Peyton (1990). "Taxation and Equal Sacrifice," American Economic Review, 80(1).

[72] Young, H. Peyton (1994).Equity: Theory and Practice. Princeton: Princeton University Press.

[73] Zelenak, Lawrence. (2006). "Taxing Endowment," Duke Law Journal.