DISCUSSION PAPER SERIES

DP18758

HOW DOES PREDISTRIBUTION AFFECT REDISTRIBUTION?

Ravi Kanbur and Matti Tuomala

PUBLIC ECONOMICS

ISSN 0265-8003

HOW DOES PREDISTRIBUTION AFFECT REDISTRIBUTION?

Ravi Kanbur and Matti Tuomala

Discussion Paper DP18758
Published 17 January 2024
Submitted 16 January 2024

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

HOW DOES PREDISTRIBUTION AFFECT REDISTRIBUTION?

Abstract

This paper presents an analysis of the consequences for redistribution policies of achieving a more equal predistribution, that is, a more equal distribution of the predetermined income earning capabilities that individuals bring to the market. We show that optimal fiscal policies are less redistributive when the predistribution is more equal. We then quantify the value of a more equal predistribution. We show that total consumption is higher with a more equal predistribution. We also develop a money metric measure of social welfare and show that a more equal predistribution increases this measure for a Utilitarian and Prioritarian Social Welfare Function but may decrease it for the Maximin.

JEL Classification: D31, H21

Keywords: Predistribution, Redistribution, Social welfare function

Ravi Kanbur - sk145@cornell.edu
Cornell Unicersity and CEPR and CEPR

Matti Tuomala - matti.tuomala@tuni.fi
Tampere University

How Does Predistribution Affect Redistribution?

Ravi Kanbur and Matti Tuomala

This Version: 11 January 2024

Abstract

This paper presents an analysis of the consequences for redistribution policies of achieving a more equal predistribution, that is, a more equal distribution of the predetermined income earning capabilities that individuals bring to the market. We show that optimal fiscal policies are less redistributive when the predistribution is more equal. We then quantify the value of a more equal predistribution. We show that total consumption is higher with a more equal predistribution. We also develop a money metric measure of social welfare and show that a more equal predistribution increases this measure for a Utilitarian and Prioritarian Social Welfare Function but may decrease it for the Maximin.

JEL: D31, H21

Keywords: Predistribution, Redistribution, Optimal Non-linear Taxation, Marginal Tax Rates, Social Welfare Function.

Introduction

The term predistribution has entered the policy discourse and is now part of the economist's lexicon. It refers broadly to the distribution of income generating capabilities which individuals bring to the market and whose returns generate the distribution of market income. It is this market distribution of income which policy tries to change through taxes and transfers to achieve a final distribution which is in line with social objectives. A contrast is then drawn between such post-market redistribution versus predistribution policies which seek to change the distribution of pre-market capabilities. Here is how former leader of the UK Labour Party, Ed Miliband, illustrated the issue for a general audience:

"Think about somebody working in a call centre, a supermarket, or in an old peoples' home. Redistribution offers a top-up to their wages. Pre-distribution seeks to go further - higher skills with higher wages." (BBC, 2012).

In the academic literature the concept of predistribution was introduced by Hacker (2011) and has now become a standard formulation (see for example Piketty et. al. (2022)). It would not be an exaggeration to say that in the last three decades there has been a distinct turn away from the post war "social democratic" consensus on the leading role of post-market redistribution in achieving egalitarian objectives. There appears to be a general disenchantment with tax and transfer policies of redistribution, and this has gone hand in hand with the increasing emphasis on addressing pre-market inequalities in labour productivity. In a recent review of the policy analytic literature, Ferreira (2023) summarizes:

"My reading is that there is a growing consensus on "pre-distribution" policies, but perhaps less so on re-distribution policies." (p. 14)

Kanbur (2023a) develops this further by specifying three types of arguments that underpin this shift away from redistribution:

"that redistribution of income has technical and economic incentive issues, that it can be challenged on moral philosophical grounds, and that political economy has turned against it in favor of other forms of intervention." (p. 2)

As shown in Kanbur (2023a, 2023b), each of these arguments can be countered and debated. Predistribution policies, for example educational interventions, also face technical and incentive issues. Equality of opportunity arguments which are the moral philosophical basis for the supremacy of predistribution, can be challenged on conceptual, empirical and policy grounds. And popular views on redistribution do not entirely eschew direct forms of income support to the very poorest. In any event, as Haaparanta et. al. (2022) argue, in a general framework with parental investment in children's education, redistribution may be a key instrument for achieving a more equal predistribution.

However, such back and forth between the normative strengths of redistribution and predistribution is not the focus of this paper. Rather, we set out to quantify the consequences of predistribution for redistribution as seen through a particular lens. We imagine two societies starting with different distributions of income earning capabilities—one more equal than the other. In each of these, governments implement tax and transfer regimes to advance the same egalitarian objective by maximizing a social welfare function, taking into account the incentive effects of taxation. We can now consider the differences in outcomes in these two societies as a comparative static exercise. How does the optimal tax regime and the extent of redistribution differ across the societies with more unequal and less unequal predistribution? And what is the value of greater equality in predistribution, as measured by its consequences for egalitarian social welfare?

We answer these questions in the framework of the Mirrlees (1971) optimal income taxation model with a distribution of labour productivity (the predistribution), individual labour supply (the incentive effects) and non-linear income taxation (the redistribution) to maximize a social welfare function (the egalitarian objective). We change the distribution of productivity in comparative static fashion and assess the consequences for and value of optimal redistribution policies. We quantify the worth of a more equal predistribution through a money metric measure of the change in maximized social welfare.

The plan of the paper is as follows. In section 2, we set up the basic Mirrlees (1971) model and highlight the role of pretax inequality in determining the shape of the optimal non-linear tax schedule. However, analytical characterisation has its limits. In section 3, we consider how to specify the model to solve it numerically. In section 4 using numerical simulations we quantify the consequences for optimal income tax schedules and calculate the worth of a more equal predistribution. Section 5 concludes.

2. The Mirrlees model and the productivity distribution

In the Mirrlees (1971) model there is pre-tax inequality because individuals differ in their labour productivities. The government chooses a nonlinear income tax and transfer schedule to maximize a social welfare function which is in principle sensitive to inequality but does so with the added constraint that individuals choose their labour supply in response to the tax/transfer function. The government must also satisfy the overall budget balance constraint, with tax revenues equal to outlays. Note that we view this model as a stylized representation of the myriad instruments the government has at its disposal. Each instrument has incentive, revenue and distributional effects. In this framework, we use numerical simulations to study how optimal taxation for redistribution might respond when pre-tax inequality is less or more unequal and what the value of lower pre-tax inequality is in terms of its impact on social welfare after optimal redistribution.

It is useful to lay out the basic model, even though it is well known.[1] There are a continuum of taxpayers, each having the same preference ordering, which is represented by a utility function $u = U(x) + V(1-y)$ defined over consumption x and hours worked y, with $U_x > 0$ and $V_y < 0$ (subscripts indicating partial derivatives). Individuals differ only in the pre-tax wage n they can earn (their productivity). There is a distribution of n on the interval $(0,\infty)$ represented by the density function f(n). Gross income is z = ny. The government cannot observe individuals' productivities and thus is restricted to setting taxes and transfers as a function only of earnings, $T(z(n))$.

The optimal income tax problem is that of finding a function $x = c(z) = ny - T(ny)$ that maximizes a social welfare function

$$S = \int_0^\infty W(u(n)) f(n) dn \tag{1}$$

subject to the resource constraint

$$\int_0^\infty (z-x) f(n) dn = R \tag{2}$$

where in the Mirrlees tradition R is interpreted as the required revenue for publicly provided goods. In addition to (2), the government faces incentive compatibility requirement that each n individual chooses x(n) and z(n) to maximize utility. Provided a single-crossing property holds, for

1 See for example Kanbur-Tuomala (1994), Tuomala (2016)

every n individual incentive compatibility is equivalent to two constraints, one that z is non-decreasing in n, and the other that

$$\frac{du}{dn} = u_n(x,z,n). \tag{3}$$

In other words, the optimal income tax problem is that of finding a function $c(z) = ny - T(ny)$ that maximizes (1) subject to the constraints (2) and (3).

From the first order conditions of government's maximization, we obtain the condition for optimal marginal tax rate. The details of the analysis are shown in the appendix A. Here we skip directly to the result. It is useful to write the ABC-formula[2] for marginal rates,[3] denoted by $t(z) = T'(z)$, in terms of traditional labour supply elasticities, E^u and E^c; [4,5]

$$\frac{t}{1-t} = \underbrace{\left[\frac{1+E^u}{E^c}\right]}_{A_n} \underbrace{\left[\frac{[1-F(n)]}{nf(n)}\right]}_{B_n} \underbrace{\left[\frac{U_x \int_n^\infty [1 - W' U_x^{(p)}/\lambda] f(p)dp}{(1-F(n))}\right]}_{C_n}, \tag{4}$$

where E^u is the uncompensated elasticity of labour supply and E^c in turn is the compensated elasticity (see appendix for derivation of elasticities).

Applying (4) it appears that there are three elements on the right-hand side of (4) that determine optimum tax rates at a given skill level: elasticity and income effects (A&C), the shape of the skill distribution (B&C) and social marginal valuation of consumption (C). B is a measure of the relative number of taxpayers at n level and above it. The C-term in (4) is a measure of the social cost of taking an euro away from everyone above that skill. C tends to favour rising marginal rates. This is

[2] As in Diamond (1998)

[3] There is an important difference between (4) and the formulation in Saez (2001, p. 215). Taking into account that the tax rules and other parameters will determine the relationship between the ability distribution and the resulting income distribution, Saez (2001) translates the results from ones in terms of the distribution of abilities to ones in terms of the distribution of income. In other words, Saez made a step forward by deriving an optimal tax formula by expressing his optimal tax formula in terms of the notion of virtual earnings distribution and verifies the consistency of his solution to the Mirrlees one. Saez (2001, p. 215) defines the virtual density at earnings level z as *'the density of incomes that would take place at z if the tax schedule T (:) were replaced by the linear tax schedule tangent to T (:) at level z'*.

[4] Revesz (1989), Atkinson (1995), Diamond (1998), Piketty (1997), and Saez (2001) formulated the Mirrlees first order condition in terms of elasticities. See also Wilson (1993) in the context of nonlinear pricing.

[5] The marginal tax rate in the formula (4) is depicted as $\frac{t}{1-t}$ instead of t. The reason is that the tax here is applied to the tax-inclusive tax base or after-tax income.

particularly so when income is low or moderate.

Let us turn now to our central question, how does the predistribution affect optimal redistribution? Formally, how does a shift in the distribution of n affect the pattern of optimal redistributive taxation as shown in (4)? It is immediately seen from (4) that this is not an easy question to answer cleanly. The distribution of n enters each of the terms A, B and C and there are complex interactions across these terms. Given this complexity there are two possible approaches in the literature. One is to draw plausible intuitions by focusing on each term separately and making further simplifying assumptions. The second is to address the complex interactions through numerical calculations. In this section we will follow the first route. Subsequent sections will take up the numerical route.

The second term B is the one which depends solely on the distribution of n and we will focus on that to develop some plausible intuitions on the consequences of changing the predistribution. B tells that the shape of the predistribution affects the optimal marginal tax rate at the productivity level n through the ratio $\frac{1-F(n)}{nf(n)}$. This ratio is the inverse hazard rate (or the Mills' ratio) divided by n. When we increase the marginal tax rate at some low n, we collect more revenue on more productive individuals, who are $1-F(n)$ in number. In other words the revenue effect of an increase in marginal rate at n depends on the proportion of the population above n. The purpose of a higher marginal rate is to increase the average tax rate higher up the scale. Hence $1-F(n)$ is in the numerator. And we distort only the behaviour of the marginal type, which explains why $f(n)$ in turn is in the denominator.

Thus a shift in the distribution of n which increases (decreases) $\frac{1-F(n)}{nf(n)}$ at a particular n will, holding terms A and C constant, increase (decrease) the marginal tax rate at that n. The change that we are particularly interested in is an inequality decreasing shift, which is what predistribtuion policies are supposed to bring about (note again that we are here conducting a comparative static exercise—we do not ask how exactly these changes are effected).

As an illustration, let the distribution F(n) be Pareto with minimum n given by n_o and the Pareto inequality coefficient denoted α:

$F(n) = 1- (n_o/n)^\alpha$

Then the term B is simply α and an increase in the Pareto inequality coefficient will raise B and through this channel increase marginal tax rates. However, while analytically sharp there are three issues with this argument. First, terms A and C will also in general change unless further assumptions are made. Second, a change in α will also change the mean of the distribution of n so there will be a "size of the pie" effect which may confound simple intuitions on inequality. Third, as is well known, the shape of the Pareto distribution does not match observed income distributions, since these typically have a "hump" shape with a mode away from the minimum value of the support.

The feature of the Pareto distribution noted above is one argument in the literature for use of the lognormal distribution and indeed this is what is done in the classic Mirrlees (1971) contribution. However, observed distributions are leptokurtic with more weight in the upper tail than shown by a lognormal distribution. They are in fact approximated better by a Pareto distribution for high incomes. As argued further in the empirical discussion in Section 4 below, a distribution which captures desirable features of the lognormal and Pareto distributions is the Champernowne distribution. The probability density function and the cumulative distribution function of the Champernowne distribution are

$$f(n) = \theta\left(\frac{r^\theta n^{\theta-1}}{(r^\theta + n^\theta)^2}\right) \qquad F(n) = 1 - \frac{r^\theta}{(r^\theta + n^\theta)} \tag{5}$$

where θ is a shape or Pareto parameter (corresponding to 1/α) and r is the median n. The ratio

$$\lim_{n\to\infty} \frac{1-F(n)}{nf(n)} = \lim_{n\to\infty} \frac{r^\theta + n^\theta}{\theta n^\theta} \to \frac{1}{\theta} \tag{6}$$

(6) confirms that the Champernowne distribution approaches asymptotically a form of Pareto distribution for large values of n.

For the Champernowne distribution

$$\frac{1-F(n)}{nf(n)} = \frac{r^\theta + n^\theta}{\theta n^\theta} \text{ and } \frac{d}{dn}\frac{r^\theta + n^\theta}{\theta n^\theta} = \frac{-r^\theta}{n^{\theta+1}} < 0 \tag{7}$$

Unlike with the lognormal n-distribution, when the median wage changes and the θ parameter does not change, the inverse hazard rates do not cross with the Champernowne-distribution. (see figure b1 and b2 in appendix B). Hence the formula (4) implies that the Champernowne distribution with a higher median, all else equal, will tend to have a higher B term, and consequently higher marginal tax rates. At the same time, the C-term tends to favour rising marginal rates, especially when the revenue requirement increases (the shadow value of revenue increase) and when income is low or moderate.

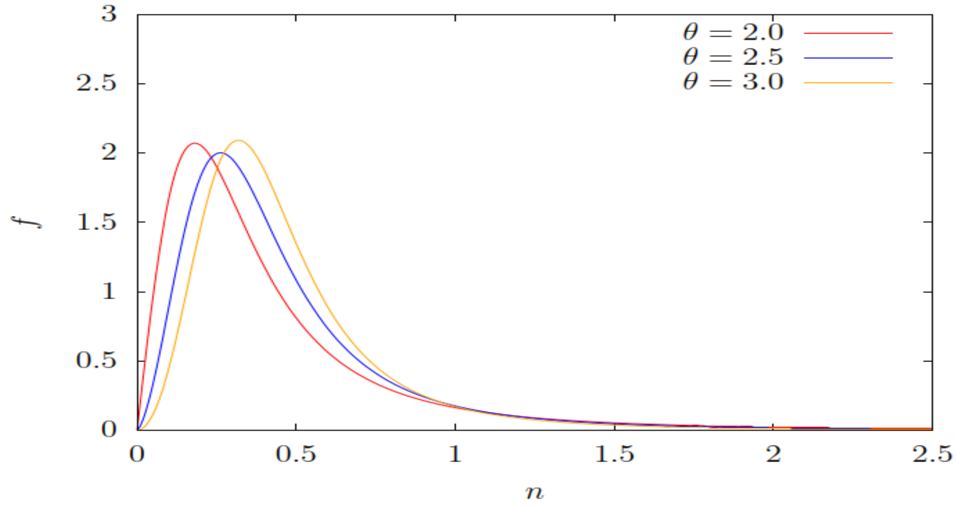

Figure 1 Champernowne density function: In all cases mean n = 0.48, (θ=2, median n =0.31, k=-1.16), (θ=2,5, median n=0.37, k=-1.0), (θ=3, median n=0.40, k=-0.91).

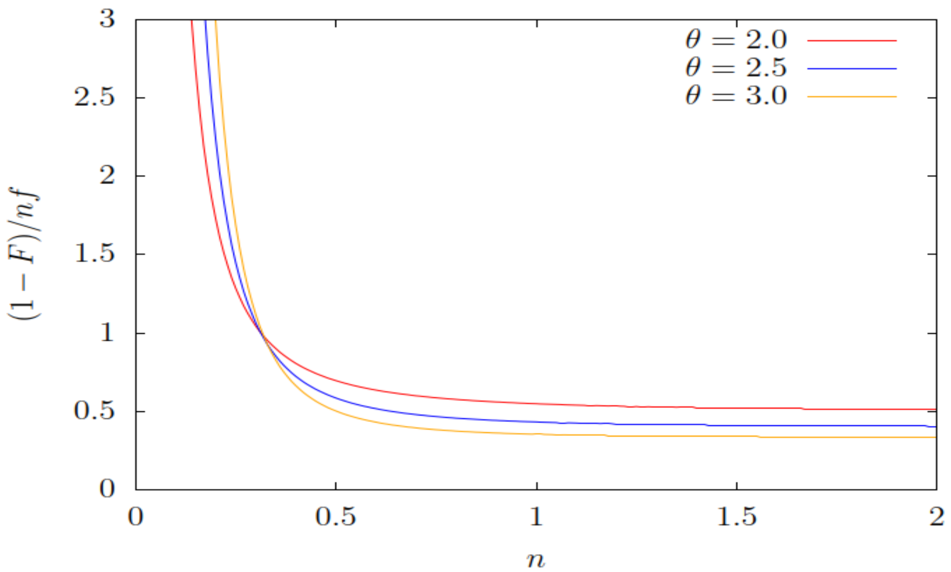

Figure 2 Inverse hazard rate: In all cases mean n = 0.48, (θ=2, median n =0.31, k=-1.16). (θ=2,5, median n=0.37, k=-1.0), (θ=3, median n=0.40, k=-0.91).

Analogously to the Pareto distribution, 1/θ tracks inequality in the Champernowne distribution. An increase in θ lowers inequality while leaving the median unchanged. It can be shown that the derivative of the B term with respect to θ is positive or negative according as $\left[c^{\theta}(\theta \log c - 1) - 1\right]$ is positive or negative, where c=r/n. Consider the case in which n is greater than median ability r (see

Figure 2) so c less than 1. Then log c is negative and the expression is definitely negative. Thus, an increase in θ, i.e. a decrease in productivity inequality, leads to a lower B term and a lower marginal tax rate on above median n. Conversely, higher productivity inequality leads to a higher optimal marginal income tax rate on above median productivity individuals through the B term. By continuity this result will hold for incomes some way below the median as well, with the signs being reversed for incomes well below the median. A more equal predistribution leads, through this particular channel, to lower marginal tax rates on the rich and higher marginal tax rates on the poor. Simply put, a more equal predistribution allows a less progressive tax and transfer system in the redistribution.

While the above characterization through term B comports with intuition, there remain the issues highlighted earlier. First, this is an argument only through term B. How terms A and C will interact and play into the final calculation is not possible to lay out analytically. Second, a change in θ will change the mean of the distribution. To address the "size of the pie" effect we would need to hold mean constant as θ changes, through a corresponding change in the r parameter.

Figures 1 and 2 illustrate the possible cases of the density functions and the B term for the Champernowne distribution with different θ parameters but constant mean of the n-distribution. For ease of presentation we use not r but k = -log r in the figures and in the discussion that follows[6]. The mean of the distribution is denoted m. We take as our base case θ = 2.5 and k = -1.0 (for empirical justification see discussion in the next section). This gives a mean m of 0.48, which is held constant for our scenarios of θ = 2.0 and k = -1.16, and θ = 3.0 and k = -0.91. Figure 1 shows that as the n-distribution becomes less unequal (θ parameter increases) the mode of the Champernowne distribution shifts to the right but the upper parts of the distributions are almost identical. Figure 2 depicts the term B with different θ parameters and the same mean value m = 0.48. The pattern described earlier is seen to be maintained in the constant mean case. The B term and thus marginal tax rates are higher for high incomes and lower for low income as θ falls ie as the predistribution becomes more unequal. Conversely, again in a comparative static sense, as the predistribution becomes less unequal marginal tax rates on higher productivities rise and those on lower productivities fall through the channel of the B term. But of course, this is only one of the three terms and we cannot arrive at a definitive overall assessment in an analytical sense.

---

6 $r = e^{-k}$

This is about as far as we can get at this level of generality[7]. In the tradition of the non-linear taxation literature, we can provide better understanding of the form of optimal policy through numerical simulations[8]. We can compute post tax income at each level of n and thus calculate inequality of pre and post tax income as well as total income, for different values of key parameters. Sections 3 and 4 take up this task of elaborating further the consequences of lower inequality in the predistribution of productivities.

3. Model Specification for Numerical Solution

There are four key elements in the optimal nonlinear income tax model we have to specify to solve it numerically. We consider each of these in turn.

The first factor is the pre-tax productivity (skill, wage) distribution. Champernowne (1952) proposed a model in which individual incomes were assumed to follow a random walk in the logarithmic scale leading to what is now known as the Champernowne distribution. Here we use the two parameter version of the Champernowne distribution. As already discussed in the previous section, this distribution approaches asymptotically a form of Pareto distribution for large values of n but it also has an interior maximum. The Champernowne distribution exhibits the following features: asymmetry, a left humpback and long right-hand tail, but it has a thicker upper tail than in the lognormal case and thus fits better observed income distributions.

Based on Finnish income distribution data Riihelä et al (2014) estimated by using maximum likelihood methods several two and three parameter distributions with corresponding measures of goodness of fit (several of them plus the log-likelihood value for estimated model). Among two parameter distributions the Champernowne is the best fitting for pre-tax earned income distribution in Finland (2002-2010). The θ-parameter varies from 2.78 to 2.4. Over the period from the latter part of 1990's to 2010 the θ-parameter was almost constant being around 2.5. Hence $\theta = 2$ reflects a low range estimate (high inequality) and $\theta = 3$ in turn a high range estimate (low inequality). The Gini coefficients estimated by this distribution (Gini=$1/\theta$) are quite close to those calculated from the data.

[7] There are also a number of asymptotic results for the unbounded case. For example, Mirrlees (1971) demonstrates that the marginal tax rate can converge to 100% under some conditions. See also Diamond (1998), Saez (2001) and Dahan, M., and M. Strawczynski(2012).
[8] Tuomala (2016) gives details of the computational procedure.

But of course the distribution of earnings is not the distribution of n--it is the distribution of z. This is a general issue with much of the numerical literature on optimal income taxation. While Kanbur-Tuomala (1994) calibrate the productivity distribution indirectly so that the income distribution inferred from the wage distribution matches the actual distribution, typically empirical estimates of the z distribution are simply applied to the n distribution, including in the original Mirrlees (1971) paper.

The second factor: Following earlier work (Kanbur-Tuomala (1994), Tuomala (2016)) we use in simulations the following special case of the constant elasticity of substitution form

$$u = -\frac{1}{x} - \frac{1}{(1-y)} \tag{8}$$

where the elasticity of substitution, denoted by ε, between consumption, denoted by x, and leisure, denoted by (l-y), ε= 0.5. In the absence of taxation x=z=ny, the labour supply function y(n) is backward bending.

The third factor is the social welfare function (SWF). 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. As in Mirrlees (1971), in the social welfare function (1) W(u) takes the form of

$$W(u) = -\frac{1}{\beta} e^{-\beta u} \quad , \tag{9}$$

where β measures the degree of inequality aversion in the social welfare function of the government (in the case of $\beta = 0$, we define $W = u$).[9] If we write $S^{-\beta} = \int e^{-\beta u} f(n) dn$ then the limit as $\beta \to \infty$ is given by $W = \min_n \left[e^u \right]$.The curvature in the utility of consumption modifies marginal social welfare weights $W'U_x$ and makes the government preferences more redistributive. $W'U_x$ is typically called the marginal social welfare weight of an individual wage or earning income 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 $W'U_x$ is made up of two components: the increase in the individual's well-

[9] (11) is sometimes called the Kolm-Pollak SWF. Rather, it should be called the Kolm-Mirrlees-Pollak (KMP) SWF. It is invariant to a translation rescaling of utility, and indeed is the only prioritarian SWF with this invariance feature. If u(.) is a well-being measure, and u*(.) = u(.) + d, then a KMP SWF is such that the ranking of outcomes is the same using u(.) and u*(.). By contrast with the Atkinson SWFs, KMP SWFs do not require well-being numbers to be non-negative.

being due to an increase in consumption U_x and the increase in social welfare that arises from an increase in that individual's well-being W'. The distinction between the two components of $W'U_x$ is important to understanding how prioritarianism differs from alternative normative principles for optimal tax theory. In a prioritarian objective[10] the second component of the $W'U_x$ 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. In formal terms, $\varphi(n) = \frac{W'\mathrm{U}_x}{\lambda}$ decreases with n in a prioritarian objective even if U_x is constant because W' is decreasing in U and thus n. In 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 non-decreasing 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 a $\varphi(n)$ in expression (4) that declines more rapidly with n, 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.

The fourth factor is the government's revenue requirement specified as a fraction of total income not used for private consumption by individuals, R=1-X/Z=$1-\int x(n)f(n)dn / \int z(n)f(n)dn$. If X/Z=1, then taxation is purely redistributive. R can be interpreted as the required revenue for essential public provided goods such infrastructure, education, health and R&D. All these are important tools for pre-distributive policies. Its size affects the cost of raising revenue to these predistributive purposes and fiscal redistribution to fund transfers to individuals[11].

4. Numerical results

We compare the productivity (wage) distribution of the base case with less and more unequal distributions in such a way that the mean of the n-distribution and the revenue requirement remains the same. The mean of n denoted by m for the base case (k = -1 and θ = 2.5) is m= 0.48. For k = -1

[10] See Tuomala-Weinzierl (2021)

[11] In this context it may be of some interest to notice an approximation result (Tuomala 1984) concerning the relationship between the marginal tax rates at the higher pre-tax income levels and the tax revenue. With an additive utility function in consumption and labour supply and utilitarian case the marginal tax rate for an individual with high n is independent of λ or revenue.

and $\theta = 3$, m = 0.44. Hence, higher θ (less inequality) leads to lower mean. We need to find the k for which with $\theta = 3$, m =0.48. Using numerical integration, we find that the value of k.= -0.91 produces a mean value of 0.48. It is for this (k, θ) pair that we calculate the optimal tax schedule and the money metric equivalent of social welfare, to compare with the base case. This will give us the worth of a more equal predistribution. We also do this for a pair with more inequality (lower θ) than in the base case. The (k=-1.16, $\theta = 2$) pair gives the same mean for n as in the base case. Throughout, the revenue requirement is kept at 10% (R = 0.1).

The output of our numerical simulations is presented in Tables 1, 2, 3 and 4, and in Figures 3a, 3b, 4a and 4b. We briefly set out the nature of this output below, before discussing its implications in detail. Tables 1, 2 and 3 display the optimal tax structures for our three cases and for three social welfare functions capturing different degrees of inequality aversion in each Table. Marginal and average tax rates are shown at different points along the productivity distribution. Also shown are the absolute values of the utility levels at these points. The information in Tables 1, 2 and 3 is partially presented graphically in Figures 3 and 4. In both Figures we have set R=0.1, mean n=0.48. Figures 3a and 3b show MTRs and ATRs with β=0 for different values of θ. Figures 4a and 4b show MTRs and ATRs $\theta = 3$ for different values of β. All the calculated tax schedules take the form of a lump sum income, denoted by x_o, followed by tax rates. Note that the marginal (MTR) and average tax rates (ATR) are for all taxes that vary with income and should be compared with the schedules for total of taxes on income and expenditures in real economies.

The range of information we present in our numerical simulations is completed by Table 4. The first row of the Table presents total post tax income (pre tax income can be calculated simply by the 10% revenue requirement). The second row is the lump sum payment component of the optimal tax schedule. The third and fourth rows are, respectively, the ratio of income at its 90th percentile to income at its 50th percentile (P90/P50) for pre-tax (z) and post-tax (x) income distributions. The fifth row, RD, is the percentage decline in the P90/P50 between pre-tax and post-tax income, as a measure of the degree of redistribution brought about by the tax regime.

The sixth row of Table 4 presents a money metric measure of the level of welfare for the particular case being considered. To assess this, we compute the level of consumption which, if equally distributed at zero work hours[12], would generate that same level of social welfare. This is denoted by

[12] Stern (1976) and (1982)

x^0. The seventh and final row of Table 4 shows this money metric measure as a ratio of per capita consumption.

What, then, are the consequences of a more equal distribution of productivity for the pattern of optimal taxation and what is the social value of this more equal pre-distribution? As seen in Tables 1-3 and Figures 3-4, optimal marginal and average tax rates fall throughout when the distribution of productivity is less unequal (holding average productivity constant). A more equal pre-distribution means there is less of a burden of redistribution, as shown by the tax rates and by the measure of redistribution RD in Table 4, which falls as θ increases from 2.0 to 2.5 to 3.0. This is so for utilitarian, prioritarian and maximin cases.

The optimum is typically characterized by a certain fraction of individuals, at the bottom end, choosing not to work. This means that there is bunching of individuals of different n. The value of n_o (an individual with no income) is considerably higher in the Maximin case with 8-16 % of the population choosing not to work. This may be contrasted with conclusion of Mirrlees (1971) that in the utilitarian case "virtually everyone is brought into the work force".

The extent of bunching is shown in the Tables 1-3 for the Maximin case with different θ in each of the Tables. It is seen that bunching increases with inequality. However, the Maximin case differs significantly from utilitarian and prioritarian ones in the number of workers who do not work in the optimum. When θ=2.0, in the Maximin case $F(n_o)$=0.16, in the utilitarian case $F(n_o)$=0.01 and in the prioritarian case $F(n_o)$=0.03. Hence the combination of high inequality aversion (β) and high pre-tax inequality (θ) determines the amount of bunching. Why is this so? The explanation is related to another seemingly surprising finding in optimal non-linear tax models, that the Maximin objective does not support increasing marginal tax rates, and it is a consequence of the fact this objective is not concerned with distribution among those other than the least advantaged individual. As we can see from tables 1-3, the utilities of these are lower in the Maximin optimum than in utilitarian and prioritarian ones. When inequality is small, θ=3.0, practically speaking bunching does not occur for utilitarian and prioritarian cases.

Is the pre-tax inequality (θ) more important than redistributional preferences (β) in determining the extent of redistribution? It turns out that increasing β has a very modest effect on the extent of redistribution. This is true for different values of θ. Hence, the extent of redistribution RD is not very different in the β=1 case and in the maximin case. It is also important to note that the extent of redistribution and rising marginal tax rate within a case may be two quite different things. The extent of redistribution is larger with β=1 than in the case of β =0, but the marginal tax schedule may be rising in the latter case (see Figures 3 and 4). How sensitive is the level of the guaranteed income or lump sum transfer component of the tax system in different cases? And what is the relationship between the lump sum transfer x_o and the progressivity of tax schedule? Table 4 displays the ratio of the guaranteed income to the average net income (x) with social welfare functions (utilitarian, prioritarian and maximin), and revenue requirement, R (=1-X/Z) =0.1. This ratio is clearly higher in the Maximin case than in the utilitarian and prioritarian ones.

What is the social value of a more equal distribution of productivities? From Table 4 we see that the total consumption X (and thus total gross income Z) is higher with greater equality of the predistribution. At least in this sense, equality is overall beneficial. From Tables 1, 2 and 3 we see that utility is higher with lower inequality at all levels of the productivity distribution except at the very top. But we also have an aggregative measure of the money metric value of the social welfare function under the different scenarios, as present by x^o in row 6 of Table 4. As θ increases from 2.0 to 2.5 to 3.0 (ie as productivity inequality decreases), x^o increases from 0.159 to 0.163 to 0.168 for the Utilitarian social welfare function (β=0) and from 0.160 to 0.162 to 0.167 for the Prioritarian case (β=1). Thus, in these cases aggregate social welfare so measured increases, and we have a money metric value of this gain.

However, for the Maximin social welfare function (β=∞), x^o goes from 0.188 to 0.181 to 0.180 as inequality decreases. Note that when inequality decreases from 2.5 to 3.0 there is not much change. But overall x^o increases when θ decreases, i.e. productivity inequality increases. Our intuition is that the bunching discussed earlier is a key explanation for this. Namely, the number of employees who do not work at the optimum increases as θ decreases. A significant number of people receive a guaranteed income (consumption) x_0, which increases as θ decreases. In order to raise the revenue for this, high productivity individuals need to pay the tax. The two forces go hand in hand to produce the seemingly paradoxical result that a more equal predistribution is not favored by the most egalitarian objective function.

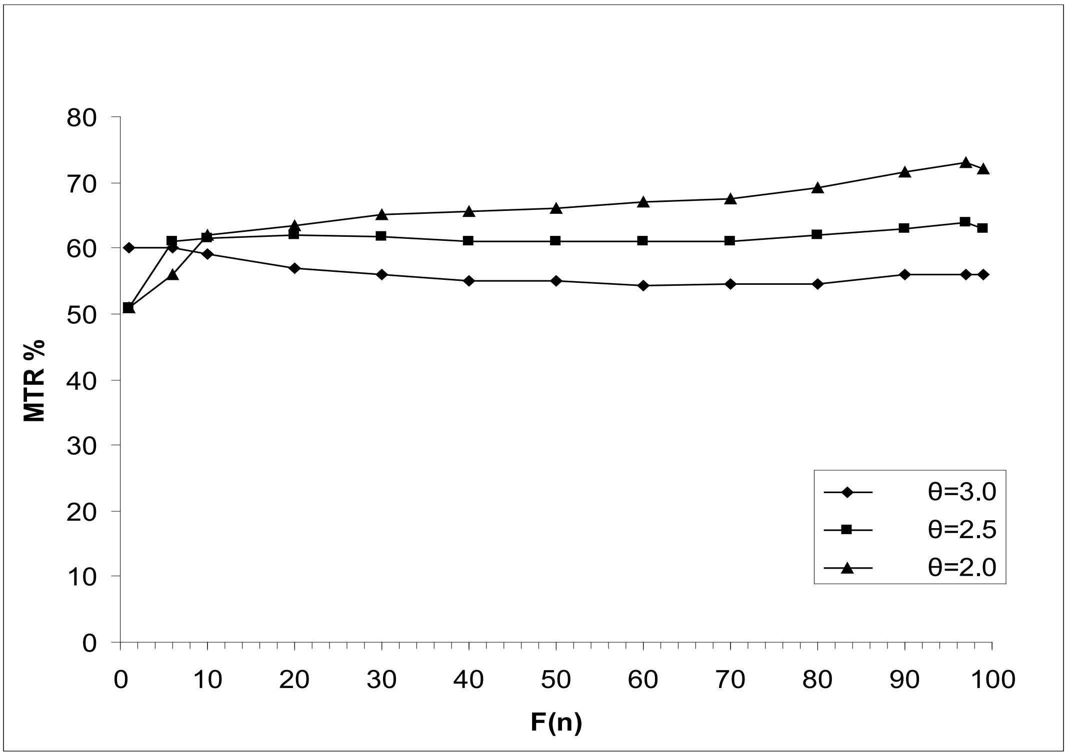

Figure 3a MTRs with different θ, β=0, R=0.1, mean n=0.48

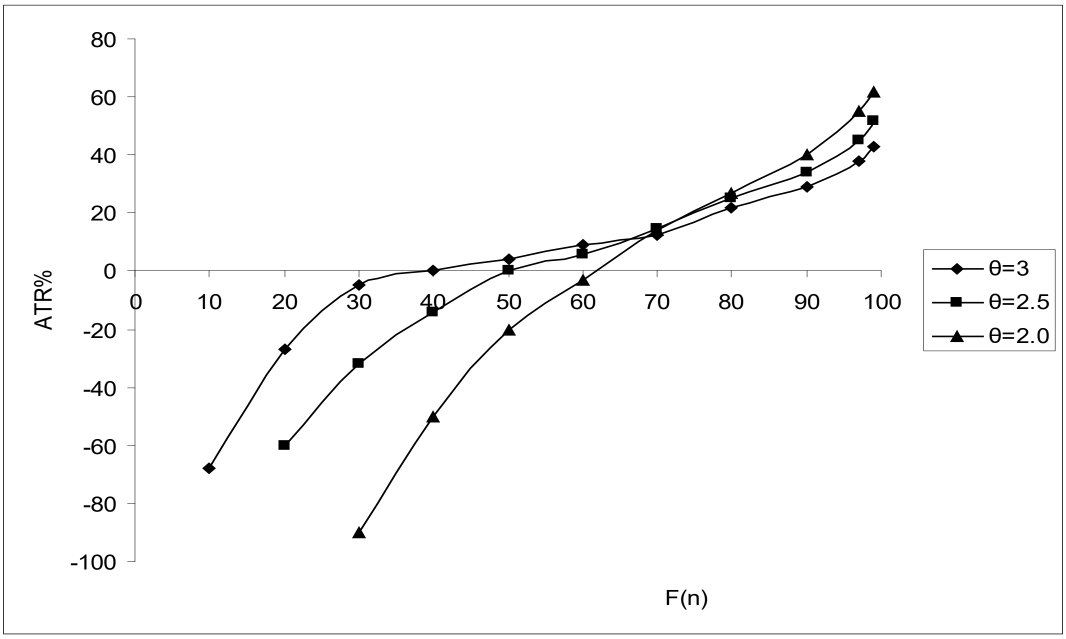

Figure 3b ATRs with different θ; β=0, θ=3, R=0.1, mean n=0.48

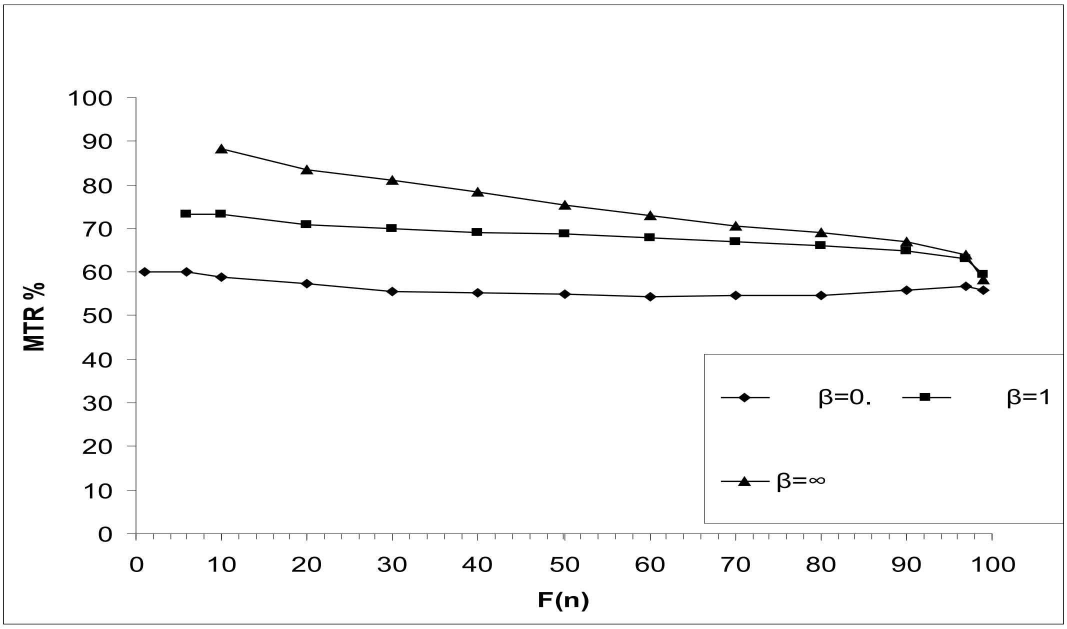

Figure 4a MTRs with different β; θ=3, R=0.1, mean n=0.48

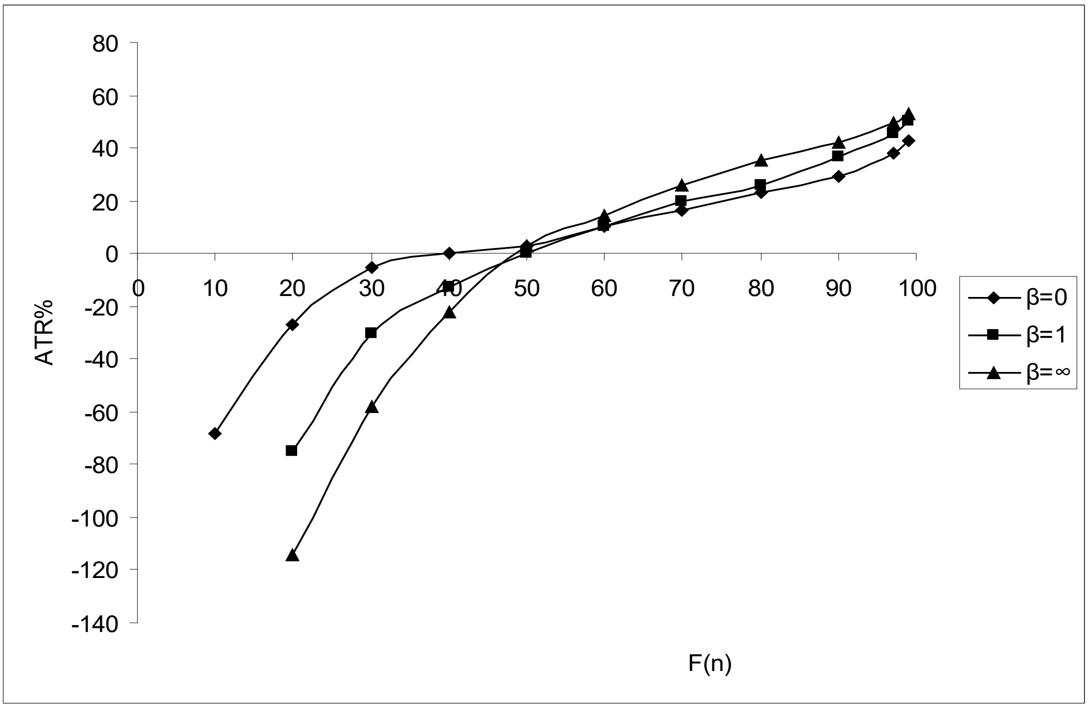

Figure 4b ATRs with different β; θ=3, R=0.1, mean n=0.48

Table 1 Base case θ=2.5, mean n =0.48, median=0.37, k=-1.0, ε= 0.5, R=0.1

	Utilitarian β=0			Prioritarian β=1			Maximin β=∞ *F(n)=0.12		
F(n)	MTR%	ATR%	Utility	MTR%	ATR%	Utility	MTR%	ATR%	Utility
0.10	62	-130	-8.67	74	-200	-8.34	88	-	-8.12*
0.50	61	-3	-7.25	73	-7	-7.49	79	-9	-7.75
0.90	63	35	-5.46	69	42	-5.89	72	46	-6.20
0.99	64	52	-3.81	56	54	-3.92	64	60	-4.32

Table 2 θ=3.0, mean n =0.48, median n = 0.40, k=-0.91,ε= 0.5, R=0.1

	Utilitarian β=0			Prioritarian β=1			Maximin β=∞ *F(n)=0.08		
F(n)	MTR%	ATR%	utility	MTR%	ATR%	utility	MTR%	ATR%	utility
0.10	59	-68	-8.50	73	-100	-8.05	88		-8.03*
0.50	55	3	-7.04	69	0	-7.19	76	2	-7.55
0.90	56	29	-5.38	65	37	-5.77	67	42	-6,10
0.99	56	43	-3.92	59	50	-4.20	58	53	-4.43

Table 3 θ=2.0, k=-1.16, mean n =0.48, ε= 0.5, R=0.1

	Utilitarian β=0			Prioritarian β=1			Maximin β=∞ *F(n)=0.16		
F(n)	MTR%	ATR%	utility	MTR%	ATR%	utility	MTR%	ATR%	utility
0.10	62	-	-8.84	72	-	-8.46	86	-	-8.37*
0.50	66	-21	-7.58	76	-32	-7.73	81	-37	-8.05
0.90	72	40	-5.65	77	47	-6.10	78	51	-6.41
0.99	72	62	-3.78	74	67	-4.09	71	67	-4.24

Table 4 In all cases mean n =0.48, ε= 0.5, R=0.1

	θ=2.5 k=-1.0	θ=2.5 k=-1.0	θ=2.5 k=-1.0	θ=3.0, k=-0.91	θ=3.0, k=-0.91	θ=3.0, k=-0.91	θ=2.0, k=-1.16	θ=2.0, k=-1.16	θ=2.0, k=-1.16
	β=0	β=1	β=∞	β=0	β=1	β=∞	β=0	β=1	β=∞
X	0.21	0.19	0.18	0.22	0.20	0.19	0.20	0.18	0.17
x_0	0.105	0.132	0.139	0.115	0.135	0.142	0.122	0.132	0.136
P90/P50(z)	2.44	2.70	2.96	1.95	2.21	2.39	3.00	3.46	4.18
P90/P50(x)	1.53	1.42	1.46	1.47	1.42	1.47	1.58	1.50	1.53
RD	34	48	51	25	32.	39	51	59	64
x^0	0.163	0.162	0.181	0.168	0.167	0.180	0.159	0.160	0.188
$\frac{x^0}{X}$	0.78	0.85	0.99	0.76	0.83	0.95	0.80	0.84	1.04

An individual with no income would get the lump sum subsidy $x(n_o) = x_o = 0 - T(z=0) = -T(0)$.

5. Concluding remarks

A major debate is under way on the normative superiority or otherwise of policies for redistribution of market incomes versus those that equalize the pre-market distribution of income earning capabilities or productivities, in other words the predistribution. This paper does not take a position on this debate, but rather presents an analysis of the consequences of achieving a more equal predistribution for redistribution policies. It does this through the lens of the classic Mirrlees (1971) model of optimal non-linear income taxation. In this model individuals bring a pre-determined productivity (ability) to the market which translates into their wage. They face a government tax schedule and maximize utility through labour-leisure choice. The government in turn maximizes a social welfare function subject to a revenue constraint.

We first conduct a theoretical exercise by examining in detail the formula for optimal marginal tax rates to maximize egalitarian social welfare functions. We compare these tax rates across distributions of productivities which have the same mean but differ in inequality. While the formulae are complex with myriad interactions across different components of the optimization problem, we provide some intuition that a more equal predistribution leads to lower marginal tax rates on the rich and higher rates on the poor. In other words, equalizing predistribution eases the burden on fiscal redistribution.

However, as is common in the optimal non-linear income taxation literature, such analytical exercises can only get us so far. Numerical solutions to the government's optimization problem are needed to provide further insight. We show that the theoretical intuitions are confirmed by numerical calculations, fiscal policies are (need to be) less redistributive when the predistribution is more equal. We then quantify the value of a more equal predistribution. First we show that average consumption is higher with a more equal predistribution. We develop a money metric measure of social welfare and show that a more equal predistribution increases this measure for a Utilitarian and Prioritarian Social Welfare Function. However, perhaps surprising but ultimately intuitive, a Maximin objective function favours a more unequal predistribution. This last result in particular, but more generally the social value of predistribution, will bear further exploration and enquiry.

References

Atkinson, A. B., (1996). Public Economics in Action: The Basic Income/Flat Tax Proposal, OUP Catalogue, Oxford University Press.

BBC (2012), "Ed Miliband unveils 'predistribution' plan to fix economy", 6 September.

Champernowne, D. G. (1937), Notes on Income Distribution, Econometrica, Report of Econometric Conference at New College in 1936.

Dahan, M. and M. Strawczynski, (2012), Optimal Asymptotic Income Tax Rates, Journal of Public Economic Theory.

Diamond, P. (1998), Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, American Economic Review, 88(1), pp.83–95.

Ferreira, Francisco. 2023. "A London Consensus on Inequality?", presented at the London Consensus Conference, London School of Economics, 4-5 May 2023.

Haaparanta P., R. Kanbur, T. Paukkeri J.Pirttilä, and M. Tuomala, 2022. "Promoting education under distortionary taxation: Equality of opportunity versus welfarism, "Journal of Economic Inequality, Springer vol. 20(2), pages 281-297.

Hacker, J (2011), "The institutional foundations of middle-class democracy", Policy Network, 6 May.

Kanbur, R. (2023a). "Resisting the Drift away from Redistribution.", paper presented at conference on London Consensus, London School of Economics, 4-5 May. https://static1.squarespace.com/static/58cc36ed03596e3341b757ac/t/650ab591c6f4366660fe64d6/1695200657737/Resisting+The+Drift+Away+From+Income+Redistribution.pdf

Kanbur, R. (2023b). "Use With Extreme Caution", *Boston Review*, May. https://www.bostonreview.net/forum_response/use-with-extreme-caution/

Kanbur, R. and Tuomala, M. 1994, Inherent inequality and the optimal graduation of marginal tax rates, Scandinavian Journal of Economics 96 (2), pp. 275-282

Mirrlees, J.A. (1971), An exploration in the theory of optimum income taxation, Review of Economic Studies 38, 175–208.

Piketty, T. ,M. Guillot, B.Garbinti, J. Goupille-Lebret and A. Bozio (2022) Pre-distribution versus redistribution: Evidence from France and the US Vox EU Column, November.

Revesz, J. T. (1989), The optimal taxation of labour income, Public Finance/Finances Publiques,44, 433-475

Riihelä, M., R. Sullström and M. Tuomala (2014), Top Incomes and Top Tax Rates: Implications for Optimal Taxation of Top Incomes in Finland, Tampere Economic Working Papers, Net series no. 88, University of Tampere.

Saez, E. (2001), Using elasticities to derive optimal income tax rates, Review of Economic Studies 68, 205–29.

Stern, N. (1976). On the specification of optimum income taxation. Journal of Public Economics, 6, 123-62.

Stern, N, 1982, Optimum taxation with errors in administration, Journal of Public economics,17,181-211.

Stiglitz, J. E. (1982). Self-selection and Pareto-efficient taxation. Journal of Public Economics,17, 213–40.

Tanninen, H., M. Tuomala & E. Tuominen (2019), Inequality and optimal redistribution. Cambridge Elements, Public Economics, Cambridge University Press.

Tuomala, M., (1984), On the Optimal Income Taxation: Some further Numerical Results Journal of Public Economics, 23, 351–366.

Tuomala, M. (2016), Optimal Redistributive Taxation. Oxford, Oxford University Press.

Tuomala, M. and M. Weinzierl (2021), Prioritarianism and Optimal Taxation. in Prioritarianism in Practice, eds. Matthew Adler and Ole Norheim, Cambridge University Press.

Wilson, R., 1993, Nonlinear pricing, New York: Oxford University Press.

Appendix A

Incorporating the resource constraint (2) into the social welfare function (1), we have a control-theory problem to maximize

$$\int_0^\infty [W + \lambda(ny - x)) f(n) dn \qquad \text{(A1)}$$

where λ is the multiplier on the resource constraint, x is a function of u, z and n given by $u = U(x) + V(1-y)$ and subject to the differential equation (3). in the text. In addition we assume that z is non-decreasing. The optimality conditions for this problem are obtained by treating u as a state variable and y as a control variable and x is determined from the utility function. Now there must exist a dual function $\mu(n)$ such that

$$\frac{d\mu(n)}{dn} = \left[W' - \frac{\lambda}{U_x} \right] f(n) \qquad \text{(A2)}$$

and differentiating with respect to y

$$\lambda t f(n) = -\mu(n)(V_y + yV_{yy}) / n^2 \qquad \text{(A3)}$$

where the marginal tax rate is $t = (1 + \frac{V_y}{nU_x})$.

(A2) satisfies the transversality conditions

$$\frac{\partial L}{\partial u(0)} = \mu(0) = 0; \frac{\partial L}{\partial u(\infty)} = \mu(\infty) = 0$$

Integrating in (A2)

$$\mu(n) = \int_n^\infty [\frac{\lambda}{U_x} - W'] f(n') dn' \qquad \text{(A4)}$$

Using (4) we obtain from (3) the following condition for optimal marginal tax rate $t(z)$; [Note: $\frac{t}{1-t} = \frac{1}{1-t} - 1 = \frac{U_x n}{V_Y} - 1$]

Differentiating the FOC of the individual maximization, $U_x n(1-\tau) + V_l = 0$, with respect to net wage, labour supply and virtual income, b, we have after some manipulation elasticity formulas; $E^u = \frac{(V_l / l) - (V_l / U_x)^2 U_{xx}}{V_{ll} + (V_l / U_x)^2 U_{xx}}$, (income effect parameter) $I = \frac{-(V_l / U_x)^2 U_{xx}}{V_{ll} + (V_l / U_x)^2 U_{xx}}$, and from the Slutsky equation $E^c = E^u - I$, then $E^c = \dfrac{(V_l / y)}{V_{ll} + (V_l / U_x)^2 U_{xx}}$.

Using formulas for E^u and E^c, E^u is the uncompensated supply of labour and E^c in turn is the compensated elasticity, we obtain the formula (5) for optimal marginal tax rate in the text.

Appendix B

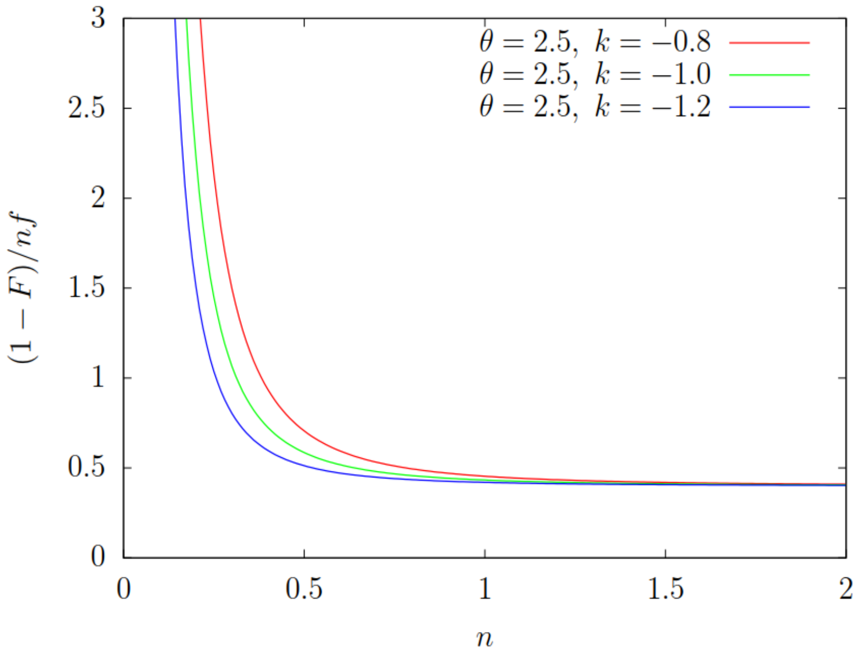

Figure b1 (1-F)/nf Champernowne distribution

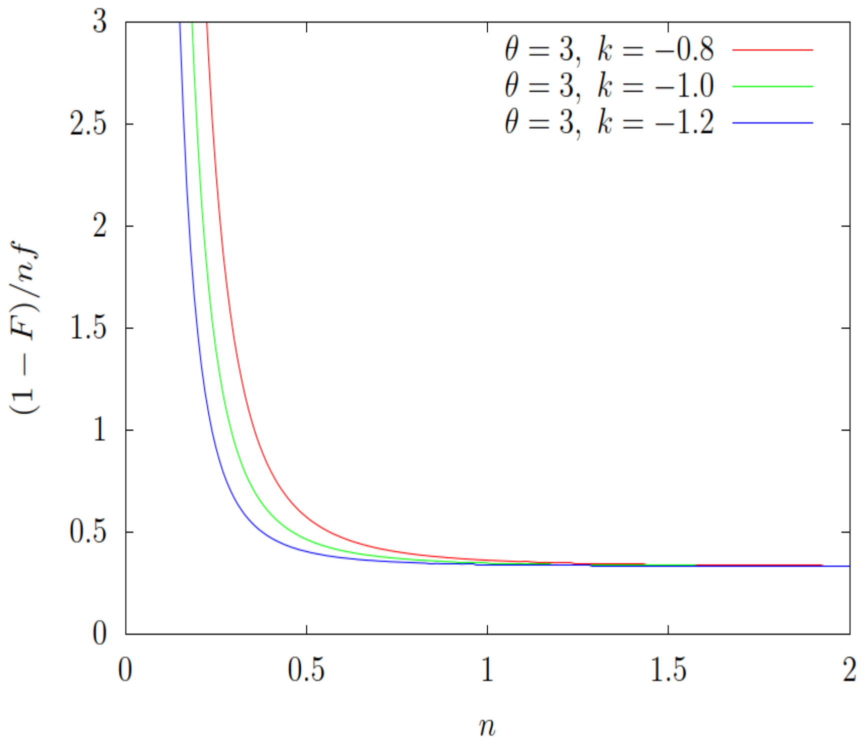

Figure b2 (1-F)/nf Champernowne distribution

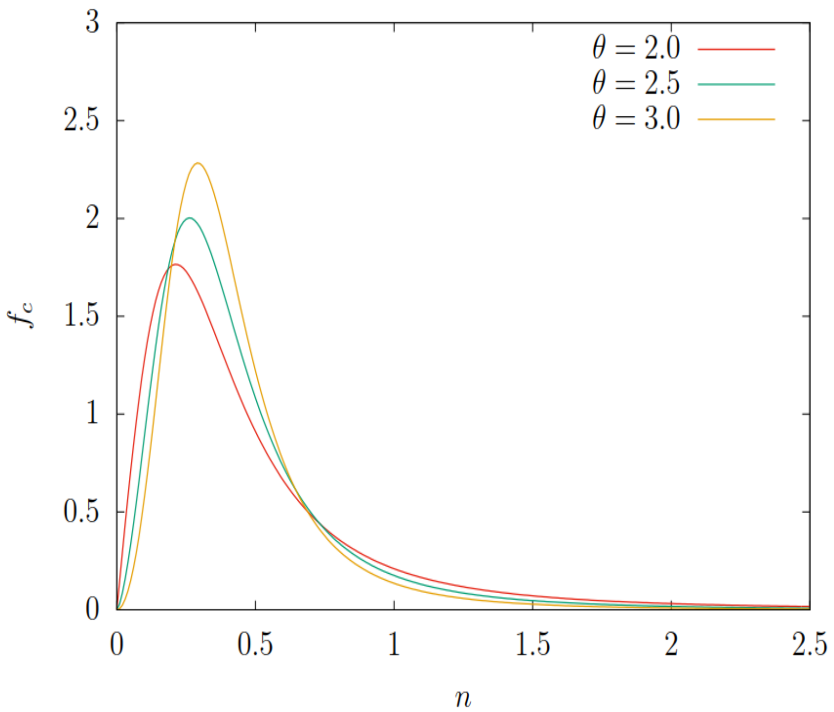

Figure b3 Champernowne distribution k=-1

Appendix C

Table C1 Utilitarian β=0, θ=3.0, k=-0.92, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.10	0.46	0.09	0.15	59	-68	-8.50
0.50	0.52	0.21	0.20	55	3	-7.04
0.90	0.51	0.42	0.30	56	29	-5.38
0.97	0.48	0.63	0.39	57	38	-4.50
0.99	0.49	0.85	0.49	56	43	-3.92

Table C2 Prioritarian β=1, θ=3.0, k=-0.91, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.10	0.33	0.06	0.15	73	-100	-8.05
0.50	0.47	0.19	0.19	69	0	-7.19
0.90	0.51	0.42	0.27	65	37	-5.77
0.97	0.50	0.63	0.34	63	46	-4.93
0.99	0.49	0.89	0.44	60	50	-4.20

Table C3 Maximin β=∞, θ=3.0, k=-0.91, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.08	0.05	0.009	0.14	88		-8.03
0.50	0.44	0.18	0.17	76	2	-7.55
0.90	0.52	0.43	0.25	67	42	-6,10
0.97	0.52	0.67	0.33	63	50	- 5.06
0.99	0.51	0.89	0.42	58	53	-4.43

Table C4 Utilitarian β=0, θ=2.0, k=-1.16, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.10	0.32	0.03	0.14	62	-	-8.84
0.50	0.46	0.15	0.18	66	-21	-7.58
0.90	0.48	0.45	0.27	72	40	-5.65
0.97	0.46	0.83	0.37	73	55	-4.56
0.99	0.44	1.33	0.50	72	62	-3.78

Table C5 Prioritarian β=1, θ=2.0, k=-1.16, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.10	0.19	0.02	0.14	72	-	-8.46
0.50	0.40	0.13	0.16	76	-32	-7.73
0.90	0.48	0.45	0.24	77	47	-6.10
0.97	0.48	0.84	0.33	77	61	-4.98
0.99	0.47	1.37	0.46	74	67	-4.09

Table C6 Maximin β=∞ θ=2.0, k=-1.16, ε= 0.5, R=1-X/Z =0.1

F(n)	y	z	x	MTR%	ATR%	utility
0.16			0.14	86	-	-8.37
0.50	0.36	0.11	0.15	81	-37	-8.05
0.90	0.50	0.46	0.23	78	51	-6.41
0.97	0.50	0.87	0.32	77	63	-5.15
0.99	0.50	1.36	0.44	71	67	-4.24

Table C7 Decile ratios P90/P50

	β=0	β=0	β=1	β=1	β=∞	β=∞
	x	z	x	z	x	z
θ	P90/P50	P90/P50	P90/P50	P90/P50	P90/P50	P90/P50
3.0	1.47	1.95	1.42	2.21	1.47	2.39
2.5	1.53	2.44	1.42	2.70	1.46	2.96
2.0	1.58	3.00	1.50	3.46	1.53	4.18