Heterogeneous Social Preferences in a Model of Voting on Redistribution

Sabine Flamand*
Universitat Autònoma de Barcelona

May 14, 2012

Abstract

We introduce heterogeneous social preferences in a standard model of voting on a redistributive parameter in a direct democracy. In particular, and in accordance with experimental evidence, we assume that selfish, rawlsian and utilitarian voters coexist with given proportions. We characterize implicitly the unique political equilibrium of this economy, and prove its existence for any positively skewed income distribution. It turns out that the level of redistribution in the heterogeneous economy may be either lower or higher than in the selfish one. Furthermore, we show that slight variations in the relative proportion of a given type may lead to very important changes in the extent of redistribution, and we illustrate the implications this may have in the context of the political economy of border formation. Finally, we investigate the theoretical implications of the model regarding the link between inequality and redistribution, and show that it yields different predictions than the standard model with self-interested voters. In particular, an increase in poverty is very likely to increase redistribution. Furthermore, this is also true for a mean-preserving spread leaving the median income unaffected, although it has no effect whatsoever on redistribution in the traditional selfish economy.

Keywords: Social Preferences, Voting, Redistribution

JEL classification: D74, H56, H77

1 Introduction

Recent experimental evidence indicates that individual concerns for fairness and altruism can explain a range of economic phenomena which are not in accordance with the traditional selfishness assumption. Other-regarding preferences may take several forms: altruism, inequity aversion, and reciprocity (Fehr and Schmidt

*sabine.flamand@uab.cat. I wish to thank my advisor Caterina Calsamiglia for very helpful discussions, comments and suggestions.

(2006)). Furthermore, there is strong evidence that individual concerns for fairness are not homogeneous. Rather, some individuals may be more altruistic than others, some are simply totally selfish, but it may also well be the case that people derive utility from different types of altruism.

In this paper, we investigate this latter possibility by introducing heterogeneous social preferences in a standard model of voting on a redistributive parameter. More specifically, we assume that three different types of voters coexist with given proportions: selfish, rawlsian, and utilitarian. Tyran and Sausgruber (2006) and Ackert et al. (2007a and b) have shown in their experiments that individuals demonstrate concerns for others when voting for a redistributive policy. Furthermore, there is also strong evidence that a significant proportion of individuals exhibit either selfish, rawlsian or utilitarian preferences.

According to the experiments conducted by Andreoni and Miller (2002), people differ on whether they care about fairness at all, and when they do, the notion of fairness they employ differs widely, ranging from rawlsian to utilitarian. They report that 44% of their subjects are completely selfish, 35% exhibit egalitarian preferences, and 21% of the subjects can be classified as surplus maximizers. Fisman et al. (2007) and Iriberri and Rey-Biel (2008) also find similar proportions[1]. Bolton and Ockenfels (2002), investigating the trade-off between efficiency and equity in a voting experiment, find that, as a social good, equity is in greater demand than is efficiency. According to their results, about twice as many people deviate from pure self-interest for equity than for efficiency. Finally, Erlei (2008, p. 436), using a simple model which combines the basic ideas of different models of social preferences, and especially allowing for heterogeneity, applies it to 43 games and concludes that '*[...] models of social preferences are particularly powerful in explaining behavior if they are embedded in a setting of heterogeneous actors with heterogeneous (social) preferences*".

The standard approach on redistribution through the voting process comes from the models of Romer (1975), Roberts (1977) and Meltzer and Richards (1981). The commonly used name for this class of models is the RRMR model. It consists in a general equilibrium model which assumes purely selfish individuals who differ with respect to their productivity, which in turn determines their income. It predicts that the extent of redistribution is determined by the median-income voter, with higher-income individuals supporting strictly less redistribution than lower-income ones. More specifically, the median voter, being selfish, supports a positive level of redistribution only to the extent that he's poorer than average —and thus benefits himself from redistribution. Consequently, the model predicts that higher inequality, as measured by the mean-to-median income ratio, translates into more redistribution in equilibrium.

In this paper, we examine the theoretical implications of assuming heterogeneous social preferences in the latter traditional framework of voting on redistribution. In particular, we assume that selfish, rawlsian and utilitarian voters coexist with given proportions, and given weights associated to altruistic motives relatively to self-interest in their utility function. Given the significant experimental evidence highlighting the coexistence of those three distinct types of individuals, we believe it is important to investigate what kind of political equilibrium would arise from the interactions between them (if any), and what are its properties. That is, we aim at determining what are the equilibrium redistribution outcomes resulting from the prevalence of different altruistic motives among the voters. Observe that we abstract from the possibility of reciprocity, as we believe that such behavior is more likely to occur in strategic settings, where players

[1] Note that there is also some evidence that people may behave jealously, and that a minority of people is competitive. However, Erlei (2008), allowing for this latter type of preferences in his analysis of 43 games, finds that it does not improve the predictive success of the approach.

can directly affect each other's payoffs (in the case of bilateral interactions, for example), than in a voting context with a large electorate.

Our findings are as follows: first, we characterize implicitly the (strictly positive) Condorcet winner tax rate of the heterogeneous economy, and prove its existence for any positively skewed income distribution. Second, we show that, while an increase in the relative proportion of rawlsian (utilitarian) voters always yields more (less) redistribution, an increase in the relative proportion of selfish voters has an ambiguous effect on the equilibrium tax rate. Indeed, given that the redistributive preferences of the selfish voters have an intermediate position between the rawlsian and utilitarian ones for any level of income, it turns out that the level of redistribution in the heterogeneous economy may be either lower or higher than in the selfish one. Third, using simulations of the model, we show that varying slightly the relative proportion of a given type can have a very large impact on the redistributive outcome when voters attach a high relative weight to altruistic motives in their utility function. In turn, we illustrate with some examples the implications of this latter fact in the context of (the political economy of) border formation. In particular, we show that, even though there are no income differences whatsoever between regions, the fact that the relative proportion and/or distribution of a given type differs across regions is very likely to prevent voluntary regional integration. Finally, we show that the standard prediction regarding the link between inequality and redistribution in the model with selfish voters may be overturned when one allows for heterogeneous preferences among the population. In particular, an increase in poverty is very likely to increase redistribution. Furthermore, this is also true for a mean-preserving spread leaving the median income unaffected, although it has no effect whatsoever on redistribution in the traditional selfish economy.

A few theoretical papers have recently introduced social preferences into the standard model of voting on redistribution. Dhami and Al-Nowaihi (2010a, b, and c) also use the RRMR framework in order to allow for fairness, using the utility function proposed by Fehr and Schmidt (1999), reflecting two-sided self-centered inequality aversion. This means that a voter dislikes both being poorer and richer than other voters, as he suffers disutility from any income difference between himself and the other individuals (i.e. there is both envy and altruism). They show that the existence of fair voters always increases the equilibrium level of redistribution as compared to an economy in which voters are self-interested. Furthermore, they show that in economies where the majority of voters are selfish, the decisive policy may be chosen by fair voters, and vice-versa. Finally, they demonstrate that increased poverty may lead to more redistribution in equilibrium, which is also true in our setup where selfish, rawlsian and utilitarian voters coexist.

Using another approach, Tyran and Sausgruber (2006) also assume inequity aversion *à la* Fehr and Schmidt (1999) in order to study voting on redistribution. They test their predictions, and find that the model with fair voters predicts voting outcomes far better than the standard model of voting assuming rationality and strict self-interest.

Galasso (2003) also introduces rawlsian altruism into Meltzer and Richard's (1981) framework, and allows for the coexistence of both selfish and rawlsian voters. However, he assumes away the possibility of some agents behaving partly as surplus maximizers (i.e. utilitarians), although experimental evidence indicates that they form a significant part of the population. Interestingly, it turns out that the presence of utilitarian voters does not alter the (qualitative) link between inequality and redistribution as compared with an economy composed exclusively of selfish and rawlsian voters.

Finally, Luttens and Valfort (2010) assume that voters care about others in a mix between maximizing the surplus (utilitarian concern) and helping the worst-off (rawlsian concern) (Charness and Rabin (2002)). They show that when altruistic preferences are desert-sensitive, that is, when there is a reluctance to redistribute from the hard-working to the lazy, the political equilibrium is characterized by lower levels of redistribution.

The rest of the paper is structured as follows: in Section 2, we describe the model, derive the preferred tax rates of the three types of voters, and characterize implicitly the Condorcet winner tax rate of the heterogeneous economy. Furthermore, we perform some simulations of the model in order to investigate further the properties of the equilibrium level of redistribution. In Section 3, we investigate the link between inequality and redistribution in the heterogeneous economy using three alternative definitions of inequality. In Section 4, we illustrate some implications of the model in the context of (the political economy of) border formation. Finally, Section 5 concludes.

2 The Model

We introduce heterogeneous social preferences in the standard RRMR framework, in which the proceeds from a linear tax are used to finance equal per-capita transfers to all voters.

There is a continuum of individuals of mass 1. The voters are differentiated by their ability level, which is also their wage rate, denoted by ω. Furthermore, the abilities are distributed in the interval $[\underline{\omega}, \overline{\omega}] \subset \mathbb{R}^+$ according to the continuous cumulative distribution function $F(.)$. The distribution of abilities has mean $\widetilde{\omega}$, and, consistently with empirical evidence, is assumed to be skewed, so that $\omega_m < \widetilde{\omega}$. A production function transforms labor into a consumption good, according to the worker's ability: $y(\omega) = \omega l(\omega)$, where $l(\omega)$ represents the amount of labor supplied by the agent with ability ω.

Each individual is endowed with a fixed time endowment of one unit and supplies l units of labor, where $0 \leq l \leq 1$. Hence, the wage rate offered to each worker-voter coincides with the marginal product (i.e. the skill level ω). Therefore, the gross tax income of a voter is given by $y = \omega l$, and its budget constraint by

$$0 \leq c \leq (1-t)y + b = (1-t)\omega l + b$$

where $t \in [0,1]$ is the tax rate, and b is the uniform transfer given to each voter that equals the average tax proceeds, that is,

$$b = t\overline{y} = t \int_{\underline{\omega}}^{\overline{\omega}} \omega l(\omega) dF(\omega)$$

We consider a two-stage game. In the first stage, voters choose a tax rate, t, anticipating the outcome of the second stage. Consumers exhibit fairness by voting for the tax rate that would maximize social welfare as seen from their own perspective (see below). In the second stage, individuals choose their labor supply l so as to selfishly maximize their private utility. This determines the vector of labor supplies and indirect utilities.

2.1 Individual Choice of Labour Supply (Second-Stage Game)

Taking the redistributive policy of the government as given (i.e. t and b), labor supply is determined on the basis of private preferences. A voter has a private utility function $U(c, 1-l)$, over own consumption, c, and own leisure, $(1-l)$. All voters have the same private utility function, and thus they differ only in that they are endowed with different ability levels, ω. Following the literature, we assume that private utility is a quasi-linear function of the form[2]

$$U\left(c, 1-l\right) = c + u\left(1-l\right)$$

The optimization problem of a voter is given by

$$\underset{l}{Max}\; U\left(c, 1-l\right) \text{ such that } 0 \leq c \leq \left(1-t\right)\omega l + b$$

The economic optimization problem yields the usual result

$$\left(1-t\right)\omega = u'\left(1-l^{*}\right)$$

which implicitly defines l^{*} as a function of ω and t, and thus $y^{*} = \omega l^{*}$.

Following the literature, we assume that the quasi-linear utility function takes the following quadratic form in leisure:

$$U\left(c, 1-l\right) = c - \frac{1}{2}l^{2}$$

The utility function $U\left(c, 1-l\right)$ is twice differentiable, strictly concave in leisure, and the marginal utility of both consumption and leisure is positive:

- $\frac{\partial U(c,1-l)}{\partial c} = 1 > 0$
- $\frac{\partial U(c,1-l)}{\partial l} = -l < 0$
- $\frac{\partial^2 U(c(l),1-l)}{\partial l^2} = -1 < 0$

The first-order condition yields

$$l^{*} = \left(1-t\right)\omega$$

$$y^{*} = \omega l^{*} = \left(1-t\right)\omega^{2}$$

Hence, private preference satisfaction is measured by the indirect utility function υ:

$$\upsilon = (1-t)\omega l^{*} + b + u\left(l^{*}\right)$$

[2]Note that assuming such linearity in consumption has an important implication regarding the utilitarian voters. When an individual's utility function is linear in c, maximizing the sum/average of utilities (i.e. utilitarianism) is independent of any distributional concerns. In contrast, if one assumes a utility function that is strictly concave in c, maximizing the sum of utilities requires perfect equality of consumption among individuals. Besides the fact that such (quasi)-linearity is convenient for the tractability of the analysis, it also reflects the assumption according to which utilitarianism may only entail efficiency concerns.

$$\Leftrightarrow \upsilon(t,b,\omega) = \frac{1}{2}(1-t)^2\omega^2 + b$$

$$\Leftrightarrow \upsilon(t,\omega) = \frac{1}{2}(1-t)^2\omega^2 + t(1-t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

2.2 Voting for a Tax Rate (First-Stage Game)

All agents take their voting decisions by maximizing their indirect utility function. There is a proportion $\alpha \in (0,1)$ of the population that is selfish (S), and that maximizes

$$V^S(t,\omega) = \upsilon(t,\omega)$$

There is a proportion $\beta \in (0,1)$ of the population that is rawlsian (R), and that maximizes

$$V^R(t,\omega) = (1-\lambda^R)\upsilon(t,\omega) + \lambda^R \upsilon(t,\underline{\omega})$$

There is a proportion $\gamma \in (0,1)$ of the population that is utilitarian (U), and that maximizes

$$V^U(t,\omega) = (1-\lambda^U)\upsilon(t,\omega) + \lambda^U \int_{\underline{\omega}}^{\overline{\omega}} \upsilon(t,\omega)dF(\omega)$$

where $\lambda^i \in (0,1)$, $i = R, U$, are the relative weights associated to rawlsian and utilitarian motives, respectively. The indirect utility function of all types of voters is quasi-concave, and so preferences are single-peaked on the tax rate dimension. Proposition 1 below gives the preferred tax rate of each type of voter for a given ability level ω.

Proposition 1 (Preferred Tax Rates)**.**

1. *The preferred tax rate of a selfish voter with ability ω is given by*

$$t^S(\omega) = \frac{\overline{y} - y(\omega)}{2\overline{y} - y(\omega)}$$

2. *The preferred tax rate of a rawlsian voter with ability ω is given by*

$$t^R(\omega) = \frac{\overline{y} - \lambda^R y(\underline{\omega}) - (1-\lambda^R)y(\omega)}{2\overline{y} - \lambda^R y(\underline{\omega}) - (1-\lambda^R)y(\omega)}$$

3. *The preferred tax rate of a utilitarian voter with ability ω is given by*

$$t^U(\omega) = \frac{\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)}{2\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)}$$

Observe that selfish and utilitarian individuals never vote for a positive tax rate when their income is above average. In contrast, the rawlsian tax rate is strictly positive whenever $y(\omega) < \frac{\overline{y} - \lambda^R y(\underline{\omega})}{1-\lambda^R}$, which is strictly higher than $\overline{y}$ for $\lambda^R > 0$. Hence, even though a rawlsian individual is richer than average —and so does not benefit from redistribution himself— he may still vote for a positive level of redistribution, since he derives utility from giving away resources towards the poorest individual in the economy.

Proposition 2 (Tax Rates' Ordering)**.** *For a given ability level ω, the following relationship holds:*

$$t^U(\omega) \leqslant t^S(\omega) \leqslant t^R(\omega)$$

Clearly, for three distinct types of voters with an identical ability level ω, the one supporting the highest tax rate is rawlsian, while the one supporting the smallest tax rate is utilitarian (see Figure 1). As the rawlsian voter enjoys utility from redistributing to the poorest individual in the economy, he's also the one valuing redistribution the most. In contrast, utilitarian voters do not care about redistribution *per se*, and thus, as selfish individuals, they only vote for a positive tax rate to the extent that they benefit themselves from redistribution (i.e. $y(\omega) < \overline{y}$). However, since they are (partly) surplus maximizers, they support a strictly smaller tax rate than selfish voters for any ability level ω.

Proposition 3 below gives some comparative statics results regarding the preferred tax rate of the three types of voters.

Proposition 3 (Comparative Statics)**.** *Suppose that $t^i > 0$, $i = S, R, U$. Then it holds that:*

1. *The three tax rates are decreasing in own income $y(\omega)$, and increasing in average income $\overline{y}$*
2. *The rawlsian tax rate is increasing in λ^R*
3. *The utilitarian tax rate is decreasing in λ^U*

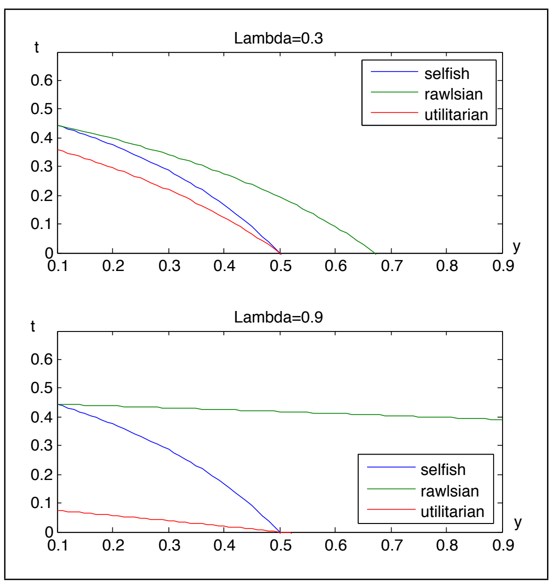

Figure 1: Tax rates versus income

As for selfish voters (and thus as in the standard Meltzer and Richard's (1981) framework), rawlsian and utilitarian altruists vote for lower tax rates when their own income increases, and for higher tax rates when the average income increases. Indeed, if $y(\omega)$ decreases and/or $\overline{y}$ increases, a given voter with ability level ω gets relatively poorer, so that he prefers a strictly higher tax rate independently of his type.

Furthermore, rawlsian altruists vote for higher tax rates when the weight they associate to the maximin criterion increases, since they derive more utility from redistributing to the poorest individual. For a utilitarian/surplus maximizer voter, the opposite holds: the higher λ^U, the smaller his preferred tax rate. Indeed, when λ^U increases, (poorer-than-average) utilitarian voters care relatively more about minimizing the distortions associated to taxation, so that they correspondingly vote for smaller tax rates.

Finally, observe that the higher the weights λ^U and λ^R, the bigger the difference between the preferred tax rates of the three types for a given ability level ω (see Figure 1).

The relationships between the incomes of the three types of individuals when they vote for the same tax rate t are the following:

$$y^S(t) = y^R(t) + \lambda^R \left[y\left(\underline{\omega}\right) - y^R(t) \right]$$

$$y^S(t) = y^U(t) + \lambda^U \left[\overline{y} - y^U(t) \right]$$

$$y^R(t) = \tfrac{(1-\lambda^U)}{(1-\lambda^R)} y^U(t) + \tfrac{\lambda^U}{(1-\lambda^R)} \overline{y} - \tfrac{\lambda^R}{(1-\lambda^R)} y\left(\underline{\omega}\right)$$

As can be seen in Figure 1, and as already noted, selfish and utilitarian individuals never vote for a positive tax rate when their income is above average. Furthermore, observe that a utilitarian individual votes at most for a tax rate of $\frac{1-\lambda^U}{2-\lambda^U}$. Therefore, we have the following lemma:

Lemma 1. *Assume that $\lambda^R = \lambda^U = \lambda$. If $t^S = t^R = t^U = t \in \left(0, \frac{1-\lambda}{2-\lambda}\right)$, then it holds that*

$$y^U(t) < y^S(t) < y^R(t)$$

Clearly, when the three types of voters support the same (positive) tax rate, the richest individual is rawlsian, while the poorest one is utilitarian. The utilitarian voter, as he cares about the maximization of the surplus, is consequently the unique poorer-than-average type bearing a cost from taxing the rich. Hence, as can be seen in Figure 1, when a utilitarian and a selfish voter have the same (below average) income, the selfish voter supports a strictly higher tax rate than the utilitarian one. We saw that the maximum value of the utilitarian tax rate is $\frac{1-\lambda}{2-\lambda}$. Clearly, this value is decreasing in the parameter λ. At the limit, when $\lambda = 0$, utilitarians become selfish, and hence vote at most for a tax rate of $1/2$. Conversely, if $\lambda = 1$, that is, if utilitarians only care about surplus maximization, they are against any positive level of redistribution, so that $t^U(\omega) = 0$ for any ability level ω.

2.3 Political Equilibrium

In order to find the political equilibrium resulting from the interaction of selfish, rawlsian and utilitarian voters, we need to identify the Condorcet winner tax rate of this economy. In other words, we need to find the tax rate corresponding to the median, that is, such that half of the preferred tax rates are below it and the other half of the preferred tax rates are above it. In order to do so, we first express the three types' incomes as a function of their preferred tax rates:

$$y^S(t) = \frac{(1-2t)}{(1-t)} \overline{y}$$

$$y^R(t) = \frac{(1-2t)}{(1-t)(1-\lambda^R)}\overline{y} - \frac{\lambda^R}{(1-\lambda^R)}y(\underline{\omega})$$

$$y^U(t) = \frac{(1-2t)}{(1-t)(1-\lambda^U)}\overline{y} - \frac{\lambda^U}{(1-\lambda^U)}\overline{y}$$

As there is a one-to-one correspondence between the preferred tax rate and the income for each type, it is equivalent to identify the median income corresponding to the median tax rate. Therefore, the Condorcet winner tax rate is given by

$$t^* = \left\{ t | \alpha F\left(y^S(t)\right) + \beta F\left(y^R(t)\right) + \gamma F\left(y^U(t)\right) = \frac{1}{2} \right\}$$

$$\Leftrightarrow t^* = \left\{ t | \Phi(t) = \frac{1}{2} \right\}$$

In order to have an interior solution for the equilibrium tax rate of this economy, it must hold that

$$\begin{cases} \Phi(t) < \frac{1}{2} \text{ for some } t \in \left[0, \frac{1}{2}\right] \\ \Phi(t) > \frac{1}{2} \text{ for some } t \in \left[0, \frac{1}{2}\right] \\ \Phi(t) \text{ is continuous and monotonic on the interval } \left[0, \frac{1}{2}\right] \end{cases}$$

From the following lemma, it is direct that monotonicity is satisfied:

Lemma 2. *$\Phi(t)$ is strictly decreasing in t for all $t < \frac{1}{2}$.*

For the two extreme values of t, we have that

1. $y^S(0) = \overline{y}$, $y^R(0) = \frac{\overline{y} - \lambda^R y(\underline{\omega})}{1-\lambda^R} > \overline{y}$, and $y^U(0) = \overline{y} \Rightarrow \Phi(0) > 0$
2. $y^S\left(\frac{1}{2}\right) = 0$, $y^R\left(\frac{1}{2}\right) < 0$, and $y^U\left(\frac{1}{2}\right) < 0 \Rightarrow \Phi\left(\frac{1}{2}\right) = 0$

Hence, we have the following result:

Proposition 4 (Existence of the Political Equilibrium)**.**
If the distribution of abilities $F(\omega)$ is continuous and positively skewed, there exists a (unique) Condorcet winner tax rate $t^* > 0$, which is implicitly defined by $\Phi(t^*) = 1/2$.

Proof. Given that $y^i(t)$, $i = S, R, U$, are continuous in their domain, and $F(.)$ is continuous, $\Phi(t)$ is also continuous. Then, given that the distribution of abilities $F(\omega)$ is positively skewed, and that gross income $y = \omega l = (1-t)\omega^2$, it follows directly that the gross income distribution $F(y(\omega))$ is positively skewed as well, so that $y_m < \overline{y}$. Therefore, we have that $F(\overline{y}) > 1/2$, and thus

$$\Phi(0) = \alpha F(\overline{y}) + \beta F\left(\frac{\overline{y} - \lambda^R y(\underline{\omega})}{1-\lambda^R}\right) + \gamma F(\overline{y}) > \frac{1}{2}$$

Hence, we have that

1. $\Phi(0) > \frac{1}{2}$

2. $\Phi\left(\frac{1}{2}\right) = 0$

3. $\Phi(t)$ is continuous and strictly decreasing in t (Lemma 2)

and so there exists a unique value of t such that $\Phi(t) = 1/2$. Finally, any individual —independently of his type— with income $y(\omega) < \overline{y}$ always supports a tax rate that is strictly positive, from which it follows that $t^* > 0$.

□

From Lemma 1, we can identify which voter, within each type, votes for the Condorcet winner tax rate t^* of the economy (i.e. the decisive income, within each type):

$$y^U(t^*) < y^S(t^*) < y^R(t^*) \text{ for } t^* \in \left(0, \frac{1-\lambda}{2-\lambda}\right)$$

and

$$y^S(t^*) < y^R(t^*) \text{ for } t^* \in \left(\frac{1-\lambda}{2-\lambda}, \frac{1}{2}\right) \text{ and no utilitarian has } t^* \text{ as a preferred tax rate}$$

In order to perform comparative statics regarding the equilibrium value of the redistributive parameter, we do the following transformation on the proportion of each type, so that the effect of a marginal increase in a given proportion can be determined: the Condorcet winner tax rate is given by

$$t^* = \left\{t|\alpha F\left(y^S(t)\right) + \beta F\left(y^R(t)\right) + \gamma F\left(y^U(t)\right) = \frac{1}{2}\right\}$$

Equivalently, we can write

$$t^* = \left\{t|\frac{\alpha}{\alpha+\beta+\gamma} F\left(y^S(t)\right) + \frac{\beta}{\alpha+\beta+\gamma} F\left(y^R(t)\right) + \frac{\gamma}{\alpha+\beta+\gamma} F\left(y^U(t)\right) = \frac{1}{2}\right\}$$

Proposition 5 (Comparative Statics)**.** *For an economy composed of selfish, rawlsian and utilitarian voters with respective proportions α, β, and γ, the Condorcet winner tax rate is increasing in λ^R, and decreasing in λ^U. Furthermore, it is increasing in β, decreasing in γ, while the effect of an increase in α is ambiguous.*

As expected, the equilibrium tax rate is increasing in the intensity of rawlsian altruism, while it is decreasing in the intensity of utilitarian altruism. Furthermore, an increase in the proportion of rawlsian voters increases the equilibrium level of redistribution, whereas an increase in the proportion of utilitarian voters has the opposite effect. Finally, an increase in the relative proportion of selfish individuals has an ambiguous effect on the equilibrium tax rate. As we saw, the preferred tax rate of a selfish voter has an intermediate position between the rawlsian and the utilitarian tax rates, for a given ability level (Proposition 2). Furthermore, an increase in the relative proportion of selfish individuals translates into an equivalent decrease in the relative proportion of one of the other two types, or both. Hence, the total effect of an increase in α depends on the distance between the selfish tax rates and the utilitarian and rawlsian ones (and thus on λ^U and λ^R), as well as on how this change in α affects the two other remaining proportions (i.e. β and γ).

2.4 Simulations

In order to investigate further the properties of the equilibrium tax rate of this economy, we simulate the model assuming that individual (gross) income is lognormally distributed, that is,

$$y(\omega) \backsim \log N\left(\mu, \sigma^2\right)$$

where

$$mean\,(y(\omega)) = e^{\mu+\frac{\sigma^2}{2}},\; var\,(y(\omega)) = (e^{\sigma^2}-1)e^{2\mu+\sigma^2} \text{ and } med\,(y(\omega)) = e^{\mu}$$

The lognormal is a good approximation of empirical income distributions, leads to tractable results, and allows for an unambiguous definition of inequality (Benabou (2000)). We set $mean\,(y(\omega)) = 0.5$ and $var\,(y(\omega)) = 0.1$, so that the ratio of median-to-mean income is around 0.85.

Figure 2 depicts the Condorcet winner tax rate for $\lambda^i \in [0.1, 0.9]$, $i = R, U$, assuming the following relative proportions of selfish, rawlsian and utilitarian voters: $(\alpha, \beta, \gamma) = (0.44, 0.35, 0.21)$. These proportions are the ones that have been found experimentally by Andreoni and Miller (2002)[3]. As can be seen in the figure, t^* reaches its maximum value for $(\lambda^R, \lambda^U) = (0.9, 0.1)$ (i.e. $t^* = 0.3$), and it reaches its minimum value for $(\lambda^R, \lambda^U) = (0.1, 0.9)$ (i.e. $t^* = 0.05$). Therefore, even though the selfish individuals constitute the biggest group in the population (i.e. $\alpha = 0.44$), variations in the altruistic weights lead to very important corresponding changes in the redistribution level.

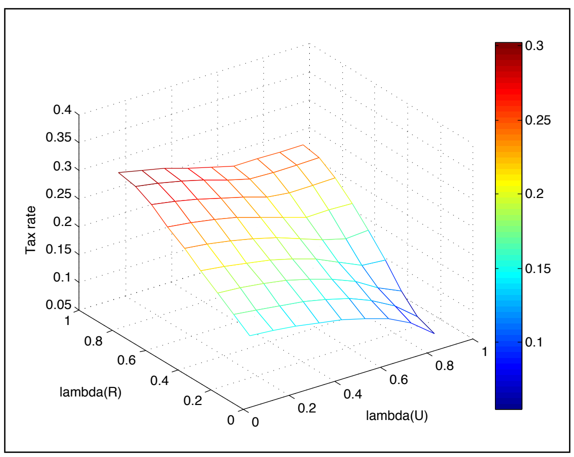

Figure 2: Condorcet winner tax rate as a function of the altruistic weights λ^R and λ^U

In Figure 3, we illustrate a rather extreme situation in which both utilitarian and rawlsian voters attach a very high relative weight to altruistic motives (i.e. $\lambda^U = \lambda^R = 0.9$). In this case, the equilibrium tax rate is equal to 0.41 when there are no utilitarian voters in the economy (i.e. $\alpha = \beta = 0.5$). Then, when the proportion of rawlsian voters decreases, the equilibrium tax rate decreases quickly to reach much lower levels, especially for low values of α. For instance, when $(\alpha, \beta, \gamma) = (0.1, 0.5, 0.4)$, the Condorcet winner tax rate of the economy is around 40%. Then, if the proportion of rawlsian voters decreases slightly to 0.35, with an

[3]The experiments conducted by Fismal et al. (2007) and Iriberri and Rey-Biel (2008) give similar proportions.

equivalent increase in the proportion of utilitarian voters to 0.55, the equilibrium tax rate falls dramatically to reach a value below 5%.

Observe that for $\lambda^U = \lambda^R = 0.9$, the poorest utilitarian individual votes (approximatively) for $t = 0.08$. When $(\alpha, \beta, \gamma) = (0.1, 0.5, 0.4)$, the fact that $t^* = 0.4$ means that the proportion of both selfish and rawlsian voters who support $t^* > 0.4$ constitutes exactly one half of the population. Then, if the relative proportions change slightly to $(\alpha, \beta, \gamma) = (0.1, 0.35, 0.55)$, utilitarians are consequently needed in order to form a majority, and thus t^* has to decrease. In particular, t^* has to be smaller than 0.08, so that at least part of the utilitarians can be "attracted".

Therefore, it turns out that slight changes in the relative proportions of the respective types (selfish, rawlsian, utilitarian) can lead to very important variations in the extent of redistribution. Furthermore, this gets increasingly likely the bigger the altruistic weights in the voters' utility function.

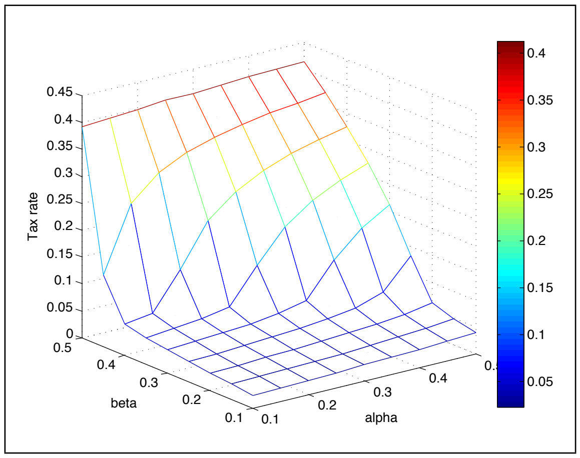

Figure 3: Condorcet winner tax rate as a function of the proportions α and β

3 The Link Between Inequality and Redistribution

The RRMR model predicts that higher inequality (lower median-to-mean income ratio) implies higher redistribution. Indeed, with selfish voters, an increase in inequality is relevant only to the extent that it concerns the relative position of the median voter. Empirically, this prediction remains controversial. Borck (2007, p. 96), in his survey on voting, inequality and redistribution, discusses this issue and concludes that " *[...] the RRMR hypothesis of a link between inequality and the size of the government has met with mixed empirical evidence*". In this section, we explore how the introduction of heterogeneous social preferences in the RRMR framework affects the predicted link between income inequality and redistribution.

We borrow the definition of inequality to Galasso (2003), who considers three types of income inequality: the poor getting poorer, the rich getting richer, and a mean-preserving spread.

Definition 1. *Let $F(.)$ be the original (positively skewed) cumulative distribution function of income, and let $G(.)$ be the (positvely skewed) cumulative distribution function of income after the change in inequality. Furthermore, let $y(\omega_1)$ and $y(\omega_2)$ be such that*

$$y(\underline{\omega}) < y(\omega_1) < \overline{y} < y(\omega_2) < y(\overline{\omega})$$

There are three types of income inequality:

1. *The poor get poorer if* $\begin{cases} F(y) = G(y) \text{ for } y > y(\omega_1) \\ F(y) \leq G(y) \text{ for } y \leq y(\omega_1) \end{cases}$ *and* $y(\underline{\omega})$ *and* $\overline{y}$ *decrease*

2. *The rich get richer if* $\begin{cases} F(y) = G(y) \text{ for } y < y(\omega_2) \\ F(y) \geq G(y) \text{ for } y \geq y(\omega_2) \end{cases}$ *and* $y(\overline{\omega})$ *and* $\overline{y}$ *increase*

3. *There is a mean-preserving spread if*

$$\begin{cases} F(y) \leq G(y) \text{ for } y \leq y(\omega_1) \\ F(y) = G(y) \text{ for } y(\omega_1) \leq y \leq y(\omega_2) \\ F(y) \geq G(y) \text{ for } y \geq y(\omega_2) \end{cases} \quad \text{and } y(\underline{\omega}) \text{ decreases, } y(\overline{\omega}) \text{ increases, with } \overline{y} \text{ unchanged}$$

In the traditional RRMR framework with selfish voters, and given that $y_m < \overline{y}$, the equilibrium level of redistribution always increases when the rich get richer (according to the above definition), since the median voter is relatively poorer. Then, a mean-preserving spread leaving the median income unaffected (i.e. $y_m > y(\omega_1)$) has no effect on the equilibrium tax rate. Finally, if the poor get poorer, the equilibrium level of redistribution decreases provided that $y_m > y(\omega_1)$, since the median voter is relatively richer (i.e. $\overline{y}$ decreases).

The following proposition describes the effect of an increase in the three types of inequality on the Condorcet winner tax rate of the economy, and shows that it yields different predictions than for the case of purely self-interested voters.

Proposition 6. *Let t^* be the equilibrium tax rate under $F(.)$, and let t^{**} be the equilibrium tax rate under $G(.)$. In an economy with selfish, rawlsian and utilitarian voters with respective proportions α, β, and γ,*

1. *If the poor get poorer, the effect on the equilibrium tax rate is ambiguous*
2. *If the rich get richer and $y^R(t^*) < y(\omega_2)$, then $t^{**} > t^*$*
3. *If there is a mean-preserving spread and $y^R(t^*) < y(\omega_2)$, then $t^{**} > t^*$*

If the rich get richer, and provided that $y^R(t^*) < y(\omega_2)$, none of the decisive voters within each type is directly affected by the rise in inequality. Therefore, in that case, the only relevant effect is the corresponding increase in the average income $\overline{y}$, which translates into a bigger mass of voters supporting $t > t^*$, so that t^* increases.

If there is a mean-preserving spread and $y^S(t^*) \leqslant y(\omega_1)$, the mass of selfish voters who supports $t > t^*$ increases, while it remains the same if $y^S(t^*) > y(\omega_1)$, since the average income $\overline{y}$ is unaffected by the rise in inequality. The same is true for utilitarians. Then, if $y^R(t^*) < y(\omega_2)$, although the rawlsian decisive voter is not affected directly by the rise in inequality, the mass of rawlsian voters supporting $t > t^*$ increases, since the income of the poorest individual has decreased and $\frac{\partial y^R(t^*)}{\partial y(\underline{\omega})} < 0$. Therefore, when there is a mean-preserving spread and $y^R(t^*) < y(\omega_2)$, t^* has to increase.

If the poor get poorer, things are slightly more complex. If $y^S(t^*) > y(\omega_1)$, the selfish decisive voter is not affected directly by the rise in inequality, and the only relevant effect is the corresponding decrease in the average income $\overline{y}$, so that the mass of selfish voters supporting $t > t^*$ decreases. Conversely, if $y^S(t^*) \leqslant y(\omega_1)$, the effect is ambiguous, since besides the decrease in $\overline{y}$, the selfish decisive voter is now poorer (in absolute terms), which makes him value strictly more redistribution. The same is true for utilitarians. Then, if $y^R(t^*) > y(\omega_1)$, the rawlsian decisive voter may prefer either a lower or a higher tax rate, since the effect of the decrease in $y(\underline{\omega})$ and $\overline{y}$ go in opposite directions. The same is true if $y^R(t^*) \leqslant y(\omega_1)$, although in this case the positive effect on t^* is reinforced, since the decisive rawlsian voter is now poorer (in absolute terms). Hence, altogether, the effect on the equilibrium tax rate t^* when the poor get poorer is ambiguous, since the mass of voters supporting $t > t^*$ may either increase or decrease.

4 An Application to the Political Economy of Border Formation

Bolton and Roland (1997) have shown that income-based redistribution has two effects on the incentives for a given region to secede from a union: a political effect, as the regional and national median incomes differ, and a tax base effect, as average income differs between regions. The political effect reflects differences in preferences for redistribution, and always induces a given region to secede, independently of the existence of interregional transfers. Such transfers arise when regional average incomes differ, and typically induce richer regions to secede (the tax-base effect).

In this section, we illustrate with a few examples the fact that even though there are no income differences whatsoever between regions (and thus no political nor tax-base effect), a region may still prefer to secede from a union provided that its distribution of types (i.e. selfish, rawlsian, utilitarian) is different than the one of the union. Furthermore, when there are income differences between regions regarding their average and/or median income, this "type distribution" effect then interacts with the two above-mentioned "income distribution" effects in non-trivial ways so as to shape the regional incentives to (voluntarily) form part of a political union.

Assume there are three regions A, B, C, and three income types (i.e. individulas) in each region given by $y_1^i < y_2^i < y_3^i$, $i = A, B, C$. Assume, furthermore, that the (gross) income distribution in each region is positively skewed, that is,

$$y_m^i = y_2^i < \frac{(y_1^i + y_2^i + y_3^i)}{3} = \overline{y}^i$$

so that there is a positive level of redistribution in all the three regions when they are independent. Finally, assume that individuals can be either selfish (S), rawlsian (R), or utilitarian (U) (with $\lambda^U = \lambda^R = \lambda$), and let the preferred tax rate of a type k individual ($k = S, R, U$) with income y_j ($j = 1, 2, 3$) in region i ($i = A, B, C$) be given by

$$t^k\left(y_j^i\right)$$

Example 1 (Selfish Economy - No Income Differences Between Regions)**.** *Suppose* $y_j^A = y_j^B = y_j^C = y_j$, $j = 1, 2, 3$. *That is, there are no income differences between regions. Suppose, furthermore, that all individuals* *are selfish. In this case, in each region, it holds that $t^S(y_1) > t^S(y_2) > t^S(y_3) = 0$, and $t^S(y_2)$ is implemented in each one of them if there are independent. If the three regions unify, then, the median tax rate is the one preferred by the median-income class, and thus coincides with the equilibrium tax rate in each region under independence. Since all regions are equally rich, and given that the median income is the same across regions, all voters are indifferent between independence and unification.*

Example 2 (Selfish and Utilitarian Economies - No Income Differences Between Regions)**.** *Suppose now that $y_j^A = y_j^B = y_j^C = y_j$, $j = 1, 2, 3$, but individuals in region A are utilitarian altruists, while individuals in regions B and C are selfish. In this case, we have the following ordering of the preferred tax rates:*

$$t^S(y_1) > t^S(y_2) > t^U(y_2) > t^S(y_3) = t^U(y_3) = 0$$

and

$$t^U(y_1) > t^S(y_2) \text{ if and only if } y_2 - (1-\lambda)y_1 > \lambda\overline{y}$$

If the three regions are independent, the median-income voter is decisive in each one of them. Now, if the three regions unify, the median tax rate is either $t^U(y_1)$ or $t^S(y_2)$, depending on which one of the two is higher. Suppose that $t^U(y_1) > t^S(y_2)$. In that case, the equilibrium (median) tax rate is $t^S(y_2)$, from which it follows that all individuals in region B and C are indifferent between independence and unification. However, A majority of individuals in region A prefer a tax rate strictly lower than $t^S(y_2)$, and so region A does not join the union if integration is voluntary.

Suppose now that $t^U(y_1) < t^S(y_2)$. In that case, the equilibrium (median) tax rate is $t^U(y_1)$, and thus the lowest income is now decisive in setting the level of redistribution in the unified country. In that case, a majority of voters in region A prefer a tax rate strictly lower than $t^U(y_1)$, while a majority of individuals in region B and C prefer a tax rate strictly higher than $t^U(y_1)$. Therefore, although there are no income differences whatsoever between the three regions, none of them is willing to form a union.

Example 3 (Heterogeneous economies - No Income Differences Between Regions)**.** *Suppose now that $y_j^A = y_j^B = y_j^C = y_j$, $j = 1, 2, 3$, and we have the following distribution of types in each region: in region A, the lowest and highest-income individuals are rawlsian, while the median-income one is selfish, that is, we have (R, S, R) in region A. Furthermore, suppose that we have (U, U, R) in region B, and (R, U, U) in region C. In this case, we have the following ordering of the preferred tax rates:*

$$t^R(y_1) > t^S(y_2) > t^U(y_2) > t^U(y_3) = 0$$

As in the previous example, $t^U(y_1)$ may be either lower or higher than $t^S(y_2)$, while $t^R(y_3)$ may have any position between $t^R(y_1)$ and $t^U(y_3)$. Suppose that we have the following ordering:

$$t^R(y_1) > t^R(y_3) > t^S(y_2) > t^U(y_1) > t^U(y_2) > t^U(y_3) = 0$$

In that case, we have that $t^R(y_3)$, $t^U(y_1)$, and $t^U(y_2)$ are respectively implemented in region A, B and C when they are independent. Observe that in region A, although a majority of voters are selfish, the redistributive outcome is controlled by a rawlsian voter. Notice, furthermore, than thanks to the presence *of the rawlsian high-income voter in region B, the low-income utilitarian voter is decisive in setting the tax rate. Therefore, in both regions A and B, the coexistence of different types of voters reduces the decisiveness of the median-income class.*

If the three regions unify, the median tax rate is then given by $t^S(y_2)$. Observe that although the selfish voters form a minority in the unified country (i.e. $1/9$), it is a selfish individual that controls the redistributive outcome. In region A, a majority of individuals prefer $t > t^S(y_2)$, while a majority of individuals in region B and C prefer $t < t^S(y_2)$. Again, although there are no income differences across regions, none of them is willing to form a union.

Example 4 (Heterogeneous economies - Income Differences Between Regions)**.** *Suppose now that there are only 2 regions A and B, and assume the following distribution of types in each region: (S, R, R) in region A and (S, R, U) in region B. Suppose, furthermore, that $y_1^A > y_1^B$, $y_2^A = y_2^B$, and $y_3^A < y_3^B$. Finally, assume that the distribution of income in B is a mean-preserving spread of the one in A, so that $\overline{y}^A = \overline{y}^B$. Under **independence**, we have the following ordering of the tax rates:*

$$t^S(y_1^B) > t^S(y_1^A) > t^R(y_2^B) > t^R(y_2^A) > t^R(y_3^A) \geqslant t^U(y_3^B) = 0$$

and thus, if A and B are independent, the median tax rates are given by $t^R(y_2^A)$ and $t^R(y_2^B)$ respectively. Observe that even though the median-income (decisive) voter in both regions is of the same type with the same income, $t^R(y_2^B) > t^R(y_2^A)$ since $y_1^A > y_1^B$.

*Under **unification** of the two regions, we have the following ordering of the tax rates:*

$$t^S(y_1^B) > t^S(y_1^A) > t^R(y_2^B) = t^R(y_2^A) > t^R(y_3^A) \geqslant t^U(y_3^B) = 0$$

The rawlsian voter in A and B now have the same preferred tax rate (since the reference income for both of them is y_1^B), which is decisive under unification. In region B, all voters are indifferent between unification and independence, since the level of redistribution is unaffected, and there are no interregional transfers taking place under unification (recall that $\overline{y}^A = \overline{y}^B$). In region A, since unification yields more redistribution, it follows that the poorest (selfish) individual is unambiguously better off under unification. Then, the median and high-income rawlsian individuals may be better off under unification if they care enough about helping the poor, that is, if λ is high enough.

Hence, differences in income and type distributions now interact to shape the incentives to unify in each region. If, in addition to that, $\overline{y}^A \neq \overline{y}^B$, there is an additional tax-base effect that will, all other things being equal, induce the richer region to prefer independence, and the poorer region to prefer unification.

5 Conclusion

We endowed individuals with heterogeneous social preferences in the traditional RRMR framework on voting on redistribution. More specifically, we assumed that selfish, rawlsian and utilitarian voters coexist with given proportions. We characterized implicitly the Condorcet winner tax rate of this economy, and proved its existence for any positively skewed distribution of income. While an increase in the relative proportion of rawlsian (utilitarian) voters always yields more (less) redistribution in equilibrium, the effect of an increase in the relative proportion of selfish voters has an ambiguous effect on the equilibrium tax rate. In fact, given that the preferred tax rates of selfish voters have an intermediate position between the ones of rawlsian and utilitarian voters, the equilibrium level of redistribution in the heterogeneous economy may be either lower or higher than the one in an economy composed exclusively of selfish voters.

By simulating the model, we showed that small variations in the relative proportions of the three types may have a very large impact on the extent of redistribution when rawlsian and utilitarian voters attach a high relative weight to altruistic motives in their utility function.

Regarding the link between inequality and redistribution, it turns out that allowing for the coexistence of selfish, rawlsian and utilitarian voters yields slightly different predictions than when all voters are purely self-interested. In particular, an increase in poverty may increase the equilibrium level of redistribution. Furthermore, in case of a mean-preserving spread leaving the median income unaffected, which has no effect on redistribution in a selfish economy, redistribution is likely to increase in the heterogeneous economy, since the decrease in the lowest income induces the rawlsian voters to vote for more redistribution, and, in addition to that, the income position of the decisive voter within each type does not necessarily correspond to the median. Finally, if the rich get richer, the effect on redistribution is the same (qualitatively) as in the case of selfish voters, provided that the rawlsian decisive voter is not directly affected by the rise in inequality.

We assumed heterogeneity with respect to the *notion* of fairness individuals include in their preferences. However, a possible extension could be to consider heterogeneity with respect to the *weight* people attach to fairness considerations. Cappelen et al. (2007) showed that "*[...] both kinds of heterogeneity matter in explaining individual behavior*". However, observe that this requires making additional assumptions regarding the distribution of fairness weights among the voters. In particular, we would like to know how the fairness weights relate to income. If they are positively related, the relationship between the preferred tax rate and income will not anymore be monotonic for the rawlsian voters. If they are negatively related, the same holds for the utilitarian voters. This obviously has implications for the resulting equilibrium. Similarly, it could also well be the case that fairness intensity, rather than being related to income, changes according to whom redistribution applies to (e.g. immigrants). Finally, the fairness weights could also evolve over time, or change according to the extent of inequality, or change according to the information available. Indeed, Iriberri and Rey-Biel (2008) have shown experimentally that knowing the distribution of types among the population might change how individuals feel about others. In particular, subjects that exhibit other-regarding preferences tend to behave more selfishly once they are provided with social information.

More generally, this raises the question of why people exhibit social preferences in the first place. In other words, it raises the issue of the endogeneity of fairness preferences in the model, which is beyond the scope of this paper. What we showed in this paper is that fairness, both in type and intensity, does matter for the equilibrium level of redistribution. Therefore, understanding who behaves fairly and why, as well as with which intensity, seems worth investigating.

6 References

1. Andreoni, J. and Miller, J. (2002), Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism, *Econometrica*, Vol. 70, No. 2, pp. 737-753.

2. Ackert, L. F., Gillette, A. B., Martinez-Vasquez, J. and Rider, M. (2007a), Voting on Tax Policy

Design: A Test of the Selfish versus Social Preferences Hypotheses, *Public Finance Review*, Vol. 35, No.2, pp. 263-284.

3. Ackert, L. F., Martinez-Vasquez, J. and Rider, M. (2007b), Social Preferences and Tax Policy Design: Some Experimental Evidence, *Economic Inquiry*, Vol. 45, No. 3, pp. 487-501.

4. Benabou, R. (2000), Unequal Societies: Income Distribution and the Social Contract, *American Economic Review*, Vol. 90, No. 1, pp. 96-129.

5. Bolton, G. E. and Ockenfels, A. (2002), The Behavioral Tradeoff between Efficiency and Equity when a Majority Rules, Discussion Papers on Strategic Interaction from Max Planck Institute of Economics, Strategic Interaction Group.

6. Bolton, P. and Roland, G. (1997), The Breakup of Nations: A Political Economy Analysis, *Quarterly Journal of Economics*, Vol. 112, No. 4, pp. 1057-1090.

7. Borck, R. (2007), Voting, Inequality and Redistribution, *Journal of Economic Surveys*, Vol. 21, No. 1, pp. 90-109.

8. Cappelen, A.W., Hole, A.D., Sorensen, E.Ø. and Tungodden, B. (2007), The Pluralism of Fairness Ideals: an Experimental Approach, *American Economic Review*, Vol. 97, No. 3, pp. 818-827.

9. Charness, G and Rabin, M. (2002), Understanding Social Preferences with Simple Tests, *Quarterly Journal of Economics*, Vol. 117, No. 3, pp. 817–869.

10. Dawes, C. T., Loewen, P. J. and Fowler, J. H. (2008), Social Preferences and Political Participation, Working Paper.

11. Dhami, S. and al-Nowaihi, A. (2010a), Redistributive Policy with Heterogenous Social Preferences of Voters, *European Economic Review*, Vol. 54, Issue 6, pp. 743759.

12. Dhami, S. and al-Nowaihi, A. (2010b), Existence of a Condorcet Winner When Voters Have Other Regarding Preferences, *Journal of Public Economic Theory*, Vol. 12, Issue 5, pp. 897-922.

13. Dhami, S. and al-Nowaihi, A. (2010c), Inequality and Redistribution When Voters Have Other Regarding Preferences, Unpublished manuscript, University of Leicester.

14. Edlin, A., Gelman, A. and Kaplan, N. (2007), Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others, *Rationality and Society*, Vol. 19, No. 3, pp. 293-314.

15. Erlei, M. (2008), Heterogeneous Social Preferences, *Journal of Economic Behavior and Organization*, Vol. 65, Issues 3-4, pp. 436-457.

16. Fehr, E. and Schmidt, K. (1999), A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-868.

17. Fehr, E. and Schmidt, K. (2006), The Economics of Reciprocity, Fairness and Altruism - Experimental Evidence and New Theories, in *Handbook of the Economics of Giving, Altruism and Reciprocity*, pp. 615–692, North-Holland.

18. Fisman, R., Kariv, S. and Markovitz, D. (2007), Individual Preferences for Giving, *American Economic Review*, Vol. 97, No. 2, pp. 153-158.

19. Galasso, V. (2003), Redistribution and Fairness: A Note, *European Journal of Political Economy*, Vol. 19, No. 4, pp 885-892.

20. Iriberri, N. and Rey-Biel, P. (2008), Elicited Beliefs and Social Information in Modified Dictator Games: What Do Dictators Believe Other Dictators Do?, Working Paper No. 1137, Universitat Pompeu Fabra.

21. Luttens, R. I. and M-A. Valfort (2010), Voting for Redistribution under Desert-Sensitive Altruism, *Scandinavian Journal of Economics*, forthcoming.

22. Meltzer, A. H. and Richard, S.F. (1981), A Rational Theory of the Size of Government, *Journal of Political Economy*, Vol. 89, No. 5, pp. 914–927.

23. Roberts (1977), Voting over Income Tax Schedules, *Journal of Public Economics*, Vol. 8, Issue 3, pp. 329-340.

24. Romer (1975), Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax, *Journal of Public Economics*, Vol. 4, Issue 2, pp. 163-185.

25. Tyran, J-R. and Sausgruber, R. (2006), A Little Fairness May Induce a Lot of Redistribution in Democracy, *European Economic Review,* Vol. 50, No. 2, pp. 469-485.

7 Appendix

Proof of Proposition 1.

1. For a **selfish** individual with ability ω, we have

$$V^S(t,\omega) = v(t,\omega) = \frac{1}{2}(1-t)^2\omega^2 + t(1-t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

Taking partial derivative with respect to the tax rate t,

$$\frac{\partial V^S(t,\omega)}{\partial t} = -(1-t)\omega^2 + (1-2t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

Setting this quantity equal to zero and solving for t yields

$$t^S = \frac{\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega) - \omega^2}{2\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega) - \omega 2} = \frac{\overline{y} - y(\omega)}{2\overline{y} - y(\omega)}$$

2. For a **rawlsian** individual with ability ω, we have

$$V^R(t,\omega) = (1-\lambda^R)\left[\frac{1}{2}(1-t)^2\omega^2\right] + \lambda^R\left[\frac{1}{2}(1-t)^2\underline{\omega}^2\right] + t(1-t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

Taking partial derivative with respect to the tax rate t,

$$\frac{\partial V^R(t,\omega)}{\partial t} = -(1-\lambda^R)(1-t)\omega^2 - \lambda^R\underline{\omega}^2(1-t) + (1-2t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

Setting this quantity equal to zero and solving for t yields

$$t^R = \frac{\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega) - \lambda^R\underline{\omega}^2 - (1-\lambda^R)\omega^2}{2\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega) - \lambda^R\underline{\omega}^2 - (1-\lambda^R)\omega^2} = \frac{\overline{y} - \lambda^R y(\underline{\omega}) - (1-\lambda^R)y(\omega)}{2\overline{y} - \lambda^R y(\underline{\omega}) - (1-\lambda^R)y(\omega)}$$

3. For a **utilitarian** individual with ability ω, we have

$$V^U(t,\omega) = (1-\lambda^U)\left[\frac{1}{2}(1-t)^2\omega^2 + t(1-t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)\right]$$
$$+\lambda^U\left[\frac{1}{2}(1-t)^2\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega) + t(1-t)\int_{\underline{\omega}}^{\overline{\omega}}\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)dF(\omega)\right]$$

Taking partial derivative with respect to the tax rate t,

$$\frac{\partial V^U(t,\omega)}{\partial t} = -(1-t)(1-\lambda^U)\omega^2 + (1-\lambda^U)(1-2t)\int_{\underline{\omega}}^{\overline{\omega}} \omega^2 dF(\omega)$$

$$-\lambda^U(1-t)\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)+\lambda^U(1-2t)\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)$$

Setting this quantity equal to zero and solving for t yields

$$t^U=\frac{\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)-\lambda^U\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)-(1-\lambda^U)\omega^2}{2\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)-\lambda^U\int_{\underline{\omega}}^{\overline{\omega}}\omega^2 dF(\omega)-(1-\lambda^U)\omega^2}=\frac{\overline{y}-\lambda^U\overline{y}-(1-\lambda^U)y(\omega)}{2\overline{y}-\lambda^U\overline{y}-(1-\lambda^U)y(\omega)}$$

□

Proof of Proposition 2. It follows directly from the analytical expression of t^S, t^R and t^U that

1. If $y(\omega)<\overline{y}$, it holds that $0<t^U(\omega)<t^S(\omega)<t^R(\omega)$
2. If $\overline{y}<y(\omega)<\frac{\overline{y}-\lambda^R y(\underline{\omega})}{1-\lambda^R}$, it holds that $0=t^U(\omega)=t^S(\omega)<t^R(\omega)$
3. If $\frac{\overline{y}-\lambda^R y(\underline{\omega})}{1-\lambda^R}<y(\omega)$, it holds that $0=t^U(\omega)=t^S(\omega)=t^R(\omega)$

Therefore, for any ability level ω, it holds that $t^U(\omega)\leqslant t^S(\omega)\leqslant t^R(\omega)$. □

Proof of Proposition 3. Suppose that $y(\omega)<\overline{y}$, so that $t^i>0$, $i=S,R,U$.

1. For a **selfish** individual, we have

$$t^S=\frac{\overline{y}-y(\omega)}{2\overline{y}-y(\omega)}$$

Taking derivatives,

$$\frac{\partial t^S}{\partial y(\omega)}=-\frac{\overline{y}}{[2\overline{y}-y(\omega)]^2}<0$$

$$\frac{\partial t^S}{\partial \overline{y}}=\frac{y(\omega)}{[2\overline{y}-y(\omega)]^2}>0$$

2. For a **rawlsian** individual, we have

$$t^R=\frac{\overline{y}-\lambda^R y(\underline{\omega})-(1-\lambda^R)y(\omega)}{2\overline{y}-\lambda^R y(\underline{\omega})-(1-\lambda^R)y(\omega)}$$

Taking derivatives,

$$\frac{\partial t^R}{\partial y(\omega)}=-\frac{(1-\lambda^R)\overline{y}}{[2\overline{y}-\lambda^R y(\underline{\omega})-(1-\lambda^R)y(\omega)]^2}<0$$

$$\frac{\partial t^R}{\partial \overline{y}}=\frac{\lambda^R y(\underline{\omega})+(1-\lambda^R)y(\omega)}{[2\overline{y}-\lambda^R y(\underline{\omega})-(1-\lambda^R)y(\omega)]^2}>0$$

$$\frac{\partial t^R}{\partial \lambda^R} = \frac{\overline{y}\left[y(\omega) - y(\underline{\omega})\right]}{\left[2\overline{y} - \lambda^R y(\underline{\omega}) - (1-\lambda^R)y(\omega)\right]^2} > 0$$

3. For a **utilitarian** individual, we have

$$t^U = \frac{\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)}{2\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)}$$

   Taking derivatives,

$$\frac{\partial t^U}{\partial y(\omega)} = -\frac{(1-\lambda^U)\overline{y}}{\left[2\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)\right]^2} < 0$$

$$\frac{\partial t^U}{\partial \overline{y}} = \frac{\lambda^U \overline{y} + (1-\lambda^U)y(\omega)}{\left[2\overline{y} - \lambda^U \overline{y} - (1-\lambda^U)y(\omega)\right]^2} > 0$$

$$\frac{\partial t^U}{\partial \lambda^U} = \frac{\left[y(\omega) - \overline{y}\right]\overline{y}}{\left[2\overline{y} - (1-\lambda^U)y(\omega) - \lambda^U \overline{y}\right]^2} < 0$$

□

Proof of Lemma 1.

1. Selfish versus Rawlsian: $t^S = t^R = t > 0$ is equivalent to

$$y^S(t) = \lambda^R y\left(\underline{\omega}\right) + (1-\lambda^R)y^R(t) = y^R(t) + \lambda^R\left[y\left(\underline{\omega}\right) - y^R(t)\right] < y^R(t)$$

   Hence, if a selfish and fair rawlsian voter support the same positive tax rate, the rawlsian agent has a higher income than the selfish one.

2. Selfish versus Utilitarian: $t^S = t^U = t > 0$ is equivalent to

$$y^S(t) = (1-\lambda^U)y^U(t) + \lambda^U \overline{y} = y^U(t) + \lambda^U\left[\overline{y} - y^U(t)\right] > y^U(t)$$

   Hence, if a selfish and utilitarian voter support the same positive tax rate, the selfish agent has a higher income than the utilitarian one.

3. Rawlsian versus Utilitarian: $t^R = t^U > 0$ is equivalent to

$$y^R(t) = \frac{(1-\lambda^U)}{(1-\lambda^R)}y^U(t) + \frac{\lambda^U}{(1-\lambda^R)}\overline{y} - \frac{\lambda^R}{(1-\lambda^R)}y\left(\underline{\omega}\right)$$

   For $\lambda^R = \lambda^U = \lambda$, this boils down to

$$y^R(t) = y^U(t) + \frac{\lambda}{(1-\lambda)}\left[\overline{y} - y\left(\underline{\omega}\right)\right] > y^U(t)$$

Hence, if a utilitarian and fair rawlsian voter support the same positive tax rate, the rawlsian agent has a higher income than the utilitarian one.

□

Proof of Lemma 2.

$$\frac{\partial\Phi(t)}{\partial t} = \alpha f\left(y^S(t)\right)\frac{\partial y^S(t)}{\partial t} + \beta f\left(y^R(t)\right)\frac{\partial y^R(t)}{\partial t} + \gamma f\left(y^U(t)\right)\frac{\partial y^U(t)}{\partial t}$$

and thus

$$\frac{\partial\Phi\left(t\right)}{\partial t} < 0 \text{ for all } t < \frac{1}{2}$$

□

Proof of Proposition 5. The function $\Phi(t)$ is given by

$$\Phi(t) = \alpha F\left(y^S(t)\right) + \beta F\left(y^R(t)\right) + \gamma F\left(y^U(t)\right)$$

Substituting yields

$$\Phi(t) = \alpha F\left(\tfrac{(1-2t)}{(1-t)}\overline{y}\right) + \beta F\left(\tfrac{(1-2t)}{(1-t)(1-\lambda^R)}\overline{y} - \tfrac{\lambda^R}{(1-\lambda^R)}y(\underline{\omega})\right) + \gamma F\left(\tfrac{(1-2t)}{(1-t)(1-\lambda^U)}\overline{y} - \tfrac{\lambda^U}{(1-\lambda^U)}\overline{y}\right)$$

Taking derivatives,

$$\frac{\partial\Phi}{\partial\lambda^R} = \beta f\left(y^R(t)\right)\left[\frac{(1-2t)\overline{y} - (1-t)y(\underline{\omega})}{(1-t)(1-\lambda^R)^2}\right] > 0$$

$$\frac{\partial\Phi}{\partial\lambda^U} = -\gamma f\left(y^U(t)\right)\left[\frac{t\overline{y}}{(1-t)(1-\lambda^U)^2}\right] < 0$$

The Condorcet winner tax rate is given by

$$t^* = \left\{t|\alpha F\left(y^S(t)\right) + \beta F\left(y^R(t)\right) + \gamma F\left(y^U(t)\right) = \frac{1}{2}\right\}$$

Equivalently, we can write

$$t^* = \left\{t|\frac{\alpha}{\alpha+\beta+\gamma}F\left(y^S(t)\right) + \frac{\beta}{\alpha+\beta+\gamma}F\left(y^R(t)\right) + \frac{\gamma}{\alpha+\beta+\gamma}F\left(y^U(t)\right) = \frac{1}{2}\right\}$$

This way, we can determine the effect on the Condorcet winner tax rate of a marginal increase in the proportion of each type.

We know that, when the three types vote for the same tax rate t, the following relationship holds:

$$y^R(t) > y^S(t) > y^U(t)$$

and thus

$$F\left(y^R(t)\right) > F\left(y^S(t)\right) > F\left(y^U(t)\right)$$

Therefore, taking derivatives, we get

$$\frac{\partial \Phi}{\partial \alpha} = \frac{1}{(\alpha+\beta+\gamma)^2}\left\{\beta\left[F\left(y^S(t)\right) - F\left(y^R(t)\right)\right] + \gamma\left[F\left(y^S(t)\right) - F\left(y^U(t)\right)\right]\right\} \lesseqgtr 0$$

$$\frac{\partial \Phi}{\partial \beta} = \frac{1}{(\alpha+\beta+\gamma)^2}\left\{\alpha\left[F\left(y^R(t)\right) - F\left(y^S(t)\right)\right] + \gamma\left[F\left(y^R(t)\right) - F\left(y^U(t)\right)\right]\right\} > 0$$

$$\frac{\partial \Phi}{\partial \gamma} = \frac{1}{(\alpha+\beta+\gamma)^2}\left\{\alpha\left[F\left(y^U(t)\right) - F\left(y^S(t)\right)\right] + \beta\left[F\left(y^U(t)\right) - F\left(y^R(t)\right)\right]\right\} < 0$$

Using the implicit function theorem, we finally get the following comparative statics results regarding the Condorcet winner tax rate of an economy with heterogeneous social preferences:

$$\frac{\partial t}{\partial \lambda^R} > 0 \text{ and } \frac{\partial t}{\partial \lambda^U} < 0$$

$$\frac{\partial t}{\partial \alpha} \lesseqgtr 0 \ , \ \frac{\partial t}{\partial \beta} > 0 \text{ and } \frac{\partial t}{\partial \gamma} < 0$$

□

Proof of Proposition 6. Let $y(\omega_1)$ and $y(\omega_2)$ be such that

$$y(\underline{\omega}) < y(\omega_1) < \overline{y} < y(\omega_2) < y(\overline{\omega})$$

Furthermore, recall that

$$y^S(t^*) = \frac{(1-2t^*)}{(1-t^*)}\overline{y}$$

$$y^R(t^*) = \frac{(1-2t^*)}{(1-t^*)(1-\lambda^R)}\overline{y} - \frac{\lambda^R}{(1-\lambda^R)}y(\underline{\omega})$$

$$y^U(t^*) = \frac{(1-2t^*)}{(1-t^*)(1-\lambda^U)}\overline{y} - \frac{\lambda^U}{(1-\lambda^U)}\overline{y}$$

From Lemma 1, we know that

$$y^U(t^*) < y^S(t^*) < y^R(t^*) \text{ for } t^* \in \left(0, \frac{1-\lambda^U}{2-\lambda^U}\right)$$

and

$$y^S(t^*) < y^R(t^*) \text{ for } t^* \in \left(\frac{1-\lambda^U}{2-\lambda^U}, \frac{1}{2}\right) \text{ and no utilitarian has } t^* \text{ as a preferred tax rate}$$

Finally, from Proposition 4, given that the distribution of abilities $F(\omega)$ is positively skewed, we know that $t^* > 0$ and thus $y^S(t^*) < \overline{y}$, $y^U(t^*) < \overline{y}$, and $y^R(t^*) \lesseqgtr \overline{y}$. Therefore, we have that

1. If the poor get poorer,

$$\begin{cases} F(y) = G(y) \text{ for } y > y(\omega_1) \\ F(y) \leq G(y) \text{ for } y \leq y(\omega_1) \end{cases} \text{ and } y\left(\underline{\omega}\right) \text{ and } \overline{y} \text{ decrease}$$

Hence, we have

(a) If $y^S(t^*) \leq y(\omega_1)$, then $G\left(y^S(t^*)\right) \lesseqgtr F\left(y^S(t^*)\right)$, since $G(y) \geqslant F(y)$ for $y \leqslant y(\omega_1)$, and the effect of the decrease in $\overline{y}$ under $G(.)$ goes in the opposite direction (i.e. $\frac{\partial y^S(t^*)}{\partial \overline{y}} > 0$).

If $y^S(t^*) > y(\omega_1)$, then $G\left(y^S(t^*)\right) < F\left(y^S(t^*)\right)$, since $\overline{y}$ is smaller under $G(.)$ than under $F(.)$.

(b) If $y^R(t^*) \leqslant y(\omega_1)$, then $G\left(y^R(t^*)\right) \lesseqgtr F\left(y^R(t^*)\right)$, since $G(y) \geqslant F(y)$ for $y \leqslant y(\omega_1)$, and the decrease in $y(\underline{\omega})$ and in $\overline{y}$ under $G(.)$ go in opposite directions (i.e. $\frac{\partial y^R(t^*)}{\partial y(\underline{\omega})} < 0$ and $\frac{\partial y^R(t^*)}{\partial \overline{y}} > 0$).

If $y^R(t^*) > y(\omega_1)$, then $G\left(y^R(t^*)\right) \lesseqgtr F\left(y^R(t^*)\right)$, since the decrease in $y(\underline{\omega})$ and in $\overline{y}$ under $G(.)$ go in opposite directions.

(c) If $y^U(t^*) \leq y(\omega_1)$, then $G\left(y^U(t^*)\right) \lesseqgtr F\left(y^U(t^*)\right)$ (with strict equality if $t^* \geqslant \frac{1-\lambda^U}{2-\lambda^U}$), since $G(y) \geqslant F(y)$ for $y \leqslant y(\omega_1)$, and the effect of the decrease in $\overline{y}$ under $G(.)$ goes in the opposite direction (i.e. $\frac{\partial y^U(t^*)}{\partial \overline{y}} > 0$).

If $y^U(t^*) > y(\omega_1)$, then $G\left(y^U(t^*)\right) \leqslant F\left(y^U(t^*)\right)$ (with strict equality if $t^* \geqslant \frac{1-\lambda^U}{2-\lambda^U}$), since $\overline{y}$ is smaller under $G(.)$ than under $F(.)$.

Therefore, when the poor get poorer, the effect on the Condorcet winner tax rate t^* is ambiguous.

2. If the rich get richer,

$$\begin{cases} F(y) = G(y) \text{ for } y < y(\omega_2) \\ F(y) \geq G(y) \text{ for } y \geq y(\omega_2) \end{cases} \text{ and } y\left(\overline{\omega}\right) \text{ and } \overline{y} \text{ increase}$$

Hence, we have

(a) $G\left(y^S(t^*)\right) > F\left(y^S(t^*)\right)$, since $\overline{y}$ is higher under $G(.)$ than under $F(.)$.

(b) If $y^R(t^*) < y(\omega_2)$, then $G\left(y^R(t^*)\right) > F\left(y^R(t^*)\right)$, since $\overline{y}$ is higher under $G(.)$ than under $F(.)$.

If $y^R(t^*) \geqslant y(\omega_2)$, then $G\left(y^R(t^*)\right) \lesseqgtr F\left(y^R(t^*)\right)$, since $G(y) \leqslant F(y)$ for $y \geqslant y(\omega_2)$, and the effect of the increase in $\overline{y}$ under $G(.)$ goes in the opposite direction (i.e. $\frac{\partial y^R(t^*)}{\partial \overline{y}} > 0$).

(c) $G\left(y^U(t^*)\right) \geqslant F\left(y^U(t^*)\right)$ (with strict equality if $t^* \geqslant \frac{1-\lambda^U}{2-\lambda^U}$), since $\overline{y}$ is higher under $G(.)$ than under $F(.)$.

and thus

$$\alpha G\left(y^S(t^*)\right) + \beta G\left(y^R(t^*)\right) + \gamma G\left(y^U(t^*)\right) > \frac{1}{2} \text{ if } y^R(t^*) < y(\omega_2)$$

Therefore, $t^{**} > t^*$ whenever $y^R(t^*) < y(\omega_2)$.

3. If there is a mean-preserving spread,

$$\begin{cases} F(y) \leq G(y) \text{ for } y \leq y(\omega_1) \\ F(y) = G(y) \text{ for } y(\omega_1) \leq y \leq y(\omega_2) \\ F(y) \geq G(y) \text{ for } y \geq y(\omega_2) \end{cases} \quad \text{and } y(\underline{\omega}) \text{ decreases, } y(\overline{\omega}) \text{ increases, with } \overline{y} \text{ unchanged}$$

Hence, we have

(a) If $y^S(t^*) \leq y(\omega_1)$, then $G\left(y^S(t^*)\right) \geqslant F\left(y^S(t^*)\right)$, by definition of $F(.)$ and $G(.)$.
If $y^S(t^*) > y(\omega_1)$, then $G\left(y^S(t^*)\right) = F\left(y^S(t^*)\right)$, since $\overline{y}$ is the same under $G(.)$ as under $F(.)$.

(b) If $y^R(t^*) \leqslant y(\omega_1)$, then $G\left(y^R(t^*)\right) \geqslant F\left(y^R(t^*)\right)$, since $G(y) \geqslant F(y)$ for $y \leqslant y(\omega_1)$, and the effect of the decrease in $y(\underline{\omega})$ under $G(.)$ goes in the same direction (i.e. $\frac{\partial y^R(t^*)}{\partial y(\underline{\omega})} < 0$).
If $y(\omega_1) < y^R(t^*) < y(\omega_2)$, then $G\left(y^R(t^*)\right) > F\left(y^R(t^*)\right)$, since $y(\underline{\omega})$ is smaller under $G(.)$ than under $F(.)$.
If $y^R(t^*) \geqslant y(\omega_2)$, then $G\left(y^R(t^*)\right) \lesseqgtr F\left(y^R(t^*)\right)$, since $G(y) \leqslant F(y)$ for $y \geqslant y(\omega_2)$, and the effect of the decrease in $y(\underline{\omega})$ under $G(.)$ goes in the opposite direction.

(c) If $y^U(t^*) \leq y(\omega_1)$, then $G\left(y^U(t^*)\right) \geqslant F\left(y^U(t^*)\right)$ (with strict equality if $t^* \geqslant \frac{1-\lambda^U}{2-\lambda^U}$), by definition of $F(.)$ and $G(.)$.
If $y^U(t^*) > y(\omega_1)$, then $G\left(y^U(t^*)\right) = F\left(y^U(t^*)\right)$, since $\overline{y}$ is the same under $G(.)$ as under $F(.)$.

Hence, we have that

$$\alpha G\left(y^S(t^*)\right) + \beta G\left(y^R(t^*)\right) + \gamma G\left(y^U(t^*)\right) > \frac{1}{2} \text{ if } y^R(t^*) < y(\omega_2)$$

Therefore, $t^{**} > t^*$ whenever $y^R(t^*) < y(\omega_2)$.

□