Prioritarianism and Optimal Taxation

Matti Tuomala[1] Matthew Weinzierl[2]

December 31, 2020

Note: This text is forthcoming as Chapter 4 in *Prioritarianism in Practice*, Matthew Adler and Ole Norheim (editors), Cambridge University Press.

Abstract

Prioritarianism has been at the center of the formal approach to optimal tax theory since its modern starting point in Mirrlees (1971), but most theorists' use of it is motivated by tractability rather than explicit normative reasoning. We characterize analytically and numerically the implications of a more explicit use of prioritarianism in optimal tax theory. We also examine prevailing tax policies and surveys on tax preferences to gauge the influence of prioritarianism in practice. We conclude that optimal policy is highly sensitive to many key modeling choices and parameter assumptions, and these choices interact in complicated ways, but that a substantial shift in policy results if the social objective moves from utilitarian to prioritarian. When looking at existing policy and preferences, we find only limited evidence of prioritarian reasoning. We conclude with suggestions on the future of prioritarianism in optimal tax theory.

Keywords

Prioritarianism, Optimal Taxation, Utilitarianism, Redistribution, Inverse-optimum

In this chapter, we trace out the implications of adopting prioritarianism as the normative standard for the design of labor income taxation, and we examine the extent to which prevailing tax policies fit with those implications. Labor income taxation embodies, perhaps more directly and with broader reach than any other policy decision, social judgments on the proper role of government and distribution of economic resources.[3] The design of "optimal" taxation therefore depends critically on what we assume optimality entails. In other words, normative choices are at the heart of the study of optimal taxation.

First, we discuss and build on over nearly five decades of research in optimal tax theory to provide both analytical and numerical results on prioritarian-optimal tax policy. Though little appreciated—even by most tax theorists—James Mirrlees (1971) specified an objective for tax policy in the founding paper of that literature which directly translated the core idea of prioritarianism into his model's formal mathematical language. Emmanuel Saez (2001), in his important and influential restatement of Mirrlees's analysis, followed his lead, as have most others. Optimal tax theorists' use of prioritarianism is typically motivated by technical convenience, however, rather than explicit normative reasoning. In fact, most optimal tax theorists prefer to avoid normative judgments altogether and evaluate only the (Pareto) efficiency of policies, focusing their

1 University of Tampere, Matti.Tuomala@tuni.fi.

2 Harvard Business School and NBER, mweinzierl@hbs.edu. Thanks to Robert Scherf for excellent research assistance. The editors of the volume for which this text was prepared and the authors of other chapters within it provided valuable comments.

3 While we focus on labor income taxation, adopting prioritarianism would have implications for the design of related policies such as social insurance and capital taxation, as well.

efforts on managing an array of positive (not normative) complexities. The objective used by Mirrlees, Saez, and most other theorists admits a prioritarian specification, but it also admits a utilitarian or maximin specification, and when forced to choose among efficient allocations, the convention in modern optimal tax theory has been to default to these extreme cases. Thus, much more can be done to develop the normative case for prioritarianism in modern optimal tax theory.[4]

Second, we examine the extent to which prevailing tax policies and attitudes toward taxation appear to embrace prioritarian goals. We provide an overview of the results of both recent so-called inverse optimum analyses, in which a society's apparent normative preferences are inferred from its chosen tax policies, and recent positive optimal taxation analyses, in which carefully designed surveys are used to elicit popular views on taxation. Although we do not find widespread evidence of the influence of prioritarianism in either case, much more work remains to be done along these lines. Our hope is that this chapter will help inspire that work.

This chapter proceeds as follows. Section 1 examines what prioritarianism can teach us about optimal taxation through theoretical analysis and numerical simulations. In Section 2 we turn from theory to practice, as we look for evidence on the influence of prioritarianism in prevailing tax policies and attitudes toward taxes. Section 3 raises and discusses prominent reservations about prioritarianism and looks to the future, considering arguments for prioritarianism that economists may find appealing and suggesting approaches that future research might take.

4.1 Prioritarianism and the theory of optimal taxation

Modern optimal tax theory is written in the formal language of constrained optimization. Society is assumed to have an objective for its tax policy and to face economic constraints on what its policies can achieve. The task of the optimal tax theorist—and the hypothetical "social planner" assumed to have authority over tax policy—is to derive the tax policy that will best achieve this objective given these constraints.

In a typical optimal tax analysis, the objective for policy is to maximize social welfare, an object that is calculated through a social welfare function that depends upon only the welfare (i.e., well-being) levels of the individuals in society. The founding modern optimal tax analysis of Mirrlees (1971) has the social planner maximize the following expression:

$$W = G\left(U\left(c, \frac{y}{n}; n\right)\right) f(n) dn \tag{4.1}$$

In expression (4.1), W denotes social welfare, $U(\cdot)$ individual utility[5], c consumption, y income, n an individual's exogenous and unobserved productivity (with $f(n)$ as the associated density), and $\frac{y}{n}$ labor effort. $G(\cdot)$ is the social welfare function that takes individual utilities as its arguments.

Prioritarianism fits seamlessly into this formal structure, requiring only that the social welfare function be strictly concave (and increasing) in individual welfare levels; that is, for $U_j > U_i$, $G'(U_j) < G'(U_i)$. In words, strict concavity of the social welfare function means that an additional unit of utility for an individual with a high level of utility increases social welfare by less than an additional unit of utility for an individual with low utility.

[4] At several points in this chapter we discuss the maximin criterion as a limit case of the standard social welfare function used in optimal tax. An alternative criterion is the leximin criterion, discussed in Chapter 2, which can diverge from the maximin when the worst off individuals are unaffected by policy and which is arguably the limiting case of prioritarianism. The near absence of discussion of leximin in the optimal tax literature, perhaps because it is analytically less convenient than maximin, leads us to focus on maximin in this chapter. But the difference between maximin and leximin imply that some critiques of the former do not apply directly to prioritarianism.

[5] Note that we assume a one-period model of utility with homogenous preferences over consumption and leisure. Both are commonplace in optimal tax research. The former reflects that the model is best understood as a lifetime model. The latter relates to a fairly extensive literature on heterogeneous preferences both in the context of labor and capital income taxation. See Tuomala (2016), Chapter 10 for a discussion. Two-dimensional problems tend to lead to difficult technical and normative problems.

In fact, Mirrlees (1971) assumes the following specific form for $G(U)$ when conducting numerical analysis:

$$G(U) = \frac{-1}{\beta} e^{-\beta U} \tag{4.2}$$

Marginal social welfare from an increase in utility is thus given by:

$$G'(U) = e^{-\beta U} \tag{4.3}$$

Consistent with a prioritarian objective, this marginal social welfare is decreasing in the level of U for $\beta > 0$. Mirrlees assumes $\beta \geq 0$, thus putting a prioritarian objective at the heart of his analysis.[6]

In his baseline quantitative analyses, however, Mirrlees uses $\beta=0$, the utilitarian specification in which $G'(U)=1$, "for simplicity," leaving the prioritarian objective for robustness checks. These choices set the stage for the next nearly five decades of research in optimal tax, where prioritarianism plays an active but mostly implicit role in both analytical and quantitative results, and the default parameterization of the social welfare function that admits prioritarianism has been, instead, utilitarian.

4.1.1 Prioritarianism's effects on analytical optimal tax results

Mirrlees (1971) derived an analytical expression for marginal tax rates that shows the central place of prioritarianism—or at least optimal tax theorists' formalization of it—in the literature. Marginal tax rates are the key determinant of the income tax structure in the Mirrlees model, and they are obtained by maximizing the objective specified in expression (4.1) subject to two sets of constraints: the government needs to raise enough tax revenue to fund its spending, and (2) taxpayers will respond to tax policy to maximize their own well-being. The details of the analysis are not important to this chapter, so we skip directly to the result.[7]

Saez (2001) provided the most widely used version of this expression for optimal marginal tax rates. It is:

$$\frac{T'(y)}{1-T'(y)} = \frac{1-F(y)}{\varepsilon y f(y)} \int_y^{\infty} \frac{1-h(z)}{1-F(y)} dF(z) \tag{4.4}$$

In this expression, $T'(y)$ is the marginal tax rate at income level y. The income distribution is $F(y)$, with density $f(y)$, and ε denotes the elasticity (responsiveness) of income to taxation.[8]

Briefly, the intuition behind this formula is as follows. The optimal marginal tax rate at income y is greater when it transfers income from a larger population that earns more than y (i.e., when $1-F(y)$ is greater), when it directly distorts a smaller tax base (i.e., when $yf(y)$ is smaller), when the responsiveness of taxable income is less (i.e., when ε is smaller), and when the welfare gains of transferring resources from those with incomes greater than y are larger (i.e., when the final integrand is larger).

For the purposes of this chapter, the key term in expression (4.4) is $h(y)$, which is typically called the marginal social welfare weight (MSWW) of an individual earning income y and is defined as the social welfare generated by a marginal increase in consumption for this individual relative to a marginal increase in consumption spread equally across the entire population. Mathematically, this MSWW is made up of two components: the increase in the individual's well-being due to an increase in consumption, and the increase in social welfare that arises from an increase in that individual's well-being.

The distinction between the two components of a MSWW is important to understanding how prioritarianism differs from alternative normative principles for optimal tax theorists. In a prioritarian objective, as shown above in the case of Mirrlees' specification in expression (4.3), the second component of the MSWW has a specific feature: the increase in social welfare due to an increase in an individual's well-being is decreasing in that individual's level of well-being. Specifically, $h(y)$ decreases with y in a prioritarian objective even if the (assumed non-increasing) marginal utility of consumption is constant, because

[6] Mirrlees (1982) wrote, in the context of this welfare function: "At time I used it, I had no intention of avoiding the addition of utilities, but rather of looking at the effect on optimal policies of having a more inequality averse utility function."

[7] For derivations, see Diamond (1998) and Saez (2001) or Tuomala (2016)

[8] This form was also found in Atkinson (1990) and Atkinson (1995).

marginal social welfare is decreasing in utility and thus income. In other words, the marginal value to society of an extra unit of consumption is greater if it goes to the worse-off even if individuals' utility from income is nondecreasing because a gain in utility for the worse-off is worth more to society than a gain in utility for the better-off.

The implication of moving to a prioritarian objective from a utilitarian objective, and thus moving to an *h(y)* in expression (4.4) that declines more rapidly with y, is that marginal tax rates are greater along the income distribution, enabling greater redistribution to individuals earning less. This result is, of course, consistent with the intention to give priority to those with less well-being. To see more specific results on its optimal policy effects, however, we need to turn to numerical simulations.

4.1.2 Prioritarianism's effects on quantitative optimal tax results

A voluminous literature has explored how a range of specifications of the social welfare function, including those that capture a prioritarian objective, translate into quantitative optimal tax results. To solve the optimal nonlinear income tax model numerically requires four key elements.[9] We consider each in turn.

The first factor is the social welfare function (SWF). As in Mirrlees (1971), we will suppose the $G(U)$ function in the SWF takes the form in expression (4.2), where β measures the degree of inequality aversion in the SWF of the government. In the case of $\beta = 0$, we define $G(U)=U$ and have the utilitarian SWF; in the limit as $\beta \to \infty$, we have the maximin criterion in which $W = \min(U)$. Recall that the curvature in the individual's utility from consumption modifies $G'(U)\, U_c$ and makes social preferences (implicitly) more redistributive.

The second factor is the distribution of income-earning abilities. Following Mirrlees (1971), much work on optimal non-linear and linear income taxation used a lognormal distribution of productivities (e.g., Atkinson (1972), Stern (1976), Tuomala (1984), Kanbur and Tuomala (1994), Mankiw et al. (2009)). As commonly known, the lognormal distribution fits data on incomes reasonably well over a large part of the income distribution but diverges from it markedly at both tails. The Pareto distribution in turn fits well at the upper tail. Here we will use the two-parameter version of the Champernowne distribution, with parameters m (scale parameter) and θ (shape parameter).[10] This distribution approaches asymptotically a form of the Pareto distribution for large values of wages but it also has an interior maximum. A small value for the shape parameter, say $\theta = 2$, reflects high inequality, while a larger value such as $\theta = 3$ reflects low inequality. Among two parameter distributions, it is the best fitting for pre-tax income distribution in Finland (1990-2010), where the θ parameter varies from 2.78 to 2.40 over the period and was approximately constant at 2.50 from the latter part of 1990s to 2010.

The third factor is the shape of individual preferences. Of course, many representations of preferences are possible, and as we will see below which representation we choose will matter for our results. Here, we consider a utility function of the following CES form, which has been frequently used in numerical simulations:

$$U = \frac{-1}{c} - \frac{1}{1-l}, \tag{4.5}$$

where the elasticity of substitution between consumption (c) and leisure ($1 - l$) equals one-half. In the absence of taxation, the labor supply function is backward-bending given these preferences.

---

[9] For details on these simulations, Appendix 5.2 in Tuomala(2016) describes the computational procedure and the FORTRAN program used. Standard optimal tax models simplify reality in many ways to focus on specific questions. Unfortunately, in many cases these simplifications may affect the results. For example, the interplay between marginal rates and public provision is missing in the standard optimal income tax model. Blomquist, Christiansen and Micheletto (2010) examine implications of public provision for tax distortions. They suppose public provision is strictly tied to working hours (e.g. daycare) and paid by the income tax. Then part of the tax is a direct payment, like a service fee or market price. Hence, not all of the marginal tax rate is distortionary, and one must also look at how tax revenue is spent as well as raised to get the full picture.

[10] It is also known as the Fisk distribution.

The fourth factor is the government's revenue requirement R, specified as the fraction of total income not used for consumption by individuals

$$R = 1 - \int \frac{c(n)f(n)dn}{y(n)f(n)dn}$$

This revenue can be interpreted as what is required to fund public goods, and its size affects the cost of raising revenue to fund transfers to poor individuals, a key distributive tool in these models.[11]

To see the effect of prioritarianism on quantitative results in optimal tax theory, we now examine what happens when we change the values for the four inputs described above.

First, the most direct effects of prioritarianism are apparent when we vary inequality aversion (i.e., through the parameter β, where prioritarianism requires $\beta > 0$). In Table 4.1 and Figures 4.1 through 4.3, we show tax schedules for different levels of β (note that these taxes should be compared with the schedules for the overall taxation of income and expenditures in real economies). All the calculated tax schedules take the form of a lumpsum credit (or basic income) followed by marginal tax rates, and we plot these schedules against incomes.

Table 4.1: Tax schedules for different social objective parameterizations

Table1 f=Champernowne distribution holding mean constant m=e^{-1} ,θ =2.5,R=0.1, $u = -\frac{1}{x} - \frac{1}{(1-y)}$

F(n) =Percentiles of wage distribution

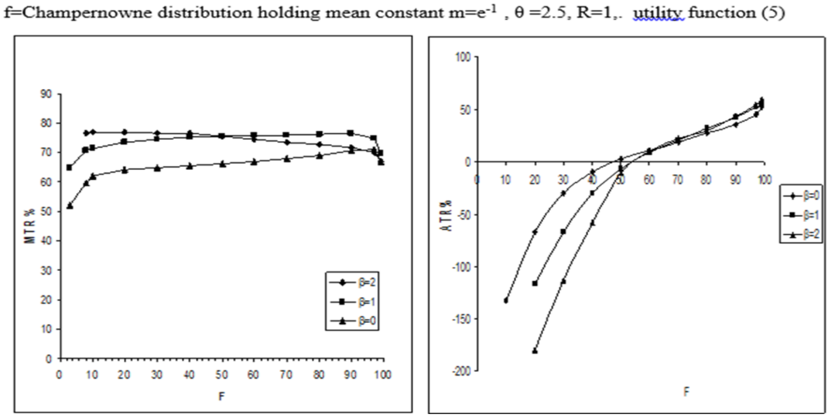
	β= 0	β= 0	β= 0.5	β= 0.5	β= 1	β= 1	β= 2	β= 2	β=∞	β=∞
F(n)	ATR %	MTR %	ATR %	MTR %	ATR %	MTR %	ATR %	MTR %	ATR %	MTR %
0.10	-133	62	-198	70	-200	74	-400	77		
0.50	3	61	-8	69	-7	73	-12	75	-6	79
0.90	35	63	39	69	42	69	43	72	47	72
0.97	45	64	51	69	52	64	55	70	57	70
0.99	52	64	57	67	54	47	59	67	60	63

Note: Table 4.1 shows the average tax rate (ATR) and marginal tax rate (MTR) schedules—along points in the income distribution F(n)—for different values of the parameter β in the social welfare function, given the parameter values shown atop the table.

Figure 4.1: Tax schedules for prioritarian and utilitarian social objectives

Marginal tax rates (MTR) **Average tax rates (ATR)**

f=Champernowne distribution holding mean constant m=e^{-1} , θ =2.5, R=1,. utility function (5)

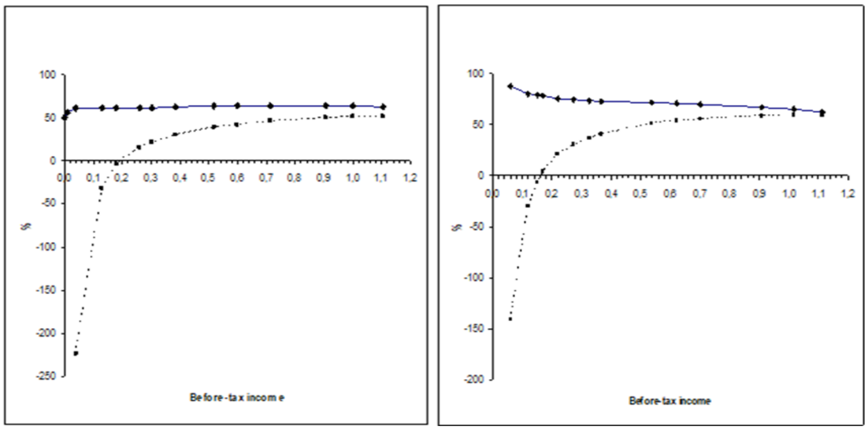


[11] Weinzierl (2018) raises the concern that how tax revenue is spent may matter for the distribution of income, a dependence largely assumed away in the literature.

Note: Figure 4.1a plots the marginal tax rate (MTR) and Figure 4.1b the average tax rate (ATR) schedules—over the income distribution F(n)—for values of the parameter β in the social welfare function, given the parameter values shown atop the table.

Figure 4.2: Tax schedules for utilitarian and maximin social objectives

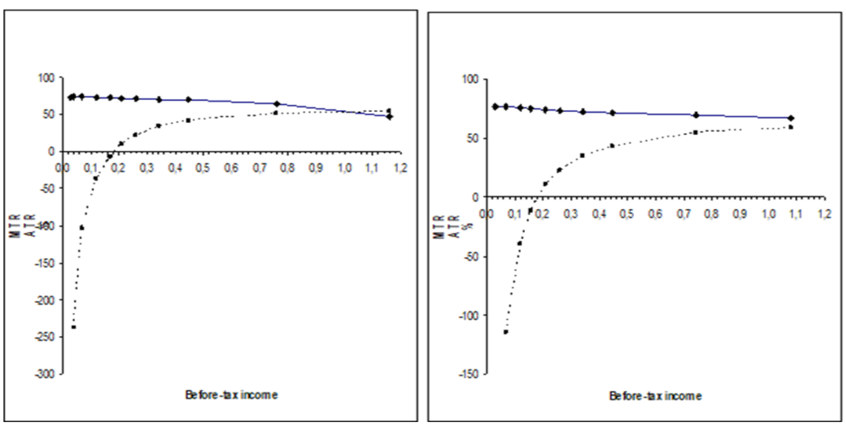
f=Champernowne distribution holding mean constant $m=e^{-1}$, θ =2.5, R=1, utility function (5)



Note: Figure 4.2a plots the MTR and ATR schedules—against before-tax labor income—for the Utilitarian social welfare function (β=0), given the parameter values shown atop the table. Figure 4.2b does the same for the maximin social welfare function (β→∞).

Figures 4.3a and 4.3b: Tax schedules for prioritarian social objectives

f(n)=Champernowne distribution holding mean constant $m=e^{-1}$, θ =2.5, R=1, utility function (5)



Note: Figure 4.3a plots the MTR and ATR schedules—against before-tax income—for the social welfare function with β=1, given the parameter values shown atop the table. Figure 4.3b does the same for the maximin social welfare function (β=2).

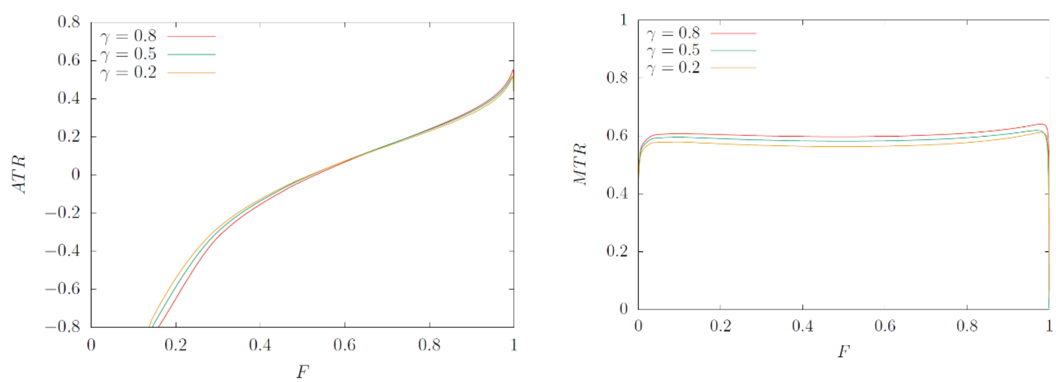
Our numerical results suggest that marginal tax rates tend to increase for all taxpayers with increasing inequality aversion. One might have expected that increasing β above 1.0 would have a particularly large effect, but our calculations show that this is not true. In fact, it can be seen that the difference in the marginal rates between the case β = 1 (or β = 2) and the maximin solution (the limit as $\beta \to\infty$) is not very large, whereas it is much more marked between the case β = 1 and the utilitarian case β = 0. Our results also seem to suggest that a sufficiently high level of inequality aversion (a large enough β) leads to a pattern of optimally declining marginal tax rates, and indeed in the maximin case we find that marginal tax rates decline monotonically with income. It may be surprising that the maximin (Rawlsian, as in Rawls (1971)) objective does not lead to increasing marginal tax rates, but this pattern follows from the fact that this objective is not concerned with inequality among those not in the "least fortunate group". Average tax rates rise with income much more steeply in the maximin case than in the utilitarian case considered in our simulations. On the other hand, it can be seen that the maximin case and the case β = 2 do not differ much in terms of average tax rates. That is, they do not differ much in the extent of redistribution. We will return to the extent of the redistribution later in this article.

In addition to exploring different parameter values within the SWF of the form in expression (4.2), called the Kolm-Pollak form or (especially in this chapter's context) the Kolm-Mirrlees-Pollak form, we also consider a constant relative utility-inequality aversion form: the so-called Atkinson form of

$$V(U) = \frac{U^{1-\gamma}}{1-\gamma}$$

The numerical simulations with the Atkinson form proved cumbersome when γ is greater than 1. We settled for numerical solutions where γ is less than 1. The results of the simulations are summarized in the following figures 4.3c and 4.3d, using the CES case (5) with a Champernowne distribution. The revenue requirement R=0.1 in all cases. As expected, an increase in ρ leads to higher marginal tax rates in both specifications. The pattern and the levels of MTR's and ATR's are similar to those we have in the Kolm-Mirrlees-Pollak case for β between 0 and 0.5.

Figures 4.3c and 4.3d: Tax schedules for Atkinson form of the social objective



Note: These figures plot the MTR and ATR schedules—over the income distribution F(n)—for the Atkinson form of the social welfare function with several values for γ, given CES utility and a Champernowne ability distribution.

The discussion of tax structures in the optimal income tax literature has been almost entirely about marginal tax rates. Notably, almost all analytical results focus on the structure of marginal tax rates to the neglect of average tax rates, the latter of which are arguably more important indicators of income tax progressivity. After all, high marginal tax rates as such perform no direct distributional function; their purpose is to increase average tax rates higher up the income scale. In fact, in all cases shown in our Tables and Figures, average tax rates are increasing in income. Analytically it is difficult to establish this, but computational techniques can demonstrate these patterns.

Second, we gauge the sensitivity of the shape of the tax schedule to the joint choice of the parameter θ in the Champernowne distribution and the parameter β, the degree of inequality aversion. In Table 4.2, we computed solutions when θ varies from 2.0 to 3.3 with the CES utility function in (5) and different β parameters. When θ=3.3, marginal tax rates are declining. When θ=2.0 and β=0, marginal tax rates are increasing with income to around the 98th percentile; when β = 1 they are increasing to the 96th percentile.

Table 4.2: Tax schedules for different social objective and ability distribution parameterizations

Table2 f=Champernowne distribution holding mean constant m=e^{-1} ,θ=3.3,2,0,R=0.1, $u=-\frac{1}{x}-\frac{1}{(1-y)}$

F(n) =Percentiles of wage distribution

	β=0	β=0	β=1	β=1	β=0	β=0	β=1	β=1
	θ=3.3	θ=3.3	θ=3.3	θ=3.3	θ=2.0	θ=2.0	θ=2.0	θ=2.0
F(n)	ATR%	MTR%	ATR%	MTR%	ATR%	MTR%	ATR%	MTR%
0.10	-50	58	-100	73	-100	62	-100	71
0.50	6	52	3	67	-18	66	-29	75
0.90	27	52	35	64	41	71	47	76
0.97	33	53	44	61	55	71	61	75
0.99	39	51	47	58	60	68	65	70

Note: Table 4.2 shows the average tax rate (ATR) and marginal tax rate (MTR) schedules—along points in the income distribution F(n)—for different values of the distributional parameter θ and the parameter β in the social welfare function, given the other parameter values shown atop the table.

As an alternative SWF we calculate solutions for rank order SWFs. Aaberge (2007) provides a parametric variant for rank order social preferences as follows

$$G_k = \begin{cases} -lnF & \text{for k} = 1 \\ \dfrac{k}{k-1}(1-F^{k-1}) & \text{for k} = 2,3,4\ \ldots' \end{cases}$$

where F denotes the percentile of the income distribution for an individual. When k approaches ∞, this weighting function approaches the utilitarian case and there is no concern for inequality. When k=2, then G_2=2(1-F), which is in effect the weighting underlying the Gini coefficient, as shown by Sen (1974) who provided an axiomatic justification for such a social welfare function. In this case, the social marginal valuation declines linearly with F from twice the average for the lowest paid taxpayer to zero for the highest paid taxpayer. The optimal tax schedules with utilitarian and Gini weights differ considerably, with marginal rates being higher in the latter (see Figure 4.4).

Third, we consider different expressions for individual preferences. Most previous simulations (see Mirrlees (1971), Atkinson (1972), Tuomala (1984), Kanbur and Tuomala (1994)) have used either logarithmic (Cobb-Douglas) utility of consumption or what is shown in expression (4.5), both of which are instances of the utility functional form in which the coefficient of relative risk aversion (the CRRA, which controls the curvature of utility over consumption) is constant (in these cases, it is 1.0 and 2.0, respectively). In light of Table 1 in Chetty (2006), those values may be too high. In addition, other preferences yield more complex—and arguably more realistic—relationships between net wage and labour supply. In empirical labor supply studies, e.g. Keane and Moffitt (2001), preferences over working time and net income are given by a utility function that is quadratic in hours and net income. To illustrate the effects of using this alternative specification, we solve numerically cases in which the utility function is quadratic in consumption:

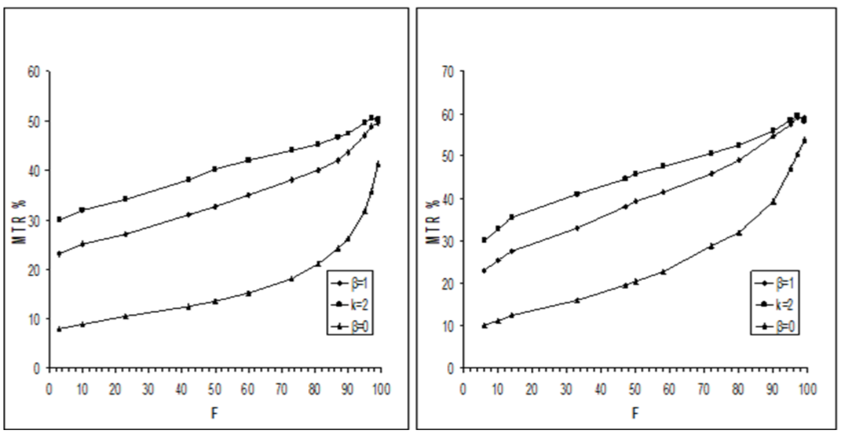
$$u = (c-1) - a(c-1)^2 - (1-l) \tag{4.6}$$

The curvature in the utility from consumption in (4.6) is smaller than that used in the previous simulations, and the coefficient of relative risk aversion varies at different values of c. With the parameterization used in our computations (including a=5), the values of the coefficient of relative risk aversion are smaller and more in line with the empirical labor supply literature. Moreover, the elasticity-based marginal tax formulas turn out to be useful because we can calculate traditional labor supply elasticities at each point of the distribution.

The striking thing in the numerical results shown in Figures 4.4 (plotted against income percentiles) and 4.5 (plotted against income levels) is that once we assume that preferences are given by the utility function that is quadratic in consumption, the shape of optimum tax schedules may be altered drastically. The marginal tax rates rise with income, practically speaking, over the whole range (up to the 99.7 percentile point), except that in cases $\beta = 1$ and $\beta = 2$ marginal rates decline slightly at the very top of the earnings distribution. The reason for this is that this utility function with an upper bound on consumption implies a concave budget constraint in the Mirrlees model.

Figure 4.4: Tax schedules for different social objectives, quadratic utility specification

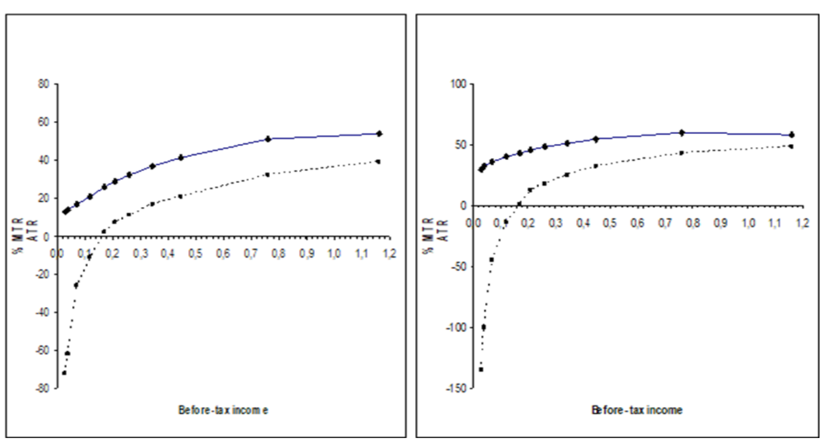
f(n)=Champernowne distribution holding mean constant $m=e^{-1}$, R=0.1In both cases R=0.1, utility function (6)



Note: Figures 4.4a and 4.4b plot the marginal tax rate (MTR) schedules—over the income distribution F(n)—for different social welfare functions, assuming the quadratic form for utility in expression (4.6) and the parameter values shown atop the figures.

Figure 4.5: Tax schedules for utilitarian and prioritarian social objectives, quadratic utility specification

f(n)=Champernowne distribution holding mean constant $m=e^{-1}$. In both cases R=0.1 and utility function (6)



Note: Figures 4.5a and 4.5b plot the MTR and ATR schedules—against before-tax incomes—for utilitarian and prioritarian social welfare functions, assuming the quadratic form for utility in expression (4.6) and the parameter values shown atop the figures.

Next, we explore how sensitive is the level of the lump sum transfer (or basic income) component of the tax system to the specification of the model. Tables 4.3 and 4.4 display the ratios of the basic income to the average net income with different welfare weights and values for θ when utility is either (4.5) or (4.6) and the revenue requirement is ten percent of total income.

Tables 4.3 and 4.4: Level of basic income under different utility and social objective specifications

Table 3. The ratio of x_o (basic income) to the average net income

$u = -\frac{1}{x} - \frac{1}{(1-y)}$ and X/Z=0.9

θ	β=0	β=1	k=2,gini weights
3.0	0.53	0.66	0.62
2.5	0.61	0.69	0.67
2.0	0.62	0.70	0.65

Table 4. The ratio of x_o (basic income) to the average net income

$u = (x-1) - a(x-1)^2 - (1-y)^{-1}$, a=5 and X/Z=0.9

θ	β=0	β=1	k=2, Gini weights
3.3	0.06	0.25	0.33
2.5	0.15	0.35	0.42
2.0	0.25	0.44	0.51

Note: Tables 4.3 and 4.4 show basic income level as a share of the average income level for different specifications of utility, the ability distribution, and the social welfare function (as shown atop each table).

This ratio is clearly higher in the case of $\beta = 1$ and Gini weights $2(1-F)$ (when $k = 2$) than in the case of the pure utilitarian $\beta = 0$. The ratio is increasing with pre-tax inequality (i.e., decreasing in θ).

Since marginal tax rates may be a poor indication of the redistribution powers of an optimal tax structure we measure the extent of redistribution, denoted by RD, as the proportional reduction between the percentile ratio P90/P50 for market income y and disposable income c.[12] In other words, RD captures the extent to which pre-tax inequality between these points in the income distribution is reduced through taxation. Tables 4.5 and 4.6 show the extent of redistribution (in terms of our measure) for the two forms of the utility function we consider and different distributional objectives.

Tables 4.5 and 4.6: Extent of redistribution under different utility and social objective specifications

[12] Unlike the scalar inequality measures, the use of fractile measures such as the percentile ratio allows us to consider changes in inequality at various different points in the distribution.

Table 5. The extent of redistribution (RD)

$u=-\frac{1}{x}-\frac{1}{(1-y)}$, R=0.1

θ	β=0	β=1	k=2, Gini weights
3.0	60.0	67.0	61.6
2.5	71.8	86.1	80.3
2.0	84.3	93	92.9

Table 6. The extent of redistribution (RD):

$u=(x-1)-a(x-1)^2-(1-y)^{-1}$, a=5 R=0.1

θ	β=0	β=1	k=2,Gini weights
3.3	31.1	49.5	28.1
2.5	52.3	55.0	44.4
2.0		61.6	57.3

Note: Tables 4.5 and 4.6 show a measure of redistribution—defined in the text—for different specifications of utility, the ability distribution, and the social welfare function (as shown atop each table).

We can ask: Is pre-tax inequality (which is decreasing in θ) more important than redistributional preferences in determining the extent of redistribution? In response, perhaps the most interesting finding in our simulations can be seen in Table 4.7.

Table 4.7: Extent of redistribution under different utility and social objective specifications

Table 7. The extent of redistribution (RD); utility function $u=-\frac{1}{x}-\frac{1}{(1-y)}$, P90/P50, R=0.1

θ	β=0	β=1	Maximin	Gini weights
3.3	22.6	34.0	36.8	33.0
2.5	37.3	47.5	50.7	44.3
2.0	47.0	60.0	61.7	60.5

Note: Table 4.7 shows a measure of redistribution—defined in the text—for different specifications of utility, the ability distribution, and the social welfare function (as shown atop the table).

It turns out that increasing β above 1 has a very modest effect to the extent of redistribution. This is true for all three values of θ. Hence, the extent of redistribution is about the same in the β =1 case as in the maximin case. It is important to note that the extent of redistribution and rising marginal tax rates may be quite different things. The extent of redistribution is larger with higher β, but the marginal tax schedule may be rising in the case of β =0.

Finally, if larger tax revenues are to be raised, would it mean significant changes in the shape of the tax schedule? Our earlier numerical results seem to suggest that marginal tax rates tend to increase for most taxpayers with increasing net government expenditure. The shape of the tax schedule remains quite similar, however, in spite of changes in government expenditure. As expected in the following tables we see the extent of redistribution is much smaller in the utilitarian case than in the prioritarian case (with β =1). Perhaps slightly surprising, in the utilitarian case the extent of redistribution declines dramatically with an increased revenue requirement, whereas in the prioritarian case it changes little (see Tables 4.8 and 4.9).

Tables 4.8 and 4.9: Extent of redistribution under different revenue requirements

Table 8 The extent of redistribution (RD), utility function, $u = -\frac{1}{x} - \frac{1}{(1-y)}$, P90/P50, θ=3

R	β=0	β=1
0.1	22.9	33.0
0.2	9.1	32.0

Table 9. The extent of redistribution (RD), utility function $u = -\frac{1}{x} - \frac{1}{(1-y)}$, P90/P10, θ=3

R	β=0	β=1
0.1	51.6	69.1
0.2	14.0	67.4

Note: Tables 4.8 and 4.9 show a measure of redistribution—defined in the text—for different values of β and the revenue requirement *R*, given the parameters shown atop each table.

In sum, the optimal income tax schedule is very sensitive to many key modeling choices and parameter assumptions. Moreover, these choices interact in complicated ways. For example, consider the choice of the form of utility from consumption. Unlike in most of cases with the CES utility function (5), marginal tax rates rise with income when we assume that utility is quadratic in consumption. Hence, the interaction between two components of a MSWW plays a central role in determining the shape of tax schedule.

4.2 Prioritarianism and the practice of taxation

We now turn from the task of designing prioritarian-optimal taxation to that of examining prioritarianism's place in prevailing tax policies and attitudes toward them. After all, the scholarly study of taxation must be in dialogue with the reality of taxation if it is to be seen as relevant to those who vote on and are affected by it.

4.2.1 Are existing taxes prioritarian? An overview of inverse-optimum research

A recent literature tries to extract the implicit judgments of society—e.g., society's MSWWs in the Mirrlees model discussed above—from existing tax policy. Called the "inverse optimum" approach, the idea of this research is that the optimal tax model can be run "in reverse" to back out society's objective function, once we assume a set of constraints and have data on outcomes. That is, an implied schedule of marginal social welfare weights can be inferred from optimal tax formulas by populating them with existing tax schedules and values for all non-normative parameters, leaving the welfare weights as the only unknowns.

Inverse optimum researchers have generally found that MSWWs are non-negative and greater on low earners than high earners, but consensus on more specific normative implications—such as the policies' fit with prioritarianism—remains elusive. An inventory of results would include, among others, the following. A very early contribution was Christiansen and Jansen (1978), who study Norway and find that its implicit inequality aversion parameter is quite far from the utilitarian value, giving more social weight to low income people. A similar study was carried out in India by Ahmad and Stern (1984); they also infer substantial inequality aversion. Stern (1978) finds, however, that the U.K. system implicitly fits with an equal sacrifice principle quite far from prioritarianism (and discussed below). More recent studies have used micro data, such as the pioneering work of Bourguignon and Spadaro (2010), who consider the revealed social preferences of the French tax-benefit system. Spadaro et al. (2012) study the tax and transfer systems of 26 European countries, finding evidence of Rawlsian policies in some countries but less redistributive policies in others. Bargain et al. (2014) consider the tax and transfer systems of 17 EU countries and the United States and find that social welfare weights are always positive, though they are not monotonically declining for low income groups, and that there are significant differences in social welfare weights between, but not within, groups of countries (e.g., the U.S. vs. Continental Europe, Scandinavia, and Southern Europe). Lockwood and Weinzierl (2016) examine the U.S. tax system over time, highlighting the relative stability of implicit weights over the last several decades. Jacobs, Jongen, and Zoutman (2017) use this method to find the redistributive preferences implicit in the reform proposals of Dutch political parties and find that all parties give a higher social weight to the poor than the rich.

Such an inventory of results is difficult to synthesize, however, and (as far as we are aware) no meta-analysis of the inverse-optimum literature exists. Complicating a synthesis is that existing results span a range of countries, time periods, and estimation approaches, and they are communicated in many different ways. While most authors discuss whether their estimated weights are consistent with Pareto efficiency (i.e., whether they are non-negative) and their overall shape (e.g., whether they are declining in income), only some authors discuss whether their estimates are consistent with utilitarian, Rawlsian, libertarian, prioritarian, or other more specific normative principles. We hope to begin filling this gap, and in the process provide some evidence on the role of prioritarianism in the practice of taxation.

Given the diversity within the inverse optimum literature, performing a meta-analysis requires us to impose some structure if we hope to say something informative about prevailing policy and prioritarianism. We will rely on the functional forms discussed above: in particular, we will use the form

$$G(U) = \frac{-1}{\beta} e^{-\beta U}$$

for social welfare as a function of individual utility and

$$U(c) = \frac{(c - \bar{c})^{1-\lambda} - 1}{1 - \lambda}$$

for utility, which is a generalization of the CES form (i.e., where $\lambda = 2$). The only unfamiliar term in these expressions is $\bar{c}$, which we include to represent a "minimum acceptable" level of consumption. While this term is somewhat unconventional, it plays an important role in making our numerical analysis suitable to existing policies because, fortunately, no households in the advanced economies for which inverse-optimum analyses have been completed face the prospect of zero consumption. Rather, these societies appear to have established a minimum level of consumption below which individuals are not allowed to fall. By including the $\bar{c}$ term, we assign the high marginal utilities of consumption implied by $\lambda > 0$ to the individuals whose after-tax incomes place them nearest to this minimum acceptable level.[13]

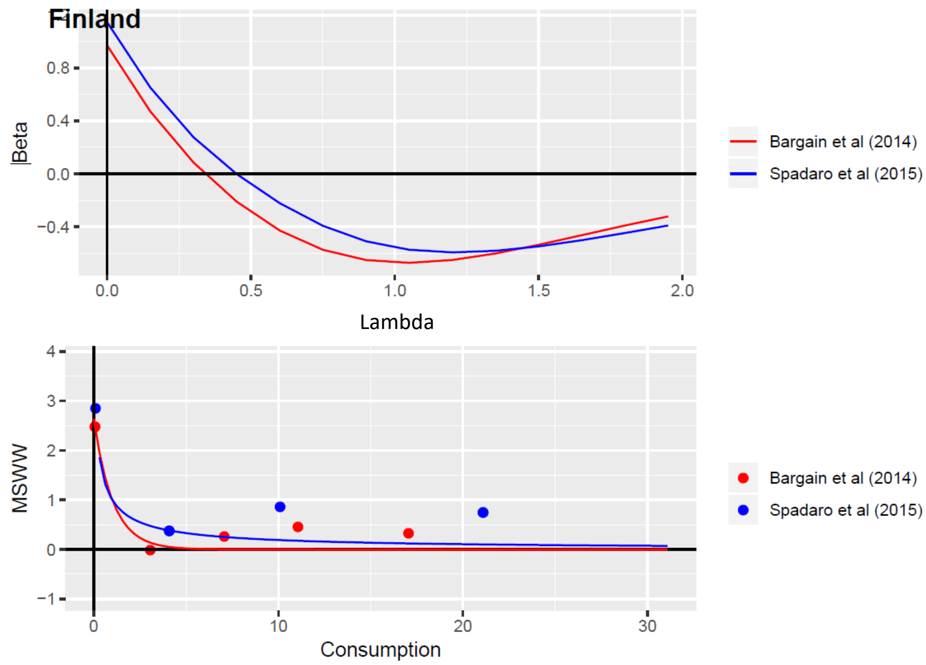
Our procedure is as follows. We survey the inverse optimum literature by country. For each analysis, we record the estimated marginal social welfare weights at a set of income percentiles (0, 10, 25, 50, 75, and 90 where available). Each of these points corresponds to a level of disposable (after-tax and -transfer) income. We then numerically find the value of β for a range of values of λ that, assuming our functional forms above, imply marginal social welfare weights that minimize the unweighted sum of squared errors from these recorded estimates.

For reference, recall that $\beta = 0$ is a utilitarian SWF, and prioritarianism corresponds to $\beta > 0$. Similarly, $\lambda = 0$ implies utility that is linear in consumption, while concave utility corresponds to $\lambda > 0$ (with $\lambda = 1$, log utility, a standard benchmark).

To illustrate our results, consider the following figures for Finland:

[13] Note that this modification to the utility function is not a positive affine transformation—so that the transformed utility function is not CRRA—and it is not the kind of rescaling proposed in Chapter 2 and undertaken in Chapter 8.

Figure 4.6: Inverse-optimum results for tax policy in Finland



Note: Figure 4.6 shows the pairs (λ, β) that best explain existing tax policy (top graph) and the marginal social welfare weights implied by the best-fitting (λ, β) pair (bottom graph).

The top figure plots the pairs of β and λ that provide the best fit to two separate, well-known, inverse-optimum estimates: those by Bargain et al. (2014) and Spadaro et al. (2015). The bottom figure shows the estimated MSWWs from these papers at the discrete points in the income distribution (the dots) along with an example of the fitted MSWW curves for each (in fact, it shows the best fitting curves among the possible parameter pairs for each paper). Note that consumption of zero in the bottom figure is the point at which $c = \bar{c}$ not c = 0.

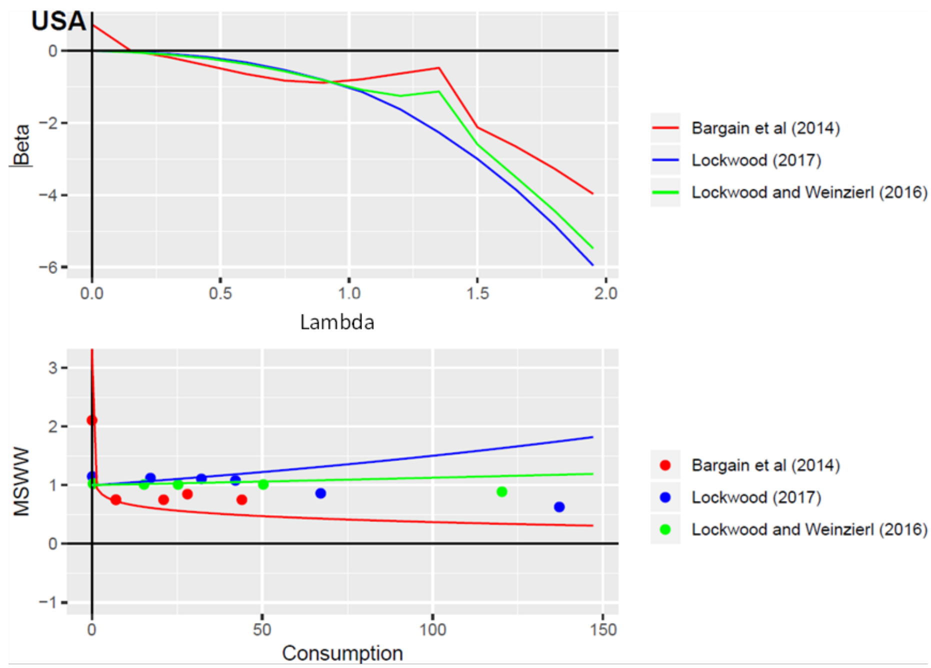
These results for Finland reveal two lessons that appear, from our early steps in this analysis, to apply across a wide range of countries.

First, the estimated MSWWs are "too flat for too long" to be well-described by a SWF with substantial concavity over consumption. The dots for both papers lie above the fitted curves at higher income percentiles, as the SWF has a difficult time assigning sizeable positive weights to the relatively well-off, explaining why this “too flat for too long result” appears to apply to the inferred values, not the fitted curve.

Second, a prioritarian objective for policy can be inferred from the data only if the concavity of individual utility is well below levels typically assumed. That is, the tradeoff between concavity in the social welfare function ($\beta > 0$) and in the individual utility function (λ > 0) is severe, and $\beta > 0$ obtains only at values for λ far below the standard benchmark of λ = 1. The utilitarian case ($\beta = 0$) is consistent with a value of λ=0.5, larger than for prioritarianism but still substantially lower than typically estimated.

Other countries yield similar lessons. Take, for example, the United States.

Figure 4.7: Inverse-optimum results for tax policy in the United States



Note: Figure 4.7 shows the pairs (λ, β) that best explain existing tax policy (top graph) and the marginal social welfare weights implied by the best-fitting (λ, β) pair (bottom graph).

Here, the two lessons are even more stark. The estimated welfare weights are far too flat at too high incomes for $\beta > 0$ and $\lambda > 0$ to hold. Strikingly, the estimates for Lockwood (2017) and Lockwood and Weinzierl (2016) lie below their best-fit lines only because these lines are upward-sloping, using a $\beta < 0$ to generate increasing MSWWs. Moreover, the scope for $\beta > 0$ is extremely limited in the case of the United States (as shown in the top panel). For virtually all values of λ, the SWF must give greater weight to those with greater utility levels—a sort of anti-prioritarianism—to be consistent with prevailing policies.

Similar figures can be generated for a range of economies. These two lessons apply in nearly all cases.

Of course, our analysis could be improved in many ways in future research. Most important would be to have additional inverse-optimum results to fit. Experimenting with the MSWWs implied by different underlying functional forms, as our simulations from earlier in the chapter suggested they can affect optimal rates substantially. And an important step will be to clarify the role of different policy components (e.g., taxes versus medical insurance) in driving the MSWW estimates. Perhaps these and other improvements in the elicitation or analysis of inverse optimum estimates will revise the implication that is hard to avoid from the current results: we lack strong evidence that prioritarianism is a prominent guiding principle of existing tax policy.

4.2.2 Are tax policy preferences prioritarian? Positive optimal taxation

A recent literature, referred to as positive optimal tax theory, studies prevailing views on distributive justice and thus taxation among the general public. Using both theory and empirics, this new literature attempts to complement philosophical introspection with the use of empirically verifiable evidence for social preferences. In this way, the positive optimal tax theory project can be seen as a way to accommodate both discomfort with imposing normative judgments and a desire to go beyond Pareto efficiency.

Early papers in positive optimal tax theory are Weinzierl (2014, 2017) and Saez and Stantcheva (2015), each of which finds evidence suggesting that prevailing normative priorities in the United States are not well-captured by prioritarianism. Weinzierl (2014) asks U.S. survey respondents to choose between two tax policy alternatives: one based on the standard utilitarian criterion and the other based on the principle of Equal Sacrifice, a less redistributive alternative supported by John Stuart Mill (1871) in which each individual's tax burden imposes the same sacrifice in absolute utility terms. He finds that a majority of the respondents choose the latter, suggesting that support for the more redistributive prioritarian objective (as the sole guide to policy) would be quite limited. In subsequent work, Weinzierl (2017) finds similar enthusiasm for the principle of taxation preferred by Adam Smith (1776), referred to by Richard Musgrave (1959) as Classical Benefit-Based Taxation. As with Equal Sacrifice, popular support for this principle is sharply at odds with any claim that prioritarianism captures the public's priorities for tax policy. Saez and Stantcheva (2014) introduce the formal tool of Generalized Social Marginal Welfare Weights to the modern optimal tax model, where prioritarianism could be easily captured formally (i.e., by weights that decline even if individual marginal utility from consumption is constant). They gather survey evidence on preferred values for those weights in the United States, and they find substantial support for values reflecting a mix of principles: "...social preferences are in between the polar utilitarian and libertarian cases," they write. In sum, recent evidence points away from prioritarianism (or utilitarianism) as a good guide to how the public thinks about tax policy, at least in the U.S. context. (Also see Chapter 10 for related evidence)

What can this evidence tell us about which normative principles are consistent with existing policy? One possibility is that policies reflect a mix of priorities that, together, generate a strong interest in helping the worst off but a weak interest in redistribution across the rest of the income distribution. Perhaps the preferences of Richard Musgrave (1998) are as good a guide as any: "Moreover, observers such as myself who tend to be egalitarian should not rule out the norm of Lockean entitlement to earnings (Locke 1690, Nozick 1974) as an alternative criterion that also deserves consideration. Most people, I suggest, would wish to assign some weight to both norms...I would give it, say, one-quarter weight with three-quarters to the Rawlsian concept."

4.3 The future of prioritarianism in optimal tax theory and practice

Consider this striking fact: none of five of the most important, authoritative book-length treatments of modern optimal tax theory includes "prioritarianism" in its index (nor, aside from minor exceptions that we have found, in their texts). In order of publication, these are Louis Kaplow (2008), The Mirrlees Review (2010, 2011), Bernard Salanie (2011), Robin Boadway (2012), and Matti Tuomala (2016). Of course, all of these books prominently feature the implicitly prioritarian social welfare function discussed earlier, but none considers the case for a prioritarian specification of it by that name.

The lack of explicit engagement with prioritarianism may reflect the youth of prioritarianism as a named concept (it was first introduced into published philosophical works around 2000) or economists' naivete regarding normative considerations, but it may also reflect an underlying discomfort with the principle. After all, the bar for embracing a particular normative standard is extremely high among most economists, including optimal tax theorists, who are hesitant to wade into philosophical debate and have long been tempted to limit their analysis to determining only the (Pareto) efficiency of tax policy.

As discussed above, utilitarianism serves as something like a default normative principle for many economists. To prioritarianism's supporters, its delineation from utilitarianism is a selling point, but to most optimal tax theorists it is an obstacle. Skepticism toward prioritarianism can be traced to the tremendously influential—among economists, at least—arguments of William Vickrey and John Harsanyi, as well as Kenneth Arrow's defense of them, in which utilitarianism was derived as a social analogue to individual expected utility maximization. Many economists view debates over the normative objective of policy as having been essentially resolved by their contributions—resolved, that is, in favor of utilitarianism.14 As Louis Kaplow (2008) notes, when choosing among possible allocations from behind a veil of ignorance, a prioritarian social welfare function will recommend policies that deliver lower expected utilities to everyone in society, violating an ex ante version of the Pareto criterion that economists tend to hold in high esteem. A utilitarian social welfare function avoids this concern. Of course, respecting ex ante Pareto comes at a cost: utilitarianism cannot respect the Pigou-Dalton principle at the heart of prioritarianism (see Chapter 2 of this volume for a more detailed discussion of this tradeoff).

A closely related critique of prioritarianism (also discussed in Chapter 2) is that it may invite confusion over the root of concavity in distributional judgments. In the first several decades following Mirrlees (1971), most optimal tax theorists explicitly stated a separate utility function and social welfare function. More recently, however, and often in numerical applications, it has become common to adopt a "reduced form" implicit composition of these two functions and, thus, conceptually combine their concavities (over utility levels and income, respectively). Louis Kaplow (2010) strongly objects to this approach, arguing that these two layers of concavity are conceptually distinct and showing that they are not mathematically additive in the way some theorists assume. Prioritarians agree with Kaplow, separating clearly these two layers of concavity. But anti-prioritarians prefer avoiding the confusion altogether by simply omitting concavity in the social welfare function, in part because the concavity of the individual utility function reflects declining marginal utility of income that is (arguably) empirically verifiable while the concavity of the social welfare function involves a value judgment that is not.

Perhaps the most direct path for prioritarianism to gain acceptance is for it to be seen as a plausible—even preferable—alternative to utilitarianism. After all, it shares the tractability that economists so value in utilitarianism, so if philosophers were to give economists good reasons to prefer it on normative grounds, the latter would likely be receptive. For example, prioritarianism might be put forward as a viable candidate for a sufficiently flexible and general single objective that can capture the normative pluralism behind existing tax policies and prevailing tax policy preferences identified in this chapter. As another example, the work of Fleurbaey and Maniquet (2006) provides an axiomatic derivation of "fair taxation" designed to incorporate a careful response to preference heterogeneity-a long-standing concern in the optimal tax literature[15]-and derives the following result: "the optimal tax should give the greatest subsidies to the

[14] Economists have tended to see the prioritarian view as a less extreme version of John Rawls's difference principle under which priority is given to the worst off in society. (The use of maximin in much optimal tax research is driven—at least in part—by its tractability and its position opposite utilitarianism on a redistributive spectrum). Skepticism toward Rawls's claims, therefore, has contaminated economists' views of prioritarianism. Kenneth Arrow (1973) famously compared Rawls's principle to the analysis of William Vickrey and John Harsanyi, who employed similar thought experiments to Rawls's veil of ignorance but obtained strikingly different results. Vickrey and Harsanyi approached the problem of choosing an SWF as formally equivalent to the problem of individual choice under uncertainty. They argued that rational individuals choosing economic policy with an equal probability of becoming any realized person-e.g., from behind what Rawls called the veil of ignorance-would choose a utilitarian social welfare function so as to maximize their expected level of well-being. Rawls (1972) thought this logic for utilitarianism failed to adequately account for the degree of risk aversion that agents behind the veil of ignorance would surely feel when faced with such high-stakes consequences. But Arrow disputed that argument by noting that Rawls in turn failed to account for Vickrey and Harsanyi's representation of utility, which already embedded risk aversion. Arrow then demonstrated that, far from being opposed to utilitarianism, the difference principle could be viewed as "a limiting case of it." That is, if individual preferences were represented by von Neumann-Morgenstern utility and the degree of risk aversion approached infinity, a properly scaled utilitarian social welfare function would approach the maximin criterion. Rawls (1972) objected to Arrow's interpretation of his difference principle as a limiting case of utilitarianism. He thought it was wrong to suggest that we can "shift smoothly from the moral conception to another simply by varying the parameter" (p.664). In fact, Rawls thought that an important feature of a distributive criterion is that it should serve as a public principle. He wrote that "citizens generally should be able to understand it and have some confidence that it is realized" (Rawls, 1972 p 143). He claimed that the maximin, unlike utilitarianism, satisfies this criterion of sharpness or transparency.

[15] While earlier theorists such as Mill (1871) and Edgeworth (1897) were comfortable using cardinal measures of utility across individuals, the rise of the New Welfare Economics of the 1930s caused many economists to view utility as an ordinal rather than a cardinal object. If utility levels and changes have no cardinal meaning for individuals, then redistributive preferences-e.g., prioritarian preferences-that compare levels and changes in utilities across individuals are similarly meaningless. Instead, all one can say is whether individuals gain or lose from a policy change.

To economists, the attractiveness of analyses limited to ordinal comparisons survives to the present day. For example, Ivan working poor (the agents having the lowest skill and choosing the largest labour)."[16] This conclusion is resonant with the ideals of prioritarianism, which is perhaps not surprising given that one of Fleurbaey and Maniquet's axioms is the Pigou-Dalton principle.

At this point in the optimal tax literature's development, perhaps the best summary of the relationship between optimal tax theory and prioritarianism is the following: optimal tax theorists can and sometimes do use prioritarian objectives to define optimal policy, but they don't give reasons—theoretical or empirical—why they should. One implication of this state of affairs is that those committed to prioritarianism ought to guide optimal tax theorists toward a more satisfying way to use it in their work and toward analyses that would demonstrate its appeal through some combination of its normative force, its desirable implications for taxation, and its real-world relevance.

Werning, a leading modern optimal tax theorist, provides conditions for Pareto efficient bounds on tax rates in Werning (2007) and notes, at the outset of the paper, "It is worth remarking that, given the focus on Pareto efficiency, no interpersonal comparisons of utility will be needed. Thus, the cardinality of preferences is completely irrelevant and only the ordinal features of preferences matter."

Compelling as it remains, resistance to interpersonal comparisons of utility was weakened in the 1970s with the rise of what Stiglitz (1987) called the New New Welfare Economics. While economists of this era retained a respect for the Pareto criterion, they sought a way around the limits it imposed. As Stiglitz noted, "it still remains the case that many of the critical choices necessitate interpersonal trade-offs, choices among alternative Pareto-efficient allocations." In response, these economists-led by Mirrlees-advocated the use of the social welfare function to evaluate and even compare different Pareto-efficient allocations. This shift made room for the use of normative principles such as prioritarianism in the leading models of optimal taxation.

While modern theorists have thus embraced interpersonal utility comparisons of the sort required by prioritarianism, following through on this choice raises a number of difficulties. To cite just one example, most modern optimal tax analyses assume all individuals have the same utility function in order to aggregate utilities. If preferences across goods are heterogeneous, however, the way these preferences are represented in the utility function will matter for optimal policy and requires a normative judgment.

[16] There are well-known technical difficulties related to incentive constraints to study multidimensional optimal tax problems including both of the elements. Another problem is how to incorporate heterogeneous preferences into a social welfare function (SWF) in analyzing optimal tax policy. There is a growing body of literature which studies multidimensional optimal tax problem by avoiding the technical complications by assuming multidimensionalities can be represented with one-dimensional aggregation of the multidimensional characters. Lockwood and Weinzierl (2015) take this approach.

In the spirit of Roemer (1998) and Van de Gaer (1993) another approach is to apply a compromise between the principle of compensation and the principle of responsibility. For individuals with the same preferences but different wage rates, the maximin criterion is applied (of course we can apply less extreme prioritarianism). In other words a zero aversion of inequality can be applied along the dimension of responsibility preference) whereas a high aversion to inequality is acceptable along the dimension of circumstances (in skill). Ravaska, Tenhunen, and Tuomala (forthcoming in ITAX) applied this approach.

References

[1] Adler, Matthew (2012). *Well-being and fair distribution: beyond cost-benefit analysis*. Oxford: Oxford University Press.

[2] Ahmad, Ehtisham, and Nicholas Stern (1984). "The theory of reform and Indian indirect taxes." *Journal of Public Economics*, 25: 259-298.

[3] Arrow, Kenneth J. (1972). "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice by John Rawls," *The Journal of Philosophy*, 70: 245-263.

[4] Atkinson, Anthony B., and Joseph E. Stiglitz (1972). "The structure of indirect taxation and economic efficiency." Journal of Public Economics, 1: 97-119.

[5] Atkinson, Anthony B. (1990) “Public Economics and the Economic Public,” *European Economic Review* 34,225-248

[6] Atkinson, A.B. (1995), *Public Economics in Action*. Oxford: Oxford University Press.

[7] Bargain, Olivier, et al. (2014). "Tax-benefit revealed social preferences in Europe and the US." Annals of Economics and Statistics/Annales d'Economie et de Statistique, 113/114: 257-289.

[8] Blomquist, Soren, Vidar Christiansen, and Luca Micheletto (2010). "Public provision of private goods and nondistortionary marginal tax rates." American Economic Journal: Economic Policy, 2: 1-27.

[9] Boadway, Robin (2012). *From Optimal Tax Theory to Tax Policy*. Cambridge: MIT Press.

[10] Bourguignon, Fran;ois, and Amedeo Spadaro (2012). "Tax-benefit revealed social preferences." *The Journal of Economic Inequality*, 10: 75-108.

[11] Chetty, Raj (2006). "A new method of estimating risk aversion." *American Economic Review*, 96: 1821-1834.

[12] Christiansen, Vidar, and Eilev S. Jansen (1978). "Implicit social preferences in the Norwegian system of indirect taxation." *Journal of Public Economics*, 10: 217-245.

[13] Diamond, Peter A. (1998). “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates.” *American Economic Review*, 88: 83-95.

[14] Edgeworth, Francis Y. (1897). "The pure theory of taxation." *The Economic Journal,* 7: 46-70.

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

[16] Fleurbaey, Marc, and Francois Maniquet (2018). "Optimal income taxation theory and principles of fairness." Journal of Economic Literature, 56: 1029-79.

[17] Jacobs, Bas, Egbert L. Jongen, and Floris T. Zoutman (2017). "Revealed social preferences of Dutch political parties." Journal of Public Economics, 156: 81-100.

[18] Kanbur, Ravi, and Matti Tuomala (1994). "Inherent inequality and the optimal graduation of marginal tax rates." The Scandinavian Journal of Economics, 96: 275-282.

[19] Kaplow, Louis (2008). *The Theory of Taxation and Public Economics*, Princeton: Princeton University Press.

[20] Kaplow, Louis (2010). "Concavity of utility, concavity of welfare, and redistribution of income." *International Tax and Public Finance*, 17: 25-42.

[21] Keane, Michael, and Robert Moffitt (1998). "A structural model of multiple welfare program participation and labor supply," International Economic Review, Aug:553-589.

[22] Locke, John (1690). *Second treatise of government: An essay concerning the true original, extent and end of civil government*. John Wiley & Sons, 2014.

[23] Lockwood and Weinzierl (2015). "De Gustibus non est Taxandum," *Journal of Public Economics*, 124: 74-80.

[24] Lockwood and Weinzierl (2016). "Positive and Normative Judgments Implicit in U.S. Tax Policy, and the Costs of Unequal Growth and Recessions." *Journal of Monetary Economics*, 77: 30-47.

[25] Lockwood, Benjamin B. (2017). "Optimal income taxation with present bias," Working Paper, University of Pennsylvania.

[26] Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan (2009). "Optimal taxation in theory and practice." *Journal of Economic Perspectives*, 23: 147-74.

[27] Mill, John Stuart (1871). *Principles of Political Economy*. Oxford University Press, 1994.

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

[29] Mirrlees, James A. (1982). "The economic uses of utilitarianism," in *Utilitarianism and Beyond,* Sen, A., Williams, B. and Williams, B.A.O. eds., Cambridge University Press.

[30] Mirrlees, James A., and Stuart Adam (2010). *Dimensions of tax design: the Mirrlees review.* Oxford University Press.

[31] Mirrlees, James A., et al. (2011). "The Mirrlees Review: conclusions and recommendations for reform." *Fiscal Studies*, 32: 331-359.

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

[33] Musgrave, Richard A. (1998). *Public Finance and Public Choice: Two Contrasting Visions of the State*, Cambridge: MIT Press.

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

[35] Ravaska, Terhi, Sanna Tenhunen, and Matti Tuomala (2018). "On the optimal lifetime redistribution and social objectives: a multidimensional approach." *International Tax and Public Finance* 25: 631-653.

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

[37] Roemer, John (1998). *Equality of Opportunity*, Cambridge: Harvard University Press.

[38] Saez, Emmanuel (2001). "Using elasticities to derive optimal income tax rates." *The Review of Economic*

Studies, 68: 205-229.

[39] Saez, Emmanuel and Stefanie Stantcheva (2015). "Generalized Social Welfare Weights for Optimal Tax Theory." *American Economic Review*: 24-45.

[40] Salanie, Bernard (2011). *The economics of taxation*. Cambridge: MIT press.

[41] Sen, Amartya (1974). "Informational bases of alternative welfare approaches: aggregation and income distribution." *Journal of Public Economics*, 3: 387-403.

[42] Smith, Adam (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations.* Homewood, Ill: Irwin.

[43] Spadaro, Amedeo, Luca Piccoli, and Lucia Mangiavacchi (2015). "Optimal taxation, social preferences and the four worlds of welfare capitalism in Europe." *Economica*, 82: 448-485.

[44] Stern, Nicholas (1976). "On the specification of models of optimum income taxation." Journal of Public Economics, 6: 123-162.

[45] Stern, Nicholas (1977). "Welfare weights and the elasticity of the marginal valuation of income." In: Artis, M. and Nobay, R., (eds.) *Proceedings of the Aute Edinburgh Meeting of April 1976*. Basil Black- well.

[46] Stiglitz, Joseph (1987). "Pareto Efficient and Optimal Taxation and the New Welfare Economics," in A. Auerbach and M. Feldstein (eds.) *Handbook of Public Economics*, 991-1041, Amsterdam: North Holland.

[47] Tuomala, Matti (1984). "On the optimal income taxation: Some further numerical results." *Journal of Public Economics*, 23: 351-366.

[48] Tuomala, Matti (2016). *Optimal redistributive taxation*. Oxford: Oxford University Press.

[49] Van De Gaer, D. (1995). "Equality of opportunity and investment in human capital." Katholieke Universiteit Leuven.

[50] Weinzierl, Matthew (2014). "The Promise of Positive Optimal Taxation: Normative Diversity and a role for Equal Sacrifice," *Journal of Public Economics*, 118: 128-142.

[51] Weinzierl, Matthew (2017). "Popular Acceptance of Inequality Due to Innate Brute Luck and Support for Classical Benefit-Based Taxation." *Journal of Public Economics*: 54-63.

[52] Weinzierl, Matthew (2018). "Revisiting the Classical View of Benefit-Based Taxation." *Economic Journal*, 128: F37-F64.

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