University of Leicester

DEPARTMENT OF ECONOMICS

INEQUALITY AND SIZE OF THE GOVERNMENT WHEN VOTERS HAVE OTHER REGARDING PREFERENCES

Sanjit Dhami, University of Leicester, UK
Ali al-Nowaihi, University of Leicester, UK

Working Paper No. 08/23
July 2008

Inequality and size of the government when voters have other regarding preferences*

Sanjit Dhami† **Ali al-Nowaihi**‡

22 July 2008

Abstract

The celebrated relation between *inequality* and *redistribution* is based on *selfish voters* who care solely about own-payoffs. A growing empirical literature highlights the importance of *other regarding preferences* (ORP) in voting over redistribution. We reexamine the relation between inequality and redistribution, within a simple general equilibrium model, when voters have ORP. Our contribution is five-fold. First, we demonstrate the existence of a Condorcet winner. Second, poverty can lead to increased redistribution (which implies a countercyclical social spending to GDP ratio). Third, we show that disposable income 'strongly median-dominates' factor income. Fourth, we show that fair voters respond to an increase in 'strong median-dominance' by engaging in greater redistribution. Fifth, an illustrative empirical exercise using OECD data points to the importance of fairness in explaining redistribution.

Keywords: Redistribution, Other regarding preferences, Single crossing property, Income inequality, Difference dominance, Median dominance, American Exceptionalism.

JEL Classification: D64 (Altruism); D72 (Economic Models of Political Processes: Rent-Seeking, Elections, Legislatures, and Voting Behavior); D78 (Positive Analysis of Policy-Making and Implementation).

*We are grateful to helpful comments from participants at the Public Economics Weekend at the University of Leicester, June 2007. In particular we would like to thank Ian Preston and Hamish Low.

†Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: Sanjit.Dhami@le.ac.uk.

‡Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522898. Fax: +44-116-2522908. E-mail: aa10@le.ac.uk.

1. Introduction

Meltzer and Richard (1981), drawing on the work of Romer (1975) and Roberts (1977), set up a simple general equilibrium model with endogenous labour supply and *selfish voters* (i.e. who only derive utility from 'own' payoffs). Voters vote directly over alternative redistributive policies and the median voter is decisive (*direct democracy*).[1] We call this work collectively as the 'RRMR' model. The lasting legacy of Meltzer and Richard (1981) is that they derived the testable prediction that the extent of redistribution directly depends on the ratio between mean and median income. The intuition is that as inequality increases, the median voter is relatively poorer and, hence, chooses greater redistribution.

The RRMR framework treats only a specific kind of transfer, namely, the intra-generational transfers of income. In actual practice, the growth of government transfers in recent decades has been driven by a range of other considerations. For instance, the increased (inter-generational) transfers to the old could possibly reflect their growing political clout in Western democracies. Regional transfers could arise from special interest group considerations. Unemployment and health insurance can only be understood within the model insofar as these entail intra-generational transfers. These issues are possibly better analyzed within a dynamic model[2].

The evidence on the relation between inequality and redistribution is mixed, however. Positive support is found by Meltzer and Richard (1981), Easterly and Rebelo (1993), Alesina and Rodrick (1994), Persson and Tabellini (1994), and Milanovic (2000). However, Lindert (1996), Perotti (1996) do not find any support.

Most empirical work on the relation between inequality and redistribution uses disposable (after tax) income data. Factor (pretax) income data has recently been made available in Milanovic (2000). Milanovic argues that it is *factor* income, rather than *disposable* income, that should be used in empirical work on the relation between inequality and redistribution. Milanovic's argument is that people first observe their own factor (pretax) income, then decide on the redistributive tax rate. In our model, the tax rate, factor income and disposable income are all endogenous variables and our model predicts relationships between all three. In our model, the tax and redistribution policy of the government results in lower disposable-income inequality relative to that measured by factor-income inequality. Thus, in our model, government policy removes some, but not all, income in-

[1] In actual practice, redistributive policies are chosen by the elected representatives of the citizens (i.e. *representative democracy*). Furthermore, the political process is complicated by issues of political agency, information asymmetries, and legislative logrolling etc.; see, for instance, Persson and Tabellini (2000). However, in a range of applications in political economy, one often needs to abstract away from many of these issues. Recent experience in Western democracies suggests that direct democracy is more than a useful benchmark; see for instance, Matsusaka (2005a,b).

[2] See Persson and Tabellini (2000) for a range of such models.

equality. Milanovic's empirical work shows that almost a third of factor income inequality is removed by government tax and transfer programs. Hence, from an econometric point of view, there is likely to be a greater association between factor incomes and redistribution. This is also borne out by our own illustrative empirical exercise, which we present towards the end of the paper.

Subsection 1.1, below, gives an illustrative example. However, the evidence from relatively large cross section samples of countries show that this illustration is not an isolated case (see the references cited in the previous paragraph).

1.1. European versus American redistribution: A Case Study

Consider the comparison between Sweden and the USA, given in Table I. Each of the columns gives the ratio of some relevant economic magnitude for Sweden and the US. DII is the ratio of *disposable* income inequality for Sweden relative to that of the USA. By contrast, FII is the ratio of *factor* income inequality for Sweden relative to that of the USA. SS/GDP is the ratio of social spending to gross domestic product for Sweden as a fraction of the corresponding ratio for the USA. Multi-Aid/GDP is the ratio of multilateral aid to GDP for Sweden as a fraction of the corresponding ratio for the USA.

Table I: Sweden versus USA: An illustrative comparison

	DII	SS/GDP	FII	Multi-Aid/GDP
Sweden/US	0.6	2	1	9

Disposable income inequality in Sweden is about 60 percent that of the USA, while the Swedish social spending to GDP ratio is about twice that of the USA. This raises the following problems for the classical theory which posits that "greater inequality creates greater redistribution."

1. If one uses disposable income inequality (DII), which is the norm in applied work, then the USA should undertake greater redistribution than Sweden, but we observe the exact opposite.

2. What if we were to use the variable suggested by theory i.e. factor income inequality? In this case, it turns out that the factor income inequality is almost identical in the two countries (FII = 1). Theory would then predict identical redistribution, which is not the case either.

1.2. Some recent explanations

The problem, presented in the above example (Sweden versus USA), has been addressed by some more recent work, which is primarily theoretical. We must await the outcome of additional empirical work before we can be sure if any of these explanations suffice. To illustrate, we outline two recent explanations.

Alesina and Angeletos (2003) have a novel explanation. The key to understanding their model is the beliefs of individuals on the source of poverty (or affluence). It is claimed that more Americans than Europeans believe that poverty (or affluence) is caused by individual effort rather than luck. A crucial assumption of the model is that voters expect there to be greater public redistribution if income outcomes are governed by luck rather than effort. Hence, in the European equilibrium, more people believe that income is caused by luck, so put in less effort and, hence, actual outcomes are indeed governed more by luck rather than effort. Given the assumption on public redistribution, there is greater redistribution in equilibrium. The American, high-effort, low-redistribution equilibrium can be understood analogously.

Benabou (2000) develops a stochastic growth model with incomplete asset markets and heterogeneous agents who vote over redistributive policies. He shows that multiple equilibria can exist, some featuring low inequality and high redistribution, while others exhibit high inequality and low redistribution. Thus countries with similar preference, technologies and political systems can feature very different levels of inequality and redistribution.

1.3. Other regarding preferences (ORP) and voting over redistribution

Despite the standard assumption of self-interested voters, it seems plausible that explanations for redistribution should be underpinned by the inherent human desire to care directly for the wellbeing of others. In other words, in the domain of redistribution, it seems reasonable to postulate that individuals have *other regarding preferences* (or ORP for short). We describe voters who have ORP as *fair voters*.

An emerging empirical literature is strongly supportive of the role of ORP specifically in the domain of voting models; see, for instance, Ackert et al. (2007), Bolton and Ockenfels (2006) and Tyran and Sausgruber (2006). These papers establish that voters often choose policies that promote equity/fairness over purely selfish considerations. Bolton and Ockenfels (2006), for instance, examine the preference for equity versus efficiency in a voting game. Groups of three subjects are formed and are presented with two alternative policies: one that promotes equity while the other promotes efficiency. The final outcome is chosen by a majority vote. About twice as many experimental subjects preferred equity as compared to efficiency. Furthermore, even those willing to change the status-quo for efficiency are willing to pay, on average, less than half relative to those who wish to alter the status-quo for equity.

1.4. Which model of ORP?

There are several models of ORP in the literature. We choose to use the Fehr-Schmidt (1999) (henceforth, FS) approach to ORP[3]. In this approach, voters care, not only about their own payoffs, but also about their payoffs relative to those of others. If their payoff is greater than that of other voters then they suffer from *advantageous-inequity* (arising from, say, *altruism*). If their payoff is lower than that of other voters then they suffer from *disadvantageous-inequity* (arising from, say, *envy*). Several reasons motivate our choice of the FS model.

1. The FS model is tractable and explains the experimental results arising from several games where the prediction of the standard game theory model with selfish agents yields results that are not consistent with the experimental evidence. These games include the ultimatum game, the gift-exchange game, the dictator game as well as the public-good game with punishment[4].

2. The FS model focusses on the role of inequity aversion. However, a possible objection is that it ignores the role played by intentions that have been shown to be important in experimental results (Falk et al., 2002) and treated explicitly in theoretical work (Rabin, 1993; Falk and Fischbacher, 2006). However, experimental results on the importance of intentions come largely from bilateral interactions. Economy-wide voting, on the other hand, is impersonal and anonymous, thereby making it unlikely that intentions play any important role in this phenomenon.

3. Experimental results on voting lend support to the use of the FS model in such contexts. Tyran and Sausgruber (2006) explicitly test for the importance of the FS framework in the context of direct voting. They conclude that the FS model predicts much better than the standard selfish-voter model. In addition, the FS model

[3]Bolton and Ockenfels (2002) provide yet another model of inequity averse economic agents, referred to as ERC (short for equity, reciprocity and cooperation). However, it cannot explain the outcome of the public good game with punishment, which is a fairly robust experimental finding. Charness and Rabin (2002) provide two successive versions of their model. In the first version, economic agents do not care directly about outcome differences or the role of 'intentions'. This model is unable to explain the results of the public good game with punishment. A second version of the model introduces the role of intentions. However, voting is anonymous and involves very large number of voters, hence, intentions, in all possibility have a minor, if any, role to play. For a survey of theoretical models of ORP, its neuroeconomic foundations, and the empirical results, see Fehr and Fischbacher (2002).

[4]In the first three of these games, experimental subjects offer more to the other party relative to the predictions of the Nash outcome with selfish preferences. In the public good game with punishment, the possibility of ex-post punishment dramatically reduces the extent of free riding in voluntary giving towards a public good. In the standard theory with selfish agents, bygones are bygones, so there is no ex-post incentive for the contributors to punish the free-riders. Foreseeing this outcome, free riders are not deterred, which is in disagreement with the evidence. Such behavior can be easily explained within the FS framework.

provides, in their words, "strikingly accurate predictions for individual voting in all three income classes." The econometric results of Ackert et al. (2007), based on their experimental data, lend further support to the FS model in the context of redistributive taxation. The estimated coefficients of altruism and envy in the FS model are statistically significant and, as expected, negative in sign. Social preferences are found to influence participant's vote over alternative taxes. They find evidence that some participants are willing to reduce their own payoffs in order to support taxes that reduce advantageous or disadvantageous inequity. In the context of voting experiments, Bolton and Ockenfels (2006) conclude that "...while not everyone measures fairness the same way, the simple measures offered by ERC or FS provide a pretty good approximation to population behavior over a wide range of scenarios that economists care about."

1.5. The Sweden v. US example revisited in the context of ORP

Although our paper does not directly deal with Swedish versus American redistribution, it is illustrative to see how a model based on ORP could provide an alternative explanation for greater European redistribution relative to America. The explanation hinges on using the argument that Europeans, on average, are relatively more inequity averse (or *fair* in the Fehr-Schmidt sense) and that there is basis for '*American Exceptionalism*'[5]. Our first explanation requires us to construct some empirical measure of fairness. Measuring fairness presents a challenging and contentious set of issues.

One possible candidate for measuring inequity aversion or fairness is *charitable giving per capita*. However, charitable contributions are endogenous in a model where voters have ORP. So, for instance, if government redistribution is perceived to be inadequate, citizens might attempt to compensate by donating more to charity. For that reason we do not believe that the relatively greater per capita giving of Americans necessarily indicates that they are more inequity averse/fair as compared to Europeans.

A better candidate for measuring fairness/inequity aversion is 'aid given to other countries, particularly, developing countries.' As in the case of charitable contributions made to own country residents, aid given by a country to any particular developing country could well reflect the low volume of aggregate giving to that country in the first place. However, crucially, this applies equally to all giving countries. Hence, relative giving of countries potentially reflects relative fairness/inequity aversion.

The main focus of this paper is theoretical. However, in a illustrative empirical exercise

[5] A good summary (with some historical background) of American Exceptionalism is provided in Glaeser (2005). The reasons include, among others, proportional versus majoritarian representation, greater ethnic heterogeneity in America, and the US tradition of federalism that gives redistributive powers to the individual states.

(in section 6 below), and also to illustrate the Sweden versus US comparison, we will use 'aid given to developing countries' as our measure of fairness. We will use those measures of aid that remove the effects of strategic political considerations and other such confounding factors. The resulting measure of aid (described more fully in section 6) is called 'multilateral aid to GDP ratio'.

The US is the single largest contributor to development aid. However, in per capita terms its contribution is lower than most European countries. According to OECD figures, the US contributed only 0.15 percent of its GDP to development assistance, placing it last in a list of 21 western (mostly European) countries. The Center for Global Development estimated that US development assistance per capita is one eighth that of Norway, one sixth that of Denmark and close to half of the average contributions of Belgium, France, Finland and Britain.

In terms of our Swedish versus US comparison in Table I, our proxy for inequity aversion, multilateral aid to GDP ratio (this corrects for strategic giving), for Sweden is about 9 times that of the US. We argue that this reveals greater inequity aversion of the Swedes relative to the US and potentially helps to explain the differences in the social spending to GDP ratios between the two countries; a difference that factor income inequality is unable to explain.

1.6. A critique of the literature on voting and fairness

There is a relatively small theoretical literature that considers voters with ORP. We concentrate below on the papers that are directly relevant to our work. Tyran and Sausgruber (2006), reviewed above, do not analyze the relation between inequality and redistributive taxes which is important in the RRMR framework (and ours'). Their's is not a general equilibrium model, does not analyze the efficiency costs of redistribution and, probably most importantly, does not provide existence results for there to be a decisive median voter. Furthermore, they consider a more restricted tax policy choice than us. While we consider changes in a linear progressive income tax that affect all taxpayers, they focus attention only on redistributions from the rich to the poor that leave the middle income voters unaffected[6].

Galasso (2003) modifies the RRMR model to allow for fairness concerns. However, his notion of fairness is not only one-sided but it is of a very specific form; it is not fully consistent with any of the accepted models of ORP. In particular, fair voters care about their own payoffs but suffer disutility through a term that is linear in their payoffs relative to the worse off voter in society.[7] Since this concern for fairness arises from a linear term,

[6] They do introduce a cost of such redistribution to the middle income voters, but it is not an integral part of the redistributive fiscal package considered.

[7] The latter term captures some notion of social justice. Others have included such a term to incorporate

preferences continue to be strictly concave and a median voter equilibrium exits. Within this framework there is greater redistribution when there is a mean preserving spread in inequality. However, this leaves open the question of whether a median voter equilibrium will exist in a standard model of fairness, such as the FS model, and what the properties of the resulting equilibrium will be.

1.7. The main questions and results

We now summarize the main questions that we pose and our findings in each case.

1. Does a Condorcet winner exists in a model with ORP? The current situation is analogous to the period of time before the median voter theorem was discovered by Duncan Black (1948) (and later popularized by Anthony Downs, 1957) for the case of self-interested voters. However, once known, and popularized, the median voter theorem opened up the domain of modern political economy as we know it today. Some of the most successful applications of the median voter theorem have relied on quasi-linear preferences and a constant elasticity of labour supply (we review some of the literature below). Under these preferences, we show that a Condorcet winner exists when voters have ORP in a RRMR framework.

2. What is the effect of increased poverty on redistribution? In the standard selfish voter model, an *increase* in poverty *reduces* redistribution. The reason is that an increase in poverty reduces average incomes, so making it less worthwhile for the median selfish voter to increase the redistributive tax rate. Insofar as periods of increased poverty are also associated with unemployment shocks, the prediction is that the ratio of social spending to GDP is pro-cyclical. However, when voters have ORP, poverty can lead to increased redistribution (which implies a countercyclical social spending to GDP ratio). The reason is that if the inequity aversion of the fair median voter is high enough, then he/she derives a utility loss from increased poverty. An increase in the redistributive tax rate, under these circumstances, reduces post-tax poverty and makes the median voter better off.

3. What is the relation between inequality and redistribution when one pretax income distribution is more unequal than another? To answer this question, we need a formal definition of 'unequal' (see Definition 2, below). Consider two (pretax or

social justice e.g. Charness and Rabin (2002). However, they posit preferences, different from Galasso (2003), that are a convex combination of the total payoff of the group (this subsumes selfishness, in so far as one's own payoff is part of the total, and altruism) and a Rawlsian social welfare function. These sorts of models are able to explain positive levels of giving in dictator games, and reciprocity in trust and gift exchange games. However, they are not able to explain situations where an individual tries to punish others in the group at some personal cost, for instance, punishment in public good games.

after-tax) income distributions, **x** and **y**, with the same mean. In both distributions, median income is lower than the mean income. We say that **x** *median-dominates* **y** if, and only if, median income under **x** is no less than median income under **y**, regardless of other incomes. We also say that **x** *strictly median-dominates* **y** if, and only if, median income under **x** is strictly higher than median income under **y**, regardless of other incomes. These concepts are clearly relevant for the case of a selfish median voter. For the case of a fair median voter advantageous and disadvantageous inequalities also become important. We say that **x** *strongly median-dominates* **y** if, and only if, the following criteria hold: (i) income of the median voter under **x** is no less than that under **y**, (ii) advantageous inequality for the median voter under **x** is no less than that under **y**, (iii) disadvantageous inequality for the median voter under **x** is no worse than that under **y**; and the median voter under **x** is strictly better-off than under **y** according to at least one of these three criteria. Several concepts of dominance are available in the literature. Ours is closest to the concept of *difference-dominance* due to Marshall and Olkin (1979) (see Remark 3, below). We show that fair voters respond to an increase in strong median-dominance by engaging in greater redistribution. Although median-dominance is developed for the purposes of this paper, it may prove to be a useful concept in other applications as well.

4. Is fairness empirically important in explaining actual redistribution? For reasons that we have already outlined in sections 1.1, 1.5 above, the argument is difficult to settle at the moment. Lack of data on factor income inequality, the relevant inequality variable, hampers a satisfactory empirical exercise from being carried out. Furthermore, it is not clear what should be an appropriate measure of fairness or inequity aversion. We present an illustrative empirical exercise based on OECD data, using as our measure of fairness, multilateral aid to GDP, and the figures on factor income inequality made available in Milanovic (2000). We find that fairness is the most important variable in explaining redistribution. We hope that more satisfactory empirical exercises with much larger data sets (and controlling for other factors) will throw further light on the importance of ORP in this context.

1.8. Plan of the paper

Section 2 describes the theoretical model and derives some preliminary results. Section 3 establishes the existence of a Condorcet winner for fair voters. Comparative static results along with some calibration exercises are derived and discussed in Section 4. Section 5 considers the relationship between income distribution and the tax rate using a discrete analogue of second order stochastic dominance. Section 6 presents our illustrative empirical exercise. Finally, section 7 concludes. The results of regression analysis are presented in Appendix 1. Proofs are relegated to Appendix 2.

2. Model

We consider a general equilibrium model as in Meltzer and Richard (1981). Let there be $n = 2m - 1$ voter-worker-consumers (henceforth, voters). Let the skill level of voter j be s_j, $j = 1, 2, ..., n$, where

$$0 < s_i < s_j < 1, \text{ for } i < j, \tag{2.1}$$

Denote the median skill level by s_m and the skill vector by $\mathbf{s} = (s_1, s_2, ..., s_n)$. Each voter has a fixed time endowment of one unit and supplies l_j units of labor and so enjoys $L_j = 1 - l_j$ units of leisure, where

$$0 \leq l_j \leq 1. \tag{2.2}$$

Labour markets are competitive and each firm has access to a linear production technology such that production equals $s_j l_j$. Hence, the wage rate offered to each voter coincides with the marginal product, i.e., the skill level, s_j. Thus, the before-tax income of voter j is given by

$$y_j = s_j l_j. \tag{2.3}$$

Note that 'skill' here need not represent any intrinsic talent, just ability to translate labour effort into income[8]. Let the average before-tax income be

$$\overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j. \tag{2.4}$$

We make the empirically plausible assumption that the pretax income of the median-skill voter, y_m, is less than the average pretax income,[9]

$$y_m < \overline{y}. \tag{2.5}$$

The government operates a linear progressive income tax that is characterized by a constant marginal tax rate, t,

$$t \in [0, 1], \tag{2.6}$$

[8] For example, a highly talented classical musician may be able to earn only a modest income, while a merely competent 'pop' musician may earn millions. In our model, the former would be classified as having a low s while the latter would be classified as having a high s. Similarly, in recent years, there has been a record level of skilled (in the ordinary sense of the word) migration into Britain from Eastern Europe. However, since they are predominantly accepting low pay work, they would be classified in our model as having low s.

[9] The assumption, that $y_m < \overline{y}$, is needed for Propositions 4, 5 and 6 but not for Propositions 1, 2 or 3. The condition $y_m < \overline{y}$ is a restriction on the endogenous variables y_m and $\overline{y}$. The equivalent restriction on the exogenous variables, $\mathbf{s}$, is given by (2.25).

and a uniform transfer, b, to each voter that equals the average tax proceeds,

$$b = t\overline{y} = \frac{t}{n}\Sigma_{i=1}^{n} y_i = \frac{t}{n}\Sigma_{i=1}^{n} s_i l_i \geq 0. \tag{2.7}$$

Thus, the tax rate is also the ratio of social spending to aggregate income,

$$t = \frac{nb}{\Sigma_{i=1}^{n} y_i}. \tag{2.8}$$

Remark 1 : *From (2.8), changes in the tax rate can equivalently be viewed as changes in the ratio of social spending to aggregate income.*

The budget constraint of voter i is given by

$$0 \leq c_i \leq (1-t)\, y_i + b. \tag{2.9}$$

In view of (2.3), the budget constraint (2.9) can be written as

$$0 \leq c_i \leq (1-t)\, s_i l_i + b. \tag{2.10}$$

2.1. Preferences of Voters

We define a voter's preferences in two stages. First, let voter j have an *own-utility* function, $\widetilde{u}(c_j, 1-l_j)$, defined over own-consumption, c_j, and own-leisure, $1-l_j$. In common with the literature, we assume that all voters have the same own-utility function. Hence, voters differ only in that they are endowed with different skill levels, s_i. Furthermore, we assume that the own-utility function is quasi-linear, with constant elasticity of labour supply, which is the most commonly used functional form in various applications of the median voter theorems.

$$\widetilde{u}(c_j, 1-l_j) = c_j - \frac{\epsilon}{1+\epsilon} l_j^{\frac{1+\epsilon}{\epsilon}}, \tag{2.11}$$

where ϵ is the constant elasticity of labour supply, and satisfies[10]

$$0 < \epsilon \leq 1, \tag{2.12}$$

The case $\epsilon = 1$ has special significance in the literature. In this case,

$$\widetilde{u}(c_j, 1-l_j) = c - \frac{1}{2} l_j^2 \tag{2.13}$$

[10] A large number of studies suggest labour supply elasticities consistent with (2.12) (see, for example, Pencavel (1986) and Killingworth and Heckman (1986)). Those that do not (for example, negative labour supply elasticities) may be due to estimating misspecified models (see Camerer and Loewenstein (2004), Chapter 1, 'Labor Economics', pp33-34).

It is well known that (2.13) is sufficient to derive the celebrated result that the extent of redistribution varies directly with the ratio of the mean to median income.[11] Piketty (1995) restricts preferences to the quasi-linear case with disutility of labour given by the quadratic form, (2.13). Benabou and Ok (2001) do not actually consider a production side and their model has exogenously given endowments which evolve stochastically. Benabou (2000) considers the additively separable case with log consumption and disutility of labor given by the constant elasticity case, (2.11).

From (2.11), $\frac{\partial \widetilde{u}}{\partial c_j} > 0$. It follows that the budget constraint (2.9), and also (2.10), holds with equality. These give:

$$c_i = (1-t)\, y_i + b. \tag{2.14}$$

$$c_i = (1-t)\, s_i l_i + b. \tag{2.15}$$

From (2.7) and (2.14), we get

$$\Sigma_{i=1}^{n} c_i = \Sigma_{i=1}^{n} y_i. \tag{2.16}$$

Thus total pretax income is equal to total post tax income. This is simply a consequence of the fact that all the tax revenue is returned to consumers.

Substituting $c_j = (1-t)\, s_j l_j + b$, from (2.15), into (2.11), gives the following form for own-utility

$$u_j = u\left(l_j; t, b, s_j\right) = \widetilde{u}((1-t)\, s_j l_j + b, 1 - l_j) = (1-t)\, s_j l_j + b - \frac{\epsilon}{1+\epsilon} l_j^{\frac{1+\epsilon}{\epsilon}}. \tag{2.17}$$

Second, and for the reasons stated in the introduction, voters have *other-regarding preferences* as in Fehr-Schmidt (1999). Let $\mathbf{l}_{-j}$ be the vector of labour supplies of voters other than voter j. Under Fehr-Schmidt preferences the *FS-utility* of voter j, $U_j\left(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s}\right)$, is as follows.

$$U_j\left(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) = u_j - \frac{\alpha}{n-1} \sum_{k \neq j} \max\left\{0, u_k - u_j\right\} - \frac{\beta}{n-1} \sum_{i \neq j} \max\left\{0, u_j - u_i\right\}, \tag{2.18}$$

where u_j is defined in (2.17), and

$$\text{for } \textit{selfish} \text{ voters } \alpha = \beta = 0, \text{ so } U_j\left(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) = u_j \tag{2.19}$$

$$\text{for } \textit{fair} \text{ voters } 0 < \beta < 1,\ \beta < \alpha, \text{ so } U_j\left(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) \neq u_j \tag{2.20}$$

Thus, $u_j = u\left(l_j; t, b, s_j\right)$ is also the utility function of a selfish voter, as in the standard textbook model. From (2.18), the fair voter cares about own payoff (first term), payoff

[11] See, for instance, Hindriks and Myles (2006). Not much is known about a more general class of utility functions for which this result holds. Meltzer and Richard (1981), who first report this result need to assume the constancy of a complicated function of endogenous variables (see their equation (15)). In this sense, we generalize this class of results somewhat by showing that they hold when utility is given by (2.11).

relative to those where inequality is disadvantageous (second term) and payoff relative to those where inequality is advantageous (third term). The second and third terms which capture respectively, *envy* and *altruism*, are normalized by the term $n-1$. Notice that in FS preferences, inequality is *self-centered*, i.e., the individual uses her own payoff as a reference point with which everyone else is compared to. Also, while the Fehr-Schmidt specification is directly in terms of monetary payoffs, it is also consistent with comparison of payoffs in utility terms. These and related issues are more fully discussed in Fehr and Schmidt (1999). From (2.20), β is bounded below by 0 and above by 1 and α. On the other hand, there is no upper bound on α.[12]

2.2. Sequence of moves

We consider a two-stage game. In the first stage, all voters vote directly and sincerely on the redistributive tax rate. Should a median voter equilibrium exist, then the tax rate preferred by the median voter is implemented. In the second stage, all voters make their labour supply decision, conditional on the tax rate chosen by the median voter in the first stage. On choosing their labour supplies in the second stage, the announced first period tax rate is implemented and transfers made according to (2.7).

In the second stage, the voters play a one-shot Nash game: each voter, j, chooses his/her labour supply, l_j, given the vector, $\mathbf{l}_{-j}$, of labour supplies of the other voters, so as to maximize his/her FS-utility (2.18). In the first stage, each voter votes for his/her preferred tax rate, correctly anticipating second stage play.

The solution is by backward induction. We first solve for the Nash equilibrium in labour supply decisions of voters, conditional on the announced tax rates and transfers. The second stage decision is then fed into the first stage FS-utilities to arrive at the indirect utilities of voters, which are purely in terms of the tax rate. Voters then choose their most desired tax rates which maximize their indirect FS-utilities, with the proposal of the median voters being the one that is implemented.

2.3. Labour supply decision of taxpayers (second stage problem)

Given the tax rate, t, and the transfer, b, both determined in the first stage (see Section 2.4, below), the voters play a one-shot Nash game (in the subgame determined by t and b). Each voter, j, chooses own labour supply, l_j, so as to maximize his/her FS-utility (2.18), given the labour supplies, $\mathbf{l}_{-j}$, of all other voters.

[12] $\beta \geq 1$ would imply that individuals could increase utility by simply destroying all their wealth; this is counterfactual. The restriction $\beta < \alpha$ is based on experimental evidence. Finally, the lack of an upper limit on α implies that 'envy' is unbounded.

Proposition 1 *: In the second stage of the game, voter j, whether fair or selfish, chooses own labour supply, l_j, so as to maximize own-utility, $u(l_j; t, b, s_j)$, given t, b and s_j.*

The FS utility of any voter is separable in own labour, l_j, and the labor supplies of others, $\mathbf{l}_{-j}$. Hence, the result in Proposition 1 should not be surprising. This is yet another attractive feature of FS preferences. Namely, that there is no essential difference in computing the labor supplies of selfish voters and voters with ORP.

We list, in lemmas 1, 2, below, some useful results.

Lemma 1 *(Labour supply): Given t, b and s_j, the unique labour supply for voter j, $l_j = l(t, b, s_j)$, that maximizes utility (2.17), is given by*

$$l_j = l(t, b, s_j) = (1-t)^{\epsilon} s_j^{\epsilon},$$

and is independent of b.

Substituting labour supply, $l(t, b, s_j)$, given by Lemma 1, into (2.3) gives pretax income:

$$y_j = y(t, b, s_j) = (1-t)^{\epsilon} s_j^{1+\epsilon}. \tag{2.21}$$

Suppose $t < 1$ and $i < j$. From (2.1), (2.14) and (2.21), we get:

$$0 < y_i < y_j, \tag{2.22}$$

$$0 < c_i < c_{j.} \tag{2.23}$$

Hence, at any tax rate (below 100%) a higher skill worker earns higher pretax (y_i) and post tax (c_i) incomes than a lower skill worker. It follows the the median skill worker earns median factor and disposable incomes.

Define $\overline{S}$ to be the 'weighted average of skills' when there are n voters, in the following sense.

$$\overline{S} = \frac{1}{n}\Sigma_{i=1}^{n} s_i^{1+\epsilon} > 0. \tag{2.24}$$

From (2.4), (2.21) and (2.24), we get that, for $t < 1$, (2.5) holds (i.e., $y_m < \overline{y}$) if, and only if,

$$s_m^{1+\epsilon} < \overline{S}, \tag{2.25}$$

Substituting labour supply in (2.7) we get,

$$b(t, \mathbf{s}) = t(1-t)^{\epsilon}\,\overline{S}. \tag{2.26}$$

Substituting labour supply in (2.17) we get the indirect utility function corresponding to the own-utility of voter j:

$$v_j = v(t, b, s_j) = u(l(t, b, s_j); t, b, s_j) = b + \frac{(1-t)^{1+\epsilon}}{1+\epsilon} s_j^{1+\epsilon}. \tag{2.27}$$

Lemma 2 *(Properties of the indirect utility function):*
(a) $\frac{\partial v(t,b,s)}{\partial b} = 1$,
(b) $\frac{\partial v(t,b,s)}{\partial s} = (1-t)^{1+\epsilon} s^{\epsilon}$. *Hence,* $\left[\frac{\partial v(t,b,s)}{\partial s}\right]_{t=1} = 0$ *and* $t \in [0,1) \Rightarrow \frac{\partial v(t,b,s)}{\partial s} > 0$.

Lemma 2 shows that an increase in transfer payment, b, increase utility one for one and that for any interior tax rate, indirect utility is strictly increasing in the level of skill.

Substitute $b(t, \mathbf{s})$, given by (2.26), into the indirect utility (2.27), to get

$$w_j(t, \mathbf{s}) = v(t, b(t, \mathbf{s}), s_j) = t(1-t)^{\epsilon} \overline{S} + \frac{(1-t)^{1+\epsilon}}{1+\epsilon} s_j^{1+\epsilon}. \tag{2.28}$$

2.4. Preferences of voters over redistribution (the first stage problem)

Given the second stage choice of labor supplies by the voters (Proposition 1 and Lemma 1), the first stage problem is to choose the redistributive tax rate, t (and, consequently, the transfer, b, given by (2.3), (2.4) and (2.7)). For this purpose, we calculate the voters' indirect utility functions corresponding to their FS-preferences.

To find the indirect utility function, for voter j, $V_j = V_j(t, b, \alpha, \beta, \mathbf{s})$, that corresponds to his/her FS-preferences, substitute $v_j = v(t, b, s_j) = u(l(t, b, s_j); t, b, s_j)$ from (2.27) into (2.18) to get,

$$V_j(t, b, \alpha, \beta, \mathbf{s}) = v_j - \frac{\alpha}{n-1} \sum_{k \neq j} \max\{0, v_k - v_j\} - \frac{\beta}{n-1} \sum_{i \neq j} \max\{0, v_j - v_i\}. \tag{2.29}$$

In the light of Lemma 2b, we can rewrite (2.29) as[13]

$$V_j(t, b, \alpha, \beta, \mathbf{s}) = v_j - \frac{\alpha}{n-1} \sum_{k > j} (v_k - v_j) - \frac{\beta}{n-1} \sum_{i < j} (v_j - v_i). \tag{2.30}$$

Substitute v_j from (2.27) into (2.30) we get,

$$V_j = b + \frac{(1-t)^{1+\epsilon}}{1+\epsilon} \left[s_j^{1+\epsilon} - \frac{\alpha}{n-1} \sum_{k>j} \left(s_k^{1+\epsilon} - s_j^{1+\epsilon}\right) - \frac{\beta}{n-1} \sum_{i<j} \left(s_j^{1+\epsilon} - s_i^{1+\epsilon}\right) \right]. \tag{2.31}$$

The model with fair voters is similar in structure to the one with selfish voters in which some weighted social welfare is maximized such that the weight placed by voter j on the i^{th} voter's indirect utility is λ_{ji}. To see this, rewrite (2.30) as

$$V_j(t, b, \alpha, \beta, \mathbf{s}) = \sum_i \lambda_{ji} v_i \tag{2.32}$$

[13] We use the standard mathematical conventions that $\sum_{i \in \emptyset} x_i = 0$, where $\emptyset$ is the empty set. In particular,

$$\sum_{k>n} (v_k - v_j) = \sum_{i<1} (v_j - v_i) = 0.$$

where, the weights are defined by:

$$\lambda_{ji} = \left\{ \begin{array}{cc} \frac{\beta}{n-1} & if \quad i<j \\ 1-\frac{(j-1)\beta}{n-1}+\frac{(n-j)\alpha}{n-1} & if \quad i=j \\ \frac{-\alpha}{n-1} & if \quad k>j \end{array} \right\} \tag{2.33}$$

Finding a tax rate to maximize (2.32) is a completely standard problem in public economics. Hence, the introduction of fair voters does not fundamentally alter the optimization problem.

In order to write the expressions in a compact manner, for any voter j, define the following four useful constants, S_j^-, S_j^+, ψ_j and $\widehat{j}$.

$$S_n^- = 0,\ S_j^- = \frac{1}{n-1}\sum_{k>j}\left(s_k^{1+\epsilon} - s_j^{1+\epsilon}\right) > 0, \text{ for } j<n, \tag{2.34}$$

$$S_1^+ = 0,\ S_j^+ = \frac{1}{n-1}\sum_{i<j}\left(s_j^{1+\epsilon} - s_i^{1+\epsilon}\right) > 0, \text{ for } j>1, \tag{2.35}$$

$$\psi_j = \overline{S} - s_j^{1+\epsilon} + \alpha S_j^- + \beta S_j^+. \tag{2.36}$$

$$\widehat{j} = \max\left\{j : \psi_j > 0\right\}. \tag{2.37}$$

From (2.19), (2.20), (2.25), (2.34), (2.35) and (2.36) we get

$$\psi_m = \overline{S} - s_m^{1+\epsilon} + \alpha S_m^- + \beta S_m^+ > 0. \tag{2.38}$$

From (2.38) and (2.37) we see that for the median skill voter, s_m,

$$m \leq \widehat{j}. \tag{2.39}$$

The constants S_j^- and S_j^+ are function of the exogenous parameters, n (number of voters), ϵ (elasticity of labour supply) and $\mathbf{s}$ (the skills vector). In addition to these, ψ_j is also a function of α (disadvantageous inequity parameter) and β (advantageous inequity parameter).

Remark 2 : *From (2.31), (2.34) and (2.35), S_j^- and S_j^+ are, respectively, directly proportional to disadvantageous and advantageous inequality experienced by voter j.*

Substitute from (2.34) - (2.36) into (2.31) to get

$$\begin{aligned} V_j(t,b,\alpha,\beta,\mathbf{s}) &= b + \frac{(1-t)^{1+\epsilon}}{1+\epsilon}\left(s_j^{1+\epsilon} - \alpha S_j^- - \beta S_j^+\right), \\ &= b + \frac{(1-t)^{1+\epsilon}}{1+\epsilon}\left(\overline{S} - \psi_j\right). \end{aligned} \tag{2.40}$$

When voter j votes on the tax rate, t, and the transfer, b, he/she takes into account the government budget constraint (2.26). Hence, substitute $b(t, \mathbf{s})$, given by (2.26), into (2.40), to get

$$W_j(t, \alpha, \beta, \mathbf{s}) = t(1-t)^{\epsilon} \overline{S} + \frac{(1-t)^{1+\epsilon}}{1+\epsilon}\left(\overline{S} - \psi_j\right). \tag{2.41}$$

Voter j votes for that tax rate, t, which maximizes social welfare from his/her own point of view, as given by his/her FS-indirect utility function (2.41).

In Proposition 2, below, we give some results on the existence of optimal (or most preferred) taxes for any individual voter who at that first stage is asked to state his/her choice of the most preferred tax rate. The next section, Section 3, will look at the equilibrium tax rate that is actually implemented by society.

Proposition 2 *(Existence of optimal tax rates):*
(a) Given α, β and $\mathbf{s}$, $W_j(t, \alpha, \beta, \mathbf{s})$ attains a maximum at some $t_j \in [0, 1)$.
(b) If $j > \widehat{j}$, then the tax rate preferred by voter j, is $t_j = 0$.
(c) If $j \leq \widehat{j}$, then the tax rate, t_j, preferred by voter j, is unique, satisfies $0 < t_j < 1$ and is given by

$$t_j = \frac{\psi_j}{\epsilon \overline{S} + \psi_j}, \tag{2.42}$$

where $\overline{S}$, ψ_j and $\widehat{j}$ are defined respectively in (2.24), (2.36) and (2.37).
(d) For $j \leq \widehat{j}$, t_j is strictly increasing in α and β.
(e) $1 > t_1 > t_2 > ... > t_{\widehat{j}} > t_{\widehat{j}+1} = t_{\widehat{j}+2} = ... = 0$. In particular, if $\widehat{j} = n$, then $1 > t_1 > t_2 > ... > t_n > 0$.

The result of Propositions 2(d) and (e) may be deserving of further comment. The interpretation of (d) is that an increase in α increases disutility arising from disadvantageous inequity. By increasing the redistributive tax rate, the voter reduces, relatively, the utility of anyone who is richer, hence, reducing disadvantageous inequity. On the other hand, an increase in β increases disutility arising from advantageous inequity. An increase in the redistributive tax benefits everyone poorer than the voter relatively more, thus, reducing advantageous inequity. The interpretation of (e) is that the system of taxes and transfers benefit low skill voters relatively more than high skill voters. Hence lower skill voters prefer higher tax rates than high skill voters.

3. Existence of a Condorcet winner

We shall show that a majority chooses the tax rate, t_m, that is optimal for the median-skill voter, in the sense that, for each $j \neq m$, a majority prefers t_m over t_j. We do this by using the *single-crossing property* of Gans and Smart (1996).

Definition 1 : *(Gans and Smart, 1996) The 'single-crossing' property holds if for tax rates* t, T *and voters* j, J,

$$t < T,\ j < J,\ W_j(t, \alpha, \beta, \mathbf{s}) > W_j(T, \alpha, \beta, \mathbf{s}) \Rightarrow W_J(t, \alpha, \beta, \mathbf{s}) > W_J(T, \alpha, \beta, \mathbf{s}).$$ [14]

Lemma 3 : *(Gans and Smart, 1996) The 'single-crossing' property holds if* $-\frac{\partial V_j}{\partial t} / \frac{\partial V_j}{\partial b}$ *is an increasing function of* j.

Lemma 4 : *(Gans and Smart, 1996) If the 'single-crossing' property holds, then the median voter is decisive, i.e., a majority chooses the tax rate that is optimal for the median voter.*

The proofs of lemmas 3, 4 can be found in Gans and Smart (1996). The intuition behind these lemmas is straighforward to illustrate in the following diagram in (b, t) space. In Figure 3.1, the aggregate budget constraint of the economy, given in (2.7), is shown by the straight upward sloping line, BB', that has slope $\overline{y}$. We show two indifference curves belonging to a poor (I_pI_p) and a rich (I_rI_r) voter respectively. Lemma 3 requires that $-\frac{\partial V_j}{\partial t} / \frac{\partial V_j}{\partial b}$ be an increasing function of j i.e. I_pI_p is relatively flatter (we show this in section 3.1 below). The most preferred tax rate of the poor, t_p, is greater than the most preferred tax rate of the rich, t_r. Hence, the preferred tax rates can be uniquely ordered from the rich to the poor. This monotonicity property gives rise to the result in Lemma 4.

3.1. Existence of a Condorcet winner

Proposition 3 : *A majority prefers the tax rate that is optimal for the median-voter.*

In light of the emerging evidence, it increasingly appears that issues of fairness and concern for others are important human motivations that play a significant part in the actual design of redistributive tax policies. Insofar as actual applications of a direct democracy framework largely use quasi-linear preferences and constant elasticity of labour supply, Proposition 3 establishes the existence of a Condorcet winner. Hence, the result in Proposition 3 is potentially of major importance for political economy models that seek to incorporate social preferences.

[14] Here we use "$<$" to denote the usual ordering of real numbers. In the more general setting of Gans and Smart (1996), "$<$" is used to denote several (possibly different) abstract orderings. In particular, a literal translation of Gants and Smart (1996) would give: $T < t,\ j < J,\ W_j(t, \alpha, \beta, \mathbf{s}) > W_j(T, \alpha, \beta, \mathbf{s}) \Rightarrow W_J(t, \alpha, \beta, \mathbf{s}) > W_J(T, \alpha, \beta, \mathbf{s})$, where "$j < J$" has the usual meaning "j is less than J" but "$T < t$" means "t is less than T".

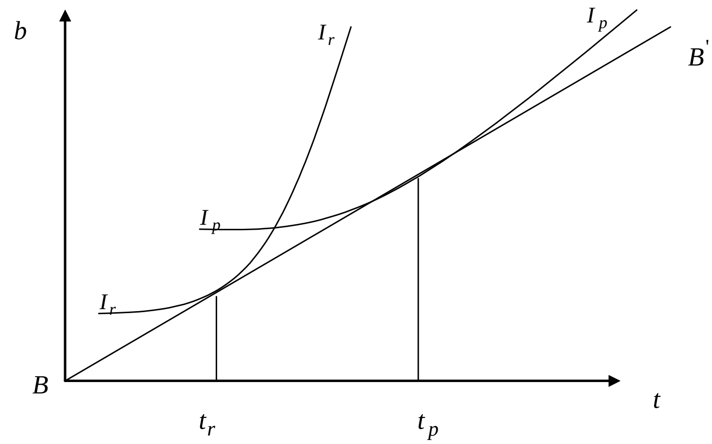

Figure 3.1: An illustration of the Gans-Smart single crossing property

4. Comparative static results

Proposition 4, below, gives the change in the tax rate chosen by the median voter, t_m, as various parameters in the model are changed.

Proposition 4 : *(a) The tax rate, t_m, chosen by the median voter, is given by*

$$t_m = \frac{\psi_m}{\epsilon \overline{S} + \psi_m}. \tag{4.1}$$

where $\overline{S}, \psi_m$ are defined respectively in (2.24), (2.36).
(b) For fair voters, $\frac{\partial t_m}{\partial \alpha} > 0$, $\frac{\partial t_m}{\partial \beta} > 0$.
(c) A fair median voter chooses a higher tax rate than a selfish median voter.
(d) For selfish and fair voters, $j > m \Rightarrow \frac{\partial t_m}{\partial s_j} > 0$.
(e) For selfish and fair voters, $\frac{\partial t_m}{\partial s_m} < 0$.
(f) For fair voters, for $j < m$,

$$\frac{\partial t_m}{\partial s_j} \lesseqqgtr 0 \Leftrightarrow \psi_m \gtreqqless \left(1 - \frac{n\beta}{n-1}\right) \overline{S}.$$

(g) For selfish voters, $j < m \Rightarrow \frac{\partial t_m}{\partial s_j} > 0$.
(h) For fair voters, for $j < m$,

$$\alpha \geq \frac{(n-1)\, s_m^{1+\epsilon}}{\sum\limits_{k>m} \left(s_k^{1+\epsilon} - s_m^{1+\epsilon}\right)} \Rightarrow \frac{\partial t_m}{\partial s_j} < 0.$$

(i) For fair voters, for $j < m$,

$$\beta \geq \frac{(n-1)\, s_m^{1+\epsilon}}{\Sigma_{i=1}^n s_i^{1+\epsilon} + \sum_{i<m} \left(s_m^{1+\epsilon} - s_i^{1+\epsilon} \right) + \sum_{k>m} \left(s_k^{1+\epsilon} - s_m^{1+\epsilon} \right)} \Rightarrow \frac{\partial t_m}{\partial s_j} < 0.$$

From part (b), the tax rate (equivalently, the ratio of social spending to GDP, see (2.8)) is increasing in α, β. The intuition is that an increase in α increases disutility arising from disadvantageous inequity. By increasing the redistributive tax rate, the median voter reduces, relatively, the utility of anyone who is richer, hence, reducing disadvantageous inequity. On the other hand, an increase in β increases disutility arising from advantageous inequity. An increase in the redistributive tax benefits everyone poorer than the median voter relatively more, reducing advantageous inequity.

Selfish median voters would like to redistribute because they are poorer than the average voter. Part (c) follows by simply noting that fair median voters have an additional tendency to redistribute on account of their fairness.

From Remark 1 and Proposition 4(d), selfish and fair voters alike, respond to increased affluence of the rich by redistributing more and so also raising the ratio of social spending to GDP. Selfish voters would like to redistribute more when the rich get richer because average incomes increase and so the lumpsum available for redistribution is higher. Fair voters have an additional motive to redistribute more, namely, that it reduces disadvantageous inequity.

Parts (f)-(i) point out to an important difference in the predictions of the fair and selfish voter models. From part (g), for selfish voters, an increase in poverty reduces the tax rate and the ratio of social spending to GDP. The intuition is that poverty reduces average income available for redistribution, hence, reducing the marginal benefit of increasing the tax rate.

For fair voters, however, the results can go either way; part (f) gives the appropriate condition. The reason is that the fair voter, like the selfish voter, cares about own payoff. However, in addition, the fair voter also cares about the income of poorer voters. The interplay between these two opposing factors determines if the fair voter will respond, unlike the selfish voter, by redistributing more in response to poverty. From parts (h), (i), for fair voters, if α or β is sufficiently high, then empathy for the poorer voters (as well as envy towards richer voters) becomes stronger, which increases the tax rate and the ratio of social spending to GDP in response to increased poverty.

4.1. Social spending in recessions

The selfish and fair voter models alike, predict that a reduction in the skill of voters *above* the median will reduce the ratio of social spending to GDP (Remark 1 and Proposition

4(d)). However, the two models differ on what would happen to this ratio in response to a decline in the skill of voters *below* the median (equivalently, an increase in *poverty*).

The selfish voter model predicts that the ratio of social spending to GDP would *decline* (Proposition 4(g)). On the other hand, the fair voter model predicts that this ratio will *increase*, if α and/or β is sufficiently high (Proposition 4(h) and 4(i)).

Recall that, in our model, 'skill' is just a measure of the ability of a voter to translate labour time into income. We may, therefore, identify periods of high unemployment with episodes where the 'skills' of below median voters receive strong negative shocks. The selfish voter model would then predict a *decline* in the ratio of social spending to GDP, while the fair voter model would predict an *increase* in this ratio. Thus, the selfish voter model predicts *pro*cyclic movement of the social spending to GDP ratio, while the fair voter model predicts a *counter*cyclical movement. For the US data, the prediction of the selfish voter model is inconsistent with the evidence, while the prediction of the fair voter model is consistent with the evidence; see, for instance, Auerbach (2003).

5. Income distribution and the tax rate

In section 4, above, we investigated the effect of a change in the level of skill of one voter on the tax rate chosen by the median-skill voter (and, hence, also chosen by society). Such a change will, necessarily, change mean income. We saw the striking difference between a selfish median-skill voter and a fair median-skill voter in response to an increase in poverty as measured by a decline in the skill level of workers below the median. The former responds by choosing a *lower* tax rate while the latter, if he/she is fair enough, responds by choosing a *higher* tax rate (Proposition 4 (f)-(i)).

We are now interested in the effect on the tax rate of an increase in pretax income inequality that leaves mean income unchanged (the analogue of a mean preserving spread, when income is given by a probability distribution). We will see that a selfish median-skill voter responds to an increase in inequality by increasing the tax rate; and that a fair median-skill voter does so to an even greater degree (Proposition 6, below).

There are a number of measures of income inequality when the income distribution is discrete. For discussions of these, see, for example, Atkinson (1970), Marshall and Olkin (1979), Preston (1990, 2006) and Zheng (2007). Here we shall introduce a new measure which is more suitable for our purposes. We shall call it *median-dominance*.

Definition 2 *: (Median Dominance): Consider the set of vectors:*

$$\mathbf{I} = \left\{ \mathbf{x} : 0 < x_1 < x_2 < \ \dots \ < x_n, \ \frac{1}{n} \sum_{i=1}^{n} x_i = \mu \text{ and } x_m < \mu \right\}. \tag{5.1}$$

Let $\mathbf{x},\mathbf{y}\in\mathbf{I}$.
(a) If $x_m \geq y_m$*, we say that* $\mathbf{x}$ *median-dominates* $\mathbf{y}$*. If the inequality is strict, we say* $\mathbf{x}$ *strictly median-dominates* $\mathbf{y}$*.*
(b) Suppose $x_m \geq y_m$*,* $\sum\limits_{k>m}(x_k - x_m) \leq \sum\limits_{k>m}(y_k - y_m)$*,* $\sum\limits_{i<m}(x_m - x_i) \leq \sum\limits_{i<m}(y_m - y_i)$ *and, at least, one of these inequalities is strict. Then we say that* $\mathbf{x}$ *strongly median-dominates* $\mathbf{y}$*.*

Consider two distributions, $\mathbf{x}$ and $\mathbf{y}$, with the same mean and where the median is lower than the mean. Then $\mathbf{x}$ median-dominates $\mathbf{y}$ if, and only if, the median under $\mathbf{x}$ is no less than the median under $\mathbf{y}$, regardless of the other components of $\mathbf{x}$ and $\mathbf{y}$. $\mathbf{x}$ strictly median-dominates $\mathbf{y}$ if, and only if, the median under $\mathbf{x}$ is strictly higher than the median under $\mathbf{y}$, regardless of the other components of $\mathbf{x}$ and $\mathbf{y}$. These concepts are clearly relevant for the case of a selfish median voter. For the case of a fair median voter advantageous and disadvantageous inequalities also become important. Thus $\mathbf{x}$ strongly median-dominates $\mathbf{y}$ if, and only if, the following criteria hold: (i) the median under $\mathbf{x}$ is no less than that under $\mathbf{y}$, (ii) advantageous inequality, relative to the median, under $\mathbf{x}$ is no less than that under $\mathbf{y}$, (iii) disadvantageous inequality, relative to the median, under $\mathbf{x}$ is no worse than that under $\mathbf{y}$; and at least one of these three criteria holds with strict inequality.

Remark 3 *: Let* $\mathbf{x},\mathbf{y}\in\mathbf{I}$ *but not, necessarily,* $x_m < \mu$*. According to Marshall and Olkin (1979)* $\mathbf{x}$ *difference-dominates* $\mathbf{y}$ *if* $i<j \Rightarrow x_j - x_i \leq y_j - y_i$*,* $i,j=1,2,...,n$*. If one of these inequalities is strict, then* $\mathbf{x}$ *strictly difference-dominates* $\mathbf{y}$*. Comparing strict difference-dominance with strong median-dominance, we see that strict difference-dominance is weaker in the sense that it does not require that* $x_m < \mu$ *or that* $x_m \geq y_m$*. However, it is much stronger in the sense that it requires* $x_j - x_i \leq y_j - y_i$*, for each* $i,j=1,2,...,n$*,* $i<j$*. Thus if* $\mathbf{x}$ *strictly difference-dominates* $\mathbf{y}$ *and if, in addition,* $x_m < \mu$ *and* $x_m \geq y_m$*, then* $\mathbf{x}$ *strongly median-dominates* $\mathbf{y}$*.*

Example 1 *: Consider the three sets:*

$$\mathbf{x}=\{0.2,0.3,0.7\}\,,\ \mathbf{y}=\{0.1,0.25,0.85\}\,,\ \mathbf{z}=\{0.05,0.35,0.8\}\,.$$

Note that these three sets have the same mean:

$$\frac{1}{3}\Sigma_{i=1}^{3}x_i=\frac{1}{3}\Sigma_{i=1}^{3}y_i=\frac{1}{3}\Sigma_{i=1}^{3}z_i=0.4,$$

then
(a) $\mathbf{x}$ *strictly difference-dominates* $\mathbf{y}$ *and* $\mathbf{z}$*,*

(b) neither $\mathbf{y}$ *difference-dominates* $\mathbf{z}$ *nor does* $\mathbf{z}$ *difference-dominate* $\mathbf{y}$*,*
(c) $\mathbf{x}$ *strictly and strongly median-dominates* $\mathbf{y}$*,*
(d) $\mathbf{z}$ *strictly, but not strongly, median-dominates* $\mathbf{x}$ *and* $\mathbf{y}$*.*

Proposition 5 : *At any tax rate* $t < 1$*, the disposable (post tax) income vector,* $\mathbf{c}$*, strongly median dominates the factor (pretax) income vector,* $\mathbf{y}$*.*

Lemma 5, below, will be used in the proof of Proposition 6 that immediately follows it.

Lemma 5 : *(a) Let* $\mathbf{y}$ *be the vector of pretax incomes*[15] *(2.21). Let*

$$y = \frac{1}{n}\Sigma_{j=1}^{n} y_j - y_m + \frac{\alpha}{n-1}\sum_{k>m}(y_k - y_m) + \frac{\beta}{n-1}\sum_{i<m}(y_m - y_i),$$

then

$$t_m = \frac{1}{1 + \frac{\epsilon}{ny}\Sigma_{j=1}^{n} y_j}.$$

(b) Let $\mathbf{c}$ *be the vector of post tax incomes, i.e., the consumption vector. Let*

$$c = \frac{1}{n}\Sigma_{j=1}^{n} c_j - c_m + \frac{\alpha}{n-1}\sum_{k>m}(c_k - c_m) + \frac{\beta}{n-1}\sum_{i<m}(c_m - c_i),$$

then

$$t_m = \frac{nc}{\epsilon\Sigma_{j=1}^{n} c_j}.$$

Proposition 6 : *Let* $\mathbf{x}, \mathbf{y} \in \mathbf{I}$*, where* $\mathbf{I}$ *is a set of pretax or post tax incomes, as in Definition 2. Let* t_m^S *and* T_m^S *be the tax rates associated with* $\mathbf{x}, \mathbf{y}$*, respectively, when the median voter is selfish* ($\alpha = \beta = 0$)*. Let* t_m^F *and* T_m^F *be the tax rates associated with* $\mathbf{x}, \mathbf{y}$*, respectively, when the median voter is fair* ($0 < \beta < 1$, $\beta < \alpha$)*. Then*
(a) $t_m^S < t_m^F$ and $T_m^S < T_m^F$.
(b) If $\mathbf{x}$ *strictly median-dominates* $\mathbf{y}$*, then* $t_m^S < T_m^S < T_m^F$.
(c) If $\mathbf{x}$ *strongly median-dominates* $\mathbf{y}$*, then* $t_m^S < t_m^F < T_m^F$.

The existing literature ignores issues of fairness. To illustrate the pitfalls that this could lead to, we plot in Figure 5.1, the choice of the optimal tax rate chosen by the median voter (vertical axis) against both, fairness and inequality. In this exercise, an increase in inequality is generated by making the rich even richer. Section 4 shows that such an increase in inequality gives the same comparative static results in the fair and the selfish

[15] Pretax income is the income before the tax is paid. However, labor supply depends on the tax rate. Hence, for that reason, the right hand side of pretax income, in (2.21), depends on the tax rate.

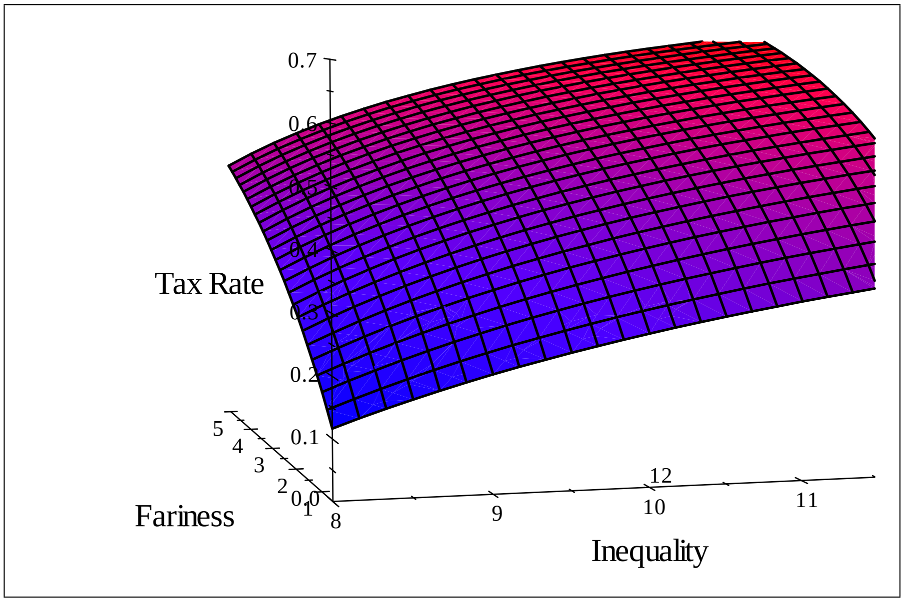

Figure 5.1: **The relation between redistribution, fairness and inequality**

voter models (see Proposition 4(d)). An increase in fairness is captured in this exercise by an increase in the parameter α of the FS preferences (see Proposition 4).

In actual practice, empirical researchers could be picking up any sequence of points along the surface in Figure 5.1. This practice is likely to lead to mixed and possible contradictory results. As the figure clearly shows, low-inequality and high-fairness countries have a similar level of redistribution as high-inequality and low-fairness countries. Not controlling for fairness would then lead to absurd results. However, controlling for fairness, a prediction of the model is that one should have greater redistribution where inequality is higher, although depending on the parameters, it is quite possible that such an effect turns out to be quite weak. To the best of our knowledge this test has not been carried out. This issue, we believe, could have seriously contaminated the existing literature's attempt at finding an empirical relation between inequality and the extent of redistribution.

6. An illustrative empirical exercise

In this section we test if redistribution is affected by (1) fairness, and (2) inequality when we use factor incomes rather than disposable incomes to generate the inequality measure. We use cross-country data for OECD economies for 2003[16]. The empirical exercise has no value beyond being merely illustrative because data on factor incomes is available only over

[16] The list of 20 countries that we use is as follows. Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and the United States.

a very short duration. This prevents one from conducting a more satisfactory econometric exercise that would, for instance, control for country-specific effects. Nevertheless, the exercise is highly suggestive of the important role played by fairness.

6.1. Description of the data

The list of variables and their explanation is as follows.

Dependent Variables: The dependent variable, denoted by SS/GDP, is social spending as a percentage of GDP at current prices taken from Lindert (2003)[17]. We have also tried as our dependent variable, tax revenues as a percentage of GDP, TR/GDP.

Fairness Variables: We have already discussed our motivations behind our fairness variable in sections 1.1, 1.5 in the introduction. The first fairness variable, denoted by ODA/GDP, is the ratio of 'Official development assistance' to GDP for each country. However, such assistance might also reflect motives other than fairness such as 'strategic giving', motivated by political considerations. For that reason, we use a second measure of aid, namely, 'multilateral aid as a percentage of GDP; this is denoted by MA/GDP. Data on these is available through OECD statistics[18]. Our third measure, denoted by QAA/GDP is 'quality adjusted aid' to GDP ratio drawn from an index compiled by Roodman (2005).[19]

Inequality Variables: We have noted above the argument against using disposable or post-tax incomes to construct inequality measures. However, factor income data is very difficult to obtain and till recently has been unavailable. Hence, most existing empirical work has used data from disposable income to measure inequality. There have been two main sources of data for measures of inequality, based on disposable income. The main source is the 'Luxembourg Income Study' (LIS), which collates micro-data from various OECD countries based on survey information.[20] The second main source of data is from the World Bank (e.g. the 2005 World Development Indicators, Table 2.7). We indicate the Gini coefficients obtained from these two data sources, respectively, by $Gini(LIS)$ and $Gini(World\ Bank)$. Disaggregated data is also available. LIS provides data on income by percentiles and the World Bank breaks down the income distribution into 5 parts. Using each of these more disaggregated data has its problems[21].

[17]The underlying source is OECD, *Social Expenditure Database* 1980-1996.

[18]Our source is 'OECD in Figures: Statistics on the Member Countries', 2004 edition.

[19]Aid is adjusted for quality in that it takes account, among other things, of (1) the recipients of aid (relatively prosperous eastern European countries or abysmally poor sub-saharan African countries) (2) tied versus untied aid (3) whether cancelled interest payments are counted as aid (4) quality of governance.

[20]The information is not contemporaneous. So, for instance, while the data for the Scandinavian countries, United Kingdom and Italy dates from 1995, that for Germany and France dates from 1994. See for instance, Figure 1 in Smeeding (2002). However, this might not be a particularly serious problem because income inequality moves relatively slowly.

[21]LIS provides data on the ratio of 90th percentile to the median percentile and the 10th percentile

Our point of departure from the existing literature in the use of inequality variables is to rely instead on the newly created dataset on factor incomes that has been made available in Milanovic (2000). Denote the Gini calculated on the basis of factor incomes as $Gini(Factor\ Income)$. We shall compare alternative regression specifications based on the various Gini coefficients.

Control Variables: The final two variables are control variables that are not necessarily related to inequality or to fairness. These are as follows. In line with several other empirical studies, our first variable is the proportion of population aged 65 or over, denoted by Pop_{65}. This takes account of transfers to the old. The second variable, denoted by D_{US}, is a dummy variable that takes a value 1 for United States and zero for all other countries in the sample. The reasons for including a US dummy (relative to Europe) can be found in several places in the literature and is referred to as 'American Exceptionalism'; for instance, Glaeser (2005).

6.2. Results

Tables I, II in Appendix 1 report the regression results. There are 20 observations and we report the results of robust OLS estimation in Stata[22]. The Akaike information criteria, AIC, and the Schwarz Bayesian information criteria, BIC, are used as specification tests (lower values of these two indicators reflect a better specification).

Table-I reports the relation between inequality and redistribution *in the absence of fairness concerns*. The three columns for results in the table correspond respectively to the use of $Gini$ (LIS), $Gini$ $(World\ Bank)$ and $Gini$ $(Factor\ income)$ as alternative proxies for income inequality. It is clear that the AIC and the BIC unambiguously pick the regression with $Gini$ $(Factor\ income)$ as the best specified. This is what one expects (see the introduction). In this regression (the last column), inequality has an insignificant effect on redistribution, while both controls are significant and have the correct signs.

Table-II uses the $Gini$ $(Factor\ income)$ as our inequality measure but *introduces three alternative notions of fairness*, respectively, MA/GDP, ODA/GDP and QAA/GDP.

to the median percentile. However, each of these variables are highly correlated (the correlation is about -0.85) hence they cannot be used simultaneously. The World Bank provides disaggregated information on the bottom 10% , top 10%, bottom 20%, next 20% and so on. However, where does the median voter lie among these? What would be an objective agglomeration of these figures when one imagines, say, society as comprised of three broad groups: poor, middle class and rich? To us, the answer to these questions is not clear. Hence, we focus only on the Gini coefficients. This might not be a bad approximation because the Gini is very highly correlated with the ratio of 90th percentile to the median percentile and with the 10th percentile to the median percentile.

[22] The Stata regress command includes a robust option for estimating the standard errors using the Huber-White sandwich estimators. This allows one to deal with problems about normality, heteroscedasticity, or some observations that exhibit large residuals. With the robust option, the point estimates of the coefficients are exactly the same as in ordinary OLS, but the standard errors take into account the issues mentioned above.

The regression corresponding to each is reported in columns 2, 3 and 4. The specification tests unambiguously pick out the first regression, using MA/GDP, as the best specified. This result is also along expected lines, since MA/GDP corrects for the strategic motive in giving aid. Relative to Table-I, the AIC and BIC are substantially lower for all regressions in Table-II. The fairness variables are all highly significant and have the correct signs. The intercept is no longer significant and the two controls have the correct sign and are generally significant. However, Pop_{65} is more significant across the regressions. The inequality variable is not significant in any of these regressions[23].

The results of our illustrative empirical exercise suggest the following. First, inequality measured on the basis of factor incomes (rather than disposable incomes) leads to a better specification. Second, the negative (and mostly significant) coefficient on the dummy variable for the US supports the idea of 'American Exceptionalism'. Third, the fairness variable is a very important determinant of redistribution. Certainly, based on the evidence, it is more important than the inequality variable, which the literature has focussed on so far.

7. Conclusions

We replace the self interested voters in the Romer-Roberts-Meltzer-Richard (RRMR) framework with voters who have a preference for fairness (as in Fehr-Schmidt, 1999) and ask the following questions. First, does a median voter equilibrium exist? Second, what is the affect on redistribution when the income distribution changes? Third, conditional on data limitations, is fairness an important factor in determining redistribution?

Our findings are as follows. The single crossing property of Gans-Smart (1996) can be used to demonstrate the existence of a Condorcet winner when voters are fair. Increased fairness leads to a more redistributive outcome. Fair voters, if they are very fair, will respond to poverty by redistributing more (and not less as the selfish voter model predicts). The ratio of social spending to GDP moves countercyclically in the fair voter model but pro-cyclically in the selfish voter model. The latter is not consistent with the evidence.

Our illustrative empirical exercise suggests that factor income inequality, which is the appropriate variable suggested by theory, outperforms disposable income inequality, which is largely used in existing empirical work. Fairness is a very significant variable in explaining redistribution in OECD economies.

[23] This contrasts our results with Milanovic (2000) who finds a positive and significant effect of factor income inequality on redistribution to the poor. However the spirit of the theoretical predictions is that everyone (rich, middle class and poor) receive a lumpsum transfer. For instance, expenditure on health that everyone benefits from but the poor possibly benefit more than the rich because they might not have access to private medical care. Hence, the measure of redistribution often used in the empirical literature i.e. social spending to GDP is the one we prefer to use.

8. Appendix 1: Fairness, Inequality and Redistribution

Table-I: Inequality and Redistribution

	1	2	3
Constant	35.96^{***} 2.87	24.79 1.29	13.22 0.74
Gini (LIS)	-110.21^{***} -4.43	-	-
Gini (World Bank)	-	-0.63^{*} -1.80	-
Gini (Factor Income)	-	-	-0.34 -1.13
D_{US}	3.03 1.28	-0.14 -0.06	-5.47^{***} -3.87
Pop_{65}	1.23^{**} 2.37	1.19^{*} 1.96	1.91^{***} 3.78
R^2	0.72	0.48	0.59
F	95.33	31.85	37.32
AIC	110.19	122.39	92.76
BIC	113.18	125.37	95.08
n	20	20	16

Note: $t-$values in parentheses. Superscript $*$, $**$, $***$ denotes significance at the 10%, 5% and 1% level respectively.

Table-II: Fairness, Inequality and Redistribution

	1 (MA/GDP)	2 (ODA/GDP)	3 (QAA/GDP)
Constant	7.54 0.40	4.70 0.21	6.00 0.30
Fairness	37.58^{***} 3.34	9.4^{*} 2.10	20.59^{***} 2.85
Gini (Factor Income)	-0.15 -0.42	-0.18 -0.45	-0.19 -0.51
D_{US}	-2.90 -1.75	-3.61^{**} -2.73	-3.85^{***} -3.33
Pop_{65}	1.23^{***} 4.82	1.63^{***} 5.79	1.58^{***} 6.03
R^2	0.80	0.72	0.71
F	92.41	65.20	0.55
AIC	85.20	88.93	89.10
BIC	89.06	92.01	92.19
n	16	16	16

Note: $t-$values in parentheses. Superscript $*$, $**$, $***$ denotes significance at the 10%, 5% and 1% level respectively.

9. Appendix 2: Proofs

Proof of Proposition 1: Consider voter j. Let $\mathbf{l}_{-j}$ be the vector of labour supplies of all other voters. Hence $u(l_i; t, b, s_i)$, $i \neq j$, are fixed numbers. Since $u(l_j; t, b, s_j)$ is continuous in l_j, and since $\max\{0, x\}$ is continuous in x, it follows that $U_j(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s})$, as given by (2.18), is a continuous function of $l_j \in [0,1]$. Hence, $U_j(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s})$ attains a maximum at some $l_j^* \in [0,1]$. We shall argue that l_j^* must maximize own-utility, $u(l_j; t, b, s_j)$. Let

$$\begin{aligned} A_j &= \{i : i \neq j \text{ and } u(l_i; t, b, s_i) \leq u(l_j; t, b, s_j)\}, \\ D_j &= \{k : k \neq j \text{ and } u(l_k; t, b, s_k) > u(l_j; t, b, s_j)\}, \end{aligned}$$

$$\begin{aligned} &\omega_{ji} = \tfrac{\beta}{n-1} > 0, \text{ for } i \in A_j, \\ &\omega_{jj} = 1 - \tfrac{|A_j|\beta}{n-1} + \tfrac{|D_j|\alpha}{n-1} > 0, \\ &\omega_{jk} = -\tfrac{\alpha}{n-1} < 0, \text{ for } k \in D_j. \end{aligned}$$

Then $U_j(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s})$ can be written as

$$U_j(l_j; \mathbf{l}_{-j}, t, b, \mathbf{s}) = \omega_{jj} u(l_j; t, b, s_j) + \sum_{i \in A_j} \omega_{ji} u(l_i; t, b, s_i) + \sum_{k \in D_j} \omega_{jk} u(l_k; t, b, s_k),$$

(note that A_j and D_j, and hence also ω_{ji}, ω_{jj} and ω_{jk}, are functions of l_j). In particular,

$$U_j\left(l_j^*; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) = \omega_{jj} u\left(l_j^*; t, b, s_j\right) + \sum_{i \in A_j} \omega_{ji} u(l_i; t, b, s_i) + \sum_{k \in D_j} \omega_{jk} u(l_k; t, b, s_k).$$

Suppose l_j^* does not maximize own-utility $u(l_j; t, b, s_j)$. Then we can find an l_j^{**}, sufficiently close to l_j^*, so that $u\left(l_j^{**}; t, b, s_j\right) > u\left(l_j^*; t, b, s_j\right)$ and the sets A_j and D_j are unchanged. Then

$$U_j\left(l_j^{**}; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) = \omega_{jj} u\left(l_j^{**}; t, b, s_j\right) + \sum_{i \in A_j} \omega_{ji} u(l_i; t, b, s_i) + \sum_{k \in D_j} \omega_{jk} u(l_k; t, b, s_k).$$

Hence, $U_j\left(l_j^{**}; \mathbf{l}_{-j}, t, b, \mathbf{s}\right) > U_j\left(l_j^*; \mathbf{l}_{-j}, t, b, \mathbf{s}\right)$, which cannot be, since l_j^* maximizes $U_j\left(l_j^*; \mathbf{l}_{-j}, t, b, \mathbf{s}\right)$. ■

Proof of Lemma 1: From (2.17) we see that, given t, b and s_i, $u(l_i; t, b, s_i)$ is a continuous function of l_i on the non-empty compact set $[0,1]$. Hence, a maximum exists. From (2.17), we get

$$\frac{\partial u}{\partial l_i}(l_i; t, b, s_i) = (1-t) s_i - l_i^{\frac{1}{\epsilon}}, \tag{9.1}$$

$$\frac{\partial u}{\partial l_i}(0; t, b, s_i) = (1-t) s_i, \tag{9.2}$$

$$\frac{\partial u}{\partial l_i}(1; t, b, s_i) = (1-t) s_i - 1, \tag{9.3}$$

$$\frac{\partial^2 u}{\partial l_i^2}(l_i; t, b, s_i) = -\frac{1}{\epsilon} l_i^{\frac{1-\epsilon}{\epsilon}}. \tag{9.4}$$

First, consider the case $t = 1$. From (2.17), or (9.1), we see that $u(l_i; 1, b, s_i)$ is a strictly decreasing function of l_i for $l_i > 0$. Hence, the optimum must be

$$l_i = 0 \text{ at } t = 1. \tag{9.5}$$

Now, suppose $t \in [0, 1)$. From (2.1), (2.2), (2.6), (9.2) and (9.3) we get that $\frac{\partial u}{\partial l_i}(0; t, b, s_i) > 0$ and $\frac{\partial u}{\partial l_i}(1; t, b, s_i) < 0$. Hence an optimal value for l_i must lie in $(0, 1)$ and, hence, must satisfy $\frac{\partial u}{\partial l_i}(l_i; t, b, s_i) = 0$. From (9.1) we then get

$$l_i = (1 - t)^{\epsilon} s_i^{\epsilon}, \tag{9.6}$$

which, therefore, must be the unique optimal labour supply (this also follows from (9.4)). For $t = 1$, (9.6) is consistent with (9.5). Hence, for each consumer, i, (9.6) gives the optimal labour supply for each $t \in [0, 1]$. ■

Proof of Lemma 2: The proof follows from (2.27) by direct calculation. ■

Lemma 6 *:* $\psi_1 > \psi_2 > ... > \psi_n$.

Proof of Lemma 6: (2.34), (2.35) and (2.36) give:

$$\psi_j - \psi_{j+1} = \left[1 - \frac{j\beta}{n-1} + \frac{(n-j)\alpha}{n-1}\right]\left(s_{j+1}^{1+\epsilon} - s_j^{1+\epsilon}\right) > 0,\ 1 < j < n-1, \tag{9.7}$$

$$\psi_{n-1} - \psi_n = \left(1 - \beta + \frac{\alpha}{n-1}\right)\left(s_n^{1+\epsilon} - s_{n-1}^{1+\epsilon}\right) > 0, \tag{9.8}$$

$$\psi_1 - \psi_2 = \left(1 - \frac{\beta}{n-1} + \alpha\right)\left(s_2^{1+\epsilon} - s_1^{1+\epsilon}\right) > 0, \tag{9.9}$$

(9.7), (9.8) and (9.9) can be combined to produce

$$\psi_j - \psi_{j+1} = \left[1 - \frac{j\beta}{n-1} + \frac{(n-j)\alpha}{n-1}\right]\left(s_{j+1}^{1+\epsilon} - s_j^{1+\epsilon}\right) > 0,\ 1 \leq j \leq n-1, \tag{9.10}$$

from which it follows that $\psi_1 > \psi_2 > ... > \psi_n$. ■

Proof of Proposition 2 (existence of optimal tax rates): (a) From (2.41), we see that $W_j(t, \alpha, \beta, \mathbf{s})$ is a continuous function of $t \in [0, 1]$, for $i = 1, 2, ..., n$. Hence, $W_j(t, \alpha, \beta, \mathbf{s})$ attains a maximum at some $t_j \in [0, 1]$. From (2.41), for the case $0 \leq t \leq 1, \epsilon = 1$ we get:

$$\frac{\partial W_j}{\partial t}(t, \alpha, \beta, \mathbf{s}) = (1 - t)\psi_j - t\overline{S} \tag{9.11}$$

whereas for the case $0 \leq t < 1, \epsilon < 1$ we get

$$\frac{\partial W_j}{\partial t}(t, \alpha, \beta, \mathbf{s}) = (1 - t)^{\epsilon}\psi_j - \epsilon t (1 - t)^{\epsilon - 1}\overline{S} \tag{9.12}$$

From (9.11) we see that, for $\epsilon = 1$, $\frac{\partial W_j}{\partial t}(1, \alpha, \beta, \mathbf{s}) = -\overline{S} < 0$. Hence, the optimal tax rate, t_j, for voter j satisfies $t_j < 1$. For $\epsilon < 1$, we see, from (9.12), that the limit of $\frac{\partial W_j}{\partial t}$, as $t \to 1$ (from below) is $-\infty$. Hence, again, the optimal tax rate, t_j, for voter, j satisfies $t_j < 1$.

(b) Suppose $j > \widehat{j}$. From (2.37) we get $\psi_j \leq 0$. Hence, from (9.11) and (9.12), we see that $\frac{\partial W_j}{\partial t} < 0$ for all $t \in [0, 1)$. Hence, the optimal tax rate, t_j, for voter j, must be $t_j = 0$.

(c) Suppose $j \leq \widehat{j}$. From (2.37) we get $\psi_j > 0$. Hence, from (9.11) and (9.12), we see that $\frac{\partial W_j}{\partial t} > 0$ at $t = 0$. Hence, the optimal tax rate, t_j, for voter j satisfies $t_j > 0$. Combining this with $t_j < 1$ (from (a)), we get that, necessarily, $\frac{\partial W_j}{\partial t} = 0$ at $t = t_j$. From (9.11) and (9.12) we then get (2.42). Since an optimum, t_j, exists (from (a)), since it must satisfy $\frac{\partial W_j}{\partial t} = 0$ and since the latter has the unique solution (2.42), it follows that (2.42) gives the unique global optimum (this can also be derived by showing that $\frac{\partial^2 W_j}{\partial t^2} < 0$ for $t \in (0, 1)$).

(d) Suppose $j \leq \widehat{j}$. From (2.36) we get

$$\frac{\partial \psi_j}{\partial \alpha} = S_j^- > 0, \tag{9.13}$$

$$\frac{\partial \psi_j}{\partial \beta} = S_j^+ > 0, \tag{9.14}$$

and, hence, from part (c), (2.42), (2.34) and (2.35), we get

$$\frac{\partial t_j}{\partial \alpha} = \frac{\epsilon \overline{S} S_j^-}{\left(\epsilon \overline{S} + \psi_j\right)^2} > 0, \tag{9.15}$$

$$\frac{\partial t_j}{\partial \beta} = \frac{\epsilon \overline{S} S_j^+}{\left(\epsilon \overline{S} + \psi_j\right)^2} > 0. \tag{9.16}$$

Hence, t_j is strictly increasing in α and β.

(e) From Lemma 6 and (2.42)

$$t_j - t_{j+1} = \frac{\epsilon \overline{S}\left(\psi_j - \psi_{j+1}\right)}{\left(\epsilon \overline{S} + \psi_j\right)\left(\epsilon \overline{S} + \psi_{j+1}\right)} > 0,\ j + 1 \leq \widehat{j}. \tag{9.17}$$

Combining (9.17) with parts (a), (b) and (c) gives $1 > t_1 > t_2 > ... > t_{\widehat{j}} > t_{\widehat{j}+1} = t_{\widehat{j}+2} = ... = 0$. In particular, if $\widehat{j} = n$, then $1 > t_1 > t_2 > ... > t_n > 0$. ■

Proof of Proposition 3: From (2.31) and Lemma 2(c) for $j = 1, .., n - 1$,

$$\frac{\partial V_{j+1}(t, b, \alpha, \beta, \mathbf{s})}{\partial t} - \frac{\partial V_j(t, b, \alpha, \beta, \mathbf{s})}{\partial t} = -(1 - t)^\epsilon \left[1 - \frac{j\beta}{n - 1} + \frac{(n - j)\alpha}{n - 1}\right]\left(s_{j+1}^{1+\epsilon} - s_j^{1+\epsilon}\right). \tag{9.18}$$

From (2.1), (2.19), (2.20) and (9.18), it follows that

$$\frac{\partial V_{j+1}(t,b,\alpha,\beta,\mathbf{s})}{\partial t}-\frac{\partial V_j(t,b,\alpha,\beta,\mathbf{s})}{\partial t}<0,\ t\in[0,1),\ j=1,2,...,n-1. \tag{9.19}$$

From (2.30)

$$\frac{\partial V_j(t,b,\alpha,\beta,\mathbf{s})}{\partial b}=\frac{\partial v_j}{\partial b}-\frac{\alpha}{n-1}\sum_{k>j}\left(\frac{\partial v_k}{\partial b}-\frac{\partial v_j}{\partial b}\right)-\frac{\beta}{n-1}\sum_{i<j}\left(\frac{\partial v_j}{\partial b}-\frac{\partial v_i}{\partial b}\right) \tag{9.20}$$

From Lemma 2(a) and (9.20)

$$\frac{\partial V_j(t,b,\alpha,\beta,\mathbf{s})}{\partial b}=1. \tag{9.21}$$

From (9.19) and (9.21) we get

$$-\frac{\partial V_j}{\partial t}/\frac{\partial V_j}{\partial b}\text{ is strictly increasing in }j. \tag{9.22}$$

From Lemma 3 and (9.22) we get that 'single-crossing' holds. Hence, from 4, we get that the median-voter is decisive, i.e., a majority chooses the tax rate that is optimal for the median-voter. ■

Proof of Proposition 4: Parts (a) and (b) follow from (2.39) and Proposition 2 (c) and (d). Part (c) then follows from part (b).

From (2.24), (2.34), (2.35), (2.36) and (4.1) we get the following:

$$\left(1+\frac{n\alpha}{n-1}\right)\overline{S}-\psi_m>0 \tag{9.23}$$

$$j>m\Rightarrow\frac{\partial t_m}{\partial s_j}=\frac{\left[\left(1+\frac{n\alpha}{n-1}\right)\overline{S}-\psi_m\right]\epsilon(1+\epsilon)s_j^{\epsilon}}{n\left(\epsilon\overline{S}+\psi_m\right)^2}>0, \tag{9.24}$$

$$[2(n-1)+n(\alpha-\beta)]\overline{S}+2\psi_m>0, \tag{9.25}$$

$$\frac{\partial t_m}{\partial s_m}=-\frac{\left\{[2(n-1)+n(\alpha-\beta)]\overline{S}+2\psi_m\right\}\epsilon(1+\epsilon)s_m^{\epsilon}}{2n\left(\epsilon\overline{S}+\psi_m\right)^2}<0, \tag{9.26}$$

$$j<m\Rightarrow\frac{\partial t_m}{\partial s_j}=-\frac{\left[\psi_m-\left(1-\frac{n\beta}{n-1}\right)\overline{S}\right]\epsilon(1+\epsilon)s_j^{\epsilon}}{n\left(\epsilon\overline{S}+\psi_m\right)^2}, \tag{9.27}$$

$$\begin{aligned}&\psi_m-\left(1-\frac{n\beta}{n-1}\right)\overline{S}\\=\ &\frac{\alpha}{n-1}\sum_{k>m}\left(s_k^{1+\epsilon}-s_m^{1+\epsilon}\right)+\frac{\beta}{n-1}\left[\Sigma_{i=1}^{n}s_i^{1+\epsilon}+\sum_{i<m}\left(s_m^{1+\epsilon}-s_i^{1+\epsilon}\right)\right]-s_m^{1+\epsilon}.\end{aligned} \tag{9.28}$$

From (9.24), we see that an increase in the skill of voter-workers above the median will increase the tax rate, whether the median voter is selfish or fair. This establishes part (d).

From (9.25) and (9.26), we see that an increase in the skill of the median skill voter-workers will reduce the tax rate, whether the median voter is selfish or fair. This establishes part (e). Part (f) follows from (9.27). For the special case of selfish voters, $\alpha = \beta = 0$ and $\phi_m = s_m$, from which part (g) follows. From (9.27) and (9.28) we see that a reduction in skill of a worker-voter below the median results in an increase in the tax rate if either α or β (or both) is sufficiently large. In particular, parts (h) and (i) follow as special cases from (9.27) and (9.28). ■

Proof of Proposition 5: Let $t < 1$. Let $\mathbf{y}$ be the factor (pretax) income vector and let $\mathbf{c}$ be the disposable (post tax) income vector. From (2.5), (2.16), (2.22), (2.23) and (5.1), we see that $\mathbf{c} \in \mathbf{I}$ and $\mathbf{y} \in \mathbf{I}$, with $\mu = \frac{1}{n}\Sigma_{i=1}^{n} c_i = \frac{1}{n}\Sigma_{i=1}^{n} y_i$. From (2.1), (2.5), (2.14), (2.21) and (2.25), we get:

$$c_m - y_m = t\left(1-t\right)^{\epsilon}\left(\overline{S} - s_m^{1+\epsilon}\right) > 0,$$

$$\sum_{k>m}\left(y_k - y_m\right) - \sum_{k>m}\left(c_k - c_m\right) = t\left(1-t\right)^{\epsilon}\sum_{k>m}\left(s_k^{1+\epsilon} - s_m^{1+\epsilon}\right) > 0,$$

$$\sum_{i<m}\left(y_m - y_i\right) - \sum_{i<m}\left(c_m - c_i\right) = t\left(1-t\right)^{\epsilon}\sum_{i<m}\left(s_m^{1+\epsilon} - s_i^{1+\epsilon}\right) > 0.$$

Hence, $\mathbf{c}$, strongly median dominates $\mathbf{y}$. ■

Proof of Lemma 5: Let

$$y = \frac{1}{n}\Sigma_{j=1}^{n} y_j - y_m + \frac{\alpha}{n-1}\sum_{k>m}\left(y_k - y_m\right) + \frac{\beta}{n-1}\sum_{i<m}\left(y_m - y_i\right) \tag{9.29}$$

and

$$c = \frac{1}{n}\Sigma_{j=1}^{n} c_j - c_m + \frac{\alpha}{n-1}\sum_{k>m}\left(c_k - c_m\right) + \frac{\beta}{n-1}\sum_{i<m}\left(c_m - c_i\right) \tag{9.30}$$

From (2.24), (2.34), (2.35) and (2.38), we get

$$\psi_m = \frac{1}{n}\Sigma_{i=1}^{n} s_i^{1+\epsilon} - s_m^{1+\epsilon} + \frac{\alpha}{n-1}\sum_{k>m}\left(s_k^{1+\epsilon} - s_m^{1+\epsilon}\right) + \frac{\beta}{n-1}\sum_{i<m}\left(s_m^{1+\epsilon} - s_i^{1+\epsilon}\right) \tag{9.31}$$

and, hence,

$$\begin{aligned}
\left(1-t_m\right)^{\epsilon}\psi_m &= \frac{1}{n}\Sigma_{j=1}^{n}\left(1-t_m\right)^{\epsilon} s_j^{1+\epsilon} - \left(1-t_m\right)^{\epsilon} s_m^{1+\epsilon} \\
&\quad + \frac{\alpha}{n-1}\sum_{k>m}\left[\left(1-t_m\right)^{\epsilon} s_k^{1+\epsilon} - \left(1-t_m\right)^{\epsilon} s_m^{1+\epsilon}\right] \\
&\quad + \frac{\beta}{n-1}\sum_{i<m}\left[\left(1-t_m\right)^{\epsilon} s_m^{1+\epsilon} - \left(1-t_m\right)^{\epsilon} s_i^{1+\epsilon}\right].
\end{aligned} \tag{9.32}$$

From (2.21), (9.29) and (9.32), we get

$$\left(1-t_m\right)^{\epsilon}\psi_m = y \tag{9.33}$$

From (2.21), (2.24), (4.1) and (9.33), we get

$$\begin{aligned} t_m &= \frac{\psi_m}{\frac{\epsilon}{n}\Sigma_{j=1}^n s_j^{1+\epsilon} + \psi_m} \\ &= \frac{(1-t_m)^\epsilon \psi_m}{\frac{\epsilon}{n}\Sigma_{j=1}^n (1-t_m)^\epsilon s_j^{1+\epsilon} + (1-t_m)^\epsilon \psi_m} \\ &= \frac{y}{\frac{\epsilon}{n}\Sigma_{j=1}^n y_j + y} = \frac{1}{1+\frac{\epsilon}{ny}\Sigma_{j=1}^n y_j}. \end{aligned} \tag{9.34}$$

This establishes part (a).

From (9.34) we get

$$t_m = \frac{n(1-t_m)\,y}{\epsilon\Sigma_{j=1}^n y_j}. \tag{9.35}$$

From (2.14), (9.29) and (9.30), and substituting t_m for t, we get

$$(1-t_m)\,y = c. \tag{9.36}$$

From (2.16), (9.35) and (9.36), we get

$$t_m = \frac{nc}{\epsilon\Sigma_{j=1}^n c_j},$$

which establishes part (b). ■

Proof of Proposition 6: Let

$$x^S = \mu - x_m, \tag{9.37}$$

$$y^S = \mu - y_m, \tag{9.38}$$

$$x^F = \mu - x_m + \frac{\alpha}{n-1}\sum_{k>m}(x_k - x_m) + \frac{\beta}{n-1}\sum_{i<m}(x_m - x_i), \tag{9.39}$$

$$y^F = \mu - y_m + \frac{\alpha}{n-1}\sum_{k>m}(y_k - y_m) + \frac{\beta}{n-1}\sum_{i<m}(y_m - y_i), \tag{9.40}$$

(I) First, we consider the case where $\mathbf{x}, \mathbf{y} \in \mathbf{I}$ are pretax incomes.

Since $\Sigma_{i=1}^n x_i = \Sigma_{i=1}^n y_i = \mu$ we get the following, from Lemma 5(a):

$$t_m^S = \frac{1}{1+\frac{\epsilon\mu}{nx^S}}, \tag{9.41}$$

$$T_m^S = \frac{1}{1+\frac{\epsilon\mu}{ny^S}}, \tag{9.42}$$

$$t_m^F = \frac{1}{1+\frac{\epsilon\mu}{nx^F}}, \tag{9.43}$$

$$T_m^F = \frac{1}{1 + \frac{\epsilon\mu}{ny^F}}. \tag{9.44}$$

From (9.37) and (9.39) it is clear that $0 < x^S < x^F$. Hence, from (9.41) and (9.43), it follows that $t_m^S < t_m^F$. Similarly, from (9.38), (9.40), (9.42) and (9.44), it follows that $T_m^S < T_m^F$. This establishes part (a).

Suppose $\mathbf{x}$ strictly median-dominates $\mathbf{y}$ (Definition 2a), it follows that $x_m > y_m$. Hence, from (9.37) and (9.38), it follows that $0 < x^S < y^S$. Hence, from (9.41) and (9.42), it follows that $t_m^S < T_m^S$. That $T_m^S < T_m^F$ has already been established in part (a). Hence, $t_m^S < T_m^S < T_m^F$. This establishes part (b).

Suppose $\mathbf{x}$ strongly median-dominates $\mathbf{y}$ (Definition 2b). It follows that $x_m \geq y_m$, $\sum_{k>m}(x_k - x_m) \leq \sum_{k>m}(y_k - y_m)$, $\sum_{i<m}(x_m - x_i) \leq \sum_{i<m}(y_m - y_i)$ and, at least, one of these inequalities is strict. Hence, from (9.39) and (9.40), it follows that $0 < x^F < y^F$. Hence, from (9.43) and (9.44), it follows that $t_m^F < T_m^F$. From this and part (a), we get $t_m^S < t_m^F < T_m^F$. This establishes part (c).

(II) Second, we consider the case where $\mathbf{x}, \mathbf{y} \in \mathbf{I}$ are post tax incomes. The proof in this case is similar to to that in case (I), except that (9.41)-(9.44) are replaced by $t_m^S = \frac{nx^S}{\epsilon\mu}$, $T_m^S = \frac{ny^S}{\epsilon\mu}$, $t_m^F = \frac{nx^F}{\epsilon\mu}$ and $T_m^F = \frac{ny^F}{\epsilon\mu}$, respectively, and part (b) of Lemma 5 is used instead of part (a). ■

References

[1] Ackert, L. F., J. Martinez-Vazquez, M. Rider (2004). Tax policy design in the presence of social preferences: Some experimental evidence. Georgia State University International Studies Program. Working Paper No. 04-25.

[2] Alesina, A. and G-M. Angeletos (2003). Fairness and Redistribution, mimeo. Harvard University.

[3] Atkinson A (1970) On the measurement of inequality. Journal of Economic Theory 2:244-263.

[4] Auerbach, A. (2003). Is there a role for discretionary fiscal policy? University of Michigan Law School, mimeo.

[5] Benabou R. (2000) Unequal societies: Income distribution and the social contract. The American Economic Review, vol.90, no.1, 96-129.

[6] Benabou R. and E A Ok (2001) Social mobility and the demand for redistribution: The POUM hypothesis. The Quarterly Journal of Economics, May, 447-487.

[7] Black, D. (1948). On the rationale of group decision making. Journal of Political Economy 56: 23-34.

[8] Bolton, E. B. and Ockenfels A. (2002) ERC: A theory of equity, reciprocity, and competition, American Economic Review, Vol. 90 pp. 166-193.

[9] Bolton, E. B. and Ockenfels A. (2006) Measuring Efficiency and Equity Motives - A Comment on "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments, American Economic Review, Vol. 96 pp. 1906-1911.

[10] Camerer, C. (2003). Behavioral Game Theory. Russell Sage Foundation.

[11] Dhami, S. and al-Nowaihi (2007). Fairness and direct democracy. University of Leicester Discussion Paper in Economics 06/11.

[12] Easterly, W. and S. Rebelo (1993). Fiscal policy and economic growth: An empirical investigation. Journal of Monetary Economics, 32: 417-58.

[13] Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition and cooperation. Quarterly Journal of Economics, 114, 817-868.

[14] Galasso, V. (2003). Redistribution and fairness: a note. European Journal of Political Economy, 19 (4), 885-892.

[15] Gans, J. S. and M. Smart (1996). Majority voting with single-crossing preferences. Journal of Public Economics, 59, 219-237.

[16] Glaeser, E. L. (2005). Inequality. NBER working paper no. 11511.

[17] Hindriks, J. and Myles, G. (2006) Intermediate Public Economics, MIT Press.

[18] Lindert, P. H. (2003). Whence the welfare state. Paper presented at the December 2003 conference on political institutions and economic policy, Harvard University.

[19] Lindert, P. H. (1996). What limits social spending? Explorations in Economic History, 33, 1-34.

[20] Marshall A, Olkin I (1979) Inequalities: theory of majorization and its applications. Academic, New York.

[21] Matsusaka, J. G. (2005a). Direct democracy works. Journal of Economic Perspectives. 19, 185-206.

[22] Matsusaka, J. G. (2005b). The eclipse of legislatures: Direct democracy in the 21st century. Public Choice, 124, 157-178.

[23] Meltzer, A. H. and S. F. Richard (1981). A rational theory of the size of government. Journal of Political Economy, 89, 914-927.

[24] Milanovic, B. (2000). The median voter hypothesis, income inequality, and income redistribution: an empirical test with the required data. European Journal of Political Economy, 16, 367-410.

[25] OECD Social expenditure database 1980-1996.

[26] OECD in figures: statistics on the member countries, 2004 edition.

[27] Perotti, E. C. (1996). Growth, income distribution and democracy: What the data say. Journal of Economic Growth, 1, 149-188.

[28] Persson, T. and G. Tabellini (2000). Political Economics. Cambridge, MA: MIT Press.

[29] Piketty, T. (1995). Social mobility and redistributive politics. Quarterly Journal of Economics, 110, 551-584.

[30] Preston, I. (1990) Ratios, differences and inequality indices. Mimeo., Nuffield College, Oxford.

[31] Preston, I. (2006) Inequality and Income gaps. Institute of Fiscal Studies Working Paper No. 06/25.

[32] Rabin, M. (1993). Incorporating fairness into game theory and economics. American Economic Review, 83, 1281-1302.

[33] Roberts, K. W. (1977). Voting over income tax schedules. Journal of Public Economics, 8, 329-340.

[34] Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. Journal of Public Economics, 4, 163-185.

[35] Roodman, D. (2005). An index of donor performance. Centre for Global Development Working paper No. 67.

[36] Smeeding, T. (2002). Globalization, inequality and the rich countries of the G-20: Evidence from the Luxembourg income study (LIS). Luxembourg Income Study Working Paper No. 320.

[37] Tyran, J-R. and R. Sausgruber (2006). A little fairness may induce a lot of redistribution in a democracy. European Economic Review, 50, 469-85.