NBER WORKING PAPER SERIES

WHY DO WE REDISTRIBUTE SO MUCH BUT TAG SO LITTLE? THE PRINCIPLE OF EQUAL SACRIFICE AND OPTIMAL TAXATION

Matthew C. Weinzierl

Working Paper 18045
http://www.nber.org/papers/w18045

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

I am grateful to Felix Bierbrauer, Kim Clausing, Raj Chetty, Mihir Desai, Rafael di Tella, Amy Finkelstein, Victor Fleischer, John Friedman, Alex Gelber, Caroline Hoxby, Bas Jacobs, Louis Kaplow, Wojciech Kopczuk, Camille Landais, Ben Lockwood, Greg Mankiw, Yoram Margalioth, Joe Mazor, Jean-Baptiste Michaud, Je¤ Miron, Alex Raskolnikov, Meg Rithmire, Julio Rotemberg, Emmanuel Saez, Florian Scheuer, Ali Shourideh, Stefanie Stantcheva, Alain Trannoy, Aleh Tsyvinski, Glen Weyl, John Weymark, and Danny Yagan 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.

Why do we Redistribute so Much but Tag so Little? The principle of equal sacrifice and optimal taxation
Matthew C. Weinzierl
NBER Working Paper No. 18045
May 2012, Revised August 2012
JEL No. D63,H2,H21

ABSTRACT

The workhorse model of optimal taxation strongly recommends tagging, but its use in policy is limited. I argue that this puzzle is a symptom of a more fundamental problem. Conventional theory neglects the diverse normative criteria with which, as extensive evidence has shown, most people evaluate policy. In particular, if the classic principle of Equal Sacri…ce augments the standard Utilitarian criterion, optimal tagging is limited. Calibrated simulations of optimal policy with normative diversity of this type simultaneously match three features of U.S. policy: substantial income redistribution; rejection of gender, race, and height tags; and acceptance of a blindness tag. Additional implications increase the appeal of this revision to conventional theory.

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 entitled *The Mirrlees Review* (2010).

It is a puzzle, then, that one of this conventional theory's clearest implications stands in stark contrast with real-world policy. "Tagging" is the dependence of taxes on personal characteristics, such as gender. Conventional theory recommends widespread tagging because it can achieve redistribution of income without distorting incentives to work. In so doing, tagging sidesteps the tradeoff between equality and efficiency that is at the heart of conventional theory. Real-world taxes, however, make only limited use of tagging. In other words, tagging seems like a free lunch that real-world policy is largely leaving on the table.

This paper argues that the puzzle of limited tagging is a symptom of a more fundamental problem: conventional optimal tax theory evaluates policy based on a criterion that is unrealistically narrow. That criterion, Utilitarianism, is powerful and compelling to many. But extensive evidence has shown that no single criterion, however appealing, can claim to be *the* criterion of the tax design problem as society perceives it. Different people find different criteria compelling, and most people find multiple criteria partially compelling. Specifically, not all people are Utilitarians and, perhaps more important, most people are not all Utilitarian.

The natural solution to this problem is to incorporate into optimal tax theory a realistically diverse normative perspective that includes compelling alternative criteria.[1] This paper develops an approach for doing so and then shows that an application of that approach can resolve the puzzle of limited tagging.

In particular, if the classic principle of Equal Sacrifice augments the standard Utilitarian criterion, tagging is no longer a free lunch, even in theory. 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 will determine whether to use a given tag by weighing the costs of such violations against the gains it generates according to Utilitarianism.

Equal Sacrifice is a compelling additional criterion of optimal taxation for at least two reasons, aside from its ability to resolve the puzzle studied here. First, its connection to the philosophical framework of Libertarianism implies that it may serve as the main criterion of optimal policy for a portion of society. Second, and arguably more important, it is likely to capture one component of the mix of criteria used by most individuals. No less a Utilitarian than John Stuart Mill, for example, endorsed it as the proper criterion for taxation.[2] I develop the case for Equal Sacrifice in detail below, but the key to its broad appeal can be succinctly stated. Equal Sacrifice rejects the assumption implicit in the conventional approach that, as summarized in a critique by Martin Feldstein (1976), "all property and individual abilities should be regarded as society's common resource."

[1] Examining the degree of correspondence between policy and theory is a central task of optimal tax research (e.g., Saez 2001). To some, that task uncomfortably blurs the line between normative and positive analysis. In my opinion, that is a false dichotomy. The most valuable goal of optimal tax research is not to impose a "correct" normative perspective on policy, but rather to uncover and apply the normative perspective that members of society agree upon through the political process. Given that goal, gaps between a model and policy present opportunities to find reasonable modifications to the theory or to reject a policy as suboptimal. See Feldstein (1976) for an early statement of a similar perspective.

[2] See the discussion in Section 1 below.

The payoff from recognizing and incorporating this[3] realistic form of normative diversity in the optimal tax model goes beyond simply providing a reason to avoid tagging. A role for Equal Sacrifice can improve the match between the recommendations of theory and the reality of policy in several ways.

A role for Equal Sacrifice can explain not only why tagging is limited but also the circumstances under which tagging is likely to be optimal. Tags that are strongly correlated with underlying income-earning ability, such as blindness and disability status, will generate smaller losses according to Equal Sacrifice and therefore be more acceptable to a society that balances that principle and Utilitarianism when evaluating policy.[4]

Moreover, incorporating Equal Sacrifice can explain the coexistence of limited tagging and substantial income redistribution, a challenge for conventional theory. The Equal Sacrifice principle is consistent with progressive taxation to pay for government spending if a given rate of taxation causes a smaller utility sacrifice for a higher-income individual than a lower-income one.[5] Thus, even if Equal Sacrifice is sufficiently important to severely limit tagging, it may have much milder effects on the conventional theory's implications for progressivity and redistribution.

In this paper, I show that the potential for the Equal Sacrifice principle to explain these features of tax policy is not merely theoretical. Using microeconomic data on earnings and personal characteristics, I find calibrations of the optimal policy model with normative diversity of this type in which society rejects the use of three prominently-proposed tags–height, gender, and race–and accepts both sizeable tagging of blindness and substantial redistribution through progressive income taxes, all as in the U.S. tax code.

Finally, if correct, this paper's explanation for limited tagging has additional implications. In fact, it may contribute to resolving two additional high-profile puzzles in optimal tax research.

First, Equal Sacrifice offsets an implication of Utilitarianism that has long been a cause for concern: rank reversals in the first best. In the standard model, if information is complete so that types are observable, higher-ability types are generally left with less utility than lower ability types despite their ability to achieve greater utility absent government intervention. This result has struck many as normatively undesirable.[6] The Equal Sacrifice principle discourages rank reversals because it recommends reducing each individual's utility from a laissez-faire starting point by the same absolute quantity. In the calibrated simulation described above, I show that an optimal tax policy with a role for Equal Sacrifice dramatically reduces rank reversals present in the Utilitarian first-best.

Second, Diamond and Saez (2011) show that the top marginal tax rate for the United States, according to standard theory, is substantially (30 percentage points) higher than in current U.S. policy. Again using the same calibration of the model with normative diversity described above, I show that the optimal top marginal tax rate falls by seven percentage points relative to a conventional Utilitarian model. While the magnitude of this effect would vary across calibrations, the lesson is clear: if we care about Equal Sacrifice and avoid most tags because of it, we are likely to moderate our use of high marginal income tax rates at the top of the income distribution.

[3] Using the approach described by this paper to include other normative frameworks, such as the Rawlsian priority on the least fortunate, would likely lead to further insights on existing policies.

[4] Note that this result contrasts with conventional theory, which endorses tags on inelastic and easily enforced traits regardless of the strength of their correlation with ability.

[5] Equal Sacrifice rejects redistribution, but to the extent that Utilitarian reasoning causes society to want to support low-ability individuals with redistribution, the optimal sharing of that burden according to Equal Sacrifice, much less to Utilitarians, is likely to be progressive.

[6] For example, Zelenak (2006) writes: "Few people may remain utilitarians, however, if that requires accepting the conclusion that the ideal tax-and-transfer system would make the most talented members of society the least well-off." In the second-best solutions to conventional models, rank reversals are avoided only because the incentive constraints guarantee the higher types more desirable allocations.

The revision to conventional theory proposed in this paper preserves the core of the Mirrlees (1971) approach to optimal tax, an approach many (including this author) consider fundamentally convincing as a basis for policy evaluation.[7] At the same time, because alternative normative criteria imply different standards for measuring policy optimality, this revision has a potentially wide range of implications beyond those explored here.[8]

This paper proceeds as follows. Section 1 presents the case for diversifying the normative objective of the optimal tax model and for including Equal Sacrifice as one component of that objective. Section 2 briefly reviews prior results and discussions of tagging in optimal tax research, including a discussion of the relationship between this paper and the concept of horizontal equity. Section 3 generalizes the conventional optimal tax model to include multiple normative criteria. Section 4 applies this generalized model in the case of two criteria, Utilitarianism (i.e., maximal aggregate utility) and Equal Sacrifice (i.e., uniform utility losses). Section 5 derives conditions on optimal tagging and marginal distortions to labor supply in that model, formally establishing that a role for the principle of Equal Sacrifice reduces optimal tagging. Section 6 performs calibrated numerical simulations of optimal policy and discusses their implications, and section 7 concludes.

1 The case for normative diversity and Equal Sacrifice

Starting with Mirrlees (1971), conventional optimal tax analysis has assumed the straightforward normative criterion of generalized Utilitarianism. According to Utilitarianism, a social planner ought to maximize the sum of the utilities of a population of individuals, in some cases applying a concave transformation to the utilities before summing. Combined with the assumptions that individuals differ in their innate ability to earn income and that preferences over consumption and leisure are common, this Utilitarian criterion powerfully recommends income redistribution.[9]

Economists, and especially optimal tax theorists, have been largely united around this Utilitarian perspective.[10] The canonical justification for it is due to John Harsanyi (1953, 1955), who argued that "value judgments concerning social welfare and the cardinal utility maximized in choices involving risk may be regarded as being fundamentally based upon the same principle." In other words, expected utility maximization, when the expectation is taken over all individuals in society, is pure-sum Utilitarianism.

[7] This feature of the approach taken here can be seen as answering the question raised in Mankiw and Weinzierl (2010) of whether limited tagging could be explained without jettisoning conventional theory.

[8] For example, the mixed normative perspective detailed below may affect optimal commodity taxes (e.g., capital taxes) in the presence of preference heterogeneity. More broadly, this paper's critique may apply to any research, on taxes or other policies, that implicitly adopts the Utilitarian approach when making reduced-form welfare calculations.

[9] See Lockwood and Weinzierl (2012) for a treatment of optimal taxation with preference heterogeneity.

[10] Though generalized Utilitarianism dominates optimal tax research, especially when the theory is made quantitative in numerical simulations, some important exceptions exist. Feldstein (1976), as noted elsewhere in this paper, was an early critic of the purely Utilitarian approach in the optimal tax literature. Sen and Williams (1982) collect and analyze a number of critiques and defenses of Utilitarianism, including from Mirrlees. Stiglitz (1987) and Werning (2007) describe Pareto-optimal taxation. Their efforts are similar in spirit to mine, in that they widen the model's normative perspective. They differ, however, in that my approach provides a way to include a specific combination of normative perspectives held by society, while these authors remain agnostic and, therefore, are able to provide less specific guidance to or explanation of policy. Related, recent research by Saez and Stantcheva (2012) focuses on marginal social welfare weights through which tax reforms may be evaluated. They allow these weights to take any positive values, including values based on principles or priorities at odds with Utilitarianism. Their approach is complementary to mine, in that they focus on the welfare weights that one might derive from, at least in part, the normative reasoning I model directly. Finally, 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 could accommodate these concerns.

1.1 Normative diversity

As even a casual observer of policy debates can attest, discussions of taxes by members of the public, policymakers, and scholars do not reflect such a pure normative perspective. While some individuals find the Utilitarian criterion appealing, others are drawn to sharply opposing criteria. For example, Milton Friedman wrote in 1962: "I find it hard, as a liberal, to see any justification for graduated taxation solely to redistribute income. This seems to me a clear case of using coercion to take from some in order to give to others and thus to conflict head-on with individual freedom." Statements, shown below, in which John Stuart Mill and Henry Sidgwick endorse the principle of Equal Sacrifice further demonstrate that the current scholarly convergence on a Utilitarian perspective is at odds with important and long-lived strains of thinking on the topic.

More generally, decades of research in psychology, political science, and economics has shown that most individuals, whether supportive of Utilitarianism or not, are not normative purists. Those who are fully convinced by a single criterion are best seen as outliers occupying the extremes of a continuum, the interior of which is populated by those for whom multiple normative criteria have appeal. For example, Frohlich, Oppenheimer, and Eavey (1987) 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]." Similarly, Scott, Matland, Michelbach, and Bornstein (2001) write: "Experimental research reveals that distributive justice judgments usually involve several distinct allocation principles."[11] Policy driven by individuals (i.e., voters) with this normative ambivalence will therefore balance competing criteria.

The prevalence of normative ambivalence begs the question of why most people appear to find Utilitarianism less appealing than standard optimal tax analysis implies. Perhaps the most prominent critique of the purely Utilitarian perspective is that, in the words of John Rawls (1971), Utilitarianism "does not take seriously the distinction between persons." Rawls, and others, are concerned that the Utilitarian's willingness to trade the losses of some for greater gains of others may, in some cases, compromise individual liberty.[12]

Strikingly, the specific context in which this concern has been seen as most forceful is "endowment" taxation, where individuals would be taxed on their potential to earn income rather than their actual earned income. Of course, endowment taxation is exactly the preferred policy of the conventional Utilitarian optimal tax model. Rawls (2001) argued 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."[13] The broad

[11] In addition to works cited in the text, see: Deutsch (1985), Feldman and Zaller (1992); Free and Cantril (1968); Frohlich and Oppenheimer (1992); Gainous and Martinez (2005) ; Hochschild (1981); Konow (2001); Miller (1976); and Mitchell, Tetlock, Mellers, and Ordonez (1993). In a valuable review of empirical findings, Konow (2003) argues that "each category [of justice theories] captures an element that is important to crafting a positive theory of justice but that no single family or theory within a family suffices to this end." Other relevant findings include the following. Feldman and Zaller (1992) state: "Our results offer strong support to studies, especially that of Hochschild, that have identified ambivalence as a fundamental feature of political belief systems...Even those who take consistently pro- or consistently antiwelfare positions often cite reasons for the opposite point of view." The Hochschild (1981) study to which they refer consisted of long-term, in-depth interviews of a group of individuals across a wide range of socioeconomic status. It concluded: "Some people...hold beliefs that are predominately clear and sharp–but even they express some ambivalence. Others...hold beliefs that are predominately ambivalent and blurred–but even they express the dominant pattern much of the time." Similarly, Gainous and Martinez (2005) conclude that "a sizable chunk of the American public is, in fact, ambivalent to some degree about social welfare."

[12] 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."

[13] 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,

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.[14]

Tagging is a clear, though perhaps more mild, example of a case in which the conventional model's Utilitarianism raises this concern. Tags treat an individual as a collection of characteristics whose statistical relationships with innate ability affect the individual's tax treatment. In other words, tagging neglects an individual's specific circumstances and, instead, taxes him or her based on patterns that hold across the population in aggregate. To tagging's critics, this seems an unjust basis for taxation.

1.2 Equal Sacrifice

As a prominent criterion that respects the individuality of taxpayers, Equal Sacrifice has a strong claim to being an important component of a realistically diverse normative framework for optimal taxation.

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

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

Mill's vision of 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." More recently, the late 1980s and 1990s saw a temporary resurgence of interest in Equal Sacrifice as a basis for policy, especially through the work of H. Peyton Young (1987, 1988, 1990, 1994) but also including Yaari (1988), Moyes (1989), Berliant and Gouveia (1993), Ok (1995), Mitra and Ok (1996), and D'Antoni (1999). That literature established conditions on the progressivity of taxes designed in accordance with Equal Sacrifice, and it argued for the centrality of that principle from both normative and positive perspectives.[15]

One channel through which Mill's principle of Equal Sacrifice enters the diverse normative perspective apparent in modern tax policy debates is through its connection to the influential philosophical framework of Libertarianism. Under standard versions of Libertarianism, taxes are justified to pay for public goods only. While Libertarian writers are frustratingly imprecise about how they would allocate required taxes,[16] one natural benchmark for doing so is to share the (utility) costs of taxation equally across individuals. Martin Feldstein (1976) made this link explicit in his critique of the Utilitarian focus of early optimal tax research: "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."[17] Similarly,

for instance, Hasen (2007), Markovits (2003), Rakowski (2000), Shaviro (2002), Stark (2005), Sugin (2011), and Zelenak (2006).

[14] 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 man, in your own person as well as in that of anyone else, always as an end, never merely as a means."

[15] Lambert and Naughton (2009) is a recent contribution that reviews much of this literature.

[16] Friedman (1962) does endorse a flat-rate income tax, but on intuitive rather than rigorous grounds.

[17] Robert Nozick is an influential modern expositor of Libertarianism.

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. Public opinion surveys estimate the proportion of individuals with traditional Libertarian views to be 10 to 20 percent in the United States (Boaz and Kirby 2007). 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."[18] The connection between equal sacrifice and Libertarianism therefore implies a sizeable portion of society may reasonably be described as using the equal sacrifice principle as its main criterion for optimal tax policy.

A second and, arguably, more important channel of influence for Equal Sacrifice is its place in the moral reasoning of the majority of individuals who feel normative ambivalence. Research has shown that even those predisposed toward redistribution feel some pull toward normative principles that, like Equal Sacrifice, avoids Rawls' and Nozick's concerns about the treatment of individuals under Utilitarianism. Feldman and Zaller (1992) conclude: "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."[19] As Feldstein (1976) noted near the start of the modern era of Mirrleesian tax theory: "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. Once again, optimal tax design involves a balancing of conflicting criteria."

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.

[18] Konow (2003) reports results consistent with these magnitudes.

[19] 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.

2 Prior work on tagging

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 Blumkin, Margalioth, and Sadka 2009).

In the modern theory of optimal taxation, tagging is a free lunch. That theory starts with the assumption that individuals differ in their unobservable abilities to earn income but are equally able to enjoy consumption. If social welfare is a weakly concave function of all individual utilities, income ought to be redistributed from those with high ability to those with low ability. But, there is a tradeoff. Taxing endogenous income rather than exogenous ability discourages effort, reducing economic activity overall. Tags carry information about ability but are hard to modify, so taxing them allows for redistributive gains without efficiency losses.

According to this theory, 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..."[20] 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.[21]

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.[22] 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 main reasons why tagging may be unappealing in practice have been discussed from the beginning.[23]

[20] 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.

[21] 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.

[22] 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.

[23] 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

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. 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.[24] Certainly a characteristic such as gender is highly inelastic and could be cheaply incorporated into the tax system.

2.1 Horizontal equity

Akerlof's remaining disadvantage of tagging is that it could violate horizontal equity: the notoriously difficult-to-define principle that "equals ought to be treated equally". This is a prominent concern: 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) and Auerbach and Hassett (1999).

On its own, the principle of horizontal equity offers an unsatisfying explanation for the limits to tagging. First, it is a tautological solution: it literally assumes that tagging is costly.[25] It is also an unreliable solution, as the choice of which characteristics are to be treated as "horizontal" is, at heart, arbitrary. For example, if "equals" are defined by income, they cannot also be defined by income-earning ability if preferences over consumption and leisure are heterogeneous. In that case, the principle of horizontal equity gives no guidance as to how to resolve this contradiction. Most important, once one chooses a definition of "horizontal," it remains to be explained why that definition is appropriate. 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?"

The core of the problem for horizontal equity is that, as Kaplow (2008) writes, it "lacks affirmative justification." It offers no defense of its limited form of equal treatment. In fact, its incompleteness as a normative criterion for optimal tax is made clear by its inability to offer guidance on how taxes required for public goods ought to be assigned–for that, its supporters must turn to vertical equity, a principle with an entirely distinct normative basis.

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.[26] 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. One contribution of this paper, therefore, can be seen as providing a normatively rigorous foundation for the concern over horizontal equity long intuited as the obstacle to greater tagging and, then, examining the broader consequences of that foundation for income taxation.

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.

[24] Mirrlees (1971) makes the same point on I.Q. See page 208.

[25] Related to this point is that the narrowness of horizontal equity prevents it from capturing the broader critique of Utilitarianism discussed above.

[26] The distaste for tagging under Equal Sacrifice comes from such personal characteristics being irrelevant to the sacrifice an individual bears to pay a given tax.

3 Generalizing the optimal tax model for normative diversity

Appealing as it may be to generalize the optimal tax model's normative perspective to capture the diverse moralities that drive public and scholarly debate over taxes, there is a methodological obstacle: many of the most prominent 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.[27]

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 judment, and some may object to their use altogether. In the end, the appeal of my analysis will depend on how closely the optimal allocations and loss functions I use align with the priorities of the normative criteria. These loss functions can be specified in a way that respects Pareto efficiency, as the examples below illustrate, avoiding the problem with non-welfarist criteria noted by Kaplow and Shavell (2001).

In this paper, a 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).[28] 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." Now I develop this generalized optimal tax model formally.

3.1 The model

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 $\sum_{i=1}^{I} p^i = 1$. An individual of type i derives utility from consumption c and disutility from exerting labor effort y/w to earn income y. Denote the utility function $U(c, y/w)$.[29]

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:

[27] 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.

[28] 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."

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

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 $\mathcal{L}_\phi$ 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 novel component of this planner's problem is its objective, captured in expression (1), to minimize the weighted sum of criterion-specific losses across a set Φ of normative criteria.

The weights $\{\alpha_\phi\}_{\phi \in \Phi}$ applied to each loss function represent the importance each normative criterion plays in society's evaluations of policy. A number of models of the policymaking process could be used to generate such weights. The most straightforward is that the median (pivotal) voter has his or her own weights on each normative criterion, adopted by policymakers as a result of electoral competition.[30] 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 the 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.[31]

[30] 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.

[31] No incentive compatibility constraints are imposed on the set of feasible allocations because we want to compare allocations to a constant ideal for each criterion when varying the constraints on the planner's information sets.

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.

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 $\left\{\mathcal{L}_\phi\right\}_{\phi\in\Phi}$ that I use in this paper satisfy the following three conditions. The first two are straightforward. The third, Pareto Efficiency, may be more controversial but is generally viewed as a reasonable requirement in the optimal taxation literature.[32]

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*[33] *when the equilibrium allocation equals the ϕ-optimal feasible allocation.*

3. *Weak Pareto Efficiency:*

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

   which can be converted into Strong Pareto Efficiency if desired.[34]

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.[35]

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.

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

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

[34] 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 $i' \in \{1, 2, ..., I\} \Rightarrow \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_1^i, y_1^i\right\}_{i=1}^I\right) < \mathcal{L}_\phi\left(\left\{c_\phi^i, y_\phi^i\right\}_{i=1}^I, \left\{c_2^i, y_2^i\right\}_{i=1}^I\right)$

[35] 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.

4 A two-criterion case: Utilitarianism and Equal Sacrifice

In this section, I apply the previous section's approach to the case of two criteria: the Utilitarian criterion of maximal aggregate utility and the principle of Equal Sacrifice.

4.1 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 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.[36] 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 taxation and, therefore, no government spending.

A well-known conceptual issue with the idea of the laissez-faire allocation is that any economy is, in reality, inseparable from the existing set of taxes that fund the government and state institutions. The laissez-faire allocation is, therefore, not well-defined, because $G = 0$ implies a very different economy than that the status quo. In other words, if $G > 0$ is required for the status quo economy to function, the laissez-faire allocation is not in the feasible set $\mathbb{F}$.

[36] 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..." The caveat in Sidgwick's earlier statement raises the important issue of whether the laissez-faire allocation is, in fact, a just starting point. One alternative to the laissez-faire allocation as a starting point is to provide redistribution prior to assessing equal sacrifice. Though a detailed analysis of that case is not pursued here, so long as the starting allocation does not use tagging, Equal Sacrifice will continue to discourage its use.

Fortunately, the Equal Sacrifice principle provides a natural way to convert the hypothetical laissez-faire allocation into a preference relation over feasible allocations and an Equal-Sacrifice-optimal (ES-optimal) feasible allocation. 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.[37]

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

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

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:

$$\left\{c^i_1, y^i_1\right\}_{i=1}^I \succeq_{ES} \left\{c^i_2, y^i_2\right\}_{i=1}^I \Leftrightarrow U\left(c^i_{lf}, y^i_{lf}/w^i\right) - U\left(c^i_1, y^i_1/w^i\right) \leq U\left(c^i_{lf}, y^i_{lf}/w^i\right) - U\left(c^i_2, y^i_2/w^i\right), \tag{7}$$

for $\left\{c^i_1, y^i_1\right\}_{i=1}^I, \left\{c^i_2, y^i_2\right\}_{i=1}^I \in \mathbb{ES}$ 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 $\left\{c^i_{ES}, y^i_{ES}\right\}_{i=1}^I$ as follows:

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

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

The next step is to specify the loss functions for the planner.

4.2 Loss functions

The Utilitarian loss function $\mathcal{L}_{Util}$ converts the familiar goal of aggregate utility maximization into aggregate sacrifice minimization:

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

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.

Unlike Utilitarianism, the Equal Sacrifice criterion does not rank allocations that deviate from its preferred allocation. Instead, I will specify a loss function that is designed to reflect the priorities of the Equal Sacrifice principle. In particular, the Equal Sacrifice loss function $\mathcal{L}_{ES}$ satisfies the following three proper-

[37] Equal Sacrifice has at times been interpreted to mean equal proportional sacrifice. Both interpretations share a distaste for tagging, as sacrifice depends only on income-earning ability.

ties:[38] 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[39]; 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. I formalize these properties as follows:

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

where

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

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.

4.3 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)*

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

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 $\alpha_{ES} = 0$.

In the next two sections, the optimal policy generated by this planner's problem will be analyzed in depth. First, however, 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.

[38] In addition to the three properties named above: Ordinality, Normalization, and Weak Pareto Efficiency.

[39] 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.

4.4 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 normative diversity of this type on the allocation of utility across individuals. Figure 1 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 allocations chosen by each criterion.

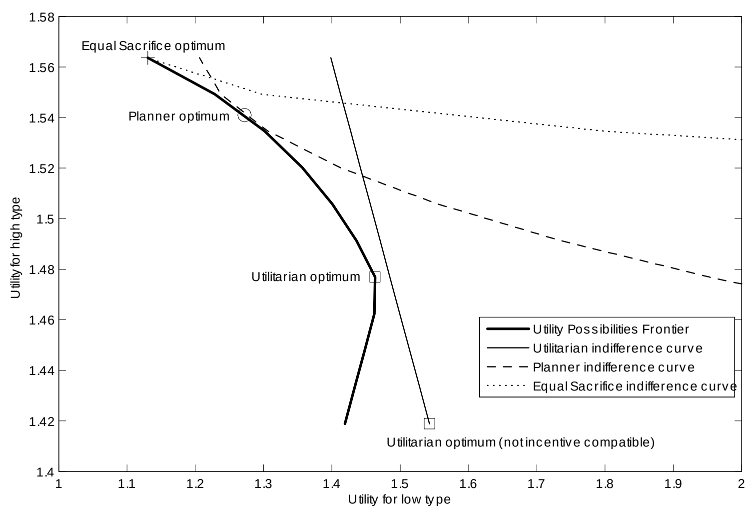

Figure 1: The Utility Possibilities Frontier and Indifference Curves

Figure 1 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.

5 Analysis of optimal policy with a role for Equal Sacrifice

In this section, I examine analytically the characteristics of optimal policy with normative diversity as formalized in Section 4, that is, with both Utilitarianism and Equal Sacrifice in the planner's objective. I consider two effects of increasing the weight on Equal Sacrifice in the social loss function. First, I show that it reduces the optimal extent of redistribution through tagging. Second, I show that it has a theoretically ambiguous impact on the pattern of optimal marginal tax rates.

5.1 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 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*

$$\min_{\left\{c_*^{i,m}, y_*^{i,m}\right\}_{i=1,m=1}^{I,M} \in \{\mathbb{F} \cap \mathbb{IC}\}} \left\{ \begin{array}{c} \alpha_{Util} \sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m} \left[U\left(c_{Util}^{i,m}, y_{Util}^{i,m}/w^i\right) - U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)\right] \\ + \ \alpha_{ES} \sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m} V\left(U\left(c_{ES}^{i,m}, y_{ES}^{i,m}/w^i\right), U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)\right) \end{array} \right\}, \tag{12}$$

where

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

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

$$V\left(U_{ES}^{i,m}, U_*^{i,m}\right) = \left\{ \begin{array}{ll} -\left(\delta\left[U_*^{i,m} - U_{ES}^{i,m}\right]\right)^{\theta} & \textit{if } U_{ES}^{i,m} < U_*^{i,m} \\ \left[\lambda\left(U_{ES}^{i,m} - U_*^{i,m}\right)\right]^{\rho} & \textit{if } U_{ES}^{i,m} \geq U_*^{i,m} \end{array} \right., \tag{13}$$

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

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

$$\mathbb{F} = \left\{ \left\{c^{i,m}, y^{i,m}\right\}_{i=1,m=1}^{I,M} : \sum_{i=1}^{I}\sum_{m=1}^{M} p^{i,m}\left(y^{i,m} - c^{i,m}\right) \geq G \right\}, \tag{14}$$

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

$$\mathbb{IC} = \left\{ \left\{c^{i,m}, y^{i,m}\right\}_{i=1,1}^{I,M} : U\left(c^{i,m}, y^{i,m}/w^i\right) \geq U\left(c^{j,m}, y^{j,m}/w^i\right) \textit{ for all } i,j \in \{1,2,...,I\} \textit{ and } m \in \{1,2,...,M\} \right\}. \tag{15}$$

In this problem the incentive constraints (15) 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:*

$$\frac{E_i\left[\left(U_{c_*^{i,m}}\right)^{-1}\right]}{E_i\left[\left(U_{c_*^{i,n}}\right)^{-1}\right]} = \frac{E_i\left[\alpha_{Util} - \alpha_{ES}\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U_*^{i,m}}\right]}{E_i\left[\alpha_{Util} - \alpha_{ES}\frac{\partial V\left(U_{ES}^{i,n}, U_*^{i,n}\right)}{\partial U_*^{i,n}}\right]}, \tag{16}$$

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 (16) is the ratio of the expected inverse marginal utilities of consumption across tagged types. This equals the ratio of the cost in consumption units of an incentive-compatible marginal 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 (16) simplifies to:*

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

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. 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 (17) for positive α_{ES}. I make two mild assumptions to provide a clean benchmark case.[40] Importantly, both of these assumptions hold in the numerical simulations of Section 6.

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*

$$U_*^{i,m} \geq U_*^{i,n} \text{ for all } i = \{1, 2, ..., I\}, \tag{18}$$

and

$$U_*^{j,m} > U_*^{j,n} \text{ for at least one } j = \{1, 2, ..., I\}. \tag{19}$$

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

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

(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 below (in Section 6).

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

$$U_{ES}^{i,m} \neq U_{*}^{i,m} \text{ for all } i = \{1, 2, ..., I\} \text{ and } m \in \{1, 2, ..., M\}. \tag{20}$$

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-optimal feasible allocation.[41] 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 of Section 6.

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*

$$\frac{E_i\left[\left(U_{c_*^{i,m}}\right)^{-1}\right]}{E_i\left[\left(U_{c_*^{i,n}}\right)^{-1}\right]} < 1. \tag{21}$$

Corollary 2 is the main analytical result of the paper. It 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 (17), 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.[42]

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.

[41] 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 $\delta = 0$.

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

5.2 Optimal marginal distortions

To analyze marginal distortions to labor supply in this paper's generalized model, I return to the Social Planner's Problem in (11), where individuals differ only in ability w. Denote with $\mu^{j|i}$ the multiplier on the incentive constraint indicating that type i does not prefer type j's allocation. We can show the following:

Proposition 2 *The optimal marginal distortion to the labor supply decision of an individual with ability type i, denoted τ_*^i, satisfies the following condition.*

$$1-\tau_*^i=\frac{U_{y_*^i}\left(c_*^i,y_*^i/w^i\right)}{w^iU_{c_*^i}\left(c_*^i,y_*^i/w^i\right)}=\frac{p^i\left(1+\alpha_{ES}\left(\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}-1\right)\right)+\sum_{j=1}^{I}p^j\left(\mu^{j|i}-\mu^{i|j}\right)}{p^i\left(1+\alpha_{ES}\left(\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}-1\right)\right)+\sum_{j=1}^{I}p^j\left(\mu^{j|i}-\frac{w^i}{w^j}\frac{U_{y_*^i}\left(c_*^i,y_*^i/w^j\right)}{U_{y_*^i}\left(c_*^i,y_*^i/w^i\right)}\mu^{i|j}\right)},\tag{22}$$

where U_x is the partial derivative of individual utility with respect to x and U_*^i denotes $U\left(c_*^i,y_*^i,w^i\right)$.

To interpret condition (22), start with the conventional case in which $\alpha_{ES}=0$. Then, because the term $\frac{w^i}{w^j}\frac{U_{y_*^i}\left(c_*^i,y_*^i/w^j\right)}{U_{y_*^i}\left(c_*^i,y_*^i/w^i\right)}$ is less than one for $w^i<w^j$, binding incentive constraints on higher skill types (i.e., $\mu^{i|j}>0$) drive the optimal distortion τ_*^i above zero in the conventional model.

A positive marginal distortion on type i has a benefit and a cost in conventional theory. The benefit of such a distortion is that it allows the planner to offer a more generous tax treatment to i without tempting higher-skilled individuals to claim it. The greater the gain in social welfare due to this redistribution (measured by $\mu^{i|j}$ for $w^j>w^i$), the greater is the optimal distortion to i. The conventional cost of such a distortion is the reduced effort and, therefore, output from type i. The size of this cost increases with the share of i in the population, p^i, so τ_*^i falls with larger p^i.

If $\alpha_{ES}>0$, both the benefits and costs of optimal marginal distortions are affected.

First, with $\alpha_{ES}>0$ marginal distortions have a second cost because they cause deviations from the ES-optimal allocations. The social cost of this deviation for individual i is measured by the expression $\alpha_{ES}\left(\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}-1\right)$. A larger α_{ES} will increase this expression and decrease the optimal distortion on i if $\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}>1$. Note that $\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}$ measures the marginal reduction in social loss from raising the allocated utility of type i. Starting from the Utilitarian allocation, this reduction in loss will tend to be greater for the high-skilled, as their utilities will be far below the laissez-faire allocation. Formally, $\frac{-\partial V\left(U_{ES}^i,U_*^i\right)}{\partial U_*^i}$ is likely to be increasing in type because losses increase more than proportionally with deviations below the laissez-faire allocation and gains are concave in deviations above it. This effect of increasing α_{ES} will tend to be a decrease in the optimal distortions on higher-skilled workers relative to lower-skilled workers.

Second, the benefits of redistribution change when the planner puts weight on Equal Sacrifice. In particular, the social value of redistributing from higher-skilled individuals ($\mu^{i|j}$ for $w^j>w^i$) is less, because the planner places less value on individuals (i.e., low earners) enjoying greater utility that in the laissez-faire. With smaller benefits from redistributing to the low- and moderate-skilled individuals, the required distortions on them are smaller. Therefore, this effect of increasing α_{ES} will tend to decrease the optimal distortions on low- and moderate-skilled workers relative to higher-skilled workers.

The ambiguity in the effects of positive α_{ES} on optimal marginal tax rates indicates the general difficulty in obtaining definitive results from condition (22). For a more comprehensive characterization of optimal income taxes with normative diversity, I turn to calibrated numerical simulations in the next section.

6 Numerical results

In this section I use numerical simulations calibrated to micro-level data for the United States to show the quantitative effects and explanatory power of including Equal Sacrifice in the objective function of the optimal tax model. First, I consider three prominent potential tags–height, gender, and race–and show that parameter values for the model of Section 4 exist in which the optimal policy rejects the use of these tags but accepts redistributive income taxes driven by differences in income-earning ability. Second, I choose one such parameterization and show that the model simultaneously endorses a sizeable and empirically reasonable tag on blindness, one of the few personal characteristics explicitly tagged in the U.S. tax code.[43]

These results show that the revision of the conventional theory proposed above can resolve the puzzle in the title of this paper.

It turns out that the same revision can help address two additional puzzles in optimal tax research. First, I use the same parameterized model and a more detailed ability distribution to show that normative diversity of this type substantially reduces the extent of a controversial feature of the first-best allocations in the conventional optimal tax model–rank reversals in utility. Second, I show that optimal income taxes under the same parameterization include a profile of marginal tax rates resembling that in Saez (2001) but with a lower maximum rate. This finding relates to the puzzle identified in Diamond and Saez (2011) that the conventional model implies substantially higher peak marginal rates than those prevailing in current policy.

6.1 Optimal tagging with normative diversity

I use the following parameter values in the Social Planner's Problem with Tagging from Section 4:

Table 1: Parameter values

α_{ES}	ρ	θ	δ	λ	γ	$\frac{1}{\sigma-1}$	G
$\{0.10,\ 0.20\}$	2.0	1.0	0.01	$\{10,\ 20\}$	1.5	0.5	20

The parameter $\alpha_{ES} = 1 - \alpha_{Util}$ is the weight on the Equal Sacrifice loss function in the social objective function. Two plausible interpretations of α_{ES} are as either the share of the population that uses Equal Sacrifice as its primary normative criterion or the weight of the Equal Sacrifice principle in the median voter's normative preferences. Though I am not aware of any direct evidence on either of these interpretations, evidence is available on opinions toward the Libertarian normative framework to which Equal Sacrifice has been linked (for instance by Murphy and Nagel, 2004). As noted earlier, public opinion surveys and lab experiments estimate the proportion of individuals with traditional Libertarian views to be 10 to 20 percent in the United States (Boaz and Kirby 2007, Cappelen et al. 2011) Consistent with that finding, Konow (2003) reports the results of a survey testing an updated version of Robert Nozick's famous Wilt Chamberlain example, finding that the case in which a talented basketball player earns (and keeps) an extraordinarily high income due to voluntary payments by his fans is considered "fair" by at least 24 percent of respondents. These results provide some support for the assumed values of α_{ES} shown in Table 1.

The next four parameters determine the shape of the Equal Sacrifice loss function as specified in expression (13): ρ and θ determine its concavity, while δ and λ determine the extent of loss aversion. A larger λ relative

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

to δ means that the social loss function interprets downward deviations from Equal Sacrifice as more costly.[44] The final three parameters are familiar from conventional models: γ and $\frac{1}{\sigma-1}$ are the coefficient of relative risk aversion and the labor supply elasticity, and G is required government revenue. I choose γ and $\frac{1}{\sigma-1}$ to match mainstream estimates and G to approximate the current value (as a share of total income) in the United States.

6.1.1 Rejecting height, gender, and race tags while accepting redistribution

The data required for the simulation of the optimal height, gender, and race taxes are 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.[45] 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 2 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.[46] 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 2: 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 2 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 2, 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 in Table 1).

For each of the four parameter vectors implied by Table 1, I report measures of the optimal extent of tagging and income tax progressivity in Table 3 and Table 4, respectively.

To measure the extent of tagging, Table 3 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.[47] More specifically, this is the ratio of total tax payments to total income for each group

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

[45] 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.

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

[47] The planner's problem when it cannot tag differs from the Social Planner's Problem with Tagging in that each individual i, m must prefer its bundle to any other bundle j, n. In that problem, tagged groups with higher wage distributions will pay greater average tax rates because the tax system is progressive. The "extra" taxes and transfers reported in Table 3 isolate the direct effects of tags on taxes.

under the optimal policy less the same ratio under the constrained-optimal policy with no tagging. If that difference is positive, the group is paying taxes in addition to what it would pay if tagging were prohibited. If that difference is negative, it is receiving an extra transfer. For reference, I report the same statistic for the policy solution when $\alpha_{ES} = 0$, the fully Utilitarian (conventional Mirrleesian) planner.

Table 3: 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}	λ	White	White	White	NW	White	NW	NW	White	White	NW	NW	NW
0	n/a	10.5	8.1	6.3	1.7	-4.3	-5.5	-3.6	-11.7	-13.4	-17.6	-22.0	-23.4
0.10	10	1.8	1.3	1.3	0.9	-0.4	-0.5	-1.3	-2.0	-2.6	-2.4	-4.9	-5.2
0.20	10	0.8	0.5	0.5	0.7	-0.1	-0.1	-0.9	-0.9	-1.2	-0.8	-2.0	-2.7
0.10	20	0.6	0.5	0.6	0.3	-0.1	-0.2	-1.0	-0.8	-0.9	-0.3	-1.4	-2.0
0.20	20	0.0	-0.1	-0.1	0.3	0.1	0.1	0.0	0.0	0.1	0.1	-0.4	-0.4

To gauge the progressivity of the optimal income tax, Table 4 reports the average tax rate paid by the members of each wage range under each parameterization.

Table 4: Extent of Progressivity (Average tax rates, in percent)

		Average wage rate in range									
α_{ES}	λ	2.81	6.50	10.03	13.82	17.80	21.70	27.28	43.25	62.06	95.96
0	10	-396	-64	-5	17	27	32	38	50	52	53
0.10	10	-300	-39	3	18	23	27	31	43	47	50
0.20	10	-258	-29	7	17	22	24	28	40	44	47
0.10	20	-228	-21	10	17	20	23	26	37	41	45
0.20	20	-183	-10	13	16	19	20	23	32	37	42

Finally, Table 5 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 5: Welfare Gain from Tagging

α_{ES}	λ	Percent of aggregate consumption
0	10	0.96
0.10	10	0.20
0.20	10	0.10
0.10	20	0.06
0.20	20	0.02

The results in these three tables show that incorporating a concern for Equal Sacrifice can explain the coexistence of limited tagging and substantial income redistribution through progressive taxes observed in policy. Table 3 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 $\alpha_{ES} = 1$ (not shown), even seemingly modest values for α_{ES} generate a steep decline in the use of tags. At the same time, for these values of α_{ES}, Table 4 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 5 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 all other cases.

As a specific example, consider the parameterization in which $\alpha_{ES} = 0.20$ and $\lambda = 10$. 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. To see this more clearly, consider Figure 2, which plots the schedule of average tax rates for this parameterization and two polar cases: the fully Utilitarian ($\alpha_{ES} = 0$) and the fully Equal Sacrifice ($\alpha_{ES} = 1.0$) policies.[48]

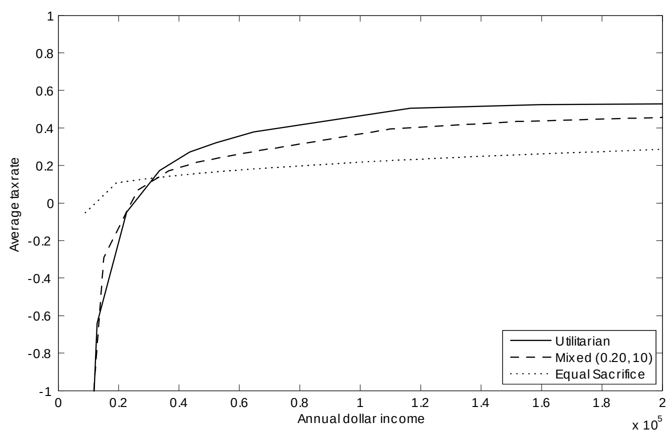

Figure 2: Optimal Average Tax Rates

As Figure 2 makes clear, the optimal policy according to the mixed social loss function is substantially redistributive and much more closely resembles the pure Utilitarian, conventional optimal policy than the policy that prioritizes only Equal Sacrifice.

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. But, that principle places little to no value on

[48] In this and the following figures that show annual dollar income, I convert the results of the simulations to annual figures as follows. The average labor effort in the more detailed simulation with $\alpha_{Lib} = 0.80$ and $\lambda = 10$ (below) is approximately 0.60 of total time available. The mean worker in the sample works approximately 2200 hours per year. That implies 2200/0.6=3667 hours as the appropriate multiplicative factor for the earnings returned by the policy simulation.

redistribution.[49] 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 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. I now turn to demonstrating this effect for blindness.

6.1.2 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.[50] 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 6. 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 6: 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

[49] 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.

[50] 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.

As Table 6 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, so both groups are treated the same.[51]

Table 7 shows the optimal extent of tagging in the conventional calibration with $\alpha_{ES} = 0$ and in the calibration with $\alpha_{ES} = 20$ and $\lambda = 10$, the case used to generate Figure 2 above. All other parameters are as in Table 1 (though G is adjusted to be a similar share of total income).

Table 7: Extent of Tagging
(Extra tax or transfer rate, in percent)

α_{ES}	λ	Not blind	Blind
0	10	0.07	-102.11
0.20	10	0.01	-19.82

As with height, gender, and race, Table 7 shows that adding 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 6, 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.

6.2 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 tax law scholars (see, for example, Zelenak, 2006). Economists, notably Feldstein (1976) and Mayshar and Yitzhaki (1995), have long noted that such rank reversals violate the "Pigou-Dalton principle," which Fleurbaey and Maniquet (2011) define as saying that "a transfer from one agent to another with lower income reduces inequality, or increases social welfare, provided it does not reverse their ranking..."

In this section, I show that a social objective function that puts weight on Equal Sacrifice reduces substantially the extent of rank reversals that would be chosen in the full-information scenario. 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.

To demonstrate this effect, I use a detailed calibration of the U.S. ability distribution[52] to simulate first-

[51] 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.

[52] 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).

best (i.e., full information) feasible income tax policies for a range of model parameterizations. Figure 3 shows the results by plotting utility as a function of ability in four scenarios. 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 $\lambda = 10$, and the dashed line is for the mixed case highlighted above in which $\alpha_{ES} = 0.20$ and $\lambda = 10$. All other parameters are as in Table 1 (though G is adjusted to be a similar share of total income).

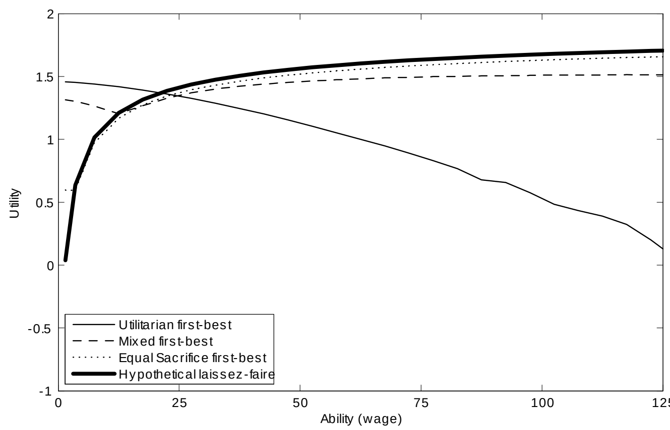

Figure 3: Utility levels by ability type under different objective functions in full-information first-best allocations.

Figure 3 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 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 mixed objective chooses a first-best allocation that substantially limits rank reversals.

6.3 Optimal income taxes

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 income tax rates. I use the same parameter values highlighted above ($\alpha_{ES} = 0.20$, $\lambda = 10$) 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 in Table 1 (though G is adjusted to be a similar share of total income).

Figure 4 shows the optimal schedule of marginal tax rates for this calibration, and Figure 5 shows the optimal schedule of average tax rates. Average tax rates are $(y - c)/y$ and marginal tax rates are as in (22). All figures show allocations for annual earnings up to \$200,000. For comparison, each figure also shows the optimal results under a pure Utilitarian criterion, that is when $\alpha_{ES} = 0$ as in the conventional model.

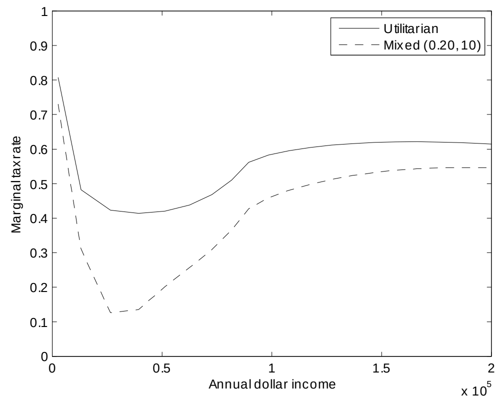

Figure 4: Optimal Marginal Tax Rates

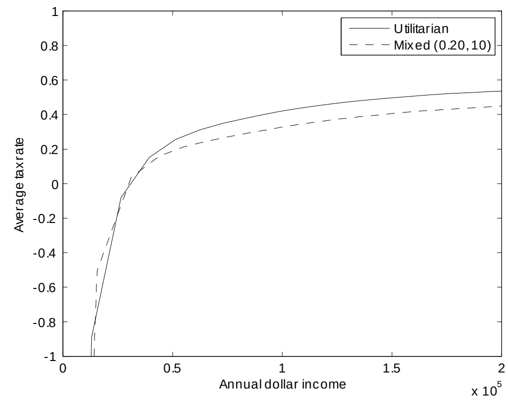

Figure 5: Optimal Average Tax Rates

Figure 4 shows that marginal tax rates have 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 effect of the second factor highlighted in the discussion of result (22) is particularly apparent, in that optimal marginal rates decline over a wider range of the ability distribution when α_{ES} is positive. The explanation for this pattern is that the planner with $\alpha_{ES} > 0$ redistributes less from high earners. This reduces the high earners' temptation to mimic moderate income earners and thus the required distortions on the latter.

The optimal marginal income tax rate at high incomes falls substantially, by about seven percentage points, with this role for the principle of Equal Sacrifice. Therefore, this paper's explanation for the limited use of tagging may help address the gap between conventional theory and existing policy noted by Diamond and Saez (2011).[53] Using the conventional model, they derive a formula for the optimal marginal tax rate on high incomes as a function of utility parameters and the shape of the ability distribution. 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 3 is 55 percent, compared to 62 percent under the conventional Utilitarian criterion.

Nevertheless, Figure 5 shows that substantial redistribution persists despite this role for the principle of Equal Sacrifice. 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 (not visible in the figures) 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 $\alpha_{ES} = 0$.

These simulations show, therefore, that variation in unobserved income-earning ability remains a powerful force for redistribution in a modification of the conventional optimal tax model that includes sufficient weight on the principle of Equal Sacrifice to reject differentiated taxation according to height, gender, and race. At the same time, this form of normative diversity does lower optimal marginal distortions to levels closer to that which we observe in reality.

[53] 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 preference heterogeneity (see Lockwood and Weinzierl, 2012).

7 Conclusion

Modern optimal tax research is inherently normative. Though controversial, being normative is also key to the literature's appeal, as it enables economists to apply their tools to a task of substantial importance: the design of the tax system.

A normative research agenda is only as relevant as its normative criterion, however, and herein lies a problem. Optimal tax research has yet to adopt a normative perspective that captures the diverse criteria with which different people and, in fact, most individuals have been shown to evaluate policy.[54]

This paper argues that one specific manifestation of this problem is the conventional optimal tax model's recommendation of substantial tagging, the tailoring of taxes to personal characteristics. Though the theoretical case for widespread tagging has been clear for nearly four decades, policy includes tagging in only limited ways.

In this paper, I propose a way to expand the normative scope of optimal tax research to include criteria other than the Utilitarianism that has dominated analysis since Mirrlees (1971). This expansion is intended to bring optimal tax research closer to, and therefore make it more relevant for, real-world debates in which normative heterogeneity plays a substantial role.

I then apply this general approach to include a specific alternative normative priority, the classic principle of Equal Sacrifice, and show that doing so can explain the limited role of tagging in policy. Equal Sacrifice is a principle of long-standing importance in tax theory and is a plausible representation of prevalent beliefs in the general population. With a social objective function that includes a role for Equal Sacrifice, the Utilitarian gains from tagging must be weighed against the costs it generates in unequal sacrifice.[55] Not all tags are equal in this model: a tag generates greater costs in unequal sacrifice the weaker is its correlation to ability. Thus, this model can explain not only the rejection of most tags but also the few cases in which tagging is used in real-world policy. At the same time, the model endorses substantial redistribution based on unobserved ability, preserving the core result of modern Mirrleesian tax analysis.

The appeal of this approach is not just theoretical. I simulate optimal policy with this form of normative diversity, calibrating the model to U.S. data. I show that optimal policy can simultaneously match three aspects of the U.S. tax code that are incompatible in conventional theory: it rejects the use of height, gender, and race as tags; it accepts the use of blindness as a tag; and it provides substantial redistribution through progressive income taxes.

This paper's model has two additional implications for income taxes that further increase its appeal. First, it substantially reduces the optimal extent of utility rank reversals in the first-best policy–a controversial feature of the conventional policy. Second, it has the potential to explain why marginal income tax rates at high incomes are not as large in reality as conventional theory would recommend, a prominent puzzle in recent research.

In sum, revising the theory of optimal taxation to include normative diversity in general, and the principle of Equal Sacrifice in particular, substantially improves the match between that theory's recommendations and the reality of tax policy.

[54] After nearly completing this paper, I rediscovered Martin Feldstein's (1976) discussion of the normative basis of optimal tax models in their early years. Feldstein makes nearly the same assertion as do I, namely: "My purpose here is not to deprecate the value of optimal tax theory but to emphasize that only a limited criterion of choice has been examined. The results of optimal tax design...are only as compelling as the criterion of social choice from which they are derived." Also, see the note in Section 1 on important exceptions to these statements in the work of Stiglitz (1987), Werning (2007), Saez and Stantcheva (2012), Fleurbaey and Maniquet (2006), and Besley and Coate (1992).

[55] As discussed in Section 2, the model's conclusions on tagging are consistent with the more familiar critique based on the idea of horizontal equity, but they are based on a comprehensive, rigorous, and precise normative justification.

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

[3] Arneson, Richard (1982), "Rawls versus Utilitarianism in the Light of *Political Liberalism*," published in Davion, Victoria and Clark Wolf, *The Idea of a Political Liberalism: Essays on Rawls,* Rowman & Littlefield: Maryland.

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

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

[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] Blumkin, Tomer, Yarom Margalioth, and Efraim Sadka (2009), "Incorporating Affirmative Action into the Welfare State," Journal of Public Economics 93, pp. 1027-1035.

[9] Boadway, Robin and Pierre Pestieau, (2006). "Tagging and Redistributive Taxation," *Annales D'Économie et de Statistique*, 83/84.

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

[11] 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).

[12] D'Antoni, Massimo (1999). "Piecewise linear tax functions, progressivity, and the principle of equal sacrifice," *Economics Letters,* 65.

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

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

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

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

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

[18] Fleurbaey, Marc and Francois Maniquet (2011). *A Theory of Fairness and Social Welfare.* Cambridge.

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

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

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

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

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

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

[25] 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.

[26] 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.

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

[28] Hasen, David. (2007). "Liberalism and Ability Taxation," *Texas Law Review,* 85, pp. 1057-1113.

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

[30] Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory," *Econometrica.*

[31] Kant, Immanuel (1785). Groundwork on the Metaphysic of Morals. 1949, Modern Library Version, Random House, page 178.

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

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

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

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

[36] Lambert, Peter J. and Helen T. Naughton (2009). "The Equal Sacrifice Principle Revisited," *Journal of Economic Surveys*, 23(2).

[37] Lockwood, Benjamin B. and Matthew Weinzierl (2012). "De Gustibus non est Taxandum: Theory and Evidence on Preference Heterogeneity and Redistribution." NBER Working Paper, No. 17784, January.

[38] Mankiw, N. Gregory and Matthew Weinzierl, (2010). "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," *American Economic Journal: Economic Policy*, 2(1), pp. 155-176.

[39] Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan, (2009). "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, 23(4), pages 147-174.

[40] Markovits, Daniel. (2003). "How Much Redistribution Should there Be?" *Yale Law Journal*, 112.

[41] Mazor, Joseph. (2012). "Who Should Pay for Entitlements," Working paper.

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

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

[44] Mirrlees, J.A., (1974). "Notes on Welfare Economics, Information, and Uncertainy," in *Essays on economic behavior under uncertainty,* M. Balch, D. McFadden and S. Wu, eds. New York : American Elsevier Pub. Co.

[45] 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.

[46] 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).

[47] Mitra, Tipra and Efe A. Ok (1996). "Personal Income Taxation and the Principle of Equal Sacrifice Revisited," *International Economic Review*, 37.

[48] Moulin, Herve, (1981). "The Proportional Veto Principle," *Review of Economic Studies* 48, pp. 407-416.

[49] Moyes, Patrick (1989). "Redistribution effects of minimal equal sacrifice taxation," *Journal of Economic Theory*, 108.

[50] Murphy, Liam and Thomas Nagel, (2002). *The Myth of Ownership.* New York: Oxford University Press.

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

[52] Ok, Efe A. (1995). "On the principle of equal sacrifice in income taxation," *Journal of Public Economics,* 58.

[53] Rakowski, Eric. (2000). "Can Wealth Taxes be Justified?" *Tax Law Review*, 53.

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

[55] Rawls, John (2001). *Justice as Fairness.* Cambridge: Harvard University Press.

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

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

[58] 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).

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

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

[61] Sen, Amartya and Bernard Williams, eds. (1982). *Utilitarianism and Beyond*, Cambridge.

[62] Shaviro, Daniel. (2002). "Endowment and Inequality," in Thorndike, Joseph J. and Dennis J. Ventry, Jr. (eds), *Tax Justice: The Ongoing Debate*. Washington D.C.: Urban Institute.

[63] Stark, Kirk J. (2005). "Enslaving the Beachcomber: Some Thoughts on the Liberty Objections to Endowment Taxation," *Canadian Journal of Law and Jurisprudence*, 18.

[64] 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.

[65] Sugin, Linda. (2011). "A Philosophical Objection to the Optimal Tax Model," *Tax Law Review,* 64.

[66] Viard, Alan D. (2001). "Some Results on the Comparative Statics of Optimal Categorical Transfer Payments." *Public Finance Review*, 29(2): 148–80.

[67] Weinzierl, Matthew (2011). "The Surprising Power of Age-Dependent Taxes," *Review of Economic Studies*, October.

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

[69] Yaari, Menahem (1988). "A Controversial Proposal Concerning Inequality Measurement," *Journal of Economic Theory,* 44.

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

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

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

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

[74] Zelenak, Lawrence (2006). "Taxing Endowments," *Duke Law Journal*, 55.

8 Appendix

8.1 Proof of Proposition 1

The first-order condition of the planner's problem with respect to $c_*^{i,m}$ is:

$$-\alpha_{Util} p^{i,m} + \ \alpha_{ES} p^{i,m} \frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} - \frac{\mu_F}{U_c\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} p^{i,m} + \mu^{j|i,m} - \mu^{i|j,m} = 0,$$

where μ_F is the multiplier on the feasibility constraint and $\mu^{j|i,m}$ is the multiplier on the incentive constraint that type i prefers its allocation to any other type j, for any group m. In deriving this condition, I used separability in the utility function to set $\frac{U_c\left(c_*^{i,m}, y_*^{i,m}/w^j, m\right)}{U_c\left(c_*^{i,m}, y_*^{i,m}/w^i, m\right)} = 1$. Taking the sum across types and simplifying yields:

$$E_i\left[-\alpha_{Util} + \ \alpha_{ES} \frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)}\right] = E_i\left[\frac{\mu_F}{U_c\left(c_*^{i,m}, y_*^{i,m}/w^i\right)}\right].$$

The analogous condition applies for $c_*^{i,n}$, where n indicates a different tagged group:

$$E_i\left[-\alpha_{Util} + \ \alpha_{ES} \frac{\partial V\left(U_{ES}^{i,n}, U_*^{i,n}\right)}{\partial U\left(c_*^{i,n}, y_*^{i,n}/w^i\right)}\right] = E_i\left[\frac{\mu_F}{U_c\left(c_*^{i,n}, y_*^{i,n}/w^i\right)}\right].$$

Combining these conditions, we can write:

$$\frac{E_i\left\{\left[U_c\left(c_*^{i,m}, y_*^{i,m}/w^i\right)\right]^{-1}\right\}}{E_i\left\{\left[U_c\left(c_*^{i,n}, y_*^{i,n}/w^i\right)\right]^{-1}\right\}} = \frac{E_i\left[\alpha_{Util} - \alpha_{ES} \frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)}\right]}{E_i\left[\alpha_{Util} - \alpha_{ES} \frac{\partial V\left(U_{ES}^{i,n}, U_*^{i,n}\right)}{\partial U\left(c_*^{i,n}, y_*^{i,n}/w^i\right)}\right]}.$$

This is the result in Proposition 1.

8.2 Proof of Corollary 2

First, I establish that $U_{ES}^{i,m} = U_{ES}^{i,n}$ for all $m, n \in \{1, 2, ..., M\}$. Recall the definition of the ES-optimal feasible allocation when individuals differ in only one dimension (ability) from the main text. The extension to two dimensions of heterogeneity is straightforward:

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

for any $i \in \{1, 2, ..., I\}$ and $m \in \{1, 2, ..., M\}$ and for all possible $\left\{c^{i,m}, y^{i,m}\right\}_{i=1,m=1}^{I,M} \in \mathbb{ES}$, where the set $\mathbb{ES}$ is defined as

$$\mathbb{ES} = \left\{ \begin{array}{c} \left\{c^{i,m}, y^{i,m}\right\}_{i=1,m=1}^{I,M} \in \mathbb{F} : U\left(c_{lf}^{i,m}, y_{lf}^{i,m}/w^i\right) - U\left(c^{i,m}, y^{i,m}/w^i\right) = U\left(c_{lf}^{j,n}, y_{lf}^{j,n}/w^j\right) - U\left(c^{j,n}, y^{j,n}/w^j\right) \\ \text{for all } i, j \in \{1, 2, ..., I\} \text{ and } m, n \in \{1, 2, ..., M\} \end{array} \right\}. \tag{23}$$

The laissez-faire allocations $\left\{c_{lf}^{i,m}, y_{lf}^{i,m}\right\}_{i=1,m=1}^{I,M}$ are defined by individual maximization, which depends only on ability as shown in the main text. Therefore, we know that

$$\begin{aligned} c_{lf}^{i,m} &= c_{lf}^{i,n}, \\ y_{lf}^{i,m} &= y_{lf}^{i,n}, \end{aligned}$$

for all $i \in \{1, 2, ..., I\}$ and $m, n \in \{1, 2, ..., M\}$. By the definition in (23), this immediately implies:

$$U\left(c_{ES}^{i,m}, y_{ES}^{i,m}/w^i\right) = U\left(c_{ES}^{i,n}, y_{ES}^{i,n}/w^i\right) \text{ for all } i \in \{1, 2, ..., I\} \text{ and } m, n \in \{1, 2, ..., M\}, \tag{24}$$

which completes this step of the proof.

Second, recall Assumption 1:

$$\begin{aligned} U_*^{i,m} &\geq U_*^{i,n} \text{ for all } i = \{1, 2, ..., I\} \text{ and } m < n \\ \text{and } U_*^{j,m} &> U_*^{j,n} \text{ for at least one } j = \{1, 2, ..., I\} \text{ and } m < n. \end{aligned} \tag{25}$$

Using (24), (25), and the technical Assumption 2 that rules out special cases in which the optimal allocation of utility equals the ES-optimal allocation for any individual, I consider three exhaustive cases.

8.2.1 Case 1

In this case, all equilibrium allocations generate at least as much utility for all individuals as in the ES-optimal feasible allocation and, for at least one individual, more utility than in the ES-optimal feasible allocation. Formally,

$$\begin{aligned} U_*^{i,m} &\geq U_*^{i,n} > U_{ES}^{i,m} = U_{ES}^{i,n} \text{ for all } i \in \{1, 2, ..., I\} \text{ and } m < n \\ \text{and } U_*^{j,m} &> U_*^{j,n} > U_{ES}^{j,m} = U_{ES}^{j,n} \text{ for some } j \in \{1, 2, ..., I\} \text{ and } m < n. \end{aligned} \tag{26}$$

The proof of the corollary for this case is as follows.

First, given (26), we can take the derivative of $V\left(U_{ES}^{i,m}, U_*^{i,m}\right)$ as specified in (13) with respect to its second argument and obtain,

$$\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} = \begin{cases} -\theta\delta^\theta \left[U_*^{i,m} - U_{ES}^{i,m}\right]^{\theta-1} & \text{if } U_{ES}^{i,m} < U_*^{i,m} \end{cases},$$
$$\text{for scalars } \{\delta \geq 0,\ \lambda > \delta,\ \theta \in (0, 1],\ \rho > 1\},$$

which directly implies

$$\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} \leq 0, \tag{27}$$

with the inequality strict if $\delta > 0$. Taking the second derivative yields:

$$\frac{\partial^2 V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\left(\partial U_*^{i,m}\right)^2} = \begin{cases} -\theta(\theta - 1)\,\delta^\theta \left[U_*^{i,m} - U_{ES}^{i,m}\right]^{\theta-2} & \text{if } U_{ES}^{i,m} < U_*^{i,m} \end{cases},$$
$$\text{for scalars } \{\delta \geq 0,\ \lambda > \delta,\ \theta \in (0, 1],\ \rho > 1\},$$

which directly implies

$$\frac{\partial^2 V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\left(\partial U_*^{i,m}\right)^2} \geq 0. \tag{28}$$

Second, combine (27), (28) , and (25) to obtain:

$$\begin{aligned}
\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} &\geq \frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,n}\right)}{\partial U\left(c_*^{i,n}, y_*^{i,n}/w^i\right)} \text{ for all } i \in \{1,2,...,I\} \text{ and } m<n \\
\text{and } \frac{\partial V\left(U_{ES}^{j,m}, U_*^{j,m}\right)}{\partial U\left(c_*^{j,m}, y_*^{j,m}/w^j\right)} &> \frac{\partial V\left(U_{ES}^{j,m}, U_*^{j,n}\right)}{\partial U\left(c_*^{j,n}, y_*^{j,n}/w^j\right)} \text{for at least one } j \in \{1,2,...,I\} \text{ and } m<n.
\end{aligned} \tag{29}$$

Third, use(24) to rewrite (29) as:

$$\begin{aligned}
\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} &\geq \frac{\partial V\left(U_{ES}^{i,n}, U_*^{i,n}\right)}{\partial U\left(c_*^{i,n}, y_*^{i,n}/w^i\right)} \text{ for all } i \in \{1,2,...,I\} \text{ and } m<n \\
\text{and } \frac{\partial V\left(U_{ES}^{j,m}, U_*^{j,m}\right)}{\partial U\left(c_*^{j,m}, y_*^{j,m}/w^j\right)} &> \frac{\partial V\left(U_{ES}^{j,n}, U_*^{j,n}\right)}{\partial U\left(c_*^{j,n}, y_*^{j,n}/w^j\right)} \text{for at least one } j \in \{1,2,...,I\} \text{ and } m<n.
\end{aligned} \tag{30}$$

which implies

$$E_i\left[\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)}\right] > E_i\left[\frac{\partial V\left(U_{ES}^{i,n}, U_*^{i,n}\right)}{\partial U\left(c_*^{i,n}, y_*^{i,n}/w^i\right)}\right]. \tag{31}$$

Finally, use expression (16) to show that (31) is a sufficient condition for Corollary 2 to hold. This completes the proof in this case.

8.2.2 Case 2:

In this case, all equilibrium allocations generate no more utility for all individuals as in the ES-optimal feasible allocation and, for at least one individual, strictly less utility than in the ES-optimal feasible allocation. Formally,

$$\begin{aligned}
U_{ES}^{i,m} &= U_{ES}^{i,n} > U_*^{i,m} \geq U_*^{i,n} \text{ for all } i \in \{1,2,...,I\} \text{ and } m<n \\
\text{and } U_{ES}^{j,m} &= U_{ES}^{j,n} > U_*^{j,m} > U_*^{j,n} \text{ for some } j \in \{1,2,...,I\} \text{ and } m<n.
\end{aligned} \tag{32}$$

The proof of the corollary for this case is as follows.

First, given (32), we can take the derivative of $V\left(U_{ES}^{i,m}, U_*^{i,m}\right)$ as specified in (13) with respect to its second argument and obtain,

$$\frac{\partial V\left(U_{ES}^{i,m}, U_*^{i,m}\right)}{\partial U\left(c_*^{i,m}, y_*^{i,m}/w^i\right)} = \left\{-\rho\lambda^\rho\left(U_{ES}^{i,m} - U_*^{i,m}\right)^{\rho-1} \text{ if } U_{ES}^{i,m} \geq U_*^{i,m}\right.,$$

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

$$\frac{\partial V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\partial U\left(c_*^{i,m},y_*^{i,m}/w^i\right)}\leq 0. \tag{33}$$

Taking the second derivative yields:

$$\frac{\partial^2 V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\left(\partial U_*^{i,m}\right)^2}=\left\{\rho\left(\rho-1\right)\lambda^{\rho}\left(U_{ES}^{i,m}-U_*^{i,m}\right)^{\rho-1}\text{ if }U_{ES}^{i,m}\geq U_*^{i,m}\right.,$$
$$\text{for scalars }\left\{\delta\geq 0,\ \lambda>\delta,\ \theta\in(0,1],\ \rho>1\right\},$$

which directly implies

$$\frac{\partial^2 V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\left(\partial U_*^{i,m}\right)^2}\geq 0. \tag{34}$$

Given (33) and (34), the remainder of the proof for this case repeats exactly the proof of Case 1 after (28). This completes the proof in this case.

8.2.3 Case 3:

In this case, the equilibrium allocations give some individuals more utility than in the ES-optimal feasible allocation and others of the same ability but different tagged values less utility than in the ES-optimal feasible allocation. Formally,

$$\begin{aligned} U_*^{i,m} &\geq U_*^{i,n}\text{ for all }i\in\{1,2,...,I\}\text{ and }m<n \\ \text{and }U_*^{j,m} &> U_{ES}^{j,m}=U_{ES}^{j,n}>U_*^{j,n}\text{ for some }j\in\{1,2,...,I\}\text{ and }m<n. \end{aligned} \tag{35}$$

The proof of the corollary in this case for any $i\in\{1,2,...,I\}$ such that either $U_*^{i,m}\geq U_*^{i,n}\geq U_{ES}^{i,m}=U_{ES}^{i,n}$ or $U_{ES}^{i,m}=U_{ES}^{i,n}\geq U_*^{i,m}\geq U_*^{i,n}$ is the same as in Case 1 or Case 2, respectively. The new scenario is the second line in (35), where the m type has been given more utility and the n type less utility than in the ES-optimal feasible allocation. Using (24) and (13) and we can derive:

$$\begin{aligned} \frac{\partial V\left(U_{ES}^{j,m},U_*^{j,m}\right)}{\partial U_*^{j,m}} &= -\theta\delta^{\theta}\left[U_*^{j,m}-U_{ES}^{j,m}\right]^{\theta-1} \\ \frac{\partial V\left(U_{ES}^{j,m},U_*^{j,n}\right)}{\partial U_*^{j,n}} &= -\rho\lambda^{\rho}\left(U_{ES}^{j,m}-U_*^{j,n}\right)^{\rho-1} \\ \text{for scalars }&\left\{\delta\geq 0,\ \lambda>\delta,\ \theta\in(0,1],\ \rho>1\right\}, \end{aligned} \tag{36}$$

for some $j\in\{1,2,...,I\}$ and $m<n$.

Expression (36) implies that $\frac{\partial V\left(U_{ES}^{j,m},U_*^{j,m}\right)}{\partial U_*^{j,m}}$ and $\frac{\partial V\left(U_{ES}^{j,m},U_*^{j,n}\right)}{\partial U_*^{j,n}}$ are both negative (i.e., non-zero). Thus, for some $\delta>0$,$\frac{\partial V\left(U_{ES}^{j,m},U_*^{j,m}\right)}{\partial U_*^{j,m}}>\frac{\partial V\left(U_{ES}^{j,m},U_*^{j,n}\right)}{\partial U_*^{j,n}}$, and we have

$$\begin{aligned} \frac{\partial V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\partial U\left(c_*^{i,m},y_*^{i,m}/w^i\right)} &\geq \frac{\partial V\left(U_{ES}^{i,n},U_*^{i,n}\right)}{\partial U\left(c_*^{i,n},y_*^{i,n}/w^i\right)}\text{ for all }i\in\{1,2,...,I\}\text{ and }m<n \\ \text{and }\frac{\partial V\left(U_{ES}^{j,m},U_*^{j,m}\right)}{\partial U\left(c_*^{j,m},y_*^{j,m}/w^j\right)} &> \frac{\partial V\left(U_{ES}^{j,n},U_*^{j,n}\right)}{\partial U\left(c_*^{j,n},y_*^{j,n}/w^j\right)}\text{for at least one }j\in\{1,2,...,I\}\text{ and }m<n, \end{aligned} \tag{37}$$

which implies

$$E_i\left[\frac{\partial V\left(U_{ES}^{i,m},U_*^{i,m}\right)}{\partial U\left(c_*^{i,m},y_*^{i,m}/w^i\right)}\right]>E_i\left[\frac{\partial V\left(U_{ES}^{i,n},U_*^{i,n}\right)}{\partial U\left(c_*^{i,n},y_*^{i,n}/w^i\right)}\right]. \tag{38}$$

The same last step as in the previous two cases completes the proof.

8.3 Wage distributions by tagged group

The following table lists the wage distributions by tagged group, adjusted for the NLSY's sample weights. The representative wage rate in each bin is the mean wage, again adjusted for sample weights, across the population within that bin range. The sample weights adjustments have only minor effects on the mean wages and the population distributions. In particular, simulations using unweighted means and distributions (reported in earlier, working paper versions of this paper) yield results on optimal tagging and income taxation that closely resemble the results with weighting and that are consistent with the qualitative conclusions reached here.

Appendix Table 1: Wage distributions by tagged group

	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
Wage	White	White	White	NW	White	NW	NW	White	White	NW	NW	NW
2.89	0.02	0.03	0.03	0.03	0.04	0.06	0.09	0.08	0.12	0.05	0.12	0.12
6.53	0.10	0.12	0.13	0.23	0.24	0.23	0.27	0.29	0.27	0.30	0.26	0.29
10.07	0.21	0.22	0.23	0.23	0.24	0.24	0.16	0.23	0.23	0.31	0.31	0.25
13.86	0.23	0.20	0.24	0.15	0.18	0.21	0.14	0.21	0.16	0.13	0.14	0.18
17.82	0.15	0.15	0.14	0.21	0.13	0.09	0.15	0.08	0.09	0.13	0.08	0.09
21.73	0.11	0.12	0.09	0.02	0.09	0.07	0.12	0.05	0.05	0.05	0.06	0.04
27.26	0.11	0.09	0.08	0.10	0.04	0.06	0.03	0.03	0.06	0.02	0.01	0.03
43.59	0.05	0.06	0.06	0.02	0.03	0.02	0.02	0.03	0.01	–	0.01	–
62.20	0.02	0.01	0.01	0.03	0.01	0.01	0.03	0.01	0.01	–	–	–
94.05	0.003	0.002	0.002	–	0.003	0.005	–	0.004	0.007	–	–	–