FALL 2019
NEW YORK UNIVERSITY
SCHOOL OF LAW

"Optimal Tax Theory as a Theory of Distributive Justice"

Lily Batchelder
NYU School of Law

September 3, 2019
Vanderbilt Hall – 202
Time: 4:00 – 5:50 p.m.
Week 1

SCHEDULE FOR FALL 2019 NYU TAX POLICY COLLOQUIUM

(All sessions meet from 4:00-5:50 pm in Vanderbilt 202, NYU Law School)

1. Tuesday, September 3 – Lily Batchelder, NYU Law School.
2. Tuesday, September 10 – Eric Zwick, University of Chicago Booth School of Business.
3. Tuesday, September 17 – Diane Schanzenbach, Northwestern University School of Education and Social Policy.
4. Tuesday, September 24 – Li Liu, International Monetary Fund.
5. Tuesday, October 1 – Daniel Shaviro, NYU Law School.
6. Tuesday, October 8 – Katherine Pratt, Loyola Law School Los Angeles.
7. Tuesday, October 15 – Zachary Liscow, Yale Law School.
8. Tuesday, October 22 – Diane Ring, Boston College Law School.
9. Tuesday, October 29 – John Friedman, Brown University Economics Department.
10. Tuesday, November 5 – Marc Fleurbaey, Princeton University, Woodrow Wilson School.
11. Tuesday, November 12 – Stacie LaPlante, University of Wisconsin School of Business.
12. Tuesday, November 19 – Joseph Bankman, Stanford Law School.
13. Tuesday, November 26 – Deborah Paul, Wachtell, Lipton, Rosen, and Katz.
14. Tuesday, December 3 – Joshua Blank, University of California at Irvine Law School.

Optimal Tax Theory as a Theory of Distributive Justice ‡

Lily L. Batchelder*

NYU School of Law

Draft of Aug. 20, 2019

The literature on taxation and transfers includes two prominent egalitarian theories of distributive justice: resource egalitarianism and welfare egalitarianism, as elaborated through optimal tax theory. With the recent explosion in empirical tax research, optimal tax theory has garnered even greater prominence. But non-welfarists remain unpersuaded, arguing it fails to adequately address a number of serious philosophical objections.

This essay considers the primary non-welfarist critiques of egalitarian optimal tax theory, paying particular attention to those of resource egalitarians. In the process, it argues the gap between these two theories is narrower than most appreciate. Indeed, once optimal tax theory is read broadly and deeply, the ideal policy design principles implied by each theory largely mimic the other.

Furthermore, on challenging issues where the stance of each theory is unsettled, the tools and objectives of optimal tax theory often imply assumptions and modeling approaches that would provide a clearer and stronger response to the concerns of non-welfarists. These include welfare-weighting externalities, modeling the optimal endowment tax as one that is risk-neutral and at most only partially based on total potential income, and treating choices as providing less information about well-being when they are further in the past or relatively predictable. More explicitly and consistently adopting these assumptions and modeling approaches would clarify that optimal tax theory does not entail the extreme limits on autonomy or unrelenting responsibility for prior choices that have troubled non-welfarists.

‡ Apologies for the many missing citations. Comments and suggestions very welcome and appreciated.

* Robert C. Kopple Family Professor of Taxation. For helpful comments, I am grateful to Joe Bankman, Ronald Dworkin, Itai Grinberg, Mitchell Kane, Adam Kern, Leandra Lederman, Yoram Margalioth, Liam Murphy, Thomas Nagel, Davud Rostam-Afschar, and participants in the Georgetown University Law Center Faculty Workshop, Indiana University Maurer School of Law Tax Policy Colloquium, National Tax Association Annual Conference, NYU Colloquium of Legal Political and Social Philosophy, Stanford Law School Tax Policy Workshop, University of Richmond School of Law Faculty Workshop, and the University of San Diego Richard C. Pugh Lecture. I thank Sarah Brodie, Ryan Fackler, Rachel Jones, and Annmarie Zell for excellent research assistance. I am also grateful to the Urban-Brookings Tax Policy Center for hosting me while working on this piece, and the Filomen D'Agostino and Max Greenberg Research Fund at NYU School of Law for generous financial support. All errors are mine.

I. Introduction

The literature on taxation and transfers includes two prominent egalitarian theories of distributive justice: resource egalitarianism and welfare egalitarianism, as elaborated through optimal tax theory. With the recent explosion in empirical tax research, optimal tax theory has garnered even greater prominence, in part because it is more amenable to empirically-grounded modeling and therefore concrete policy prescriptions. But non-welfarists remain unpersuaded, arguing it fails to adequately address a number of serious philosophical objections.

This essay considers the primary non-welfarist critiques of optimal tax theory, paying particular attention to those of resource egalitarians. In the process, it argues the gap between the two theories is narrower than most appreciate. Indeed, the ideal policy design principles implied by each theory largely mimic the other. Furthermore, on challenging issues where the stance of each theory is unsettled, the tools of optimal tax frequently offer a clearer and stronger response.

This essay aims to contribute to the literature in three ways. First, as just noted, it challenges the prevailing view that optimal tax theory and resource egalitarianism are fundamentally at odds. Optimal tax theory is a subset of welfarism that develops empirically-grounded models to estimate what combination of tax and transfers maximizes social welfare. When the social welfare function is egalitarian (as is the focus here), it weights improvements in the well-being of less-well-off individuals more heavily than those who are better-off. Optimal tax theory has a virtual monopoly in economics and is increasingly dominant in legal scholarship. It may seem odd that a theory denoted a tax theory includes transfers. But tax theorists tend to take an expansive view of their domain, considering transfers (including in-kind transfers) simply to be negative taxes.

Resource egalitarianism, in contrast, regards a just distribution as one that treats all members of society with equal concern and respect, as determined from a position of equality.[1] This has two components. First, everyone should begin their life with equal transferrable resources, as measured by the value of those resources to others.[2] Second, the state should provide insurance against inequalities arising from non-transferrable resources, such as one's talents or health. But it should only insure against the effects brute luck—not option luck, which is the result of deliberate and calculated gambles.[3] Ronald Dworkin laid out the case for resource egalitarianism (sometimes referred to as luck egalitarianism) in a famous set of articles in 1981.[4] It is largely the province of philosophers, but also some in legal academia.

Advocates of each approach deny any substantial similarity with the other. Dworkin argued that welfarism and resource egalitarianism will offer different advice in many concrete cases,[5] and nothing in his proposal "either aims at equality in any concept of welfare or makes convergence toward such equality likely."[6] Optimal tax theorists have similarly disclaimed any substantial similarity between their approach and resource egalitarianism.[7]

The concerns motivating these two theories do fundamentally differ. Resource egalitarians care about equalizing the objective resources available to individuals to further their life plans, while holding them responsible for choices within their control. Egalitarian optimal tax theorists care about maximizing subjective well-being—especially for those who are less well-off. This must be balanced against any negative effects of redistribution on efficiency, and therefore aggregate well-being. But despite these different motivations, their ideal fiscal systems would, to a large extent, try to measure and mitigate disparities of the same thing.[8]

None of this is to say that egalitarian optimal tax theory and resource egalitarianism are unimpeachable as theories of distributive justice. Each can only address deontological values,

---

[1] RONALD DWORKIN, SOVEREIGN VIRTUE: THE THEORY AND PRACTICE OF EQUALITY 313 (2000) (hereinafter, Dworkin (2000)); Ronald Dworkin, Sovereign Virtue *Revisited*, 113 ETHICS 106, 106–07 (2002) (hereinafter, Dworkin (2002)).

[2] Ronald Dworkin, *What is Equality? Part 2: Equality of Welfare*, 10:4 PHIL. & PUB. AFFAIRS 283, 285–89 (1981) (hereinafter, Dworkin (1981b)).

[3] Dworkin (1981b), *supra* note 2, at 293.

[4] Ronald Dworkin, *What is Equality? Part 1: Equality of Welfare*, 10:3 PHIL. & PUB. AFFAIRS 185, 186 (1981) (hereinafter, Dworkin (1981a)); Dworkin (1981b), *supra* note 2.

[5] Dworkin (1981a), *supra* note 4, at 186.

[6] Dworkin (1981b), *supra* note 2, at 335.

[7] *See, e.g.*, LOUIS KAPLOW, THE THEORY OF TAXATION AND PUBLIC ECONOMICS 366-68 (2008) (hereinafter, KAPLOW (2008)).

[8] Others have drawn this connection in other contexts or in less detail. *See, e.g.*, Daniel Shaviro, *Inequality, Wealth, and Endowment*, 53 TAX L. REV. 397, 402–04 (2002) (arguing that equal resources or opportunities can be treated as equal utility in the face of uncertainty); Barbara H. Fried, *Ex Ante / Ex Post*, 13 J. CONTEMP. LEG. ISSUES 123, 124, 137, 150 (2003) (noting that resource egalitarians and welfarists have converged on strikingly similar conclusions with respect to legal transitions); Kirk J. Stark, *Enslaving the Beachcomber: Some Thoughts on the Liberty Objections to Endowment Taxation*, 18 CANADIAN J. L. & JURIS. 47, 65–66 (2005); Lily L. Batchelder, *What Should Society Expect from Heirs? A Proposal for a Comprehensive Inheritance Tax*, 63 TAX L. REV. 1, 12 (2009) (hereinafter Batchelder (2009)).

In making this point, I am not just arguing that the two theories have similar practical policy implications when one takes into account political constraints and the like, which Dworkin acknowledges could be the case. Dworkin (1981a), *supra* note 460, at 187.

such as special concern for the effects of systemic prejudice, through *ad hoc* adjustments to their underlying mechanism for determining distributive justice. Both may demand larger (or smaller) contributions from those with high resources or potential earnings than some may consider reasonable. But the point is that they largely rise and fall together. Even if the fundamental concerns motivating each theory resonate with different people, both groups should agree on the general contours of the ideal tax and transfer system.

Second, this essay aims to strengthen optimal tax theory as a theory of distributive justice. Much of optimal tax theory, especially in the economic literature, disclaims any interest in normative matters. Instead, it often contends its only normative component of is the assumed social welfare function, which is the province of philosophers.[9] But optimal tax theory makes a number of foundational assumptions that are both normative and compelling. Unfortunately, it rarely takes the time to justify these assumptions, often leaving the impression that they were adopted simply for mathematical ease in formal models. In other cases, it has not explicitly responded to challenges raised by non-welfarists, leaving the impression that it cannot do so on its own terms. While optimal tax theory has been very influential, it could be even more so if it clarified and justified its normative underpinnings, and if supporters tried to reach some rough consensus on them.

To this end, this essay provides a deeper normative justification for some foundational assumptions in optimal tax theory, while still adhering to a welfarist framework. For example, following Lerner and Shaviro, it argues the assumption of identical utility functions for resources (which forecloses the possibility of so-called utility monsters and ascetics) is justifiable purely on welfarist grounds.[10]

It also advocates explicit and consistent adoption of other assumptions and modeling approaches that are implied by welfare egalitarianism, and address some common normative objections to it—and to resource egalitarianism as well. For example, contrary to the assertions of some skeptics, optimal tax theory does give "credit" for non-market-based contributions to society. But it should more consistently weight that "credit" by the contribution to social welfare, not the market value, following Micheletto, Kopczuk, and Batchelder.[11] It should also more explicitly acknowledge, following White, that the portion of the ideal endowment tax that reflects potential earnings should be partially based on one's hours worked times hourly potential earnings, not just on total potential earnings.[12] Optimal tax theory further implies an ideal tax that directly adjusts one's tax rate—not strictly one's tax base—based on one's *ex ante* endowment of objective resources. Moreover, the shorter individuals' decision horizons are, the more heavily the ideal tax should be based on current, not *ex ante*, potential income. More explicitly adopting these assumptions and modeling approaches would clarify that egalitarian optimal tax theory

[9] *See, e.g.*, Steven E. Landsburg, *The Methodology of Normative Economics*, 9(5) J. PUB. ECON. THEORY 757, 759 (2007).

[10] ABBA P. LERNER, THE ECONOMICS OF CONTROL 28–32 (1944); Shaviro, *supra* note 8, at 402-04.

[11] Luca Micheletto, *Redistribution and Optimal Mixed Taxation in the Presence of Consumption Externalities*, 92 J. PUB. ECON. 2262, 2266 (2008); Batchelder, *supra* note 8, at 30-32 (2009); Wojciech Kopczuk, *Economics of Estate Taxation: Review of Theory and Evidence*, 63 TAX L. REV. 139 (2009).

[12] Stuart White, *The Egalitarian Earnings Subsidy Scheme*, 29:4 BRITISH J. POL. SCI. 601 (1999). I use the term "potential income" here as shorthand for whatever is the best measure of economic status. As discussed below, there is a spirited debate within the optimal tax literature about whether potential income, labor earnings, personal consumption, or something else best encapsulates economic status.

does not necessarily entail the extreme limits on autonomy or unrelenting responsibility for prior choices that have troubled non-welfarists.[13]

Third, this essay aims to promote a more nuanced understanding of egalitarian optimal tax theory in the philosophical literature. This literature extensively details flaws of welfarism as a theory of distributive justice, critiquing optimal tax theory, endowment taxes, and utilitarianism specifically.[14] But this literature generally does not consider the full sweep of the theories it criticizes. Instead, it often conflates optimal tax theory with the Mirrlees model and its initial extensions, which focused strictly on labor earnings and utilitarian social welfare functions. As a result, objections to egalitarian welfarism and optimal tax theory are often predicated on, at best, a partial portrayal of what these theories actually say.

This essay seeks to address these oversights by explaining how optimal tax theory has responded to the primary objections of non-welfarists. For example, traditional concerns about how welfarists treat so-called utility monsters and ascetics[15] ignore optimal tax theory's basic assumption that individuals' utility functions for resources are identical. Objections that welfarists do not adequately account for non-market-based contributions to society overlook the rich literature on externalities. Objections that welfarism would require those with high earning potential to work nonstop to finance a life of ease for everyone else (so called "talent slavery") disregard egalitarian social welfare functions, the literature on tagging, and the implications of welfarism for how an optimal endowment tax would be structured. Meanwhile, concerns about choices that do not advance the decision maker's best interest, such as addiction, disregard the burgeoning literature on internalities. In making these claims, I admittedly define optimal tax theory broadly to include work that is sometimes classified differently, for example as the theory of optimal Pigouvian subsidies or optimal social insurance.

Given the influence of optimal tax theory in the economic and legal literature and in policy debates, philosophers of distributive justice could have more impact by understanding the many ways that it tracks major non-welfarist theories but simply uses different terms. Bridging these terminological divides would also enable philosophers to offer suggestions for how empirically-grounded optimal tax models should be improved or refined.

A few notes on the scope of this essay. The primary issue it explores is equality of what?[16] What metric should we ideally focus on if we believe justice demands some reduction in inequality? And, secondarily, how do we know when distributive justice has been achieved? In order to make the topic more manageable, I narrows the focus in several ways.

First, this essay only considers distributive justice in the context of economic institutions. While important and interrelated, it saves questions of what distributive justice implies for political

[13] LIAM MURPHY AND THOMAS NAGEL, THE MYTH OF OWNERSHIP: TAXES AND JUSTICE 104–05 (2002); David Hasen, *Liberalism and Ability Taxation*, 85 TEX. L. REV. 1059, 1064–71 (2007) (discussing conceptual objections to endowment taxation, including concerns about autonomy).

[14] *See, e.g.,* Dworkin (1981a), *supra* note 4; LIAM MURPHY AND THOMAS NAGEL, THE MYTH OF OWNERSHIP: TAXES AND JUSTICE 104–05 (2002); Linda Sugin, *A Philosophical Objection to the Optimal Tax Model*, 64 TAX. L. REV. 229 (2011).

[15] These are individuals who are unusually good or bad at converting resources into utility.

[16] This question was posed by Amartya Sen, *Equality of What*?, THE TANNER LECTURES ON HUMAN VALUES (May 22, 1979), Dworkin (1981a), *supra* note 4, at 187, and many others.

institutions and personal ethics for another day.[17] It also brackets questions about how government should finance public goods,[18] and what distributive justice implies about future generations and the ideal population size.[19] Further, it is normative, not positive. It focuses on what different theories of distributive justice imply about the ideal tax and transfer system, not what the structure of current tax or transfer policies implies about lawmakers' or the public's underlying normative principles.[20]

This essay also operates at the level of ideal policy design. On the one hand, it is unconcerned with the fundamental motivations and objectives of the theories discussed, but rather with their derivative principles for designing the ideal tax and transfer system. On the other hand, it is unconcerned with practical difficulties in measuring whatever metric we should be equalizing (sometimes referred to as the "equalisandum") so long as it is knowable and measurable to some extent. In fact, most of the metrics discussed are entirely implausible. They would require vast compliance costs, oceans of political goodwill, and a Herculean and benevolent tax authority with virtually unlimited resources. But this implausibility does not detract from the importance of identifying ideal policy design principles. Only then can policymakers consider whether and how to move towards them.[21]

Finally, in order to compare apples-to-apples, this essay focuses on theories of distributive justice that share two traits: they are egalitarian and employ a single objective function.[22] This excludes

[17] For compelling arguments of why these issues may be equally or more important than distributive justice in the context of economic institutions, *see* Elizabeth Anderson, *What Is the Point of Equality*?, 109 ETHICS 287 (1999); Liam B. Murphy, *Institutions and the Demands of Justice*, 27:4 PHIL. & PUB. AFFAIRS 251 (1999); Samuel Scheffler, *What Is Egalitarianism*?, 31:1 PHIL. & PUB. AFFAIRS 5 (2003). This essay also does not consider redistribution to be a public good, though some might disagree.

[18] For an excellent philosophical discussion of how this issue interacts with distributive justice, *see* MURPHY AND NAGEL, *supra* note 14. *See also* [ALAN PATTEN, PUBLIC GOOD FAIRNESS; DAVID MILLER, DISTRIBUTIVE JUSTICE].

[19] For discussion of these fraught issues, *see* Lawrence B. Solum, *To Our Children's Children's Children: The Problems of Intergenerational Ethics*, 35 LOYOLA OF L.A. L. REV. 162 (2001).

[20] Most of the optimal tax literature is normative, but a more recent vein adopts a positive approach. *See, e.g.*, Christina Fong, *Social Preferences, Self-Interest, and the Demand for Redistribution*, 82:2 J. PUB. ECON. 225–46 (2001); François Bourguignon & Amedeo Spadaro, *Tax-Benefit Revealed Social Preferences*, 10 J. ECON. INEQUALITY 75–108 (2012); Matthew Weinzierl, *The Promise of Positive Optimal Taxation: Normative Diversity and a Role for Equal Sacrifice*, 118 J. PUB. ECON. 128–42 (2014); Robin Jessen et al., *Optimal Taxation Under Different Conceptions of Justness* (Hohenheim Discussion Paper 27-2107, Sept., 2017).

[21] In the United States, taxes and transfers already vary to a large degree based on income, wealth, labor earnings, family structure, health status, immigration statuts, and employment. To a lesser extent, they vary by age, hours worked, state fiscal indicators, and zip code. *See, e.g.*, I.R.C. §§ 32; [PTCs; NMTC; UI]. One could imagine policymakers deciding to change the relative weight of these metrics or adding new ones.

[22] By egalitarian, I mean the objective function satisfies the Pigou-Dalton principle. The Pigou-Dalton principle says that given two possible profiles of utility across a group of individuals, the social planner should prefer the profile that reduces inequality between two individuals' utilities without changing their overall utility rank. Formally, given N individuals, let $u = (u_1, u_2, \dots, u_n)$ and $u^* = (u_1^*, u_2^*, \dots, u_n^*)$ be two utility profiles, where u_i is the utility of person i. If there exist two individuals a and b, such that: $u_a > u_b$, $\Delta > 0, and\ u_a^* = u_a - \Delta \geq u_b^* = u_b + \Delta$, and for all other individuals i (other than a and b) $u_i = u_i^*$, then the social planner should prefer profile u^*. The profile $u^*$is preferred as it reduces the inequality between persons a and b without changing a or b's utility rank and without having an impact on any other individuals. *See* Matthew D. Adler, *The Pigou-Dalton Principle and the Structure of Distributive Justice* (Working Paper

non-egalitarian theories of justice that are both welfarist (e.g., utilitarianism) and non-welfarist (e.g., Nozick's theory of entitlement). It also excludes theories, such as those of Rawls, Sen, and Nussbaum, that seek to promote multiple objectives without combining them into a single objective function.

One terminological note: The philosophical and economic literature on ideal taxes and transfers often refers to endowment, resources, or potential income as "ability" or "talent". This is unfortunate, making it sound like a person's expected market income is somehow a reflection of their skills or value as a person. But that is not at all what this literature intends. Instead, this literature generally uses the terms "ability" and "talent" to refer to the portion of an individual's economic outcomes that stems from chance—for example, the genetic and socioeconomic lottery, the whims of market preferences, or being a member of a dominant group in an prejudiced society—not choices over which an individual has control.[23] Obviously "ability" and "talent" so defined have no moral valence. But to avoid any negative connotations, I refer to this concept as potential income where possible.

This essay proceeds as follows. Section II provides a brief overview of optimal tax theory and resource egalitarianism, and defines the question explored here more precisely. Readers familiar with this literature may wish to skip to Section III. Section III explores how the two theories treat inequalities that do not stem from one's own choices, paying particular attention to their treatment of so-called utility monsters and other-regarding preferences. Section IV considers how each theory treats inequalities that arguably do arise from one's choices. This includes decisions regarding work effort and risk taking, choices that do not advance an individual's "true" preferences, and decisions in cases where one's preferences subsequently change. It also explores which actions or choices each theory considers to be morally relevant. Section V concludes.

II. A Brief Overview of Optimal Tax Theory and Resource Egalitarianism

A. Optimal Tax Theory

Optimal tax theory seeks to determine the set of taxes and transfers that maximize a social welfare function through a stylized model with a limited number of goods, and based on estimates of empirical realities.[24] It is a welfarist approach, meaning it considers distributive justice to be solely some function of a single measure of individual utility.[25] Utility, in turn, is defined as an

(2013)). *See also*, Bhaskar Dutta, *Inequality, Poverty, and Welfare*, *in* HANDBOOK OF SOCIAL CHOICE AND WELFARE 597 (eds. Kenneth Arrow et al., 2002).

[23] As discussed in Section IV, the issue gets considerably more complicated when one allows for the possibility that people can change their ability or talent through their choices over time.

[24] For excellent, short overviews of the literature, *see* N. Gregory Mankiw, Matthew Weinzierl, & Danny Yagan, *Optimal Taxation in Theory and Practice*, 23 J. ECON. PERSP. 147 (2009); Thomas Piketty & Emmanuel Saez, *Optimal Labor Income Taxation*, HANDBOOK OF PUBLIC ECONOMICS, VOL. 5, 391–474 (2013). The literature typically sets aside the question of whether to provide public goods and how to finance them. It does so by assuming either that they do not exist or that a fixed amount of revenue must be raised to finance them.

[25] *See, e.g.*, Dworkin (1981a)*, supra* note 4, at 244. Some dispute this definition. They distinguish between individualist social welfare functions, which are solely a function of individual utility, and non-individualistic social welfare functions, which can also be a function of the magnitude or distribution of specific goods independent of their effect on individual utility. *See, e.g.*, KAPLOW (2008), *supra* note 7, at 348. Others propose utility-supported theories that are not welfarist because utility is not defined as a single metric but

expression of individual preferences. This definition is the norm in modern economics but sometimes misunderstood. For example, if I choose to undertake a difficult project, I may know that the choice will not make me feel happier or more successful. In fact, I may expect that it will make me feel tired, unappreciated, and overwhelmed. But provided that I duly considered the choice and have full information, it is my preference and by definition increases my utility. I may have made the choice because I believe it will make my life more meaningful or complete, or because I will learn from it. But my rationale does not matter from a welfarist perspective. The choice is simply defined as giving me more utility, at least for the foreseeable future.

Optimal tax theory recognizes the inherent indeterminacy of comparing utility across different people. As a result, it considers the question of furthering distributive justice to be an information problem. We know that we want to maximize social welfare, which under an egalitarian social welfare function is the sum of welfare-weighted individual utilities (i.e., with improvements in the utility of the worse-off weighted more heavily).[26] But because we cannot actually measure utility across people, we must identify a proxy—or a set of proxies—for the well-being of different individuals relative to each other. We can then use those proxies to determine the level and shape of inequality, and how best to mitigate it.

1. The Mirrlees model and its extensions

One of the foundational elements of the optimal tax literature is Mirrlees' 1971 paper, *An Exploration in the Theory of Optimum Taxation*, and its extensions.[27] Like all optimal tax theory, Mirrlees aims to find the set of taxes and transfers that maximize a social welfare function through a stylized model. The Mirrlees model assumes that individuals differ only on in one exogenous characteristic—their earning ability or wage rate—which determines their utility level. Thus, we ideally want to redistribute based on earning ability. This ideal, however, is impossible. Individuals know their own earning potential, but not that of others. Instead, the only proxy for earning potential that others and government can observe is each individual's actual earnings, which is the product of their wage rate and hours worked. Because it is a one-period model, there is no savings, and actual earnings equals consumption.

In order make the math tractable, the Mirrlees model also assumes that utility is purely a product of two goods: consumption and leisure, and that all individuals have the same utility function for these two goods. This means every individual gets the same utility from a fixed amount of

---

a vector of desires. *See, e.g.*, Amartya Sen, *Plural Utility*, in 81 PROCEEDINGS OF THE ARISTOTELIAN SOCIETY, NEW SERIES 193, 201, 207 (1980-1981).

[26] Rather than giving priority to those who are worse-off, an egalitarian social welfare function can include an independent objective of minimizing inequalities of utility. This form of social welfare function is less common because it can imply "leveling down," or making everyone worse off so long as it reduces inequalities of utility.

[27] James Mirrlees, *An Exploration in the Theory of Optimum Taxation*, 38 REV. ECON. STUD. 175 (1971). Foundational work by Ramsey, Atkinson, and Stiglitz on optimal commodity taxation is also considered part of the optimal tax literature, but it is less relevant for this essay because it assumes that social welfare is maximized by maximizing efficiency, and that commodity taxes are the only possible taxes. *See* Frank P. Ramsey, *A Contribution to the Theory of Taxation*, 37 ECON. J. 47 (1927); ANTHONY B. ATKINSON & JOSEPH E. STIGLITZ, LECTURES ON PUBLIC ECONOMICS (1980). *See also* Peter A. Diamond & James A. Mirrlees, *Optimal Taxation and Public Production I: Production Efficiency*, 61 AM. ECON. REV. 8, (1971); Peter A. Diamond & James A. Mirrlees, *Optimal Taxation and Public Production II: Tax Rules*, 61 AM. ECON. REV. 261, (1971).

consumption and a fixed amount of leisure. It also means all individuals at a given earning ability level have the same utility, and choose the same mix of work and leisure. Individuals may, of course, choose to consume different things, like movies or restaurant meals. But their relative utility is purely a function of their earning potential.

The brilliance of Mirrlees' model was threefold. First, it was among the first work in economics to identify the problem of how best to further distributive justice as one of information asymmetries.[28] We want to redistribute based on utility but, because this is unknowable, we must use proxies. These proxies generate behavioral responses as individuals try to minimize taxes owed and maximize transfers received, which in turn create challenges in determining the ideal level and form of redistribution. Second, it was the first to address this problem through a mathematical model to which empirical parameters were applied.[29] Third, the math involved was incredibly difficult.

Based on this model, Mirrlees estimated that the optimal labor income tax was one that combines a universal lump-sum transfer, which is substantial but not massive, with a tax on earnings that applies lower marginal tax rates as earnings rises.[30] The lump-sum transfer is typically referred to as a "demogrant" and is considered part of the tax system because it is a negative tax.[31] The combined tax system is progressive even though the marginal tax rates decline with earnings. This occurs because the demogrant offsets the tax burden on lower-earning households so that their average tax rate (total taxes as a share of income) is low or negative. Higher-earning households, on the other hand, face higher average tax rates because the demogrant is a relatively small share of their income.

Since Mirrlees' seminal paper, there have been a number of advances in this component of optimal tax theory. Until the 2000s, these advances principally involved using more accurate empirical estimates and different social welfare functions. This generally resulted in a larger demogrant and higher top marginal tax rates.[32] But, Mirrlees' core model and findings—that the

[28] This issue was also identified by Vickrey in William Vickrey, *Measuring Marginal Utility by Reactions to Risk*, 13 ECONOMETRICA 215 (1945). He and Mirrlees shared the Nobel Prize in 1996 for this insight.

[29] Specifically, Mirrlees used empirical estimates of the distribution of wages and the responsiveness of labor supply to taxes in order to estimate the optimal tax schedule.

[30] Mirrlees' finding that declining marginal tax rates are optimal surprised many. The intuition behind it is that declining marginal rates result in high-ability workers facing relatively low tax rates at the margin; that is, when deciding whether to work a bit more or less. As a result, they work more than they would if they faced higher marginal tax rates. This increases the overall amount of taxes collected and the size of the demogrant because their wage rate is so high. While lower-ability workers face higher tax rates as part of the package, their greater tax burden is more than offset by the larger demogrant.

[31] He found that, under a utilitarian social welfare function, the optimal demogrant was 17 percent of median income, and the top marginal tax rate was 26 percent, declining to 17 percent as income rose. Mirrlees also considered an egalitarian social welfare function that maximized the product of utilities. The optimal demogrant then rose to 30 percent of median income and the top marginal tax rate rose to 34 percent, declining to 20 percent as income increased.

[32] For example, Tuomala used updated estimates of labor supply elasticities and estimated that the optimal demogrant was 40 percent of mean income under a utilitarian social welfare function, and 58 percent when the social welfare function was the product of utilities. The associated top marginal rates were 50 and 68 percent, respectively. MATTI TUOMALA, OPTIMAL INCOME TAX AND REDISTRIBUTION (1990). *See also* Christopher Heady, *Optimal Taxation as a Guide to Tax Policy: A Survey*, 14 FISCAL STUD. 15 (1993); Ravi Kanbur & Matti Tuomala, *Inherent Inequality and the Optimal Graduation of Marginal Tax Rates*, 96 SCANDINAVIAN J. ECON.

optimal income tax was one with declining marginal rates and a demogrant—remained largely intact and influential. During this period, the fiscal policies of developed countries around the world moved haltingly in the direction of Mirrlees' recommendations. Generally the marginal tax rates facing high earners declined, and income tax rate schedules became flatter.[33] A number of countries also implemented or expanded guaranteed income programs.[34]

Around the turn of the century, however, a new generation of optimal tax theorists began to expand on Mirrlees' model and question some of his core findings. For example, Saez (2001) estimated that the optimal tax rate schedule is actually U-shaped. This was based on a different mathematical model and more refined empirical estimates.[35] Many countries, including the U.S., currently exhibit this type of rate structure once transfers are included.[36]

Saez also considered optimal taxation under a maximin social welfare function,[37] and estimated that the top marginal tax rate on high earners and the optimal demogrant were then much higher. For example, he estimated that the optimal demogrant was between 50 and 60 percent of mean income.[38] While this demogrant may seem modest, a demogrant of 100 percent of mean income implies complete (and completely efficient) confiscation and redistribution. In other words, it can only occur if all individuals face a 100 percent tax rate on all of their earnings, there are no efficiency effects, and all revenues collected are used to fund the demogrant.

In a further permutation, Saez (2002) considered the implications if jobs are lumpy and people therefore tend to make decisions about how much to work at the "extensive margin"—that is,

---

275 (1994); Matti Tuomala, *On Optimal Non-linear Income Taxation: Numerical Results Revisited*, 17 INT'L TAX POL'Y FIN. 259 (2010) (finding that the optimal demogrant is especially sensitive to the degree of inequality aversion in the social welfare function).

[33] *See* Mankiw et al, *supra* note 24. For an overview of the flattening of both corporate and personal income tax schedules between 1981 and 2005, *see* Klara Sabirianova Peter, Steve Buttrick, & Denvil Duncan, *Global Reform of Personal Income Taxation, 1981-2005: Evidence from 189 Countries*, 63 NAT'L TAX J. 447 (2010) (documenting how the GDP-weighted average top marginal tax rate fell from 62% in 1981 to 36% in 2005).

[34] *See* OECD, Social Expenditures, Aggregated Data (on file with author); Jurgen de Wispelaere, *Basic Income in Our Time: Improving Political Prospects Through Policy Learning?*, 45 J. SOC. POL'Y 617 (2016) (discussing the universe of guaranteed basic income programs that have been tried); BASIC INCOME: AN ANTHOLOGY OF CONTEMPORARY RESEARCH (Karl Widerquist et al. eds., 2013).

[35] Emmanuel Saez, *Using Elasticities to Derive Optimal Income Tax Rates*, 68 REV. ECON. STUD. 205, 223–25 (2001) (estimating the optimal marginal rate schedule for the U.S. declines with income until about $75,000 in earnings and then increases until about $200,000 in earnings, after which point it becomes roughly flat). *See also* Peter A. Diamond, *Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates*, 88 AM. ECON. REV. 83 (1998); Matti Tuomala, *On Optimal Non-linear Income Taxation: Numerical Results Revisited*, 17 INT. TAX PUB. FIN. 259 (2010).

[36] This occurs because transfer programs tend to phase-out with income, which generates high implicit tax rates in the phase-out range. Once the phase-out is complete, marginal tax rates drop precipitously and then gradually rise with income. *See, e.g.*, Steven D. Holt & Jennifer L. Romich, *Marginal Tax Rates Facing Low– and Moderate–Income Workers Who Participate in Means–Tested Transfer Programs*, 60 NAT'L TAX J. 253 (June, 2007); Daniel Shaviro, *The Minimum Wage, the Earned Income Credit and Optimal Subsidy Policy*, 64 UNIV. CHIC. L. REV. 405 (1997).

[37] Technically a maximin social welfare function first maximizes the utility of the least well-off person, then, if that person's utility cannot be increased further, in maximizes the utility of the second least well-off person holding harmless the least well-off person, and so forth.

[38] Saez (2001), *supra* note 35 (also estimating an optimal top rate on high earnings of 70 to 80 percent).

whether to work full-time, or not at all.[39] If so, he estimated that the optimal tax combines a smaller demogrant with negative marginal tax rates (earnings subsidies) for low earners, much like those created by the Earned Income Tax Credit and Child Tax Credit in the U.S.[40]

In recent years, other work in the field has further altered or expanded upon some of the stylized assumptions in the Mirrlees model. For example, Lockwood, Weinzierl, and others have estimated that the optimal tax system is typically less redistributive if individuals do not have the same utility function for consumption and leisure, and if the preference for leisure falls with income.[41] Batchelder, Kopczuk, Piketty, and Saez have considered the optimal tax treatment of inheritances, which are ignored in the Mirrlees model. They estimate that the optimal tax rate on transfers between the wealthy is quite high.[42] Batchelder also concludes that heirs should optimally be taxed on inherited income, but at a rate that rises with both the earned income of the heir and the amount she inherits.[43] Blomqvist, Horn and others have estimated that the optimal redistributive system includes public health insurance if people differ in their earning potential and probability of illness.[44] Recent work by Saez and Stantcheva has begun to incorporate multiple objectives and constraints into optimal tax models through generalized social welfare weights. This opens up the field to modeling non-welfarist theories of distributive justice and positive political objectives, both of which are outside the scope of this paper.[45]

2. Tags

The second foundational element of the optimal tax literature is the concept of "tags." Tags are immutable or semi-immutable traits that are correlated with income potential, such as gender,

[39] Emmanuel Saez, *Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses,* Q. J. ECON 1039 (June, 2002).
[40] *Id*.
[41] Benjamin B. Lockwood & Matthew Weinzierl, *De Gustibus Non Est Taxandum: Heterogeneity in Preferences and Optimal Redistribution*, 124 J. PUB. ECON. 74, 75 (2015). *See also* Agnar Sandmo, *Optimal Redistribution when Tastes Differ*, 50:2 FINANZARCHIV (1993); Kenneth Judd & Che-Lin Su, *Optimal Income Taxation with Multidimensional Taxpayer Types* (July 31, 2006) (unpublished manuscript) (on file with author).
[42] *See* Batchelder (2009), *supra* note 8, at 43–44 (estimating the optimal inheritance tax rate on the highest income heirs receiving the largest inheritances is as high as 80%, based on research on bequest motives in the U.S.); Kopczuk, *supra* note 11, at 139 (suggesting that the optimal tax rate on wealth transfers between those at the top of the economic distribution is positive); Thomas Piketty & Emmanuel Saez, *A Theory of Optimal Inheritance Taxation*, 81 ECONOMETRICA 1851 (2013) (estimating that the optimal inheritance tax rate is as large as 50% based on data from France and the U.S.); Emmanuel Farhi & Ivan Werning, *Progressive Estate Taxation*, 125 Q. J. ECON. 635, 646 (2010).
[43] Batchelder (2009), *supra* note 8, at 11–44.
[44] Ake Blomqvist & Henrik Horn, *Public Health Insurance and Optimal Income Taxation*, 24 J. PUB. ECON. 352 (1984). *See also* David A. Weisbach, *A Welfarist Approach to Disabilities* (Chicago John M. Olin Law & Economics Working Paper No. 355, Aug. 2007) (arguing that the optimal tax system would transfer more to people with disabilities, but a larger share of the transfers would be in kind if the disability were unobservable).
[45] Emmanuel Saez & Stefanie Stantcheva, *Generalized Social Marginal Welfare Weights for Optimal Tax Theory*, 106 AM. ECON. REV. 24 (2016). For example, they discuss how this approach could integrate Roemer's work on equal opportunity by setting social welfare weights to zero to the extent an individual's income is attributable to luck, or political concerns about "free-loaders" by applying zero weights to those who would have worked absent transfer programs. Id. at 25.

race, height, and parental education. In an article published shortly after Mirrlees', Akerlof showed that adjusting tax rates based on tags can improve the efficiency of a redistributive scheme because of the information they provide about one's underlying potential income.[46] The concept of tags has also been justified on distributional equity grounds.[47]

Recent research has expanded on the concept of tags and incorporates multiple periods. (The Mirrlees model, and most that initially followed, assumed there was only one period or that individuals have perfect knowledge of their lifetime earning potential in advance.) The implications of this newer work, called the new dynamic public finance literature, are complex. Golosov et al argue that if the government knows an individual's earning history, this information can be used as a tag for earning ability, especially if an individual's earnings suddenly drop or rise.[48] For example, if an individual has little income in the current period but high earnings in the past, this may indicate that she has high earning potential but has decided to take time off work and live off her savings.[49] Weinzierl, Farhi, Werning, and others find that if the government does not know individuals' earnings histories, it may be desirable to incorporate other tags, such as age.[50]

Allowing for multiple periods also raises the question of whether it is appropriate to treat an individual as the same person over his or her life. The general conclusion has been that if people's economic planning horizons are shorter than a lifetime, it may be appropriate to treat them as effectively different people over each planning period.[51] As optimal tax models have incorporated multiple periods and allowed for shorter decisions horizons, the literature has begun to question another long-standing, rough consensus—that consumption taxes (or taxes on labor earnings) are

[46] George A. Akerlof, *The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning*, 68 AM. ECON. REV. 1 (1978). *See also* Albert L. Nichols & Richard J. Zeckhauser, *Targeting Transfers through Restrictions on Recipients*, 72 AM. ECON. REV. 371 (1982); N. Gregory Mankiw & Matthew Weinzierl, *The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution*, 2 AM. ECON. J.: ECON. POLICY 155 (2010). The concept of tags also underlies the well-known work on commodity taxation, which concludes it can be optimal to tax specific commodities if they are complements to leisure or correlated with earning ability. *See* W. J. Corlett & D. C. Hague, *Complementarity and the Excess Burden of Taxation*, 21 REV. ECON. STUD. 21 (1953); [Kaplow, 2000]; Emmanuel Saez, *The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes*, 83 J. PUB. ECON. 217 (2002).

[47] *See, e.g.*, Kyle Logue & Joel Slemrod, *Genes as Tags: The Tax Implications of Widely Available Genetic Information*, 61 NAT'L TAX J. 843 (2008); Zachary Liscow, Note, *Reducing Inequality on the Cheap: When Legal Rule Design Should Incorporate Equity as Well as Efficiency*, 123 YALE L. REV. 2478 (2014).

[48] *See, e.g.*, Mikhail Golosov, Narayana Kocherlakota and Aleh Tsyvinski, *Optimal Indirect and Capital Taxation*, 70 REV. ECON. STUD. 569, 570–71, 577 (2003).

[49] *Id*.

[50] Matthew Weinzierl, *The Surprising Power of Age-Dependent Taxes*, 78:4 REV. ECON. STUD. 1490 (2011); Emmanuel Farhi & Ivan Werning, *Insurance and Taxation over the Life Cycle*, 80 REV. ECON. STUD. 596, 600, 620–21, (2013) (estimating that labor income taxes should rise with age, in part because current earnings are more reflective of future earnings when one is older); [Kremer, 2002]. Others find it might be desirable to use age tags to influence human capital choices. *See e.g.*, Jean-Marie Lozachmeur, *Optimal Age Specific Income Taxation* (CORE Discussion Paper No. 2002/46 (2002)); Sören Blomquist & Luca Micheletto, *Age-related Optimal Income Taxation*, 110 SCAND. J. ECON. 45 (2008).

[51] *See, e.g.*, Daniel Shaviro, *Beyond the Pro-Consumption Tax Consensus*, 60 STAN. L. REV. 745 (2007); Lawrence Zelenak, *Tax Policy and Personal Identity over Time*, 62 TAX L. REV. 333 (2009).

more efficient and fair than income taxes.[52] While that debate is important, it is unrelated to most of the topics discussed here. As a result, I will generally use the terms earnings, consumption, and income interchangeably, except when explicitly discussing this debate.

3. Externalities and internalities

The final foundational element of the optimal tax literature involves externalities and internalities.[53] A negative externality arises whenever a choice by one person makes another worse off, but the first person does not bear those costs. Positive externalities are simply the reverse. When externalities apply to broad sections of the population, there is often no market for them, and Coasian bargaining is not a solution. Instead, Pigou showed that the government can enhance efficiency by correcting for externalities through fiscal incentives, now termed Pigouvian taxes and subsidies.[54]

The ideal Pigouvian tax corrects for a negative externality by taxing the good generating social costs so that its market price reflects its social value—not only the value the purchaser places on it. The optimal tax rate is the cost that marginal consumption of the good imposes on others.[55] This causes the marginal decision maker to fully internalize the social costs to others of his or her choice, and results in the optimal supply of the good.

Traditionally, externalities were viewed as solely an efficiency issue. But Pigouvian taxes and subsidies also raise equity questions because they raise revenue or cost money, and can alter the economic distribution. The general conclusion from work in this field is that if certain groups disproportionately generate negative externalities, this should not change the magnitude of the

[52] This consensus was based in part on the fact that consumption taxes and wage taxes generally tax individuals with the same lifetime earnings at the same rate, regardless of whether they choose to consume their labor earnings immediately or defer consumption to the future. By contrast, income taxes apply higher average tax rates to individuals who defer consumption because they tax income from savings in addition to income from labor or consumption. Historically, many viewed these higher tax rates as unfair and inefficient. *See, e.g.*, William Andrews, *A Consumption-Type or Cash Flow Personal Income Tax*, 87 HARV. L. REV. 1113 (1974); A. B. Atkinson & J. E. Stiglitz, *The Design of Tax Structure: Direct Versus Indirect Taxation*, 6 J. PUB. ECON. 55 (1976); Joseph Bankman & David A. Weisbach, *The Superiority of an Ideal Consumption Tax Over an Ideal Income Tax*, 58 STAN. L. REV. 1413 (2010).

There was never complete consensus that a consumption or wage tax was superior to an income tax, especially one that applies lower rates to capital income. For example, most agreed that some tax on capital was optimal if capital income was a complement to leisure. *See, e.g.*, Corlett & Hague, *supra* note 46.

But the contribution of the more recent literature has been to highlight many other circumstances under which taxing capital income—potentially at quite high rates—may be optimal. These include when capital income is correlated with earning ability, and when capital income (or the lack thereof) does not represent conscious decisions to defer consumption because people cannot plan over such long time periods. Similar arguments apply to wealth taxes.

[53] While a foundational topic in the optimal tax literature, externalities and internalities are not strictly issues of distributive justice, but also (or perhaps mostly) issues of corrective justice. *See, e.g.*, Solum, *supra* note 19, at 174–76.

[54] ARTHUR C. PIGOU, WEALTH AND WELFARE (1912). *See also* W.J. Baumol, *On Taxation and the Control of Externalities*, 62:3 AM. ECON. REV. 307 (1972).

[55] For example, if consumption of the good only negatively affects wealthy individuals, the optimal Pigouvian tax is lower than if it only affects low-income individuals because the social welfare function does not weight improvements for the wealthy very heavily. *See* Kopczuk, *supra* note 11, at 141; Batchelder (2009), *supra* note 8, at 30-32 (both making this point in the context of gratuitous wealth transfers).

optimal Pigouvian tax, even if those groups have disproportionately high or low potential income. Instead, the revenue the tax raises should be used to offset its general distributional effects through lump-sum transfers, assuming the tax system was otherwise optimal prior to the Pigouvian tax.[56] The net effect is that the revenue raised is returned to individuals in the same potential income class as those whose choices were generating the negative externalities, but the price of making such choices is higher. An optimal Pigouvian tax also generates efficiency gains. These gains should be redistributed in whatever way is implied by the social welfare function.

Sometimes markets fail not because of externalities, but because the decision maker does not fully account for the costs and benefits of her choice to herself. This is referred to as an internality. Internalities arise when the numerous biases and costs often associated with making decisions mean one cannot assume an individual's choices (often referred to as their "revealed preferences") necessarily represent their underlying or "fundamental" preferences. In cases where such preferences conflict, optimal tax theorists have long assumed social welfare should be assessed based on experienced utility, not decision utility.[57] Experienced utility represents what individuals would rationally choose if there were no decision making costs, they were fully informed, and they were accounting for their views both over time and in the moment. In contrast, decision utility reflects what individuals actually choose in the moment, given the costs of making decisions and limited information.[58]

The optimal Pigouvian tax for an internality is generally similar to that for an externality. It equals the cost that marginal consumption of the good imposes on the decision maker herself that she would otherwise fail to take into account.

* * *

Thus, to summarize and simplify slightly, optimal tax theory posits that the ideal tax is on endowment—including potential earned and unearned material resources—adjusted for externalities, internalities, and conditions that leave one better or worse off controlling for material resources.

B. Resource Egalitarianism

The primary non-welfarist theory on which this essay focuses is Dworkin's resource egalitarianism. It regards a just distribution as one that treats all members of society with equal concern and respect, as determined from a position of equality.[59] This has two components. First, everyone

---

[56] Louis Kaplow, *Optimal Control of Externalities in the Presence of Income Taxation*, 53 INT'L ECON. REV. 487 (2012); KAPLOW (2008), *supra* note 7.

[57] *See, e.g.*, HENRY SIDGWICK, THE METHODS OF ETHICS (1907); John H. Harsanyi, *Cardinal Welfare, Individualist Ethics, and Interpersonal Comparisons of Utility*, 63 J. POL. ECON. 309 (1955); James A. Mirrlees, *The Economic Uses of Utilitarianism*, in UTILITARIANISM AND BEYOND 63 (Amartya Sen & Bernard Williams, eds., 1982); KAPLOW (2008), *supra* note 7, at 360.

[58] *See, e.g.*, Daniel Kahneman et al., *Back to Bentham? Explorations on Experienced Utility*, 112 Q.J. ECON. 375 (1997); Sendhil Mullainathan, *A Reduced-Form Approach to Behavioral Public Finance*, 4 ANN. REV. ECON. 511 (2012).

[59] RONALD DWORKIN, SOVEREIGN VIRTUE: THE THEORY AND PRACTICE OF EQUALITY 313 (2000) (hereinafter, Dworkin (2000)); Ronald Dworkin, Sovereign Virtue *Revisited*, 113 ETHICS 106, 106–07 (2002) (hereinafter, Dworkin (2002)).

should begin their life with equal transferrable resources, as measured by the value of those resources to others.[60] But not all resources, such as one's talents or health, are transferrable. And, if initial resources were the only focus, society would quickly become unequal because of other forms of luck. Thus, he posits a second component of treating people with equal concern: a hypothetical insurance market that is the model for the community's tax and transfer system.[61] This hypothetical insurance would mitigate the effects of what he terms brute luck, but not option luck.[62] Option luck is the result of deliberate and calculated gambles that the individual might have anticipated and declined, whereas brute luck is the result of risks that were not such deliberate gambles.[63]

The amount and form of insurance in Dworkin's ideal fiscal system would be the insurance an average person would select behind a thin veil of ignorance where she knew her own tastes and preferences, but not her talents, socioeconomic status, or the preferences of others.[64] Dworkin contends the result would be partial, but not complete, insurance against brute luck. This is because making individuals with some conditions (such as severe disabilities) equally well-off would require such extraordinarily large transfers that the average person would select a less-than-full insurance system against such conditions. These high costs of achieving *ex post* equality stem from the obvious direct negative effects on the disabled individual, but also from costs endemic to insurance markets, including moral hazard.[65]

Instead of complete *ex post* equality, Dworkin's goal in employing the device of a thin veil of ignorance is equality in the resources available to individuals to pursue their life plans. While people may end up with different resources *ex post*, a less well-off person will not reasonably envy others because she will be insured against risks at a level she would have selected if she was an average person with an average likelihood of personally experiencing each form of brute luck.[66]

Resource egalitarians argue their approach is fundamentally different from welfare egalitarianism.[67] Dworkin maintains this is the case even when policymakers do not have full information about the distribution of tastes and disabilities, and are therefore operating under conditions of uncertainty.[68] This essay challenges these claims. It argues that, to a large degree, resource and welfare egalitarianism converge in their guiding policy design principles—especially under conditions of uncertainty.

---

[60] Specifically, Dworkin posits a Walrasian auction for transferrable resources at the start of life. Dworkin (1981b), *supra* note 2, at 285–89.

[61] Dworkin (1981b), *supra* note 2, at 292–99; Dworkin (2002), *supra* note 59, at 108.

[62] Dworkin (1981b), *supra* note 2, at 293.

[63] Dworkin (1981b), *supra* note 2, at 293.

[64] At points Dworkin suggests the hypothetical insurance should ideally be what each individual, not the average person, would choose behind the veil of ignorance, but he argues such full information is unrealistic. Dworkin (1981b), *supra* note 2, at 298. At other points, he argues the ideal is what the average person would select because the administrative costs of an individualized system are too high. Dworkin (2002), *supra* note 59, at 112.

[65] Dworkin (1981b), *supra* note 2, at 300; Dworkin (2002), *supra* note 59, at 108–11.

[66] Dworkin (2002), *supra* note 59, at 108.

[67] *See infra* notes __-__ & accompanying text.

[68] Dworkin (1981a), *supra* note 4, at 187.

C. Defining the Question

With this rough grounding in the two theories, we now turn to this essay's central question—how well does optimal tax theory respond to the primary critiques of it by non-welfarists, paying particular attention to those of resource egalitarians. To this end, it is necessary to define the comparison more precisely.

First, in order to compare apples-to-apples, I narrow the discussion to theories of distributive justice that are egalitarian. Welfarists are sometimes accused of not taking disparities between individuals seriously. This is a fair charge against utilitarians, who would theoretically condone killing one person to save several others.[69] But welfarists adopt many social welfare functions that value improvements in the well-being of the worse-off more heavily, especially in the optimal tax literature. These egalitarian social welfare functions include those that aim to maximize the utility of the least well-off person (maximin), the product of utilities, and so forth.[70] Similarly, I will limit the discussion of non-welfarist theories to those that are egalitarian. For example, I will not discuss Nozick's theory of distributive justice because it is concerned with entitlements, not equality.[71]

Second, I focus on theories with a single objective function. This is the norm in optimal tax theory, though some recent models permit integration of multiple objectives, including non-welfarist ones.[72] But it is not at all the norm outside of welfarist theories. In particular, two prominent, non-welfarist theories that fall outside this focus are Rawls' theory of justice[73] and Sen's and Nussbaum's capabilities approach.[74] Rawls accords special significance to primary goods and his theory is multi-layered, with certain objectives coming in to play once others are satisfied. Sen and Nussbaum focus on specific functionings and capabilities. Neither theory offers a clear weighting among these multiple objectives. Among non-welfarist theories that do employ a single objective function, I will generally focus on resource egalitarianism. But there are some related theories that may align (perhaps even more) with optimal tax theory, including weighted beneficence and equal opportunity for welfare.[75]

---

[69] *See, e.g.,* JOHN RAWLS, A THEORY OF JUSTICE 23–24 (2005).

[70] Maximin social welfare functions are akin to Rawls' difference principle (with many important differences), and other egalitarian social welfare functions are akin to weighted beneficence. *See* Liam Murphy, *Liberty, Equality, Well-Being: Rakowski on Wealth Transfer Taxation*, 51 TAX L. REV. 473, 480 (1996) (hereinafter Murphy (1996)).

[71] ROBERT NOZICK, ANARCHY, STATE AND UTOPIA 150–53 (1974, 2001 ed.). I also will not discuss libertarian theories like Barnett's, though arguably they could be characterized as egalitarian in the sense of wanting to equalize certain liberties. *See* Solum, *supra* note 19, at 185–86.

[72] *See*, *e.g.*, Saez & Stantcheva, *supra* note 45; [Fleurbaey & Maniquet].

[73] *See generally* RAWLS, *supra* note 69, at 60–83.

[74] *See generally* Sen, *supra* note 16 at 217–220; MARTHA C. NUSSBAUM, WOMEN AND HUMAN DEVELOPMENT: THE CAPABILITIES APPROACH (2000).

[75] *See, e.g.*, Murphy (1996), *supra* note 70, at 477-80 (defining weighted beneficence as concern with individual well-being that assigns greater weight to improving the well-being of those who are less well-off); Richard Arneson, *Equality and Equal Opportunity for Welfare*, 56 PHIL. STUDIES 77, 85–92 (1989); John E. Roemer, *Three Egalitarian Views and American Law*, 20 LAW & PHIL 433 (2001); Anne Alstott, *Equal Opportunity and Inheritance Taxation*, 121 Harv. L. REV. 470 (2007). An even closer cousin of resource egalitarianism is Rakowski's theory of equal fortunes. ERIC RAKOWSKI, EQUAL JUSTICE (1991).

Even with the comparison thus narrowed, at one level optimal tax theory remains fundamentally distinct. Assuming the social welfare function is egalitarian, it is concerned with improving individuals' subjective well-being, especially among those who are worse off. It is intrinsically not concerned with whether suffering has a rational or objective explanation, or whether it stems in part from the individual's choices. But it may be deeply concerned with these issues instrumentally because of the effect on incentives.

By contrast, the fundamental concern embodied in resource egalitarianism is more objective and removed. It aims to treat people as equals by reducing or eliminating inequalities that arise from objective and morally arbitrary sources. But, unlike welfare egalitarianism, it is intrinsically unconcerned with inequalities in people's subjective experience of their lives, or with inequalities stemming from choices for which they bear responsibility.

These differences between optimal tax theory and resource egalitarianism are fundamental. But they are differences of motivation and concern. My goal here is to focus on principles guiding the ideal policy. How does each theory treat different sources of inequality? How can proponents tell when distributive justice has been achieved?

To address these questions, the next section considers how each theory treats sources of inequality that are not linked to one's own choices. The following section considers those that arguably are.

III. Should Distributive Justice Focus on Welfare or Resources?

A. Utility Monsters and Ascetics

The first objection non-welfarists often levy against welfarism is that it unfairly implies that those who subjectively get more marginal utility from a given level of resources (so-called "utility monsters") should get the lion's share of society's resources.[76] They also object that welfarism implies fewer resources for those who get relatively less utility from resources.[77] This could occur because an individual is simply an ascetic by nature. Or it could be a psychological adaptation or survival strategy in an unjust world, where the individual is unlikely get many resources for morally arbitrary reasons, such as their race, gender, or socioeconomic background.[78]

While a compelling concern, these objections do not apply to optimal tax theory because it assumes identical utility for a given level of resources. Initially, Mirrlees assumed people differed only in their potential earnings, and that they had the same utility functions for potential earnings and the two goods it could buy: market consumption and leisure. However, later extensions of the Mirrlees model assumed people differed in other ways that more closely track Dworkin's definition of resources, including capital income, inherited income, educational investments, and health, to name a few.[79] While these extensions broadened the definition of resources, they

[76] *See, e.g.* Dworkin (1981a)*, supra* note 4, at 228–38; ERIC RAKOWSKI, EQUAL JUSTICE 41 (1991).
[77] *See* Dworkin (1981a)*, supra* note 4, at 239–40.
[78] *See, e.g.*, AMARTYA SEN, ON ETHICS AND ECONOMICS 45–46 (1987). Dworkin's insurance mechanism has also been criticized on these grounds; theoretically someone with adaptive preferences might easily be non-envious even with a very small endowment.
[79] *See, e.g.*, Golosov et al., *supra* note 48; Batchelder (2009), *supra* note 8; Stefanie Stantcheva, *Optimal Income, Education and Bequest Taxes in an Intergenerational Model* (NBER Working Paper no. 21177, May,

continued to assume people's utility was the same for any given level of resources they possess.[80] In short, utility monsters and ascetics simply do not exist in optimal tax theory.

Dworkin anticipated this response. He noted that if one adopted an objective conception of welfare, equality of welfare would be similar or even identical to equality of resources.[81] Moreover, the objective conception of welfare he describes is quite similar to optimal tax theory's assumption of identical utility functions for resources. Nevertheless, he argued, any theory adopting such a conception of welfare is not welfarist at all, because it is not subjective and requires "an independent theory of fair distribution."[82] Instead, it is only the same as resource egalitarianism "in the (misleading) language of welfare."[83]

Dworkin is correct that the assumption of identical utility for resources is deeply normative. And it has been adopted with little justification in the optimal tax literature. Some note in passing that it is not reasonable or feasible to adopt any other assumption.[84] Others appear to adopt the assumption simply for mathematic tractability. Moreover, intuitively it cannot be correct. If we actually could compare utility across lives, surely we would find that people differ in the relative utility they derive from the same amount of resources.[85]

But simply because the optimal tax literature has discussed this assumption briskly to date (to put it mildly), does not mean that it requires an independent, non-welfarist theory of fair distribution as Dworkin suggests. Instead, this assumption of equal utility from resources *is* welfarist because it is justified purely on welfarist grounds. It is simply the welfare-maximizing approach to operationalizing welfare egalitarianism under conditions of uncertainty.[86]

This point was first made briefly by Lerner (1944)[87] and has been rarely noted in the optimal tax literature since then.[88] The problem is not just that it is difficult to measure the relative utility of two people who otherwise appear to be identical. (If that were the case, this issue would fall outside this essay's scope, given its focus on ideals, not administrability and political constraints.) Instead, the problem is that the relative utility of two different people is fundamentally

---

2015); Weisbach, *supra* note 44; Dominique Henriet & Jean-Charles Rochet, *Is Public Health Insurance an Appropriate Instrument for Redistribution?*, 83/84 ANNALES D'ÉCONOMIE ET DE STATISTIQUE 61 (2006) (extending the Mirrlees model to include health and health insurance).

[80] Some models do not assume, however, that people have identical utility functions for different goods. For example, they might differ in how they would spend their potential income on market consumption versus leisure. *See, e.g.,* Lockwood & Weinzierl, *supra* note 41.

[81] Dworkin (1981a)*, supra* note 4, at 226.

[82] Dworkin (1981a)*, supra* note 4, at 224–26.

[83] Dworkin (1981a)*, supra* note 4, at 226.

[84] Piketty & Saez, *supra* note 24, at 405.

[85] Indeed, Kaplow asserts we eventually can and should be able to make interpersonal utility comparisons. KAPLOW (2008), *supra* note 7, at 375.

[86] The assumption could also, of course, be justified on non-welfarist grounds, as evidenced by its similarities to resource egalitarianism. But doing so would fail to respond to Dworkin's critique of welfarism on its own as a theory of distributive justice.

[87] ABBA P. LERNER, THE ECONOMICS OF CONTROL 28–32 (1944).

[88] *But see* Lily L. Batchelder, Fred T. Goldberg, Jr. & Peter R. Orszag, *Efficiency and Tax Incentives: The Case for Refundable Tax Credits*, 59 Stan. L. Rev. 23, 44–46 (2006) (hereinafter, Batchelder et al. (2006)) (making this point in the context of uncorrected externalities); Daniel Shaviro, *Beyond the Pro-Consumption Tax Consensus*, 60 STAN. L. REV. 745, 765 (2007).

immeasurable, no matter how much resources were dedicated to such a project. We could run tests on how often or intensely each person was sad or anxious, but one might prefer a frequently sad but meaningful and creative life, to a more joyful one that felt less meaningful. We could ask each person if she would prefer another's life, but it would still be unclear who has higher utility because each cannot know how it feels to live the other's life until she has done so. To be clear, this dilemma does not negate the welfarist case for redistribution in general—we can observe over the course of people's *own* lives that they are more satisfied by and prefer a life with more resources, at least to a point. But we still have no way to compare utility across *different* people.[89]

This fundamental uncertainty has important implications for a welfarist. Whenever there is complete uncertainty about how utility functions differ in form or intensity across different people, the welfare-maximizing response is to treat all people as having the same utility function.[90] This is the case because the expected utility loss from errors in assumptions about the shape and intensity of individuals' utility functions rises more than proportionately with the size of the error.[91] Assuming identical utility functions reduces the variance of errors, and the likelihood of especially costly mistakes. It therefore maximizes expected utility to assume that all individuals have identical utility functions for objective resources, such as potential income, absent evidence to the contrary. This assumption also maximizes welfare if the social welfare function is egalitarian, not utilitarian, as posited here.[92]

Another way to understand this point is to consider the circumstances in which each theory *would* redistribute to those with higher marginal utility from resources. Optimal tax theory would when there are observable and at least partially verifiable intergroup differences—i.e., when there is not fundamental uncertainty. To illustrate, if there is evidence that people in certain age groups or with certain health conditions get more utility from money (controlling for income and other resources), optimal tax theory would redistribute to people in those states. The amount of redistribution would depend, among other things, on the degree to which the state is exogenous and verifiable. Optimal tax theory would not redistribute to a person who simply says she gets more utility from money than others with no evidence. But it could entail extensive redistribution to a person with a medically-verifiable disease that shortens or reduces the quality of one's life.[93]

Likewise, resource egalitarians would insure against states like poor health that, under some views, render one a so-called utility monster. They view such states as potentially a type of negative resource. While the level of insurance is indeterminate, Dworkin argues that behind the thin veil of ignorance people would choose to insure against such states to the extent they are

---

[89] *See* Daniel M. Hausman, *The Impossibility of Interpersonal Utility Comparisons*, 104 MIND 473, 475 (1995) ("Most economists believe that interpersonal comparisons of utility are untestable, evaluative or even meaningless").

[90] *See* LERNER, *supra* note 87, at 28–32; Batchelder et al. (2006), *supra* note 88, at 44–46; Shaviro, *supra* note 88, at 765. This assumes that people generally experience declining marginal utility.

[91] If utility functions are linear and downward sloping, the expected utility loss rises with the square of the error.

[92] *See* AMARTYA SEN, ON ECONOMIC INEQUALITY 83–87 (1973) (showing that Lerner's conclusion holds when the social welfare function is concave or maxi-min).

[93] [DALY and QALY literature.] *See also* David Kamin, *When Does Money Matter? How Health Status Affects the Marginal Utility of Income* (SSRN Working Paper No. 1289602, Oct., 2008) (finding people dissatisfied with their health have higher marginal utility of income, as measured by time spent stressed, sad, or in pain, with the effect largely isolated to the bottom income quartile).

exogenous; that is, to the extent they stem from brute luck.[94] While he is less clear on this point, presumably verifiability would also matter behind the thin veil of ignorance.[95]

Thus, the assumption of identical utility from resources (absent evidence to the contrary) means optimal tax theory agrees with several of the basic principles of resource egalitarians but uses different terminology, as summarized in Table 1.[96] Neither theory supports more resources for utility monsters or less for ascetics just on their say-so. Both imply that in the hypothetical where immigrants arrive on a desert island, distributive justice requires dividing its transferrable resources equally.[97] Both further imply redistribution to adjust for differences in objective, exogenous, and non-transferrable resources—such as earning potential, disability, or social advantage—but only to a point. When redistribution becomes so extensive that many or most in society are worse off because of the associated efficiency costs, it should stop. Further, both theories give no definitive answer to where this point lies, even if we knew all the relevant behavioral elasticities. In optimal tax theory, it depends on how egalitarian the social welfare function is; in resource egalitarianism, it depends on how egalitarian the preferences are of the average person behind a thin veil of ignorance. In short, the two theories agree the ideal system entails substantial (but not full) redistribution of exogenous and observable resources.

Table 1: Analogues in Terminology of Resource Egalitarians and Optimal Tax Theory

	Resource Egalitarianism	**Optimal Tax Theory**
Thing to Be Equalized (Equalisandum)	Resources	Endowment as the welfare-maximizing proxy for utility
Extent of Redistribution Determined by	Average preferences behind thin veil of ignorance	Social welfare function
Sources of Inequality to Be Eliminated	Initial transferrable resources on a desert island	Exogenous, fixed resources
Sources of Inequality to Be Mitigated	Brute luck	Exogenous and observable conditions or states
Sources of Inequality Not Be Mitigated	Option luck	Deliberate, fully informed choices
	Utility monsters, ascetics, and expensive tastes	Unverifiable differences in marginal utility

[94] *See, e.g.,* Dworkin (1981b), *supra* note 2 at 292-94; Rakowski, *supra* note 75, at __.
[95] *See, e.g.,* Dworkin (2000), *supra* note 59, 334-35 (discussing moral hazard).
[96] Indeed, without this assumption, there are few circumstances in which optimal tax theory would have the same implications as equality of resources. For a formal explanation of why it is often impossible to equalize resources and welfare at the same time if people have different utility functions, see Marc Fleurbaey, *On Fair Compensation*, 36 THEORY & DECISION, 277, 285–87 (1994).
[97] This holds under optimal tax theory regardless of whether the social welfare function is egalitarian or utilitarian, assuming there is declining marginal utility for resources. Such resources should be fully equalized because the amount of initial resources on the island does not vary based on how it is divided; that is, there are no efficiency costs associated with redistribution. *See, e.g.*, FRANCIS Y. EDGEWORTH, THE PURE THEORY OF TAXATION (1897) (establishing that utilitarianism with fixed earnings and concave utility implies full redistribution of income if there are no behavioral responses); Piketty & Saez, *supra* note 24, at 404.

B. Altruism

A second objection often levied against optimal tax theory and welfare egalitarianism is that they deal inadequately with other-regarding preferences. According to this view, welfarism means people cannot choose to contribute to society outside of markets, disregarding all benefits provided to others unless they are delivered through a market transaction.

To make this concern concrete, suppose an investment banker wants to leave her job indefinitely in order to care for her young children full-time. A resource egalitarian might claim that optimal tax theory implies taxing her on her endowment or earning potential, and this is unfair because she will still be contributing to society in other ways.[98] She is not sitting around binging on her favorite TV shows while still benefitting from all the protections society provides. She is working hard caring for her children. And she should get credit for this hard work by owing less in taxes, regardless of whether her work is monetized or not.[99] (A resource egalitarian might further argue that such an endowment tax would be fundamentally illiberal because, given her high earning potential, she would owe so much tax that it would deny her the ability to achieve her vision of the good life.[100] But I will defer discussing this second objection, often referred to as "talent slavery," to the next section.)

While compelling, the complaint that welfare egalitarianism ignores non-monetary contributions is based on the mistaken assumption that optimal tax theory is synonymous with a tax on potential earnings.[101] As discussed, the optimal tax literature defines endowment more broadly to include things like inherited income and health. Even more relevant to the discussion here, this complaint ignores that correcting for externalities is a fundamental component to optimal tax theory.

Dworkin seems to assume externalities have no part in welfarism. He argues "under equality of welfare, people are meant to decide what sorts of lives they want independently of information… [on] how much their choices will reduce or enhance the ability of others to have what they want."[102] Sugin also assumes externalities are not part of optimal tax theory, though she explicitly defines optimal tax theory as the Mirrlees model and its extensions, which do not include them.[103] More generally, Dworkin contends that efficiency considerations have no place in discussions of the equalisandum, which is about what equality is, not how it should be compromised for efficiency's sake.[104]

But this is a false distinction. Welfarists cannot think of equality separately from efficiency because they are interrelated. For example, if a policy increases efficiency while holding the

[98] *See, e.g.*, Sugin, *supra* note 14, at 244-48.
[99] *Id*. *See also* MURPHY AND NAGEL, *supra* note 14, at 123–24.
[100] *Id*.
[101] As noted earlier, I use the terms (potential) earnings, consumption, and income interchangeably except when explicitly discussing the debate about which is the optimal tax base.
[102] Dworkin (1981a), *supra* note 4, at 288.
[103] Sugin, *supra* note 14, at 230.
[104] Dworkin (1981a), *supra* note 4, at 235. Dworkin's work is a bit contradictory on this point because, in other work, he endorses correcting for externalities as part of a theory of distributive justice, even though they are an efficiency issue. *See, e.g.*, Dworkin (2002), *supra* note 59, at 114 (supporting mandatory health insurance to correct for negative externalities arising from individuals' failure to purchase it).

distribution constant, one needs to determine how to distribute the efficiency gains, which is in part an equity question. Likewise, resource egalitarians' own discussions of the equalisandum cannot be divorced from efficiency. Dworkin acknowledges that efficiency considerations would inform the level of insurance the average person would select behind his thin veil of ignorance.[105]

Instead of simply taxing our investment banker on her potential earnings, optimal tax theory would view her situation as a case of positive externalities. Her decision to stay at home with her children potentially produces an "altruistic externality."[106] She must receive at least as much utility from being a stay-at-home mother as she did from earning a great deal as an investment banker, or she would not have made this choice. But her children may benefit as well. They grow up with a parent who is constantly available, and who is likely passing on her high skills to her children by reading to them and playing with them. When she makes the decision to leave investment banking, she takes into account the utility she derives from being a stay-at-home mother (which may be a function of the utility her children gain). But she does not take into account the completely separate utility that her children derive from this decision.

To give an example, suppose her children gain 100 utils from her being a stay-at-home mother relative to her being an investment banker, and earning the accompanying high salary.[107] She values her children's utility as if it were her own, but also likes her job. The decision to quit then generates 100 new utils for her as a result of the benefits to her children. She will only quit if her utility from being an investment banker and the accompanying salary is less than 100 utils. But, in reality, her decision generates at least 200 new utils—100 for her and 100 for her children. The 100 utils that her children gain are the altruistic externality that she does not adequately take into account.

As discussed in Section II.A.3, optimal tax theory's response would be to provide her with a Pigouvian subsidy that equals the marginal benefits her choice generates for others.[108] As Kaplow argues, if the tax system was optimal previously, the subsidy would be financed through a tax that mimics the general incidence of the subsidy.[109] That is, investment bankers and others with similarly high income potential who do not decide to become stay-at-home parents (or parents at all) would effectively share the cost of the subsidy with her.[110] (I should emphasize that by using this example, I do not intend to weigh into the debate about whether child care provided by third parties or stay-at-home parents has more positive effects on a child's well-being. Optimal tax theory implies Pigouvian subsidies for both, and even larger subsidies for whichever has greater positive effects.)

But a non-welfarist might have a further objection. Should she really receive a subsidy equal to how much she or other investment bankers value quality care for their children? Won't her

[105] *See supra* note 65 & accompanying text.
[106] [Kaplow]
[107] While utility is, in the most technical sense, purely an ordinal measure of preference, I will often operate as if it is a cardinal measure for ease of exposition. This should not change the discussion's conclusions.
[108] There is also an argument for subsidizing paid or self-provided child care under optimal commodity tax theory because child care is inversely correlated with leisure. *See* Piketty & Saez, *supra* note 24, at 448, 450. This relates to literature on tags discussed in Section II.A.
[109] [Kaplow]
[110] Technically the cost should also be shared somehow with the children who benefit from the subsidy because their parent decides to become a stay-at-home parent in response.

children be relatively well-off regardless? And, if so, shouldn't we focus subsidies on ensuring quality care for children who are less well-off, given the efficiency costs associated with taxation?

This is not just a potential objection to optimal tax theory, but to resource egalitarianism as well. Murphy argues that resource egalitarianism could be summarized as a distribution in accordance with benefits given to others (as measured by those others through market transactions) after an initially equal distribution of resources.[111] But he doubts that market returns, even in a world of initially equal resources and perfectly efficient markets, will ever track merit, or the social good.[112]

While rarely noted in the literature, optimal tax theory has a compelling response to this concern. As explained by Micheletto, Batchelder, and Kopczuk, Pigouvian subsidies should be welfare-weighted.[113] The welfare-weighted value is, roughly speaking, the dollar-denominated amount the social welfare function would give the external beneficiaries whose interests are not fully taken into account by the decision maker.

Moreover, when the welfare-weighted externality is not simply a function of average consumption of the good, but also of who consumes it, the subsidy should vary with traits of the decision maker that correlate with the welfare-weighted costs to others.[114] To provide example of negative externalities, if one unit of water pollution in Mississippi generates larger welfare-weighted negative externalities than one unit in Maryland because Mississippi residents are, on average, less well-off, the optimal Pigouvian tax should be larger on polluters in Mississippi.[115] On the other hand, if the welfare-weighted externalities associated with one unit of carbon emissions do not vary with the identity of the emitter (a so-called an atmospheric externality), the optimal Pigouvian tax should be linear, or the same for all polluters.[116] It should, however, be larger or smaller depending on the welfare-weight of the average costs to others.[117]

Returning to our investment banker, this implies she should receive a subsidy equal to, roughly speaking, the dollar-denominated amount the social welfare function would give her children to compensate them for not having a stay-at-home parent if she had stayed in her old job. In other words, it would be larger for children in low-income households and smaller for those in wealthy

---

[111] Liam Murphy, *Why Does Inequality Matter? Reflections on the Political Morality of Piketty's* Capital in the Twenty-First Century, 68 TAX L. REV. 613, 623–24 (2015) (hereinafter Murphy (2015)).

[112] Id. at 626.

[113] Luca Micheletto, *Redistribution and Optimal Mixed Taxation in the Presence of Consumption Externalities*, 92 J. PUB. ECON. 2262, 2266 (2008); Batchelder (2009), *supra* note 8, at 30-32; Kopczuk, *supra* note 11, at __. *See also* Emmanuel Farhi & Ivan Werning, *Progressive Estate Taxation*, 125 Q. J. ECON. 635, 646 (2010). These welfare weights are often referred to social marginal welfare weights or Pareto weights. *See* Piketty & Saez, *supra* note 24, at 459.

[114] Externalities that are solely a function of total consumption of the good consumed are referred to as "atmospheric," following James E. Meade, *External Economies and Diseconomies in a Competitive Situation*, 62 ECON. J. 54 (1952).

[115] *See* A. Lans Bovenberg & Lawrence H. Goulder, *Environmental Taxation and Regulation,* in HANDBOOK OF PUBLIC ECONOMICS 1535 (Alan Auerbach & Martina Feldstein, eds., 2002); [Boadway & Marchand 1995].

[116] This is not the case even with carbon emissions, which generate substantial negative local externalities in the form air pollution, in addition to contributing to global climate change. Dirk Heine et al., *Environmental Tax Reform: Principles from Theory and Practice to Date* 7 (IMF Working Paper No. 12/180, 2012).

[117] For example, global climate change is projected to disproportionately affect those who are disadvantaged, so the optimal carbon tax should be larger than if it affected all people equally.

households under an egalitarian (or even utilitarian) social welfare function because the former are more disadvantaged. The net effect would be to give our investment banker a subsidy for deciding to become a stay-at-home parent (or purchasing other quality child care), but only equal to the dollar value the social welfare function places on her doing so.

This analysis yields an important principle: at least under the interpretation I advocate, optimal tax theory *would* treat our investment banker as contributing to society—and give her "credit" for her choice—but only to the extent that she is doing something that a just government would have done otherwise.[118]

The same principle would apply to a lawyer who decides to accept a lower salary to become a human rights advocate, or a doctor who decides to become an internist in a public hospital. Each could reduce their tax liability by the amount they have saved a just government. The loss would be recouped from individuals in the same endowment class who chose occupations that did not generate as much welfare-weighted benefits for others.

This assumption that Pigouvian taxes and subsidies should be welfare-weighted should be adopted more explicitly and consistently in the optimal tax literature. Addressing Murphy's concern, optimal tax theory then would not hold that, after an initially equal distribution of resources, a just distribution is one in accordance with benefits given to others—as measured by those others through market transactions. Rather, it would hold that the distribution should be in accordance with benefits given to others *as measured by the social welfare function*. Perhaps a wealthy woman values of a fine piece of silk at $100, and a low-income mother living in a malarial region values a bed net for her child at 1 cent because that is all she can afford. Under optimal tax theory, the welfare-weighted externality associated with a gift of the bed net would be far higher, regardless of its market value.[119]

This assumption that externalities should be corrected through welfare-weighted taxes or subsidies has a further pay-off: It partially addresses concerns about how welfare egalitarians treat impersonal or political preferences.[120] Dworkin argues that if welfarists take other-regarding preferences into account when equalizing welfare, they will end up penalizing the benevolent, who will get less because they gain utility from others being well-off.[121] He further contends welfarists support subsidizing some who are disadvantaged more than others who are similarly situated, based only on which group generates more sympathy.[122]

The optimal tax theorist's response is effectively that "real" justice is what matters in the social welfare function, not our personal views about what is fair. You only get credit for furthering your

[118] This responds to Sugin's concern that the Mirrlees models and its extensions (which disregard externalities) unfairly measure social contributions only based on their market value. Sugin, *supra* note 14, at 247-49.

[119] It is unclear what principles resource egalitarians would reach in similar situations. Murphy questions whether Dworkin would give "credit" for non-market-based transactions, while Sugin clearly would. Neither address whether this "credit" would be weighted by social value, rather than the hypothetical market value.

[120] Dworkin (1981a)*, supra* note 4, at 197-204.

[121] Dworkin (1981a)*, supra* note 4, at 199.

[122] Dworkin (1981a)*, supra* note 4, at 200.

political beliefs or altruistic values if they do, in fact, advance distributive justice as defined by the social welfare function.[123]

C. Anti-Social Preferences

But perhaps our case of the stay-at-home mother is too easy. Often non-welfarists object to welfarism, not because of how it treats altruists, but because of what it implies about preferences that harm others. A broad array of non-welfarist theories maintain that such preferences should be disregarded.[124] For example, Dworkin argues doing so is a necessary predicate to his hypothetical auction establishing equality of resources.[125] The auction, he contends, would not treat people with equal concern if it failed to protect people from systemic prejudice.[126] In contrast, welfarists are often accused of condoning slavery or genocide if the benefits for the many exceed the costs to the few.[127]

These horrifying scenarios would not be supported by an egalitarian social welfare function because the utility loss would be so vast, and potentially infinite, for the victims themselves. Instead, they are really objections to utilitarianism and its indifference towards inequalities between individuals. But one still might be concerned that accounting for preferences to harm others within egalitarian social welfare functions would generate troubling and unsupportable results.

To understand how optimal tax theory would approach this issue, imagine there is an individual who loves screaming at people she doesn't know once in a while. Indeed, she would prefer to do this over many other activities. This preference is purely a choice that she has no desire to change. She was not born with an innate need to scream, which might be considered a disability (discussed in the Section III.A). Instead she cultivated this preference and enjoys seeing just how much she can upset and startle people. Suppose further that some people are only willing to be screamed at for $100 or more on a welfare-weighted basis. But others (maybe New Yorkers) don't really mind as much and are willing to be screamed at for $10 or less. Prior to any corrective tax, her victims were generally compensated by the government for the resulting discomfort in according with the social welfare function.

---

[123] This only partially address non–welfarists' objections about the potential perverse effects of taking other–regarding preferences into account. Here I am discussing the "credit" one gets for being benevolent because of the positive effects on others. But the social welfare function may also incorporate the direct effect on one's own utility of acting on such preferences, which could imply larger subsidies for disadvantaged groups that generate more sympathy. Addressing this concern requires an ad hoc adjustment to the social welfare function. But, as discussed below, this is also true for resource egalitarianism.

[124] *See, e.g.*, Dworkin (1981a), *supra* note 4; Harsanyi (1977, 1988), Nozick (1974), cited in (Kaplow 362)]

[125] Dworkin (2000), *supra* note 59, at 143–47.

[126] Dworkin (2000), *supra* note 59, at 161.

[127] *See, e.g.*, Seth Korman, *The Welfarist Approach to Human Rights Treaties: A Critique*, 58 UCLA L. REV. DISCOURSE 95, 104 (2010) (discussing how welfarist principles are potentially problematic in a "diverse society containing weakly protected minority groups" who can either be neglected or harmed through welfare maximization); Debra Satz, *Voluntary Slavery and the Limits of the Market*, 3 L. & ETHICS OF HUMAN RIGHTS 87 (discussing how welfarist philosophy endorses slavery in certain circumstances, especially situations of "voluntary" slavery or indentured servitude).

In this situation, optimal tax theory would apply a Pigouvian tax. If our screamer can control herself so she only screams at New Yorkers, it would charge her $10 each time she screams at someone, assuming that is the dollar-denominated welfare cost to her victims. Any revenue raised would be used to reduce taxes on those at the same endowment level. As a result, every endowment level would be better off. She would probably reduce her screaming in response to the tax. The government could then redistribute less to screaming victims, and could reduce taxes in general. And our screamer would still get to scream, as long as she values doing so more than the welfare-weighted costs it imposes on others. The net result is similar to a tax on pollution that results in less pollution, and therefore less government spending on environmental clean-ups, as well as lower taxes overall.

Non-welfarists probably would object to this solution. This is even more likely if the screamer's preference was, say, to punch people or, even worse, to punch people from a certain demographic group. Many views of distributive justice that are generally consequentialist incorporate deontological values when it comes to physical harm or systemic prejudice.[128] But it is instructive to consider why non-welfarists might find this approach troubling and how optimal tax theorists would respond.

One potential objection is that nobody would really consent to be punched. An optimal tax theorist might reply that some people certainly would if the price was high enough. After all, isn't this what professional boxing is? Next a non-welfarist might argue that people could be forced into accepting the payment for being punched because they have insufficient resources. But this is unlikely under an egalitarian social welfare function, which implies relatively extensive redistribution. A further worry is that people would reluctantly consent, even if they had equal resources initially, because they depleted them in the past when their preferences were different. This concern is addressed in Section IV.D. Finally, one might object that it hurts third parties to know that such punching is occurring. But optimal tax theory would apply a Pigouvian tax to correct for these harms to others as well.

A more challenging objection is that there is an important moral distinction between an individual expressly consenting in advance to be paid in exchange for being harmed, and simply receiving compensation after the fact without consenting. An optimal tax theorist could respond with a version of rule utilitarianism.[129] They could argue that preferences to harm others should be disregarded in general because, even though this might reduce welfare in specific cases, it will increase welfare overall by creating a positive norm and potentially changing such preferences over time.[130] But this response isn't very satisfying. Among other reasons, it is contingent on empirical realities, like how widespread such preferences are and how large the costs are that they impose on others.

A potentially more compelling response to claim about the importance of advance consent is that this hypothetical mixes together different fields of justice.[131] Optimal tax theory is only about distributive and corrective justice in the context of economic institutions. It is concerned with how the government should allocate material goods in society, and with issues akin to the civil justice system—what financial penalties should apply to those who injure others and how victims should

[128] *C.f.* RAWLS, *supra* note 69.
[129] [Mills, Weinzierl 2017]
[130] *See, e.g.,* KAPLOW (2008), *supra* note 7, at 352–53, 365.
[131] For a similar argument, see Rakowski, *supra* note 75, at __.

be compensated materially. It is not about what non-material sanctions should apply to people who scream at or assault others, or how people should behave ethically in their personal lives. To be sure, the criminal justice system should establish limits on behaviors that harm others, and appropriately punish those who violate these rules. But optimal tax theory has nothing to say about the optimal set of criminal laws and how to enforce them.

But what if our puncher's interest was not in punching anyone, but only members of a specific minority group? And what if she is not one person but many? As discussed, optimal tax theory would redistribute to the minority group to compensate for these harms. And it would apply a Pigouvian tax to the punchers, which should reduce their violence to some extent. But we have established that optimal tax theory does not require full *ex post* equality for those with lower endowments, which in this case would include being a member of the minority group. Thus, the preferences of the many in the majority could mean the minority is systematically worse off. This would also be true if the issue was not punching but, say, refusing to hire, or patronize establishments owned by, members of the minority group.

If one believes (as I do) that inequalities arising from systemic prejudice are more normatively troubling than inequalities arising from others types of brute luck, there is a strong argument for optimal tax theory disregarding such preferences. Indeed, some optimal tax theorists support this approach.[132] In doing so, optimal tax theory would deviate, not just from its single objective of maximizing social welfare, and from welfarism in general.[133]

But this potential flaw in egalitarian optimal tax theory does not imply resource egalitarianism offers a better approach. Instead, resource egalitarians deviate from their single objective of equalizing resources in this situation too. They generally contend the hypothetical insurance market establishing the ideal fiscal system should reflect people's tastes and preferences—including those that advantage some over others, like tastes for certain types of artistic talent. But they stipulate that when those tastes and preferences reflect systemic prejudice they should be disregarded.[134]

Thus, both theories grapple somewhat awkwardly with how to respond to discriminatory preferences. And both require *ad hoc* adjustments to their underlying mechanism for determining distributive justice in order to respond to concerns that systemic prejudice imposes uniquely heightened harms.

IV. How Should Distributive Justice Treat Inequalities Arising from One's Own Choices?

A key innovation of resource egalitarianism was its ability to couple concern for equality with a commitment to responsibility.[135] In Dworkin's view, distributive justice requires a community to

---

[132] E.g., Sen, *supra* note 25, at 201, 210 (suggesting that desires like exploitation or sexual or racial discrimination could be valued less heavily); John Harsanyi, *Rule Utilitarianism and Decision Theory*, *in* DECISION THEORY AND SOCIAL ETHICS (H. Gottinger & W. Leinfellner, eds., 1978) (suggesting that utility from sadism, resentment, and malic should be excluded from a welfarist calculus); LOUIS KAPLOW & STEVEN SHAVELL, FAIRNESS VERSUS WELFARE 427 (2002).

[133] *See* Sen, *supra* note 25, at 210 (noting that a theory de-weighting such preferences would no longer be welfarist but still could be a utility-supported morality).

[134] Dworkin (2000), *supra* note 59, at 161.

[135] G.A. Cohen, *On the Currency of Egalitarian Justice*, 99:4 ETHICS 906, 933 (July, 1989).

treat its members with equal concern.[136] But it also requires recognizing that each person is ultimately responsible for the consequences of the choices she makes that are under her control.[137] His focus on insuring people against brute luck but not option luck is similar to a more general idea in the philosophical literature that people should be compensated for circumstances over which they have no control, but held responsible for those they do.[138] Murphy refers to this as the "choices thesis."[139]

Some non-welfarists have criticized the choices thesis, arguing that holding people responsible for their choices requires a deeper justification.[140] If an individual is destitute and this is in part—or even entirely—a product of her choices, it does not necessarily follow that she *deserves* to live in poverty. And perhaps more relevant for the discussion here, it does not follow that her community has no obligation to assist her whatsoever.

Optimal tax theory also wrestles with whether and when people should be held responsible for their choices. While it does follow the choices thesis to some degree, its motivations differ. In optimal tax theory, choices matter from an equity perspective because of the information they provide about the relative utility one derives from different options. Choices also matter because they affect efficiency. Any redistributive scheme that reduces inequalities stemming in whole or part from choices will generate moral hazard costs and distort choices towards those taxed less heavily. But even though choices matter in optimal tax theory, they do not *determine* whether and when redistribution is appropriate. That is also influenced by how well correlated choices and utility are and how large the efficiency costs are of mitigating inequalities stemming from choices, among other factors.

This Section explores how optimal tax theory treats inequalities that arguably arise from one's choices, and how it defines an action for which one should be held responsible. In the process, it argues that optimal tax theory has identified—or, through logical extensions, could be used to identify—ideal policy design principles that partially or fully address some of the key objections to the choices thesis by non-welfarists.

A. Work Effort

The first potential choice that may give rise to inequality is how much and how hard to work. This is linked to the issue set aside in the prior section, often referred to as "talent slavery."

To tease out this concern, suppose individuals differ only in their potential earnings and level of ambition, but not in other ways such their health or the amount of financial assets they inherit. Suppose further that earning potential is purely innate and ambition is purely chosen. This is similar to the Lockwood and Weinzierl optimal tax model where all individuals with the same

[136] Dworkin (2000), *supra* note 59, at 5.
[137] Dworkin (2000), *supra* note 59, at 7.
[138] [Fleurbaey, Kolm.] As discussed in Section __, Dworkin differs from this literature by defining the morally important distinction as being about whether one wants one's tastes or preferences, not whether one has control over them.
[139] Murphy (1996), *supra* note 70, at 474.
[140] Murphy (1996), *supra* note 70, at 481-84.

earning potential have the same overall utility, but they differ in their preferences for market consumption versus leisure.[141]

In this situation, optimal tax theory would view potential earnings as the ideal touchstone for redistribution.[142] The rationale is that we have no basis for believing that individuals with the same earning potential have different levels of utility. Or, more precisely, and as discussed in Section III.A, it is welfare-maximizing to assume they do not absent evidence to the contrary. If one individual chooses to work for one more hour, that is because she gets more utility from the resulting income than from an hour of leisure. If another with the same potential earnings decides not to work an extra hour, she must receive at least as much utility—or imputed income—from her leisure as she would from her forgone earnings. Otherwise, she would not make that decision. Thus, potential earnings is the fairest proxy for relative utility. It is also the most efficient proxy because people cannot change their earning potential in the model and, as a result, there are no efficiency losses.

Resource egalitarians tend to approach this situation differently. One might initially think they would reach the same principle that redistribution should ideally be based on potential earnings. After all, equality of resources strives, to the extent possible, to make the distribution of resources at any moment ambition-sensitive but not endowment-sensitive.[143] It does so because it views individuals with high earning potential at birth as having no particular moral claim to their innate ability and the pre-tax income it happens to yield.[144] They have done nothing to deserve it. Likewise, individuals with low earning potential have done nothing to deserve their lot. But ambition is, by assumption, freely chosen, and resource egalitarians therefore believe people should be held responsible for its consequences. Put differently, they view the consequences of differential innate earning ability as brute luck, which should be mitigated, while the consequences of freely-chosen ambition are option luck and should not.[145]

One would think the logical conclusion for resource egalitarians would be that potential earnings is the ideal basis for redistribution in our hypothetical.[146] But they have typically declined to take this next step. Instead, they argue we should tax actual earnings, not potential earnings. For example, Dworkin contends that an income tax best advances distributive justice because it leaves intact the possibility of earning more if one is ambitious, but also recognizes the role of genetic luck in life.[147]

In part, Dworkin's objection is practical. He believes taxing endowment is unworkable because, contrary to the hypothetical, talents and ambitions are so intertwined that we cannot realistically

[141] Lockwood & Weinzierl, *supra* note 41.
[142] *See, e.g.*, DAVID F. BRADFORD AND U.S. TREASURY TAX POLICY STAFF, BLUEPRINTS FOR TAX REFORM 36 (2nd Ed., 1984).
[143] Dworkin (1981b), *supra* note 2, at 311.
[144] *See* MURPHY AND NAGEL, *supra* note 14, at 119.
[145] Dworkin (1981b), *supra* note 2, at 292–97.
[146] Others who have made this argument include Shaviro, *supra* note 8; Stark, *supra* note 8, at 5; Lawrence Zelenak, *Taxing Endowment*, 55 DUKE L. J. 1145, 1155 (2006).
[147] Dworkin (1981b), *supra* note 2, at 312–13.

separate out the effects of each.[148] But this paper is concerned with ideals, not administrative realities. Moreover, since Dworkin's initial work advancing resource egalitarianism, we have refined techniques for identifying the exogenous portion of potential income that, while still highly imperfect, are relatively sophisticated. We can estimate reasonably well an individual's future income based on their current income, parents' income, financial inheritances, race, gender, zip code, education, IQ, personality type, height, and so forth. This is precisely why optimal tax theorists advocate the use of tags to supplement actual earnings as the proxy measure of potential earnings. The administrative practicality objection to taxing earning potential has thus becoming somewhat less convincing over time.

Non-welfarists' stronger objection to endowment taxation is that it could or would generate so-called "talent slavery."[149] They argue that even if our investment banker receives subsidies for choosing to be a stay-at-home parent under a welfarist approach, these subsidies may be insufficient to cover her tax bill. She could or would owe so much tax because of her high earning potential that she will be forced to continue working as an investment banker, denying her the ability to achieve her vision of a good life.[150] At the extreme, even if she wants to work full-time as a stay-at-home parent, she will be condemned to a life of working flat out in the occupation that compensates her talents the most highly.[151]

This is powerful objection. But there are several strands to this argument and ultimately they are all flawed. They are either based on misunderstandings of what kind of endowment tax actually optimal tax theory actually implies, or simply reflect the uncomfortable reality of what egalitarianism implies for people who are lucky enough to have high potential earnings.

Instead, this sub-section argues that well-constructed endowment tax (i.e., o that includes aspects of endowment other than earning potential and is optimized for externalities and other issues described below) largely addresses the "talent slavery" objection of non-welfarists. Moreover, such a tax is not just the ideal under optimal tax theory. If one takes the concerns of resource egalitarians seriously, it is the best way to advance equality of resources as well.

To begin with the most straightforward misunderstandings, an *optimal endowment tax* (i.e., one implied by optimal tax theory) could be paid through non-monetary contributions to society, as explained in the prior section.[152]

---

[148] Dworkin (1981b), *supra* note 2, at 313–14 (dismissing as unworkable the ideas of taxation the portion of wealth attributable to different talents or fixing rates such that each person is left with the same potential after-tax income as if they had the same talents).

[149] *See, e.g.*, Dworkin (1981b), *supra* note 2, at 311–12, MURPHY AND NAGEL, *supra* note 14, at 122–23, Linda Sugin, *Let the Beachcomber Drown: Why Taxing Endowment is Unjust* (Fordham Legal Studies Research Paper No. 959710, Feb. 2008); JOHN RAWLS, JUSTICE AS FAIRNESS 158 (2001).

[150] MURPHY AND NAGEL, *supra* note 14, at 123–24.

[151] *See, e.g.*, Dworkin (1981b), *supra* note 2, at 320; MURPHY AND NAGEL, *supra* note 14, at 122.

[152] Going forward, I will refer to the endowment tax implied by optimal tax theory as an optimal endowment tax. This is distinct from the way "endowment tax" is sometimes used in the literature; for example, as a synonym for a tax on earning potential when optimal tax theory would clearly imply accounting for unearned income as well. It is also not strictly an endowment tax since it is adjusted for externalities and other concerns of optimal tax theorists that are discussed throughout.

It would also take into account the fact that work is an all-or-nothing choice in many occupations. It would be difficult, for example, for a CEO to work part-time. Optimal tax theorists acknowledge this possibility and have built it into some models.[153] The simple solution is to define endowment based on how much one could earn annually if one worked in a job that pays on an hourly basis rather than in a job that requires full-time work, or something along those lines.[154] Then our investment banker may have to continue working part-time in order to fulfill her obligations to her community, but she won't necessarily be forced to work full-time at a job she hates.

Turning to less straightforward concerns, non-welfarists may also object that an endowment tax is akin to forced labor even if it only requires one to work part-time.[155] It makes the state or others a part-owner of you. On some level, this is certainly true. Justice means we are all partly owned by something. As Nagel argues, if we believe that a person who doesn't care about anyone else has no obligations to others—that she can do absolutely whatever she wants—then there is no right or wrong.[156] But if we believe that there is right and wrong, we must believe that each person should do what someone or something else tells them, whether that is God, their conception of justice, or the democratically-elected state in which they live.[157]

But non-welfarists' concerns regarding "talent slavery" extend well beyond the state or society partially owning one's labor. While Nozick argues that any tax above that necessary to finance the minimal state is forced labor and unjust,[158] resource egalitarians and many other non-welfarists disagree. Instead, their concern is that those with high potential income will be forced into a single or narrow band of occupations. They worry that our investment banker who wants to be a stay-at-home parent (or ski instructor or poet) will be forced to work only as an investment banker or in the finance industry in general.[159] This, they argue, is illiberal and unjust. Our occupation is a fundamental part of our lives. If we cannot choose our occupation, we do not have an equal chance to pursue our vision of the good life.[160] And if we do not have an equal chance at that, we have defeated the whole point of equalizing resources in the first place.

But having to work at least part-time in a limited choice of occupations isn't servitude, as non-welfarists seem to imply with the freighted trope of "talent slavery". It is reality for many workers with low earning potential. They often have to take whatever job they can get, and work as many hours as they can get, to meet their basic needs and those of their family, let alone pay their tax liability.[161] Moreover, an endowment tax would enable the state to lower tax burdens on such

[153] *See, e.g.*, Saez, *supra* note __.
[154] *See* Stark, *supra* note 8, at 59, note 64.
[155] *See, e.g.*, NOZICK, *supra* note 71, at 172; Stark, *supra* note 8, at 57.
[156] THOMAS NAGEL, WHAT DOES IT ALL MEAN? 61–67 (1987). *See also* Stark, *supra* note 8, at 49.
[157] *Id*.
[158] NOZICK, *supra* note 71, at 150–53.
[159] MURPHY AND NAGEL, *supra* note 14, at 122-23.
[160] Sugin, *supra* note 14, at 235-37.
[161] Many critics of endowment taxes acknowledge that a tax on actual earnings constrains occupational choices too and may force one to work. *See, e.g.*, MURPHY AND NAGEL, *supra* note 14, at 123. Olson points out that, holding revenue constant, a tax on actual earnings can constrain the occupation and work choices for some even more than an endowment tax. Kristi Olson, *The Endowment Tax Puzzle*, 38(3) PHIL. & PUB. AFFAIRS, 240, 244, 250-54 (2010).

workers because the base of the tax would be broader.[162] In doing so, it would increase their liberty to choose their occupation and the amount of time they work.

To be sure, there are fewer jobs that are highly lucrative. In that sense, those with high earning potential might have relatively fewer job choices under an optimal endowment tax, at least if they only want to work part-time. But the number of jobs is not a particularly good proxy for occupational diversity or quality. To make a gross generalization, most low-paying occupations are in the service industry or involve manual labor.[163] On average, they are less creative and intellectually challenging, and more physically taxing. They are boring, tiresome, repetitive, and invisible.[164] By contrast, highly compensated occupations generally pay more because relatively few people can do them, which means they bring prestige and status. There are, of course, highly rewarding jobs that pay very little. For example, those that benefit people or causes we care about. But putting aside such altruistic jobs, which were discussed in Section III.B, it seems reasonable to assume that even if highly-compensated jobs paid the same as low-paying jobs, they would, on average, generate more utility and satisfaction.[165] As a result, it also seems reasonable to assume that most people would prefer to be condemned to working part-time in jobs that are highly-compensated, than to working the same hours in the occupations to which less well-compensated workers are restricted.

Suppose non-welfarists concede this point: highly-compensated jobs are not necessarily less fulfilling that low-paying jobs, even if the take-home pay is the same. Some nevertheless object that the fact that our investment banker, say, quits her job to become a ski instructor demonstrates that she is personally much less well-off than her fellow investment bankers. As a result, she should be taxed less than them.

Yet her quitting her job demonstrates nothing of the sort. All it reveals is that she has a different utility function for market consumption and teaching skiing than they do.[166] It doesn't tell us whether her overall utility is higher or lower than her former colleagues. For example, perhaps her colleagues stay because they don't derive much satisfaction from anything else, so they might as well continue in banking. Perhaps she leaves, not because she is dissatisfied with her job, but because she loves being a ski instructor even more.[167] In other words, the fact that she left her

[162] *See, e.g.*, Stark, *supra* note 8, at 59, note 59.

[163] *May 2017 National Occupational Employment and Wage Estimates United States*, U.S. BUREAU OF LAB. STAT., https://www.bls.gov/oes/current/oes_nat.htm#00-0000 (last visited July 12, 2018) (showing that lower paying jobs are dominated by the service industry or manual labor professions). *See also* Fatih Guvenen, Greg Kaplan, & Jae Song, *The Glass Ceiling and The Paper Floor: Gender Differences among Top Earners, 1981-2012*, (NBER Working Paper No. 20560, 2014) (discussing how finance and insurance currently dominate the highest paying jobs).

[164] *See, e.g.*, [Shipler].

[165] See, e.g., Saziye Gazioglu & Aysit Tansel, *Job Satisfaction in Britain: individual and Job Related Factors*, 38 APPLIED ECON. 1163, 1168 (2006) (finding a positive and statistically significant relationship between pay and job satisfaction measures). *But see* Greg Kaplan & Sam Schulhofer-Wohl, *The Changing (Dis-)Utility of Work*, __ J. ECON. PERSP. __ (forthcoming 2018) (showing that workers with a college educated are substantially more stressed, less happy, and find less intrinsic meaning in their work as compared with workers with at most a high school education).

[166] Shaviro makes a similar point in Shaviro, *supra* note 8, at 409.

[167] *See* Zelenak, *supra* note 146, at 1167.

job does not tell us that she has fewer resources or opportunities to pursue her vision of the good life, which is what resource egalitarians care about equalizing.

It is certainly possible that she left investment banking because it made her depressed or physically ill and could not do the work anymore. Then she might be less well-off and have fewer resources. But, as discussed in Section III.A, both optimal tax theorists and resource egalitarians would then treat her differently than her former banking colleagues, at least to the extent that her condition is exogenous and observable.

A related potential objection is that her quitting may reveal something about how she feels about working in, say, the finance industry in general.[168] Suppose she starts working as a ski instructor and asserts that she should owe the same amount of taxes as her fellow ski instructors who cannot get higher paying jobs. The fact that she gave up a lucrative job seems to suggest that working in finance simply makes her less happy (or more miserable) than it does for the average person. Therefore, she should be taxed like her fellow ski instructors because she finds finance work particularly distasteful, and this offsets any greater well-being she gets from her unusually high earning potential.

But, once again, we have no reason to believe this is the case. Because we can't compare utility across people and have no evidence apart from her choice (such as illness) that she has a stronger distaste for investment banking, we cannot know whether she is better or worse off than the average investment banker. We could *assume* she has the same level of well-being as her fellow ski instructors. But such an assumption requires that people with high earning potential, in general, are no better-off than those with low earning potential. This could be because money does not make one better-off. It could be because, putting money aside, low-earning-potential people as a group like the experience of working in finance more. Neither possibility seems particularly plausible.

The only instance where it does seem plausible that more money does not make one better-off is the extra money is a premium paid for working in a job that generates especially high disutility, perhaps working with sewage or in a physically dangerous environment.[169] But in that case, optimal tax theory implies that such wage premiums, to the extent they exist, should be disregarded.[170]

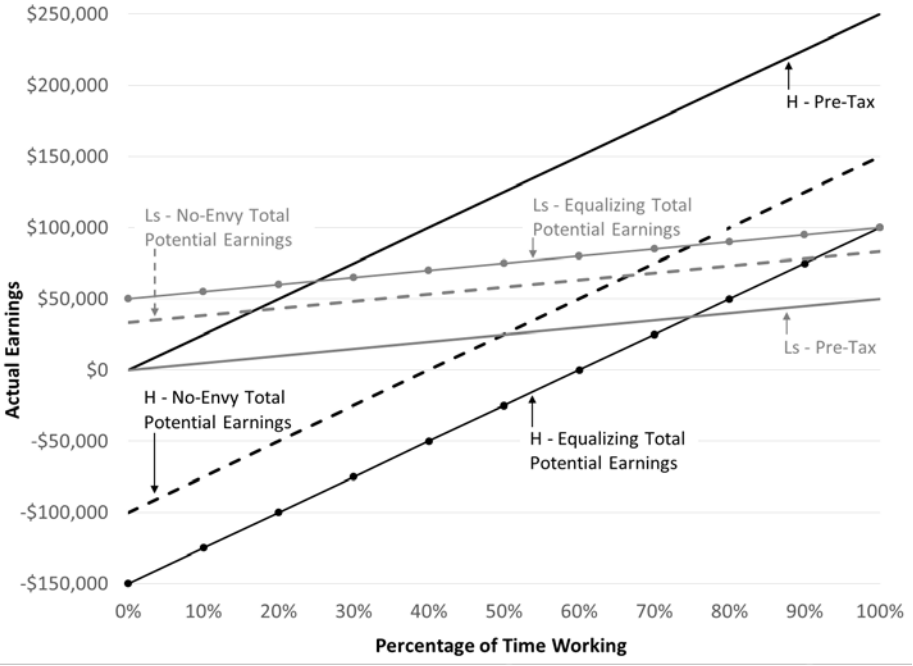
Thus far, I have argued that requiring those with high income potential to work part-time in relatively well-paid jobs in order to fulfill their responsibilities to society is far from the images that the trope of "talent slavery" evokes. Moreover, it is not at all clear that doing so would leave those with high potential income worse off than others.

If the reader agrees, the only remaining concern about an optimal endowment taxation entailing "talent slavery" would be if it requires those with high earning potential to work flat out all the time in a highly-compensated job they hate. But it would not. To see why, let's consider various ways optimal tax theory might structure an endowment tax on potential earnings.

---

[168] *See, e.g.*, Stark, *supra* note 8, at 54.
[169] *C.f.* Olson, *supra* note 161, at 261-68.
[170] *See, e.g.*, Zelenak, *supra* note 146, at 1165-69.

Figure 1: Potential Lump Sum Taxes on Potential Earnings



Suppose there are four people in the economy, illustrated in Figure 1. One person, H, can earn $250,000 annually pre-tax if she works full-time in a job that would pay her the same rate if she worked on an hourly basis. The other three people, the Ls, can only earn $50,000 annually before taxes and transfers if they work in full-time in an analogous job. Under an egalitarian social welfare function, people are generally equal if they do not envy each other's opportunity sets.[171] There are at least three ways an endowment tax could be structured to achieve this objective.

The first is lump sum tax, applying regardless of how much one works, that enables H and the Ls to have the same maximum income after tax. In this example, H would pay $150,000 in tax in order to fund a demogrant of $50,000 for each L. This would leave everyone with the same amount of potential consumption ($100,000) after taxes and the demogrant, as illustrated by the circle-marked lines in Figure 1. Let's call this approach *equalizing total potential earnings* (ETPE).

Many non-welfarists assume this is what optimal tax theorists mean by an endowment tax. But optimal tax theory would not actually consider H and the Ls to be equal under such a tax because it ignores the value of the other good in the economy: leisure.[172] Instead, the positions of H and

[171] Envy is not defined in the psychological sense of being less happy because someone else is better off. Instead, no-envy describes states when one would be equally satisfied with one's own lot and another's. Marc Fleurbaey, *On Fair Compensation*, 36 THEORY & DECISION, 277, 288 (1994).

It is possible that envy-freeness cannot be achieved or is indeterminate. For example, a successful doctor might wish she was a gifted poet, while her sister might find herself in the reverse situation. If both passionately want what they intrinsically cannot have, they might both envy each other's life no matter how much resources are transferred to them. If so, the equality goal in optimal tax theory is to get as close to envy-freeness as possible, or to select one of multiple potential envy-free allocations. For a compelling argument that envy-freeness is an insufficient standard for justice, *see* Kristi Olson, *Our Choices, Our Wage Gap?* 40 PHIL. TOPICS 45, 48–49 (2012).

[172] *See* Zelenak, *supra* note 146, at 1163.

the Ls would be reversed under this approach.[173] In all cases except one (working every waking minute), the Ls will be better off than the H because they can consume far more than H even if they work the same amount. This is the mirror image of their pre-tax situation, where H was better off than the Ls in all cases except when nobody works at all.

A second approach is to adopt a lump sum tax but set it at a lower level where the tax-demogrant combination that meets the no-envy criterion for H and the Ls, even if it involves different opportunity sets for consumption and leisure for them. Let's call this approach *no-envy total potential earnings* (NETPE). Suppose this tax was $100,000 on H, which would fund a demogrant of $33,333 for each L, as illustrated by the dashed lines in Figure 1. If H decides to work full-time in response and the Ls decide not to work, H would take home $150,000 and the Ls would take home $33,333 each. As long as neither prefers the other's bundle of potential market consumption and leisure, the policy would have achieved equality in the no-envy sense.[174]

This approach is the more common view of how an optimal endowment tax would be structured.[175] It also illustrates why egalitarian optimal tax theory is highly unlikely to require those with high earning potential to work all the time with no leisure.[176] Even if doing so would dramatically increase the income of the Ls, H's utility would then be so low that the policy would not maximize social welfare. By contrast, in this stylized example, H would have to work 40 percent of the time in order to pay the lump sum tax she owes under a no-envy total potential income—nowhere near all of her time.

But if non-welfarists are still concerned, there is a third option that has been insufficiently emphasized in the literature to date: a tax based on one's hourly potential earnings—not total potential earnings—multiplied by the hours one actually chooses to work. Under this approach, advanced by White, H would owe 60 percent of her hourly potential earnings for every hour she worked, as illustrated by the square-marked line in Figure 2.[177] This tax would fund an earnings subsidy (rather than a demogrant) for each L equal to 200% of her hourly potential earnings for

---

[173] *See* White, *supra* note 12, at 606-07.

[174] This does not necessarily mean the government has equalized utility between H and L because we do not know how to compare utility across individuals. This example also disregards disabilities and dynamic effects over time, which are discussed in Section IV.D.

One might also ask whether this point of no-envy equality is actually a single point or a range of tax-demogrant combinations. Given that individuals have the same utility functions for income potential under the standard optimal tax theory assumption, it would be a single point. All the Hs have the same income potential and therefore the same utility under the redistributive scheme, even if some choose to work part-time and some full-time in order to pay the tax and have some funds left for market consumption. All the Ls have the same utility as each other for the same reason. The point identified is the one where the Hs and Ls do not envy each other, even though individuals within each group may differ in what share of their potential income they choose to spend on market consumption and leisure.

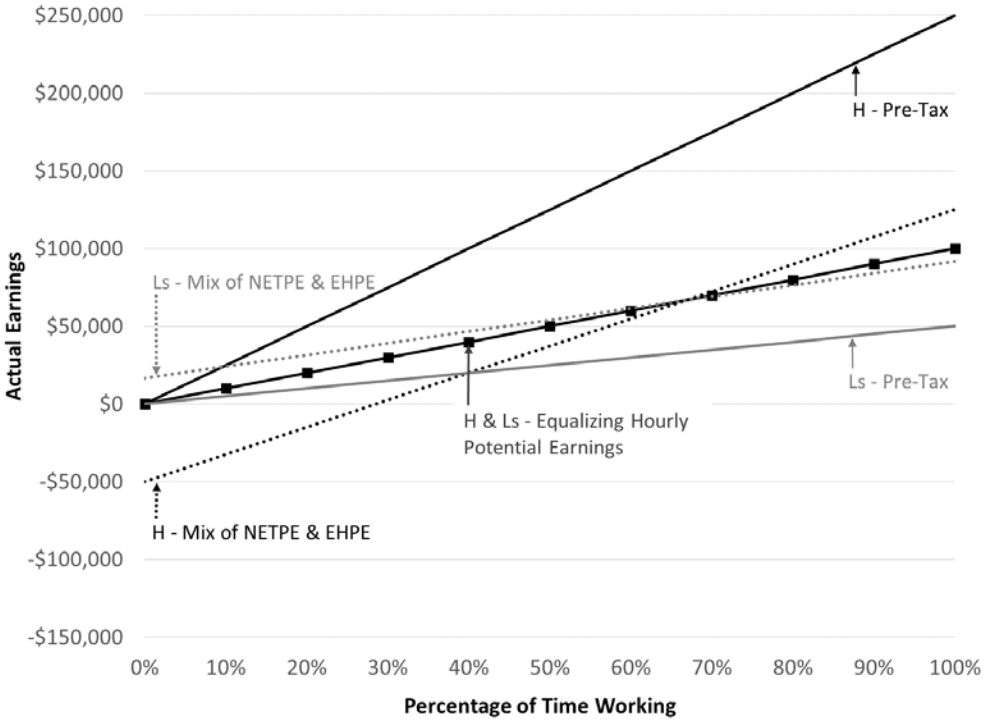
[175] *See, e.g.*, Richard Musgrave, Maximin, Uncertainty, and the Leisure Trade-Off, 88 Q.J. Econ. 625 (1974). *See also* Zelenak, *supra* note 146, at 1163-64.

[176] If the social welfare function is utilitarian, it is more possible (though still unlikely) that the optimal endowment tax would tax H so heavily that she would have to work every waking hour. Her total utility would be very low and her marginal utility for leisure would be extremely high. But if another hour of her work produced, say, $10 million, it is theoretically possible that her marginal utility for leisure would be lower than the marginal utility everyone else would gain from requiring her to work in order to fund a larger demogrant. But, as noted, this essay only considers egalitarian social welfare functions.

[177] White, *supra* note 12, at 602. White refers to this approach as an egalitarian earnings subsidy scheme.

each hour she works. As a result, H and the Ls would all be on the same line. Each would take home $50,000 if they decided to work half-time, or $100,000 if they decided to work full-time. Let's call this approach *equalizing hourly potential earnings* (EHPE).

Figure 2: Potential Hours-Based Taxes on Potential Earnings



This approach would be an even better measure of equality from a welfare egalitarian perspective than equalizing total potential earnings or its no-envy variant because everyone would have access to the same set of consumption and leisure bundles. While some might have to work in a job they dislike to achieve their preferred level of market consumption, this would be equally true for those with low and high endowments.[178] In addition, this approach would preclude any conception of "talent slavery" by not requiring those with relatively high endowments to work at all.

But *equalizing hourly potential earnings* (EHPE) would have a downside from an optimal tax perspective. Unlike *equalizing total potential earnings* (ETPE) or its no-envy variant, it is not perfectly efficient. As evidenced by the flatter slope of the *equalizing hourly potential earnings* line for H, it would reduce work incentives for those with high potential income, potentially shrinking the economic pie. As a result, optimal tax theory implies that some mix of the second and third approaches (EHPE and NETPE) is optimal. Figure 2 illustrates one such combination in the dotted lines. This mix would entail some efficiency costs by reducing work incentives for those with the highest potential income. But it would provide a more similar set of consumption-leisure bundles for H and the Ls than the second approach (NETPE), while requiring H to spend even less of her time working to meet her obligations to society – in this stylized example only 29 percent of her time.

[178] White, *supra* note 12, at 617.

In sum, optimal tax theory implies taxing those with high potential earnings *at most* at a level that would require them to work part-time in a well-paying occupation. Optimal tax models should make this clearer by explicitly and consistently acknowledging that the optimal endowment tax would be at least partially based on hours worked times one's potential hourly earnings.

Nevertheless, non-welfarists might still be concerned if optimal tax theory implies *any* work requirement for those with high earning potential. If imposing no work requirements on anyone is a bedrock principle, this would be a flaw with optimal tax theory. But it is worth emphasizing that resource egalitarianism, at least, is implicitly committed to this kind of requirement too.[179]

Dworkin contends that people would select an income tax, not a lump-sum endowment tax, behind his thin veil of ignorance. But what's to say people would not pick a tax that varies in part with hours worked and potential earnings, as implied by optimal tax theory? Dworkin proposes that those with a disability or who otherwise have low potential earnings should receive an insurance payment equal to what the average person would select if everyone had the same chance of being in such a position. The revenue for that payment has to come from somewhere.[180] Thus, resource egalitarianism also requires some or all members of society to work to finance its insurance system.

Indeed, Dworkin acknowledges that people might select a level of insurance that is so high that some members of society have to work flat out to pay for their premiums. He doesn't think people would actually decide to insure to this point because this would make the "talented" much worse off than others.[181] Markovitz makes a similar point in arguing that resource egalitarianism does not imply insuring up to the mean talent level because that implies "talent slavery."[182] But Dworkin does think people would decide to insure at a level that requires some people to work part-time, possibly in a profession they don't like.[183] This, he argues, is not a problem because such a worker "is not much differently enslaved by his talent if he insures than he would be enslaved by his lack of talent if he did not."[184] Yet that is, more or less, the view of optimal tax theorists too. Moreover, at other points he raises the possibility that the ideal tax for a resource egalitarian would be tied to potential income, and a tax on market income is just a second best approach.[185]

In short, resource egalitarians concede that requiring some people to work part-time in highly-compensated jobs may be demanded by their theory of justice and is not so-called "talent slavery"—or at least of not a normatively troubling variety. But this is the only kind of work demands on high-endowment individuals that egalitarian optimal tax theory plausibly would endorse.

---

[179] I am grateful to Mitchell Kane for pushing my thinking on this point.

[180] We have no reason to believe that people will inherit enough material resources that they can pay the tax necessary to fund the insurance payments solely with unearned income.

[181] Dworkin (1981b), *supra* note 2, at 320–22.

[182] Daniel Markovitz, *How Much Redistribution Should There Be?,* 112 YALE L. J. 2291, 2307–09 (2003).

[183] Dworkin (1981b), *supra* note 2, at 322.

[184] Dworkin (1981b), *supra* note 2, at 322–23.

[185] Dworkin (1981b), *supra* note 2, at 324–36; Dworkin (2000), *supra* note 59, at 100-102.

B. Chosen Risks

At this point, the reader might have a lingering concern. Surely potential earnings changes over time. It is not strictly innate but also the result of chosen and unchosen risks. If so, at what moment in time should potential earnings be measured for purposes of determining how well-off different people are?

Suppose our investment banker again decides to become a ski instructor. Initially she could earn $1 million annually if she continued to work full-time for her firm, but on an hourly basis. Over time, that figure falls to $250,000 because her firm and others assume her skills have atrophied. When she quits, she expects to be able to earn $75,000 working full-time as a ski instructor, though she knows this figure is subject to some risk. After several years as a ski instructor, multiple resorts shut down due to climate change. Because demand has dried up, she can only earn $50,000 as a ski instructor. What income should she be treated as having for tax and transfer purposes?

Non-welfarists criticize both optimal tax theory and resource egalitarianism for dealing inadequately with this question. It is not at all clear how a resource egalitarian would treat our investment banker. At least in theory, their goal would be to compensate her for losses from brute luck but not option luck. Because her career shift was a deliberate choice, should she be taxed as if her potential income is $1 million even though she can only earn $250,000 working as an investment banker now? Are the effects of climate change on the ski industry an instance of brute or option luck? If brute luck, should she be compensated for that and how? Or perhaps notwithstanding the option luck elements, she should just be taxed, following Dworkin, on her actual income of $50,000 as some sort of rough compromise?

Meanwhile, non-welfarists argue that because optimal tax theory ideally taxes endowment, it would often imply so-called "talent slavery" once risk is taken into account. Their argument is that *ex ante* expected potential earnings almost always differs from *ex pos*t potential earnings because of chosen and unchosen risks.[186] This is true when one deliberately places a bet in the financial markets, or when one bets on what education to obtain or what skill to cultivate. But even if one does not consciously take any risks, *ex post* potential earnings can differ, for example if market demand changes unexpectedly. Thus, if the optimal tax system is a function of *ex ante* potential earnings, one could end up owing more tax than one can pay even if working flat out.[187] Such unpayable tax burdens would be deeply troubling. And they would create inefficiently large incentives not to take risks. Moreover, they could arise under all of the endowment tax structures discussed in the prior sub-section, except one that is purely focused on equalizing hourly potential earnings.[188]

This challenge raised by non-welfarists really has two components. First, if we should hold individuals responsible for the consequences of chosen risks, does this imply so-called "talent slavery" in some cases? And second, should we actually hold individuals responsible for the

[186] *See, e.g.*, Sugin, *supra* note 14, at 253-54.

[187] Sam refers to this possibility as counterfactual talent slavery. Erick Sam, *From Each According to His Ability: An Analysis of Endowment Taxation and Potential Earnings* 56 (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3121798.

[188] Under an *equalizing hourly potential earnings* endowment tax, those with high *ex ante* potential earnings could avoid an unpayable tax burden by not working at all.

consequences of risks and, if so, when? If choice is the key, when should we consider an individual to have chosen a risk in the first place? I focus on the first question here, and the second in the following sub-sections.

On the first question, the tools of optimal tax theory offer a potential response. In particular, optimal tax theory generally aims to treat risk neutrally in order to minimize distortions to risk-taking. In light of this objective, optimal tax models can and should be explicitly structured in such a way that precludes unpayable tax burdens, even as a consequence of chosen risks. While not explicitly discussed in the literature to date, they can do so through what I will refer to as a *risk-neutral endowment tax*.

Under a risk-neutral endowment tax, our investment banker's tax *rate* would be based on her *ex ante* potential earnings ($1 million), except to the extent her current potential earnings differs because of intervening risks that she was not able to incorporate into her decisions. But this tax rate would apply to a tax *base* of her *ex post* potential earnings ($250,000). This means she would owe more tax than someone whose *ex ante* and *ex post* potential earnings were both $250,000.[189] But she would never owe more tax than she can conceivably pay.

This approach is admittedly novel.[190] It is implicit in some of the tagging literature, which occasionally posits adjusting the tax rate, and not the tax base, for the tag.[191] But none of this literature discusses whether a rate adjustment is the ideal, or simply more practical or mathematically tractable.

[189] Continuing with the example in the prior section, if the optimal tax was the no-envy total potential earnings tax (NETPE) and was $600,000 on someone with potential earnings of $1 million, this would be translated to a tax rate (60 percent) applied to her *ex post* potential earnings ($250,000), resulting in a lump sum tax of $150,000. People with *ex ante* and *ex post* potential earnings of $250,000 would only owe $100,000, like H. If the optimal tax was instead based on hours worked (EHPE) and was 70 percent of hourly potential income for those with total potential earnings of $1 million, she would pay tax at this rate, rather than 60 percent like H.

[190] Zelenak does raise the possibility of basing tax burdens entirely on ex post potential earnings. Zelenak, *supra* note 146, at 1155. While a risk-neutral endowment tax has been not explored in models of ideal endowment taxation specifically, the consequences to inherent uncertainty and risk have been incorporated into optimal tax models in many other respects. *See e.g.*, Jonathan Eaton & Harvey S. Rosen, *Optimal Redistributive Taxation and Uncertainty*, 95 Q.J. ECON. 357 (1980) (finding "that even if a relatively small fraction of the population experiences wage uncertainty, there may be a significant impact on optimal tax rates"); Helmuth Cremer & Firouz Gahvari, *Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy*, 105 ECON. J. 1995 1165, 1178 (finding that introducing uncertainty into an optimal tax model with both commodity and income taxation substantially changes optimal rates, and arguing for lower taxation on "pre-committed goods" such as housing); s*ee also*, Robin Boadway & Motohiro Sato, *Optimal Income Taxation with Uncertain Earnings: A Synthesis* (CESifo Working Paper No. 3654, 2011); Bas Jacobs, Dirk Schindler and Hongyan Yang, *Optimal Taxation of Risky Human Capital*, 114 SCAND. J. ECON. 908 (2012).

[191] *See, e.g.,* George A. Akerlof, *The Economics of 'Tagging' as Applied to Optimal Income Tax, Welfare Programs, and Manpower Planning*, 68 AM. ECON. REV. 8 (1978); Helmuth Cremer et al., *Tagging and Income Taxation: Theory and an Application*, 2 AM. ECON. J.: ECON. POL'Y 31 (2010). *But see* Logue & Selmrod, *supra* note 47 at 844 (arguing that accurate genetic information may be used as a tag, which would allow taxation not on income, but on other bases such as ability, endowment, potential income, or wage rates).

Nevertheless, a risk-neutral endowment tax, is justifiable purely on welfarist grounds as the most efficient way to account for deliberately chosen risks. To illustrate, suppose a group of individuals with high potential earnings faces a choice. Under the safe choice, they will definitely be able to earn $200,000. Under the risky choice, gaining skills in an unusual trade, there is a 50/50 chance their potential earnings will be $500,000 or $100,000. The social welfare function demands raising $100,000 on average from each person in this endowment group. To simplify the analysis, let's assume that everyone works full-time so it does not matter to what extent the endowment tax is based on hourly or total potential earnings.

As illustrated in Table 2, one option would be to charge each individual a tax of $100,000, regardless of which path they choose. This is a *fully ex ante endowment tax*.[192] But then all may choose the safe path because the risky path entails a 50 percent chance of having nothing left after tax even when working full-time. A second option would be to create insurance for the risky path so that one always ends up with $300,000 pre-tax and $200,000 after-tax. Let's call this an *insured endowment tax*. But then everyone would choose the risky path because they would be guaranteed higher after-tax income.

Table 2: After-Tax Earnings Under Different Ways of Addressing Calculated Risks

Choice	**Pre-Tax Earnings**	**After-Tax Earnings**		
		Fully *Ex Ante* Endowment Tax	**Insured Endowment Tax**	**Risk-Neutral Endowment Tax**
Safe	200K	100K	100K	120K
Risky: Win	500K	400K	200K	300K
Risky: Lose	100K	0	200K	60K

Instead, the only way to treat risk neutrally is to ensure that each individual's returns are the same on a risk-weighted basis before and after tax. This can accomplished by taxing everyone in this endowment class at a 40% rate on their *ex post* potential earnings, regardless of which path they choose. In other words, the only way to raise the required revenue and avoid distorting decisions away from the efficient mix of risky and safe choices is through a *risk-neutral endowment tax*.[193]

A *risk-neutral endowment tax* would also potentially address several challenges for resource egalitarianism. As explained in the prior section, a resource egalitarian may well conclude that the ideal tax is partially or fully a function of potential earnings. But, as Sam points out, applying an endowment tax purely on an *ex ante* basis does not seem to accord with resource egalitarianism because it would constrain life choices, potentially forcing one to get the education and skills that

[192] This result is suggested in some other contributions to the tagging literature, which generally does not delve into the normative trade-offs of a lump sum tax versus other methods of adjusting for tags. *See, e.g.*, [Mankiw and Weinzierl].

[193] There is some evidence that risk aversion falls as wealth rises and people do not measure the riskiness of a choice solely by its variance but also by the risk of loss. *See, e.g.*, John R. Brooks II, *Taxation, Risk, and Portfolio Choice: The Treatment of Returns to Risk Under a Normative Income Tax*, 66 TAX L. REV. 255, 257-58 (2013). If so, the risk-neutral endowment tax would need to be adjusted, but it still would take the form of the risk-neutral endowment tax here where the moment in time used to determine the tax rate is different from that used to determine the tax base.

would maximize one's potential earnings.[194] It could also lead to tax burdens that are so high that they are unpayable.[195] For example, *ex ante* expected potential earnings might include a very lucrative scenario that never materializes, or require a choice by the individual of which she is unaware. On the other hand, applying an endowment tax purely on an *ex post* basis will fail to adjust for choices for which resource egalitarians believe individuals should be held responsible; that is, option luck. It would also be inefficient because it would tax those who take calculated risks more lightly than those in a similar position who do not. By separating the factors determining the tax rate from those determining the tax base, a risk-neutral endowment tax avoids, or at least mitigates, these pitfalls.

Some may nevertheless view it as unfair to base an individual's tax rate on their *ex ante* potential earnings, especially if such earnings never could have materialized because the market for their talents did not develop in the way expected.[196] This concern is addressed in the following sub-sections. But it is worth noting that such an individual still could pursue any life plan they chose and would never owe more in taxes than they can earn.

In sum, a risk-neutral endowment tax uses the tools of optimal tax theory to address concerns that chosen risks will result in unpayable tax burdens under both egalitarian optimal tax theory and resource egalitarianism. Theorists in both traditions should consider adopting it.

C. Unwanted Preferences

Thus far, this section has focused on how optimal tax theorists and resource egalitarians treat inequalities that stem from choices where there is no question whether the individual truly wanted to make that choice, and whether it maximizes their well-being. This sub-section and the last consider inequalities stemming from a different source: inequality in the ability to convert resources into well-being through one's choices.[197]

Here we focus on yet another common objection to welfare egalitarianism: it can lead to perverse outcomes in the presence of addiction, internalized oppression, and other phenomena. In part, this objection questions whether our choices *define* what maximizes our well-being—or whether they are simply evidence of what maximizes our well-being that can be rebutted.[198] More fundamentally, it asks what attributes of our choices, preferences, and conditions determine whether and when we should be held responsible for them or not.

Resource egalitarians, following Dworkin, argue the key is whether a preference or condition is wanted. People should be held responsible for a preference or condition they do not want to change, regardless of whether it was actually chosen. This contrasts with Cohen who argues that choice alone is what matters.[199] Both could be considered as supporting the choices thesis.

---

[194] Sam, *supra* note 187, at 39.
[195] Sam, *supra* note 187, at 56.
[196] *Cf.*, Sam, *supra* note 187, at 57 (arguing that an endowment tax that uses ex ante potential income as the tax base and in setting the rate is unfair for these reasons).
[197] White, among others, has advanced this distinction. White, *supra* note 12, at 704.
[198] *See* Sen, *supra* note 25, at 206.
[199] [Cohen].

This distinction between what is chosen and wanted yields four possibilities. As Fried points out, this distinction between choosing and wanting would have no traction if all individuals follow the rational actor model, choosing rationally based on full information to maximize their utility.[200] Then, all choices are wanted, and the only relevant categories would be chosen and wanted preferences and unchosen and unwanted conditions (innate disabilities). But there is extensive empirical evidence challenging the assumptions of the rational actor model.

The potential disconnect between chosen and wanted preferences and conditions creates deeper challenges for resource egalitarians than they seem to acknowledge. In general, their ideal policy design principles track those of optimal tax theory on this issue. But on the most challenging questions, optimal tax theory offers more nuanced and specific responses.

As summarized in Table 3, resource egalitarians would hold individuals responsible for preferences or conditions that were chosen and (at least until any negative consequences ensued) consistently wanted. Let's call this first group "true" or *fundamental preferences*.

Table 3: Treatment of Preferences and Conditions under Resource Egalitarianism

	Chosen	**Unchosen**
Wanted	***Fundamental Preference*** (Don't Insure)	***Expensive Taste*** (Don't Insure)
Unwanted	***Acquired Disability*** (Insure)	***Innate Disability*** (Insure)

Generally optimal tax theorists also would not insure against inequalities arising from such choices because they assume such choices generate as much or more utility as any other use of the same portion of one's resources. So if our investment banker decides to become a ski instructor, she should bear the consequences of that choice under both theories. But in optimal tax theory, this conclusion is subject to numerous caveats. These include externalities and risk, as discussed in Sections III.B and IV.B, and questions about personal agency and preferences evolving over time, which are discussed in the following sub-section.

Turning to the second quadrant, both theories would also hold individuals responsible for the consequences of unchosen conditions they nevertheless want—perhaps an innately refined palate or eye for home furnishings. Let's call these conditions *expensive tastes*, which Dworkin defines as needing more income simply to achieve the same level of welfare.[201] Resource egalitarians argue individuals should be held responsible for expensive tastes, as long as the individual does not want to get rid of the taste.

Optimal tax theory also would not compensate for expensive tastes because it assumes the same utility function for resources. So if a friend of our investment banker argues he should be paid more than his colleagues doing the same work because he gets more utility from every dollar due to his fine eye for interior design, optimal tax theory would disregard his protestations. As discussed in Section III.A, it would do so because it is unknowable (even to him) whether he is

[200] Fried, *supra* note 8, at 124-25.
[201] Dworkin (1981b), *supra* note 2, at 229.

telling the truth, so it is most efficient to assume that he gets the same utility from a given level of resources.

In contrast, both theories generally *would* insure against innate, unwanted preferences or conditions affecting other well-being beyond potential income, such as a medically-verifiable health condition. Let's call these conditions, which were also discussed in Section III.A, *innate disabilities*. I will return to some of the difficulties in defining this category temporally in the final section.

Instead, I will focus here on the fourth quadrant: preferences and conditions that are chosen but unwanted. Let's call these *acquired disabilities*.

One type of acquired disability stems from inconsistent preferences over time. The quintessential examples are addiction and hyperbolic discounting, though inconsistent preferences can take many forms. For example, we discussed the possibility that once our investment banker teaches skiing for a few years, her earning potential (if employed on an hourly basis) will fall to $250,000. What if she knows this will happen and doesn't want to become a ski instructor long-term, but can't seem to restrain herself from heading to the slopes each morning instead of her job?

Resource egalitarians would address this possibility through their hypothetical insurance market, which can include regulatory approaches. For example, Dworkin argues that we generally shouldn't insure against laziness, and its effects on income, because it is part of one's tastes and ambitions, which are chosen and wanted. But he urges an exception when laziness is pathological—that is, when an individual "minded to overcome his laziness through acts of will [but] simply could not do so."[202] He also supports paternalistic mandates, such as requiring people to wear seat belts, when people make major mistakes they are very likely to regret.[203]

Thus, resource egalitarians effectively determine whether someone *wants* a preference from a temporal distance, not just based on what they choose in the moment. If, from a distance, our investment banker wants to keep banking but in any given moment she can't seem to act on this preference, she would be defined as not actually wanting to become a ski instructor in the first place. As a result, she would be treated similarly to someone who acquires a physical disability and wants to continue working at her current job but simply cannot.

Optimal tax theorists (and behavioral economists more broadly) take a similar approach. Generally they treat such phenomena as internalities. Initially, most optimal tax models assumed people were rational utility maximizers with consistent preferences. But behavioral economists have identified a wide range of ways in which people do not conform to these assumptions. Individuals' preferences may differ depending on how far a decision is in the future, how a choice is framed, or other aspects of the decision environment (so-called inconsistent preferences). Individuals may also fail to act in the ways they would if they were fully informed and there were no decision-making costs, whether due to limited time, attention, or information (so-called bounded rationality).

[202] Dworkin (2002), *supra* note 59, at 119.
[203] Dworkin (2002), *supra* note 59, at 114–15.

Optimal tax theorists argue that we should respond to internalities by assessing social welfare based on experienced utility, not decision utility, as discussed in Section II.B.[204] Thus, much like resource egalitarians, they determine what our investment banker actually prefers over time. If she prefers being a ski instructor in the moment but prefers being an investment banker at all times before and after, she essentially prefers banking over the sum of her moments.[205] The optimal response to an internality may be a Pigouvian tax or subsidy. In this example, it could involve taxing our investment banker *more* heavily if she becomes a ski instructor and using the revenue to tax her *less* heavily if she continues banking so that she appropriately accounts for her preferences outside the decision moment. But because internalities by definition involve individuals failing to maximize their experienced utility, Pigouvian taxes and subsidies may be less efficient than other approaches that leverage the underlying cause of the internality. These can include commitment devices, nudges, mandates, or other regulatory devices.[206] At the extreme, if our banker is unresponsive to such mechanisms because, for example, her myopia is so severe, she will be treated as if she has an acquired disability and can only be a ski instructor.

Another type of acquired disability stems from bounded rationality. For example, let's say our investment banker is simply unaware that her potential earnings (if employed on an hourly basis) will fall to $250,000 after a few years of working as a ski instructor. If she were aware of this fact, she would choose to continue investment banking instead.

Once again, resource egalitarians and optimal tax theorists arrive at similar principles when confronted with this possibility. Resource egalitarians assert that a key condition of the hypothetical insurance market is full information about the likelihood and consequences of different potential scenarios, and the ability to rationally calculate one's preferred level and form of insurance based on this information.[207] Put in optimal tax theory terms, they assume no bounded rationality. This implies that they would not hold our investment banker fully responsible for her lower potential income after leaving banking because it is not what she actually wants. If she had full information, she would have chosen otherwise.

Likewise, optimal tax theorists assert that redistribution should be based on one's experienced utility—one's preferences over time if one is fully informed. Focusing on experienced utility is not, as Dworkin contends, about educating people out of preferences that some objective outsider considers "wrong." To use one of his examples, it does not imply disregarding one person's belief that a certain type of life is good, even if solitary and without love.[208] Instead, it simply requires that people are fully aware of all their alternatives when forming their preferences.

In sum, both theories generally assume that an individual's choices reflect what she actually wants and what maximizes her well-being. Or, more technically, that individuals are rationally acting to maximize their fully informed preferences. But when this assumption fails—and is not amenable to social interventions correcting this failure—both will treat the choice as an acquired disability

[204] *See infra* note __.
[205] *See, e.g.*, Brian Galle, *The Problem of Intra-Personal Cost*, _ YALE J. HEALTH POL'Y L. & ETHICS _ (forthcoming). Some (potentially enlightened) behavioral economists take issue with this approach, arguing this reflects the bias of academics against living in the moment. *See, e.g.*, B. Douglas Bernheim & Dmitry Taubinsky, *Behavioral Public Economics* 27–28 (Nat'l Bur. Econ. Research Working Paper 24828, July, 2018).
[206] *See, e.g.*, Galle, *supra* note 205.
[207] *See, e.g.,* Dworkin (2000), *supra* note 59 at 212–213.
[208] Dworkin (1981a), *supra* note 4, at 225.

eligible for compensation. In the case of resource egalitarianism, an acquired disability is treated as a form of brute luck; in optimal tax theory as part of one's endowment.

With these additions and those regarding work effort, we can now update Table 1. Table 4 summarizes the analogous terms used by resource egalitarians and optimal tax theorists on issues where they largely agree.

Table 4: Analogues in Terminology of Resource Egalitarians and Optimal Tax Theory

	Resource Egalitarianism	**Optimal Tax Theory**
Thing to Be Equalized (Equalisandum)	Resources	Endowment as the welfare-maximizing proxy for utility
Extent of Redistribution Determined by	Average preferences behind thin veil of ignorance	Social welfare function
Sources of Inequality to Be Eliminated	Initial transferrable resources on a desert island	Exogenous, fixed resources
Sources of Inequality to Be Mitigated	Brute luck	Exogenous and observable conditions or states
	Pathological or unwanted tastes	Inconsistent preferences or hyperbolic discounting
	Choices behind the veil based on incomplete information	Bounded rationality
Sources of Inequality <u>Not</u> Be Mitigated	Option luck	Deliberate, fully informed choices; fundamental preferences
	Utility monsters, ascetics, and expensive tastes	Unverifiable differences in marginal utility
Maximum Required Labor Contributions by the Well-Off	Working part-time, possibly in highly-compensated job that don't like	Working part-time, possibly in highly-compensated job that don't like

There is one remaining issue: none of these neat categories should be read to imply that determining what someone wants is easy. Dworkin acknowledges that it can be difficult to distinguish between fundamental preferences and acquired disabilities—between choices that reflects one's "true" preferences versus pathologies one wishes one did not have and cannot control. The optimal tax literature grapples with this question too.[209] The difficulty is compounded when one considers all the small choices we (consciously or unconsciously) make every day, and the difficulty in isolating causal factors.[210] Should an individual who is hit by a car be treated

[209] *See, e.g.*, B. Douglas Bernheim & Antonio Rangel, *Addiction and Cue-Triggered Decision Processes*, 94:5 AM. ECON. REV. 1558 (Dec., 2004) (advancing a model of addiction where people cannot act on their underlying preferences because of cue-triggered compulsions that they view as mistakes even in the moment).

[210] For a nice discussion of these issues, see Fried, *supra* note 8, at 137-49.

differently if she chose to leave her house? To cross the street? To jaywalk? Doesn't the degree of control one has over a choice actually lie on a continuum?[211]

Here the tools of optimal tax theory once again offer a more nuanced and specific way to potentially address this conundrum. In Roemer's theory of equality of opportunity, he suggests estimating the likelihood that people in different categories will make a given choice, given all information we have about traits that are beyond their control, such as socio-economic background.[212] This likelihood can be used to estimate how much control each person has over that choice—and therefore the extent to which they truly "want" what they choose, and should be held responsible for it.[213] Roemer would then define the equalisandum as one's *ex ante* resources, adjusted for the effects of one's choices to the extent one should be considered responsible for them.[214] Effectively, the more one believes in social determinism, the more individuals should be insured against inequalities arising from their choices.

To build on one of his examples, suppose 99% of people in group A choose to smoke, while only 1% of people in group B do. Roemer would treat the smokers in group A essentially as having an innate or acquired disability. Thus, inequalities arising from their choice to smoke should be mitigated. In contrast, he would treat the smokers in group B as largely acting on their fundamental preferences. As a result, they should generally bear the costs associated with their choice.

While not yet well-integrated into the optimal tax literature, this approach offers an elegant method for addressing some of the more challenging objections to both resource egalitarianism and welfare egalitarianism by critics of the choices thesis.[215] Both theories wrestle with how to distinguish an individual who finds it difficult or impossible to act on her fundamental preferences, from one feigning difficulty who actually wanted what she chose. Part of the problem is that these categories are not in fact discrete and may, to a large extend, by the products of one's socialization. By focusing on objective correlates with one's control over a decision, Roemer's approach offers resource egalitarians a more nuanced way to disentangle option luck from brute luck than an their traditional all-or-nothing approach. And it does the same for optimal tax theorists struggling to disaggregate preferences from endowments.

D. Preferences Evolving Over Time

This brings us to a final difficulty, distinguishing between fundamental preferences and innate disabilities. As discussed in Section III.A, both theories generally would insure against innate disabilities or, more specifically, unwanted conditions. For example, suppose our investment banker has an innate chronic pain condition that dramatically improves when she spends hours skiing daily and worsens if she continues investment banking, no matter what accommodations her employer provides. Resource egalitarians would insure her against this possibility through their hypothetical insurance market because health is a non-transferrable resource. Optimal tax

[211] Murphy raises similar questions in Murphy, *supra* note 70, at 490.

[212] JOHN E. ROEMER, THEORIES OF DISTRIBUTIVE JUSTICE 276–78 (1996). While trained as an economist, Roemer has also written extensively on non-welfarist theories of justice and it is unclear whether he would view this approach as rooted in optimal tax theory.

[213] ROEMER, *supra* note 212, at 278.

[214] ROEMER, *supra* note 212, at 276-79.

[215] *See, e.g.,* Murphy (1996), *supra* note 70.

theorists would do so because her health is part of her endowment. In setting the amount of insurance, both would account for costs of insurance, such as moral hazard.

In certain cases, however, it can be difficult to distinguish innate disabilities from fundamental preferences. Perhaps the most interesting issues arise when people's fundamental preferences change over time.

For example, suppose people are only capable of optimizing over a limited period because their preferences evolve in ways they cannot anticipate. Such evolving preferences are not a case of hyperbolic discounting. In that situation, an individual always prefers X in the moment and Y before or after. Here we are positing an individual who prefers X both before, during and after the decision moment during one period of her life. She then evolves to consistently preferring Y during a later period.

Dworkin sets this possibility aside, acknowledging that his work does not discuss "what accommodation an equal distribution of resources should make for radical changes in people's minds about how they wish to spend their lives."[216] One might think that he and other resource egalitarians would respond by proposing to reinstate the hypothetical insurance market periodically over time. But he disavows this approach, arguing the tax and transfer system determined through the initial hypothetical insurance market ensures people are equal over their lifetimes.[217]

As critics of the choices thesis have pointed out, this response is somewhat dissatisfying, especially because it is plausible, and well-supported empirically, that many people's preferences do change dramatically over time.[218] For example, what if our investment banker consistently wants to be a ski instructor in her 30s and then, many years later, consistently prefers to be an investment banker but cannot do so anymore? This might not just be because she would like to make more money, but rather because she has learned from the experience of being a ski instructor that she does not enjoy it as much as she thought she would.

Theoretically, resource egalitarians would still hold her responsible for her choice to become a skiing instructor because it is an instance of option luck. But in a sense, her fundamental preference has converted over time into an unwanted preference of a former self. And a condition that one does not want and would not choose is the definition of an innate disability or brute luck.

This position of resource egalitarians is also dissatisfying because they do not assume that people should be held responsible for their choices when they are children. But as Fleurbaey argues, it is implausible that there is a mysterious moment before which one cannot meaningfully make decisions for oneself, and after which one suddenly is able to do so.[219]

---

[216] Dworkin (1981b), *supra* note 2, at 334. Rakowski appears to take a stronger view, arguing that resource egalitarianism (equality of fortune in his terminology) requires holding people permanently responsible for their actions after their reason, experience, and capabilities cross a threshold, except in extreme cases. ERIC RAKOWSKI, EQUAL JUSTICE 108 (1991).

[217] Dworkin (2002), *supra* note 59, at 120–21.

[218] At the same time, critics of the choices thesis are less troubled by holding people responsible for the choices that were made relatively recently. *See, e.g.*, Murphy, *supra* note 70, at 482.

[219] *See* Marc Fleurbaey, *Equal Opportunity or Equal Social Outcome?*, 11 ECON. & PHIL. 25, 42 (1995).

Optimal tax theory offers a different approach that would in many respects better align with the desiderata of resource egalitarians. It would consider such evolving preferences as another type of bounded rationality where one cannot think or form preferences beyond a certain time horizon, even if one knows the likelihood of different scenarios in the long-term. To simplify, suppose everyone could only optimize over 10 year periods. Then the optimal fiscal system would be one that maximizes social welfare over each 10 year period. It would treat one's self a decade from now essentially as a separate person with a different endowment, some of which is due to choices made by a different and earlier version of oneself.[220]

Interestingly, this approach implies the optimal tax system should be recalculated periodically based on the decision horizons of individuals. If our investment banker makes choices that reduce her actual or potential earnings, the fiscal system should reflect her new potential earnings once she has aged beyond her initial decision horizon. This reinforces the case for using one's *current* potential income, in whole or part, as the base of an endowment tax, as under a risk-neutral endowment tax. But unlike a risk-neutral endowment tax, it also implies that the rate should not necessarily be based on one's *ex ante* potential earnings. To the extent that individuals could not have optimized over the period encompassing the choice point and the present, they should not be held responsible for a prior choice. Instead, as Fried puts it, "moral responsibility for past actions must be systematically diluted over time in some fashion."[221]

This approach begins to bring optimal tax theory in line with the views of critics of the choices thesis, including those who support weighted beneficence. For example, Murphy generally objects to the choices thesis because he does not think people necessarily deserve to be held responsible for their choices, or that the concept of desert should necessarily play a starring role in a theory of distributive justice.[222] But he is less troubled by holding people responsible for the consequences of choices they made relatively recently.[223]

In addition, the possibility that people cannot optimize over their lifetimes is yet another reason why optimal tax theory does not entail so-called talent slavery. Shorter decision horizons imply optimal taxes that are even more focused on *ex post* outcomes than the risk-neutral endowment tax discussed above.

Optimal tax theorists increasingly subscribe to the view that decision horizons play a key role in the optimal tax and transfer system, and should do so more explicitly. Indeed, one of the primary arguments advanced in the recent literature arguing in favor of an income tax over a consumption or earnings tax is that shorter decisions horizons imply optimal taxes focused more on *ex post* income, rather than endowment or lifetime consumption possibilities. While Dworkin does not

[220] This is related to Parfit's concept of a chain of ancestral and future selves. *See* Derek Parfit, *Personal Identity*, 80 PHIL. REV. 3, 24-25 (1971).
[221] Fried, *supra* note 12, at 151.
[222] Murphy, *supra* note 70, at 484-87, 493. *See also* Elizabeth S. Anderson, *What Is the Point of Equality?*, 109 ETHICS 287, 288-89 (1999); Miranda Perry Fleischer, *Equal Opportunity and the Charitable Tax Subsidies*, 91 B.U. L. REV. 601, 642 (2011).
[223] Murphy, *supra* note 70, at 482.

explicitly make this argument, it also strengthens his case for a tax on actual, rather than potential, income.[224]

V. Conclusion

This essay has explored the primary non-welfarist critiques of optimal tax theory, paying particular attention to those of resource egalitarians. In the process, it has argued that the ideal policy design principles implied by each or these theories largely mimic the other. To be sure, the fundamental motivations and concerns of optimal tax theorists and resource egalitarians differ. But often they simply use different terminology to describe the policy design principles that should govern the ideal tax and transfer system. To use the terminology of resource egalitarians, this is generally true on the topics of so-called utility monsters and ascetics, altruistic and anti-social preferences, work effort and ambition, innate disabilities, pathological tastes, and incomplete information.

In some areas, however, the literature on optimal tax theory and resource egalitarianism both offer awkward, inconsistent or dissatisfying responses. Indeed, at times the only way they can respond to the concerns of non-welfarists is by deviating from their underlying mechanisms for determining distributive justice. For example, both require *ad hoc* adjustments to their mechanism to address concerns that systemic prejudice imposes uniquely heightened harms relative to inequalities arising from other forms of brute luck. They also both require *ad hoc* adjustments to their mechanism if one believes the ideal tax and transfer system should not require anyone to work.

But just as frequently the tools and objectives of optimal tax theory imply assumptions and modeling approaches that offer a clearer and stronger response to the concerns of non-welfarists. These assumptions and modeling approaches should be adopted more explicitly and consistently in optimal tax literature. For example, to address objections that valuations of externalities based on willingness to pay do not reflect social value, Pigouvian taxes or subsidies should be welfare-weighted. To address concerns that endowment taxes will generate so-called talent slavery, the portion of the ideal endowment tax that reflects potential earnings should be partially based on the number of hours worked times hourly potential earnings, not just on total potential earnings. The optimal endowment tax should also be risk-neutral—incorporating *ex ante* potential earnings in the tax rate, but only *ex post* potential earnings in the tax base—to address concerns about unpayable tax burdens after losing calculated gambles. To address objections that the degree of agency one has over choices often lies on a continuum, optimal tax models should consider holding people less responsible for choices that were relatively predictable based on their socioeconomic background and other immutable or semi-immutable factors. Finally, to address the fact that we are largely incapable of making choices that optimize over our lifetimes, optimal tax theory should more fully and consistently integrate individuals' decision horizons into its models.

[224] It may also have interesting implications for how the fiscal system should adjust for one's age. Potentially it should focus more on outcomes and *ex post* income later in life, because one's income at that point is more likely to be a product of one's prior selves. On the other hand, the young may have shorter decision horizons, implying that their taxation should focus more on outcomes.