ELSEVIER

Journal of Public Economics 91 (2007) 51 – 76

JOURNAL OF PUBLIC ECONOMICS

www.elsevier.com/locate/econbase

Fractionalization and the size of government ☆

Jo Thori Lind *

Department of Economics, University of Oslo, PB 1095 Blindern, 0317 Oslo, Norway

Received 25 May 2005; received in revised form 7 March 2006; accepted 7 September 2006
Available online 13 November 2006

Abstract

I study the effect of voters with a group-based social conscience. Voters care more about the well-being of those belonging to their own group than the rest of the population. Within a model of political tax determination, both fractionalization and group antagonism reduce the support for redistribution. Whereas within group inequality increases support for redistribution, inequality between groups has the opposite effect. These results hold even if a poor group forms a majority. Using a panel constructed from US micro data, I find support for the hypothesis that within race inequality increases redistribution while between race inequality decreases redistribution.

JEL classification: D31; D64; D72; E62; H11
Keywords: Fractionalization; Redistribution; Political economy

☆ I wish to thank Sam Bowles, Kjell Arne Brekke, Bård Harstad, Steinar Holden, Aanund Hylland, Tor Jakob Klette, Kalle Moene, Torsten Persson, Thumas Plümper, Ole Jørgen Røgeberg, Klaas Staal, Michael Wallerstein, and two anonymous referees as well as participants at the Econometric Society European Winter Meeting, EEA conference, Norwegian–German Seminar on Public Economics, meetings of the European Public Choice Society and the Norwegian Economic Society and seminars at MIT and the University of Oslo for help and valuable comments.
* Tel.: +47 22 84 40 27; fax: +47 22 85 50 35.
E-mail address: j.t.lind@econ.uio.no.

0047-2727/$ - see front matter
doi:10.1016/j.jpubeco.2006.09.006

First, the American public thinks that most people who receive welfare are black, and second, the public thinks that blacks are less committed to the work ethic than other Americans. There exists now a widespread perception that welfare has become a "code word" for race (Gilens, 1999: 3).

1. Introduction

The above quote from Martin Gilens's (1999) book *Why Americans Hate Welfare* is representative for a widespread view: It's impossible to understand the American welfare state without considering race, and if not racism, at least racial stereotypes. In this paper I explore how this may enhance economists' understanding of the relationship between inequality, fractionalization, and redistribution. This is important not only to understand the American welfare state, but also to understand politics in other heavily fractionalized societies, such as most African countries.

The conventional political economy approach to analysing preferences for redistribution is through the median voter model. The main result is more redistribution in societies with high inequality than in societies with less, as the median voter's preferences for redistribution are inversely related to the difference between her income and the average income (Romer, 1975; Roberts, 1977; Meltzer and Richard, 1981). The empirical support for the hypothesis is mixed. Bénabou (1996) surveys a number of older studies that mostly reject it. Milanovic (2000), however, claims this is mainly due to data problems, and using an improved data set, he finds support for the theory.[1]

A literature has emerged studying the effects of fractionalization along ethnic, linguistic, religious, and other lines on public policy and economic performance. There is now a substantial theoretical literature explaining why particularly public good provision is lower in fractionalized communities,[2] and the empirical support for the detrimental effects of fractionalization on public policy is quite strong.[3] However, this literature generally studies agents that are equal except for their group belonging, so we can't study the relationship between income distribution and public policy in fractionalized societies.

The main novelty of my approach is the joint modelling of group and income heterogeneity. I can then study how each of this influence supports for redistribution as well as how the joint impact is. It turns out that inequality may have very different effects on support for redistribution in fractionalized and non-fractionalized societies. In the latter, inequality has the usual effect of

[1] Recent research has attempted to explain this puzzle. Bénabou's (2000) model is probably the best known. He presents a model where redistribution both has beneficial effects due to credit market imperfections and distorts the labour supply decision. Under reasonable assumptions, there may be political support for two "social contracts", one with an even distribution of income and support for redistribution to reduce the effects of missing credit markets, and one with high inequality and little support for redistribution. Competing explanations have been proposed by e.g. Saint-Paul (2001), Roemer (1998, 1999), Moene and Wallerstein (2001), Bjorvatn and Cappelen (2002), and Alesina et al. (2001).

[2] Based on such factors as differentiated tastes (Alesina et al., 1999), antagonism to mixing with members of other groups (Alesina and La Ferrara, 2000), and the fact that social sanctions are more efficient within groups than between groups (Miguel and Gugerty, 2005). There is also some earlier theoretical contributions mainly based on social conflict and lack of social capital (inter alia Benhabib and Rusticini, 1996; Knack and Keefer, 1997; Keefer and Knack, 2002; Rodrik, 1999), but they are less relevant for this paper. See Alesina and La Ferrara (2005) for a more in-depth survey of this literature.

[3] Alesina and his co-workers have documented that fractionalization tends to reduce the supply of public goods, redistribution, and participation in US communities (Alesina et al., 1999, 2001; Alesina and La Ferrara, 2000). This is corroborated by similar findings in Pakistan (Khwaja, 2002) and Kenya (Miguel and Gugerty, 2005). Furthermore, comparing Kenya, where ethic conflicts are important, to Tanzania, where there is less ethnic conflict, Miguel (2004) finds that ethnic fractionalization is important in Kenya but insignificant in Tanzania.

promoting increased redistribution. In fractionalized societies, inequality between groups has the opposite effect of reducing support for redistribution.

I present a model in the tradition of Romer–Roberts–Meltzer–Richard where a tax used for redistributive transfers is determined by popular vote. Unlike the traditional model, I allow voters to have a social conscience in that they care about social welfare in addition to their private well-being. In itself, this extension does not change the main conclusions of the model. But in fractionalized societies, it is natural to assume that agents care mostly about the welfare of those belonging to their own group, that is, agents have a group or race bias in their social conscience. I label this group antagonism. It implies that two persons with the same endowment, but one belonging to a rich and one to a poor group, have different preferences for taxation. The poorer the group one belongs to, the higher is the preferred tax rate. This means that voters with the median preferred tax rate will have different endowments depending on which group they belong to. Consequently, we can no longer talk about the median voter as a single agent. Instead, there is a set of median voters, one from each group.

The model gives two key insights. First, increased group antagonism reduces the chosen tax rate. Increased group antagonism makes all individuals put more emphasis on the welfare of the other members of their group, and less on the welfare of members of other groups. Specifically, members of a rich group, putting more weight on the welfare of the rich group, will vote for lower taxes than before. This implies that the marginal voters within the rich group will now be poorer than before, as the lower individual income implying a preference for higher taxes counterbalances the group effect implying a preference for lower taxes. If the income distribution within the rich group is skewed to the right, the density grows for lower incomes, implying that the density of marginal voters in the rich group increases.

In contrast, in a poor group higher group antagonism will make members vote for higher taxes than before, so that the marginal voters now will have higher income than before (this time, the higher income involving a preference for lower taxes counterbalances the group effect involving a preference for higher taxes). If the income distribution within the poor group is skewed to the right, the density falls for higher incomes, implying that the density of the marginal voters in the poor group decreases. Thus, increased group antagonism increases the density of the marginal voters in the rich group, and decreases the density of the marginal voters in the poor group. This implies that the mass of voters in the rich group, who because of increased group antagonism now prefer lower taxes, is greater than the mass of voters in the poor group who prefer higher taxes. Thus, the upshot is that the overall marginal voter now prefers lower taxes than before. When fractionalization is high, this effect is stronger.

Second, the model also predicts that increased inequality between groups will reduce the support for redistribution. The reasoning is quite similar to the one above; when the rich group becomes richer, their preferences for redistribution decline. Hence the new median voter from the rich group is poorer and vice versa for the poor group. Again, the decline in the income of the median voter from the rich group is smaller than the rise in the income of the median voter from the poor group. Then the new political equilibrium is a lower tax rate and less redistribution. Both of these results are independent of which group is in majority.

To test some of the key insights of the model, I use two sources of data. First, I regress stated preferences for redistribution on measures of racial closeness using data from the General Social Survey. For African American respondents, preferences for redistribution is increasing in the closeness to own race and decreasing in the closeness to Whites. For White respondents, the relationship reverses. This indicates that closeness to persons from a poor group increases support for redistribution and closeness to persons from a rich group decreases support for redistribution.

Second, I use a panel of US states observed in 6 years between 1969 and 2000. As data on inequality by race are not available in preexisting sources, I constructed these data using micro data from the Luxembourg Income Study. Unlike most earlier studies, this permits focusing on pre-tax income which should be the relevant variable for determining tax preferences. In line with the predictions of the model, I find that fractionalization and between group inequality tends to reduce redistribution whereas with group inequality increases it. Although the effect of between group inequality is usually not significantly smaller than zero, it is significantly smaller than the effect of within group inequality. These conclusions remain if we include state fixed effects or use robust regression techniques.

A related work is Austen-Smith and Wallerstein (2006), who present a model of joint determination of redistribution and scope of affirmative action. They show that in divided societies, support for welfare spending is lower than in non-divided societies. Sethi and Somanathan (2004) study the inter-play of income inequality and residential segregation, but they do not study how redistribution is affected. Vigdor (2000) alludes to a theory where people are altruistic to members of their own group and discusses the effect of this on provision of public goods. Collier (2000, 2001) discusses similar questions, but his analysis of democratic regimes is somewhat brief. I will also show that his conclusions do not necessarily hold in a more general framework. Luttmer (2001) studies the relationship between group membership and preferences for redistribution. He finds a preference structure that is similar to the one I use. However, he does not study the political economic implication of these preferences. Persson and Tabellini (1994) also use a model with similarities to my model, but their focus is on centralization and regional integration. Finally there's a large literature in sociology and political science studying the impact of racial divide on policy making and political behaviour. The most comprehensive is probably Kinder and Sanders' (1996) study of a number of possible explanations of differing opinions between blacks and whites. Gilens (1999) study how racial stereotypes, mainly formed by the media, influence people's support for redistribution, while Wilson in a number of works (e.g. Wilson, 1978, 1999) has discussed class based versus racially based political segmentation and advocated a multiracial coalition of the lower- and middle-class to combat poverty. However, they refrain from using formal modelling and utility maximizing agents, so their analyses are different from mine.

2. The model

2.1. The baseline case

I consider an economy with a continuum of heterogeneous agents with mass normalized to one. Each agent has an income or endowment of a taxable good whose distribution in the economy can be described by a cumulative density function F with support $\Omega \subseteq \mathbb{R}_+$. Denote by $\bar{x}$ and x^m the mean and median endowment respectively. Utility derived from consumption of the good is given by the function u which is assumed to be increasing and concave. The model is static, so there are no credit markets. In the absence of transfers, an agent with endowment x reaches utility level $u(x)$, and under the assumption of a utilitarian social welfare function, social welfare equals $\int_\Omega u(x)\mathrm{d}F(x)$.

There is a government that redistributes resources before consumption takes place. Every agent faces a linear tax rate t and receives a transfer $T(t)\,\bar{x}$ where T is a function that represents the outcome of taxation. The function takes account of a possible deadweight loss. This could be modelled explicitly, for instance as a choice of labour supply, but this would add little to the model and make it more cumbersome. I make the standard assumptions that the deadweight loss is absent at $t=0$ and increases as t increases. This implies that T satisfies $T(0)=0$, $T'(0)=1$, $T'(t)\leq 1$, and $T''(t)<0$, that is, a concave Laffer curve. For simplicity, I will also assume that $T'(1)<0$ so T is maximized for a tax rate strictly below unity. The tax rate t is determined as the outcome of a political process where the chosen tax rate corresponds to the one preferred by the median voter.[4]

All agents care about their own utility. However, they also have a social conscience which implies that they care about the social welfare level.[5] For a given mean income (tax base) $\overline{x}$, social welfare is given by

$$S(t,F) = \int_{\Omega} u[(1-t)x + T(t)\overline{x}]\mathrm{d}F(x). \tag{1}$$

The last argument of S is an element from the space of income distributions, i.e. social welfare depends on the tax rate t and the society's income distribution F. Hence for a given tax rate, social welfare will change if we change the income distribution. Notice that S is linear in the income distribution in the sense that for two functions F_1 and F_2 and two constants a_1 and a_2, $S(t, a_1F_1+a_2F_2)=a_1S(t, F_1)+a_2S(t, F_2)$.

Agents weight their private utility by $1-\alpha$ and social welfare by α. Then an agent with initial endowment x maximizes

$$U(x,t) = (1-\alpha)u[(1-t)x + T(t)\overline{x}] + \alpha S(t,F) \tag{2}$$

where α is a coefficient of social conscience. Throughout the paper, I assume $\alpha\in[0, 1]$.[6] The assumption of social conscience may seem ad hoc. However, the decision to vote at all is hard to justify by a purely selfish oriented argument as the probability of being decisive is small so the expected gain from voting is likely to be smaller than the cost of voting (Downs, 1957: Ch. 14). For instance Knack (1992) and Mueller (1987) argue that voting may be the outcome of "social behaviour". If the decision to vote is based on non-egoistic reasoning, it seems rather implausible that the political preferences should be purely egoistic. There is also overwhelming experimental evidence to support "social preferences" (Charness and Rabin, 2002), which corresponds closely to a utility function of the form (2). See Galasso (2003) for another approach to incorporating this into political economy models.

To simplify expression (2), consider the class of step functions

$$D_x(y) = \begin{cases} 0 & \text{if } y < x \\ 1 & \text{if } y \geq x, \end{cases} \tag{3}$$

that is, the distribution of a degenerate random variable that equals x with probability one. Now, it is seen that U can be rewritten as

$$U(x,t) = (1-\alpha)S(t,D_x) + \alpha S(t,F) = S(t,(1-\alpha)D_x + \alpha\ F), \tag{4}$$

where the last equality follows from the linearity of S. The second argument in the S-function, $(1-\alpha)D_x+\alpha F$, is the subjective weighting function for the individual, i.e. the weight the agent puts on

[4] For simplicity, I adhere to a Downsian party system throughout the paper. In an Appendix available upon request, I discuss to what extent we can expect the results to hold in a more plausible model of politics.

[5] Instead of valuing social welfare, one could also imagine the case where agents value their own income relative to average income in society. This would essentially be the opposite of the case considered here as one in that case would value the others being poor to have a higher relative income. The model of fractionalized societies developed below could also be captured by assuming that agents compare their income to a weighted average of their group's income and average income in the whole society.

[6] We could also have $\alpha<0$, which implies that the agent derives utility from consumption and superiority to the average of the economy, and also $\alpha>1$ where the agent willingly accepts martyrdom. However, these cases are rather unrealistic.

persons from different income groups. The first term, $(1-\alpha)\,D_x$, generates a mass point at the agent's income level x, whereas the second term gives weight to different income levels according to their frequency in society. If $\alpha=0$, she only cares about agents with her income; if $\alpha=1$ she uses the true distribution in society. For any such weighting function, the agent's preferred tax rate is found by maximizing S with regard to t. Since S is globally concave in t for any weighting function, the maximum is given by the first order condition[7]. It follows that preferences are single-peaked, so the median voter theorem applies. Furthermore, for $\alpha<1$, the optimal t is decreasing in x as the post tax individual income is maximized at a lower tax rate for higher pre-tax income.

I assume that income distributions are continuous, i.e. contains no mass points, so that fractiles always are well-defined. The case of discontinuous distribution functions is discussed in an Appendix available upon request. In the continuous case, the tax rate chosen by the median voter satisfies the system

$$\begin{cases} S_t(t,(1-\alpha)D_{x^m}+\alpha F)=0 \\ F(x^m)=\dfrac{1}{2} \end{cases}. \tag{5}$$

where S_t is the derivative of S with regard to the tax rate t.

2.2. Fractionalized societies

Assume now that the society is divided into a number of mutually exclusive groups where an agent belonging to one group cares more about the welfare of her group than that of other groups. For simplicity, assume that there are only two groups, A and B. The main results hold for multiple groups and overlapping group dimensions, but the model gets more cumbersome. A proportion q of the population belongs to group A and the remaining $(1-q)$ to group B. The income distribution[8] within the groups are described by F_A and F_B which are assumed to have common support Ω.[9] The overall income distribution F satisfies $F(x)=qF_A(x)+(1-q)\,F_B(x)$ for all x. I will say that one group is richer than the other if the two groups' income distributions can be ranked by first order stochastic dominance. Throughout the paper, group A is the rich group and B the poor.

The case with *full group antagonism* is when agents completely ignore the welfare of other groups. Then the utility of a member of group $i\in\{A,B\}$ with endowment x is given by

$$U_i(x,t)=S(t,(1-\alpha)D_x+\alpha\,F_i). \tag{6}$$

A less extreme case that is also easier to analyse is one where agents put some weight on their group and some on the society as a whole. I will label this *partial group antagonism*. Here, agents from group i with endowment x have preferences

$$U_i(x,t)=S(t,(1-\alpha)D_x+\beta\alpha F_i+(1-\beta)\alpha F). \tag{7}$$

[7] Given the characteristics of T, S is always maximized for a $t<1$. If we require $t\geq 0$, there may be corner solutions for some agents. Although negative redistribution is unrealistic I will not exclude it to maintain analytic simplicity.

[8] We may also allow agents to put different weights on agents with different endowments in their welfare calculi. The analysis so far has assumed that F_A and F_B correspond to actual income distributions but this is not necessary. If we keep $\bar{x}$ fixed, these cumulative income distributions may also include a subjective weighting of the different income groups.

[9] The assumption of a common support is stronger than necessary. The necessary condition is that the incomes of both pivotal agents lies in the interior of the support of the group's income distribution.

I will restrict attention to $\beta\in[0, 1]$. When $\beta=1$, we have the full antagonism case whereas the case without group antagonism corresponds to $\beta=0$.[10] I will label the parameter β the degree of group antagonism. An increase in β implies that agents put more emphasis on their own group and less on society as a whole. Notice that β is not the Herfindahl index of fractionalization used in empirical analyses.

As shown above, preferences are single-peaked and within one group, the desired tax rate is decreasing in x. However, two persons with identical endowments, but belonging to different groups, have in general different preferred tax rates.[11] Hence it is insufficient to look at the initial endowments to find the median voter. In fact, we will have two median voters, one from each group.[12] I denote these as pivotal agents. They have a common preferred tax rate, but in general their endowments differ. Call the endowment of the A pivotal voter x_A^m and that of the B pivotal voter x_B^m. Then the tax rate t chosen by the pivotal voters satisfies the system

$$\begin{cases} S_t(t,(1-\alpha)D_{x_A^m}+\beta\alpha F_A+(1-\beta)\alpha F)=0 \\ S_t(t,(1-\alpha)D_{x_B^m}+\beta\alpha F_B+(1-\beta)\alpha F)=0 \\ qF_A(x_A^m)+(1-q)F_B(x_B^m)=\dfrac{1}{2} \end{cases}. \tag{8}$$

In general, we have $F_A(\cdot)\neq F_B(\cdot)$. Then normally the group-wise socially optimal tax rates differ, so two agents from different groups with the same income x will have different preferred tax rates for any x. When there is some degree of group antagonism, the person belonging to the richest group prefers a lower tax rate than the one belonging to the poorest group. Then it follows that in the system (8), $x_A^m\neq x_B^m$, and the endowment is lowest for the one belonging to the richest group. Notice also that x_A^m and x_B^m do usually not correspond to the median endowment of the respective group, but is determined by the system (8) and corresponds to the incomes of the agents with median tax preference. How should we understand this group-restricted social conscience? First, it is closely related to Sen's (1965) concept of sympathy between individuals. However, unlike his study, I impose a stronger symmetry on the structure of social consciousness. It may also arise if we view social conscience as a result of reciprocity (Bowles et al., 2006; Bowles and Gintis, 2000; Charness and Rabin, 2002). In a situation where one person's caring for another is conditional on the second caring for the first as well, an equilibrium and focal point is that everybody cares about their fellow group members and no others. Secondly, a highly group-based social conscience corresponds closely to the sociological concept of group self-interest which finds strong empirical support in studies of preferences for welfare spending (Bobo and Kluegel, 1993). For instance Kinder and Sanders (1996) find virtually no support for self-interest affecting political opinions, but conclude that group self-interest plays an important role. In the model set out above, this would mean both a high degree of social conscience α and a high degree of group antagonism β. Group antagonism can also be interpreted as a belief that people from one's own group are more deserving of public transfers than others, as found by e.g. Gilens (1999). Finally, this restricted social conscience may also be seen as an extension of Barro's (1974) dynastic utility function where the family now also includes the group, although possibly with a smaller weight.

[10] We could also have $\beta>1$, which is the racist agent who wants to hurt the other group, and $\beta<0$, which could be a "militant anti-racist" who wants to punish her own group. Both cases are rather extreme.

[11] This is a quite general result in models where agents differ by income and other characteristics, such as overlapping generations-models (Persson and Tabellini, 2000: Section 6.2.2).

[12] More precisely, we have two sets of median voters, both with measure zero. Speaking of these sets as particular voters is an abuse of language, but it makes the analysis more readable.

In the current model, the only objective of the government is to transfer income between individuals. In a dynamic setting, there could also be demand for a social insurance scheme. Consider for a moment the following reinterpretation of the model: First, a tax rate is chosen by direct voting, and society keeps this tax rate forever. Agents are subject to income shocks arriving by some Poisson process, and if they are hit by a shock their income is redrawn from their group's income distribution. With an appropriate discount rate below unity, this will give a utility function of the form (7).[13]

3. The size of government

3.1. The effect of group antagonism

First, I will study the effects of group antagonism, measured as an increase in β, on the size of government, measured as the amount of redistribution the government gives. Call the marginal density functions associated to F_A and F_B f_A and f_B respectively. Assume that there are no holes or mass points so that $0<f_i\ (x)<\infty$ for all $i\in\{\text{A, B}\}$ and $x\in\Omega$. Some deviations from these assumptions are discussed in an Appendix available upon request. Differentiation of the system (8) yields

$$S_{tt}^{A}\mathrm{d}t + (1-\alpha)\frac{\partial S_t(t, D_{x_A^m})}{\partial x_A^m}\mathrm{d}x_A^m + \alpha S_t(t, F_A - F)\mathrm{d}\beta = 0 \tag{9a}$$

$$S_{tt}^{B}\mathrm{d}t + (1-\alpha)\frac{\partial S_t(t, D_{x_B^m})}{\partial x_B^m}\mathrm{d}x_B^m + \alpha S_t(t, F_B - F)\mathrm{d}\beta = 0 \tag{9b}$$

$$qf_A(x_A^m)\mathrm{d}x_A^m + (1-q)f_B(x_B^m)\mathrm{d}x_B^m = 0 \tag{9c}$$

where $S_{tt}^{i} = S_{tt}\ [\mathrm{t}, (1-\alpha)\ D_{x_i^m} + \alpha\beta F_i + \alpha\ (1-\beta)F] < 0$, $i\in\{A, B\}$. Define

$$\hat{w}_A = \frac{s_A q f_A(x_A^m)}{s_A q f_A(x_A^m) + s_B(1-q)f_B(x_B^m)} \qquad s_A = -\left(\frac{\partial^2 u((1-t)x_A^m + T(t)\bar{x})}{\partial x_A^m \partial t}\right)^{-1}$$

$$\text{and}$$

$$\hat{w}_B = \frac{s_B(1-q)f_B(x_B^m)}{s_A q f_A(x_A^m) + s_B(1-q)f_B(x_B^m)} \qquad s_B = -\left(\frac{\partial^2 u((1-t)x_B^m + T(t)\bar{x})}{\partial x_B^m \partial t}\right)^{-1} \tag{10}$$

Then the implicit function theorem yields

$$\frac{\mathrm{d}t}{\mathrm{d}\beta} = -\alpha q(1-q)\frac{s_A f_A(x_A^m) - s_B f_B(x_B^m)}{\hat{w}_A S_{tt}^{A} + \hat{w}_B S_{tt}^{B}} S_t(t, F_A - F_B). \tag{11}$$

[13] Even if agents have the same degree of social conscience for both their own group and other groups, a segmented labour market may be sufficient to make it appear as if the agent had a group biased social conscience. To see this, consider a case where agents receive income shocks with some hazard rate λ, whereby their income is redrawn from a group specific distribution F_i. Then the individual term in their utility function is no longer $(1-\alpha)\ D_x$ but $(1-\alpha)\left[\frac{1}{1+\lambda}D_x + \frac{\lambda}{1+\lambda}F_i\right]$. So we could reinterpret the whole model as an analysis of the consequences of a segregated labour market.

From this expression, we see that the tax rate is decreasing in β if $s_A f_A\ (x_A^m) > s_B f_B\ (x_B^m)$, maintaining the assumption that the As are the richer so that $S_t\ (t,\ F_A - F_B) < 0$.

Consider first the case where $\beta = 0$, so that $s_A = s_B$ and $S_{tt}^A = S_{tt}^B$. Then the incomes of the pivotal voter from the two groups both correspond to the median income in society x^m, and

$$\frac{\mathrm{d}t}{\mathrm{d}\beta} = -\alpha q(1-q)\frac{s_A}{S_{tt}^A}[f_A(x^m) - f_B(x^m)]S_t(t, F_A - F_B). \tag{12}$$

This expression is negative if $f_A\ (x^m) > f_B\ (x_m)$. When β rises marginally from $\beta = 0$, the pivotal A-voter cares less about group B, and consequently prefers a lower tax rate whereas the pivotal B-voter now cares less about group A and therefore prefers a higher tax rate. Consequently, as β increases, the pivotal voters will be an A-agent with endowment $A < x_A^m < x^m$ and a B-agent with endowment $x_B^m > x^m$. If $f_B\ (x^m)$ is small relative to $f_A\ (x^m)$, $|x_B^m - x^m|$ will be large relative to $|x_A^m - x^m|$, so the income of the new B-pivotal voter will be high relative to x^m, the income of the former pivotal voter. Although she has a tendency to prefer high tax rates since group B is poorer than group A, this tendency is weakened by her wish to have low transfers because she is rich.

To understand how tax preferences change when β rises, notice that the weighting function for an A-agent with income x can be written as $(1-\alpha)\ D_x + \alpha\beta F_A + \alpha\ (1-\beta)\ F = (1-\alpha)\ D_x + \alpha(\beta + (1-\beta)\ q)F_A + \alpha\ (1-\beta)\ (1-q)\ F_B$. Here it is seen that the effect of a change in β on the weighting function is greater the smaller q is. If q is close to unity, then F already gives group A a large weight, and a change in β has less effect than if the group has a smaller weight in F. A similar argument holds for group B. Hence the smaller a group is, the larger are the changes in tax preferences within the group.

The effect of a rise in β is determined by two factors: How much tax preferences changes within each group, and the measure of voters the group has close to the decisive agents. If tax preferences change a lot within a group, this decreases that group's power in the political struggle as their pivotal voter is quickly swapped with a new pivotal voter that to a large extent accommodates the preferences of the other group. The measure of agents at a given income level in each group determines the number of voters that has to be swapped, and hence increases political influence. This measure may be divided into two factors, the size of the groups q_A and q_B and the measure of each income level within the group given by the density $f_i\ (x)$. Hence there are a total of three factors to take into account. When we have a continuous income distribution and $\beta = 0$, the effect of group size exactly offsets the effect of changes in preferences. Then what matters is the relative size of each income level within the group. If the density is high close to the median income of society, the group is influential.

Throughout the paper, I make the following assumption:

Assumption 1. $s_A\ f_A\ (x_A^m) > s_B\ f_B\ (x_A^m)$.

Recall that at $\beta = 0$, Assumption 1 simplifies to $f_A\ (x_A^m) > f_B\ (x_B^m)$. Whether $f_A\ (x^m) - f_B\ (x^m)$ is positive or negative will depend on the shape of the income distributions and the endowments of the pivotal voters. At $\beta = 0$, both pivotal voters have the same endowment x^m. However, since the As are richer than the Bs, the pivotal voter from group A is in a lower income fractile than the one from group B. If the shape of the distribution for the As and the Bs are relatively similar and skewed, this usually implies that $f_A\ (x^m) - f_B\ (x^m)$ is positive. Although it is not difficult to find distributions such that $f_A\ (x^m) - f_B\ (x^m)$ is not positive, it is probably at worst only slightly negative in most real world cases. When β increases, x_A^m will decrease, and x_B^m decrease. In most cases, this will increase $f\ (x_A^m)$ and decrease $f\ (x_B^m)$, hence increasing $f_A\ (x^m) - f_B\ (x^m)$ so the requirement for a negative effect on the tax rate is more likely to be satisfied. However, whether it is positive is an empirical question. For most of the families of distributions conventionally used to model income

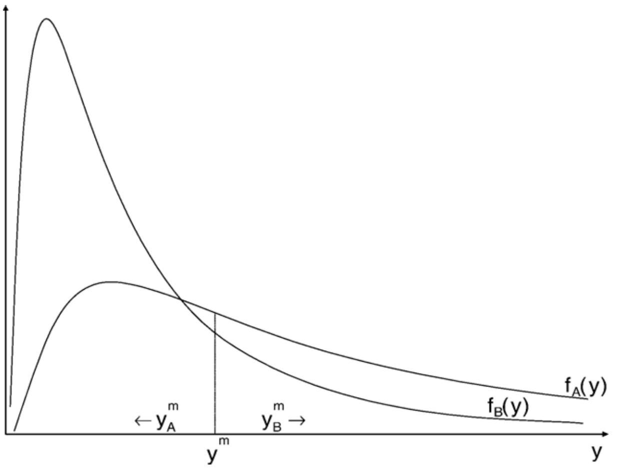

Fig. 1. Example of densities at the mean.

distributions, it is possible to both find cases where Assumption 1 holds and doesn't hold. For instance, if both F_A and F_B are log normal, Assumption 1 holds if $q>1/2$.

When $\beta>0$, the group weights s_A and s_B will also play a role. s_i capture the effect of changes in tax preferences through changes in the marginal valuation of consumption. These variables give the change in the effect of increased income on tax preferences, and their relative magnitudes depend on the third derivative of the utility function. Unless u''' is strongly positive, which is unlikely, we have $s_A<s_B$ which tends to make Assumption 1 less likely. We have

$$\frac{\partial[s_A f_A(x_A^m)-s_B f_B(x_B^m)]}{\partial x_A^m}=\frac{f_A}{1-q}\left[(1-q)s_A\left(\frac{f_A'}{f_A}+\frac{s_A'}{s_A}\right)+qs_B\left(\frac{f_B'}{f_B}+\frac{s_B'}{s_B}\right)\right], \tag{13}$$

where s_i' is the derivative of s_i wrt. x. This expression is negative if the elasticity of f_i wrt. x, which is negative, is larger in absolute value than the elasticity of s_i. As x_A^m is decreasing in β, Assumption 1 is more likely to hold when Eq. (13) is negative. This is the case if agents are not too risk averse.[14] In the special case of risk neutrality, s_i will be independent of x and Assumption 1 will hold as long as $f_A(x^m)>f_B(x^m)$ and we are on the decreasing parts of f_i. Consequently, although there may be conditions under which Assumption 1 does not hold, we know for sure that if group A is the largest group, both income distributions F_A and F_B are log normal, and risk aversion is sufficiently low, Assumption 1 will hold.

To summarize the discussion so far, we can state the following first main result:

Proposition 1. *Assume group A is richer than group B in the sense of first order stochastic dominance and that Assumption 1 holds. Then a rise in the degree of group antagonism* β *decreases the politically chosen tax rate.*

[14] To see this, notice that for a CRRA utility function $u(c)=\frac{c^{1-\sigma}-1}{1-\sigma}$, we have

$$\frac{s'}{s}=\sigma(1-t)\left(\frac{1}{c}+\frac{\tau}{\sigma\tau+(1-\sigma)\tau}\right)$$

where c is post-tax consumption $(1-t)\,x+T\,(t)\,\bar{x}$ and $\tau=[T\,(t)+T'\,(t)]\,\bar{x}$.

The effect of fractionalization and group antagonism, given by Eqs. (11) and (12), depends crucially on the difference between the densities at the median for the groups.

I argued that the density would be higher for the richest group. The effect of increased group antagonism staring at $\beta=0$ (no group antagonism) will depend on the difference at the median income of the population as a whole. An example is depicted in Fig. 1, where the density for group A is higher than for group B at the overall median y^m. As β increases, the relevant densities are to the left of the median for group A and to the right of the median for the group B, re-enforcing the effect.

I have performed some simple calculations on the densities for the US income distributions for Blacks and Whites for the years 1967 to 2001. The detailed results are presented in Appendix A. The finding is that for all these years, the density for Whites is higher than that for Blacks. Hence for the US, the models quite clearly predict that a rise in the degree of group antagonism should lower the support for redistribution. At the median, the marginal distribution function is also decreasing for both Blacks and Whites. Hence when β increases, the Black pivotal voter is pushed to the right, increasing the density and vice versa for the White pivotal voter.

So far we have assumed that the group antagonism is identical in both groups. It may be plausible that one group is more self-centered than the other, though. Let members of group A have group antagonism $\pi\beta$, so their group antagonism is larger than that of group B if $\pi>1$. An increase in π would mean that group A becomes more group-biased relative to group B. Solving a system similar to Eqs. (9a) (9b) (9c), we get

$$\frac{\mathrm{d}t}{\mathrm{d}\pi} = -\frac{\alpha\beta(1-q)}{S_{tt}^A + \frac{1-q}{q}\frac{f_B(x_B^m)}{f_A(x_A^m)}\frac{s_B}{s_A}S_{tt}^B} S_t(t, F_A - F_B) < 0$$

as $S_{tt}^i<0$, $i\in\{A, B\}$ and $S_t(t, F_A-F_B)<0$ when group A is richer than group B. Hence if the rich group shows a large degree of group antagonism or the poor group a low degree of group antagonism, redistribution decreases. The explanation in both cases is that more weight is put on the members of the rich group and less on the members of the poor group.

3.2. The effect of diversity

We see from Eq. (11) that the magnitude of the effects of group antagonism β on the tax rate depends on $q\ (1-q)$, the Herfindahl measure of fractionalization. Hence increased fractionalization increases the effect of changes in group antagonism.

But it is also interesting to study the effect of increased diversity for a given level of group antagonism. As the overall income distribution F, and hence the mean income $\bar{x}$, will depend on q, we cannot study the effect of a change in diversity by simply changing q or the Herfindahl index $q\ (1-q)$. What we need is to change the fractions q and $1-q$ in a way so F remains unchanged. One possibility would be to re-label some agents in group B as members of group A to increase q. However, this will at the same time make the groups more equal, which works in the opposite direction. Hence we should also require that the new group-wise distributions, call them $\tilde{F}_A$ and $\tilde{F}_B$, keep the same distance as F_A and F_B in the sense that $\tilde{F}_A(x)-\tilde{F}_B(x)=F_A(x)-F_B(x)$ for all x. This only holds if

$$\begin{cases} \tilde{F}_A = F_A + (q-q_0)(F_B - F_A) \\ \tilde{F}_B = F_B + (q-q_0)(F_B - F_A), \end{cases} \tag{14}$$

where q_0 is the original fraction of As and q the fraction of As after the change in diversity. The equilibrium tax rate is now determined by the system

$$S_t(t,(1-\alpha)D_{x_A^m}+\beta\alpha\ \widetilde{F}_A+(1-\beta)\alpha F)=0 \tag{15a}$$

$$S_t(t,(1-\alpha)D_{x_B^m}+\beta\alpha\ \widetilde{F}_B+(1-\beta)\alpha F)=0 \tag{15b}$$

$$q\widetilde{F}_A(x_A^m)+(1-q)\ \widetilde{F}_B(x_B^m)=\frac{1}{2}. \tag{15c}$$

To see the effect of increased diversity, we want to study the effect of equilibrium t by a change in q. Assume group A is the largest group, so an increase in diversity $q\,(1-q)$ is equivalent to a decrease in q. Then the effect of diversity on the tax rate is the opposite of the effect of

$$\left.\frac{\mathrm{d}t}{\mathrm{d}q}\right|_{q=q_0}=-\frac{(1-\alpha)\Lambda+\alpha\beta(\omega_A+\omega_B)S_t(t,F_B-F_A)}{\omega_A S_{tt}^A+\omega_B S_{tt}^B}$$

where

$$\Lambda=(1-q)[F_A(x_A^m)-F_A(x_B^m)]+q[F_B(x_A^m)-F_B(x_B^m)]<0,$$

$$\omega_A=s_A q f_A(x_A^m)>0$$

$$\omega_B=s_B(1-q)f_B(x_B^m)>0,$$

and where s_A and s_B were defined in Eq. (10).

As $S_t\,(t, F_B-F_A)>0$ and $S_{tt}^i<0$, we have two effects that pull in opposite directions. Notice first that if $\alpha=0$ or $\beta=0$, $\mathrm{d}t/\mathrm{d}q=0$ as the model reduces to the standard Meltzer and Richard (1981) model. Also, if $\alpha\beta$ is large, the second effect will dominate. However, there could be cases with α and β both close to zero where the first effect dominates, and where increased diversity, so a reduction in the fraction of As, will increase redistribution. However, we do have the following result:

Proposition 2. *For sufficiently large $\alpha\beta$, increased diversity a defined by* Eq. (14) *reduces the equilibrium tax rate.*

This extends Alesina et al.'s (1999) and Miguel and Gugerty's (2005) results on public good provision to redistribution.

4. Fractionalization and total welfare

Let us now consider the case of a general income distribution studied in Section 3. In the case of partial group antagonism, the first order condition from the optimal choice of taxes for a pivotal voter from group A is $S_t\,(t, (1-\alpha)\,D_{x_A^m}+\alpha\beta F_A+\alpha\,(1-\beta)\,F)=0$, which we may rewrite

$$S_t(t,F)+\Psi(t)=0$$

where

$$\Psi(t)=(1-\alpha)S_t(t,D_{x_A^m}-F)+\alpha\beta(1-q)S_t(t,F_A-F_B), \tag{16}$$

and of course a similar expression holds for a pivotal voter from group B. Ψ is the deviation from social welfare in the agent's maximand.[15] The first term of Ψ is the effect of the pivotal voter caring more about herself than other individuals in society and the last term stems from the pivotal voter caring more for group A than group B. It is clear that the absolute value of the second term is increasing in β. It is seen that for $\alpha=1$, the first term disappears and it follows that group antagonism necessarily decreases social welfare. For $\alpha=0$, on the other hand, antagonism does not matter.

One can show that $\Psi(t)$ is decreasing in β when Assumption 1 holds.[16] If $S_t(t, D_{xm}-F)>0$, i.e. the original pivotal voter privately prefers a tax rate above the social optimum, then at least some antagonism enhances the economic efficiency by lowering the tax rate. If the pivotal voter prefers a tax rate that is too low, then antagonism is detrimental to social welfare.

The intuition for these results is very simple. Group antagonism will push the desired tax rate towards what is beneficial for the group to which the individual belongs. Whether this will improve social welfare, will depend on the combination of individual and group. For a poor agent in a poor group, group antagonism will push his chosen tax rate up, further away from the social optimum. In contrast, for a poor agent in a rich group, group antagonism will push the tax rate down. Starting from a situation without group antagonism, $\beta=0$, a decrease in a tax rate preferred by a poor agent will always improve social welfare. We may say that group antagonism works as a counter weight to the poor agent's extreme private preferences.

5. Income distribution and the size of government

We can use the results obtained above to study the effects of increased inequality in fractionalized societies. I first look at the effect of increased intra-group inequality. Consider the case of full group antagonism, found by setting $\beta=1$ in the system (8). An increase in inequality may be studied as a mean preserving spread which is equivalent to second order stochastic dominance. If the income distribution of group i changes from F_i^0 to F_i^1, inequality has increased if F_i^0 second order stochastically dominates F_i^1. Under general conditions, this implies that the pivotal voter of group i now prefers a higher tax rate. Consider first a purely altruistic agent, so $\alpha=1$ and the optimal tax rate of the pivotal voter before the shift is determined as the solution to $S_t(t, F_i)=0$. $\frac{\partial}{\partial t}u[(1-t)x + T(t)\overline{x}]$, the integrand of S_t, is decreasing and convex in x when $u''<0$ and $u'''>0$.[17] Hence when F_i^0 second order stochastically dominates F_i^1, $S_t(t, F_i^0)<S_t(t, F_i^1)$, a parallel to the well-known result on second order stochastic dominance in the theory of choice under uncertainty (Rotschild and Stiglitz, 1970). As S_t is decreasing in t, this implies that the pivotal voter prefers a higher tax rate under F_i^1 than she used to under F_i^0. A similar result holds as long as $\alpha>0$. When $\alpha=0$, we are back to the classic result that the tax rate increases if the mean to median ratio increases. To summarize, if inequality increases in one or both groups, the size of government increases. It is easily seen that if inequality increases in one group, it also increases in society as a whole. Hence the pivotal voter in group i also prefers a higher tax rate in cases with less than full group antagonism. These results are very similar to those found in the ordinary Romer–Roberts–Meltzer–Richard model.

[15] Notice that social welfare $S(t, F)$ is the same whether we add the private utilities u or the utility functions U incorporating their social conscience.

[16] This is done by differentiation (16) with regard to β and inserting from Eqs. (9a) and (12).

[17] I haven't made any assumptions on u''' so far. However, it is positive for most common specifications of u. Particularly, it holds for the class of utility functions yielding hyperbolic absolute risk aversion, hence more particularly CARA and CRRA.

An increase in inter-group inequality is more interesting. Assume that initially, both groups have the same income distribution F. An increase in inter-group inequality is a situation where the income distribution of groups A and B move to income distributions F_A and F_B with the properties that $F = qF_A + (1-q)\ F_B$ and where F_A first order stochastically dominates F_B. For analytical simplicity, I will concentrate on a continuous transition between the two states where group $i \in \{A, B\}$ has the income distribution $\tilde{F}_{\gamma i} := \gamma Fi + (1-\gamma)\ F$. Denote by $\tilde{f}_{\gamma i}$ the marginal density of $\tilde{F}_{\gamma i}$. For all γ, the economy-wide income distribution remains fixed, but as γ increases, the difference between the groups increases. When we limit our attention to the case of full group antagonism, the politically chosen tax rate t satisfies the following system, similar to the equations studied in Section 3:

$$S_t(t, (1-a)D_{x_A^m} + \gamma\alpha F_A + (1-\gamma)\alpha F) = 0 \tag{17a}$$

$$S_t(t, (1-\alpha)D_{x_B^m} + \gamma\alpha F_B + (1-\gamma)\alpha F) = 0 \tag{17b}$$

$$q\widetilde{F}_{\gamma A}(x_A^m) + (1-q)\widetilde{F}_{\gamma B}(x_B^m) = \frac{1}{2}. \tag{17c}$$

As γ enters (Eq. (17c)), the analysis of this system is slightly more involved than of Eq. (8). However, the results are almost identical. In Appendix B I prove the following result:

Proposition 3. *Assume that group A is richer than group B in the sense of first order stochastic dominance and that Assumption 1 holds. Then a mean preserving increase in between group inequality decreases the politically chosen tax rate.*

The intuition for this result is analogue to that of Proposition 1: An increase in between group inequality γ will induce all the A-voters to prefer lower and the B-voters higher tax rates. The outcome of these changes in preferences again boils down to who has the highest density of voters close to the pivotal voters, weighted by the preference weights s_i. If we assume that the weighted density for the rich group is higher than for the poor, the new equilibrium is a lower tax rate.

When a society is fractionalized, there is a tendency towards reduced tax rates when the inter-group inequality rises. If the rate of social conscience is not too low, we can expect a rise in inter-group inequality to reduce the size of government, also if there is a rise in inter-group inequality at the same time.

6. A test of the model

In this section I report results from estimations to study the validity of some of the model's predictions. It would be interesting to study the effect of group antagonism β on equilibrium redistribution, but as β is not easily observed for a geographical area, the test of Proposition 1 is based on an intermediate result that is the driving force of the conclusion. I will test the following predictions:

1. Increased group antagonism leads to decreased support for redistribution for members of rich groups and increased support for redistribution for members of poor groups (Proposition 1).

2. For a given level of group antagonism, a higher degree of fractionalization leads to less redistribution (Proposition 2).
3. Within group inequality increases the support for redistribution (Proposition 3).
4. Between group inequality reduces the support for redistribution (Proposition 3).

6.1. Group antagonism and the support for redistribution

To test the first prediction, that increased group antagonism leads to reduced support for redistribution for members of rich groups and increased support for redistribution for members of poor groups, I use data from the US General Social Survey (GSS). This is a survey where a large number of respondents have been asked a number of questions including opinions on redistribution and racial matters. To measure support for redistribution, I use a dummy for whether respondents answer that the government should spend more on welfare. Estimation is by a simple linear probability model, but probit and ordered probit analyses give qualitatively similar results. Column (1) of Table 1 shows a baseline regression where support for redistribution is regressed on race dummies and a number of control variables. As expected, African American respondents are more supportive of redistribution than White respondents. Other races are also somewhat more supportive of redistribution than Whites, but this effect is much weaker and only significant at the 10% level. This supports the

Table 1
Relationship between preferences for redistribution and racial relations

	(1)	(2)	(3)	(4)
African American	0.142	−0.082		
	(0.005)***	(0.056)		
Other race	0.017			
	(0.010)*			
Close feeling to race:				
Respondent black, how close to white		−0.010		
		(0.006)*		
Respondent black, how close to black		0.030		
		(0.006)***		
Respondent white, how close to black		0.004		
		(0.003)		
Respondent white, how close to white		−0.002		
		(0.003)		
Not object to African American at home			0.030	
			(0.006)***	
Had African American at home recently				0.027
				(0.005)***
Sample	All	African American, White	White	White
Period	1972–2002	1996–2002	1972–2002	1972–2002
Observations	36948	3738	30932	30932
R^2	0.08	0.05	0.04	0.04

Dependent variable is a dummy for preferring to spend more on welfare. All regressions include log household income, age, age squared, years of education, years of education squared, and dummies for sex, marital status, region of residence, and year.
Standard errors in parentheses. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence.

intermediate result that members of poor groups prefer more redistribution than members of rich groups, ceteris paribus.

The next step is to study the effect of group antagonism on the support for redistribution. Column (2) of Table 1 uses two questions on "How close do you feel to Blacks?" and "How close do you feel to Whites?" The answer is given on a scale from 1 (not close) to 9 (close). A reasonable interpretation is that if someone feels close to the other group, it indicates low group antagonism whereas it is a sign of high group antagonism if someone feels close to their own group. African American respondents are more supportive of redistribution if they feel close to their own group and less supportive if they feel close to Whites, although the second effect is weaker. For white respondents, there are indications that they are more opposed to redistribution if they feel close to their own group and more supportive if they feel close to Blacks. Neither of these results are significant at usual levels of significance, but this is probably to some extent due to the smaller sample size as the questions on racial closeness have only been included in the surveys from 1996 to 2002. Taken together, these results indicate that increased group antagonism increases the support for redistribution among African Americans and reduces the support for redistribution among Whites.

To further study the preferences of white respondents, columns (3) and (4) of Table 1 report estimations from the effect of two other measures of closeness to African Americans on support for redistribution. In column (3), a dummy for answering that they would not object if a member of the family would bring an African American friend home to dinner is used, and in column (4) a dummy for whether anyone in the family has brought an African American friend home for dinner during the last few years is used. Both dummies can probably be taken for being more closely related to African Americans. As predicted by the model, both variables tend to increase support for redistribution, and both estimates are significant at all reasonable levels of significance. This strengthens the conclusion that reduced group antagonism in the rich group increases their support for redistribution.

6.2. Data on equilibrium redistribution

To test the last three hypotheses on the determinants of equilibrium redistribution, I use a panel of US states with six observations per state.[18] The main reason for using a single country is that the definition of groups and the collection of data on groups are more homogeneous. The District of Columbia is not included as it is ruled directly by the US Congress, so local characteristics to a lesser extent determine its policy. We need measures of inequality both between and within groups. As such data are not readily available, the measures had to be constructed from micro data. Income data are taken from March Current Population Survey, made available through the Luxembourg Income Study (LIS).[19] For purposes of politically determined tax rates, the relevant measure of income is pre-tax factor income. Household incomes are normalized according to the square root equivalence scale. As we want to decompose inequality into within- and between-group inequality, it is desirable to use a decomposable inequality measure. Recall that an inequality measure I is

[18] The states are observed in 1969, 1974, 1986, 1991, 1994, 1997, and 2000. Although 1979 is also available from the LIS, these data lack information about state of residence, rendering them useless. Furthermore, I do not have data on average share of transfers to disposable income for 1969.

[19] See http://www.lisproject.org for details.

said to be decomposable if for some vector of incomes $\mathbf{y}$ and some partition of it $\mathbf{y}_1,\ldots,\mathbf{y}_G$, we have

$$I(\mathbf{y}) = \sum_{g=1}^{G} w_g I(\mathbf{y}_g) + I(\bar{\mathbf{y}}_1,\ldots,\bar{\mathbf{y}}_G) \tag{18}$$

where $\bar{\mathbf{y}}_g$ is the vector where all members of group g have the group's mean income and w_g is a set of weights. Hence total inequality is the sum of within group inequality and between group inequality, where within group inequality is calculated as a weighted sum of inequality within each group. See e.g. Cowell (2000) for further details.

Requiring the transfer principle and independence of scale to hold, we are left with the class of generalized entropy measures

$$I_{GE}^{\kappa} = \frac{1}{\kappa(\kappa-1)} \int \left[\left(\frac{x}{\mu} \right)^{\kappa} - 1 \right] \mathrm{d}F(x), \tag{19}$$

where F is the CDF of the income distribution, μ the mean income, and κ a parameter (Bourguignon, 1979; Shorrocks, 1981). The higher is κ, the more weight the measure puts on inequality in the upper range of the income distribution. I concentrate on $\kappa = 0$, which should capture the inequality close to the median reasonably well. Then we have $I_{GE}^{0} = -\int \ln\left(\frac{x}{\mu}\right) \mathrm{d}F(x)$, the mean logarithmic deviation.

I use two measures of redistribution. The first measure is state expenditure on public welfare as a share of state personal income. Data on public welfare is taken from *Government finances* (US Department of Commerce, various years) whereas state personal income is from the Bureau of Economic Analysis.[20] This measure is probably rather narrow, so I also use a second measure, the average share of transfers received by households as a share of disposable income, calculated from the LIS data. This measure includes federal transfers, so it is broader than what we actually want as a measure of redistribution. However, it is unlikely that the transfers a state receive should depend more on within than between race inequality, so this should not be an obstacle for the relevant tests. Ideally, we should have a measure that is broader than the first and less broad than the second. However, such data do not exist. Consequently, I use both, and probably the truth is somewhere between the two.

To measure group fractionalization, I use the conventional Herfindahl measure which gives the probability that two randomly selected persons belong to different groups.[21] The fractionalization index is calculated from the LIS data used for calculating the between-group inequality measure. This is to avoid the inequality measure picking up elements of the fractionalization measure. Comparing my fractionalization values with values obtained from the 2000 Census, I get an overall correlation of .84, ranging from .77 in 1969 to .90 in the 1990s. This indicates that my measure should be appropriate. Data on the fraction of the population above 65 is also derived from the LIS data.

Table 2 gives basic descriptive statistics of the data and Fig. 2 shows the geographical distribution of fractionalization, total, within, and between group inequality, and the two measures of redistribution. All numbers are measured in 2000. The degree of fractionalization

[20] Available at http://www.bea.doc.gov/bea/regional/spi/.

[21] The groups differ between surveys. To get consistent data, I use three groups, Black, White, and Other, for all years.

Table 2
Descriptive statistics

Variable	Observations	Periods	States	Mean	Standard deviation		
					Overall	Between	Within
Total inequality	350	7	50	0.57	0.13	0.05	0.12
Between group inequality	350	7	50	0.010	0.025	0.011	0.022
Within group inequality	350	7	50	0.56	0.13	0.05	0.12
Fractionalization	350	7	50	0.18	0.13	0.12	0.04
Fraction above 65	350	7	50	0.09	0.03	0.02	0.02
Log per capita income	350	7	50	9.69	0.68	0.14	0.66
Fraction expenditure on welfare	350	7	50	0.02	0.01	0.01	0.01
Average share of transfers to disp. income	300	6	50	0.15	0.03	0.02	0.02

Inequality is measured by the generalized entropy measure with coefficient 0. Between standard deviations are standard deviations of the state averages and within the average within state standard deviation. District of Columbia not included.

has high values in the South and to some extent the South West. There are no very clear patterns for total and within group inequality, but between group inequality is highest in the South like the degree of fractionalization. Redistribution is highest in the Mid-West and the North-East. Within group inequality is essentially uncorrelated with between group inequality (the correlation coefficient is −.12, and is insignificant if we control for year) and does not seem to follow any strong geographical patterns. Finally, transfers are generally high in the Mid-west and the North-East.

6.3. Empirical results

Tables 3 and 4 show the main results. Table 3 uses the fraction of state welfare expenditure in state personal income as dependent variable, whereas Table 4 uses the average share of transfers in household disposable income. The first thing we notice is that overall factor income inequality seems to induce higher transfers, as predicted by the Romer–Roberts–Meltzer–Richard model.[22] A one standard deviation increase in inequality increases the fraction of state expenditure on welfare by .2 percentage points or about .2 of a standard deviation. However, when the dependent variable is share of state expenditure on welfare, this result is not very robust to the introduction of state fixed effects, so it seems that between state variation is mostly driving the result. Fractionalization seems to have a negative effect on transfers, but this effect is not very strong and it is not robust to the introduction of state fixed effects. As fractionalization changes little over time, this is not surprising.

According to the results discussed in Section 5, within group inequality should increase redistribution whereas between group inequality should reduce it. This is contrary to the conventional wisdom from the Romer–Roberts–Meltzer–Richard model where inequality within and between groups have the same effect. In column (3) I split inequality into within and between inequality. Fig. 3 shows the partial regression coefficients on the two measures of inequality from this regression. Only year 2000 is shown, but this seems to be a quite representative year. We see

[22] This is contrary to the findings in the older papers surveyed by Bénabou (1996) as well as the results of e.g. Moffitt et al. (1998). However, they use a restrictive measure of inequality and study hourly wages whereas I use total earnings.

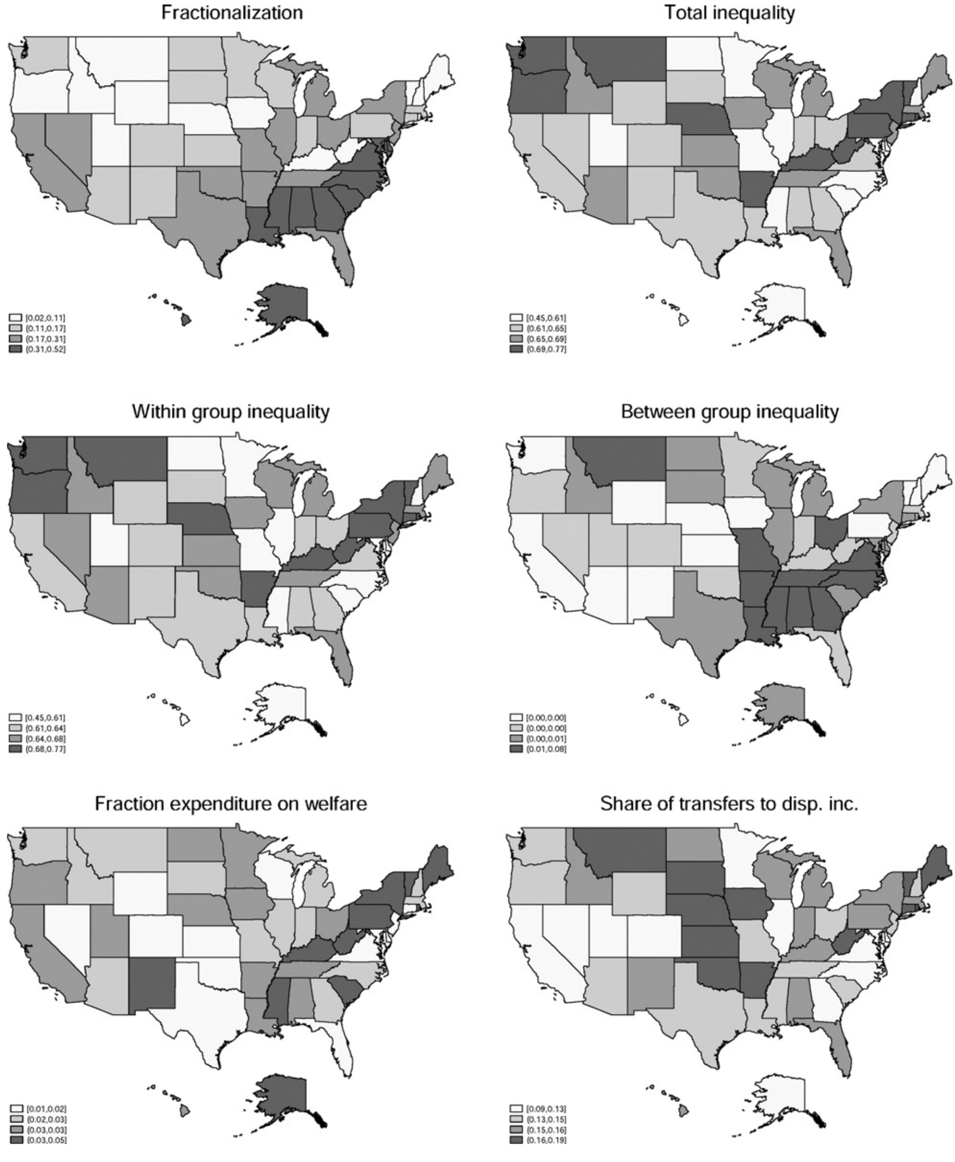

Fig. 2. Geographical presentation of the data. All measured in 2000, inequality measure is generalized entropy measure with parameter 0.

that the estimates conform to the expectations from the theoretical model, although the coefficient on between group inequality is not significantly different from zero. Notice, however, that Proposition 3 predicts that increased inequality reduces the support for redistribution, keeping the overall income distribution constant. When we observe empirically that the measure of between group inequality increases, this entails increased overall inequality as well. A more appropriate test of point 3 is therefore to see whether the effect of between group inequality is significantly

Table 3
Inequality and redistribution measured by

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction above 65	−0.026	−0.015	−0.037**	−0.010	−0.029**	−0.034***	−0.050**
	(0.017)	(0.014)	(0.018)	(0.015)	(0.014)	(0.010)	(0.022)
Log per capita income	−0.001	−0.006**	−0.001	−0.006**	−0.001	−0.001	−0.013***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.003)
Fractionalization	−0.006*	−0.013*	−0.004	−0.015*	0.000	0.001	−0.002
	(0.003)	(0.008)	(0.003)	(0.008)	(0.003)	(0.002)	(0.005)
Total inequality	0.015***	0.005			0.021***		
	(0.005)	(0.003)			(0.004)		
Within group inequality			0.019***	0.003		0.023***	0.009
			(0.005)	(0.004)		(0.003)	(0.006)
Between group inequality			−0.018	0.018*		0.003	0.010
			(0.016)	(0.011)		(0.006)	(0.012)
Constant	0.034	0.095***	0.033	0.098***	0.033	0.030*	0.164***
	(0.027)	(0.032)	(0.027)	(0.032)	(0.023)	(0.016)	(0.038)
Different			4.62	1.63		9.39	0.01
			[0.03]	[0.20]		[0.00]	[0.91]
Observations	350	350	350	350	350	350	350
R^2	0.46	0.72	0.46	0.72	0.32	0.32	0.42
Ind. effects		States		States			Regions
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	LS	LS	LS	LS	Median	Median	Median

All inequalities refer to the generalized entropy measure with parameter 0. Estimator is either least squares (LS) or least absolute deviations (Med). Different is the F-test of the parameters on between and within group inequality being different. R^2 is pseudo-R^2 for median regressions. District of Columbia not included.
Standard errors in parenthesis, p-values in square brackets. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence.

lower than the effect of within group inequality, as we are not be able to keep the overall distribution constant.[23]

The parameters on between and within group inequality are significantly different from each other at the 5% level of confidence. Hence there is support for Proposition 3, and we can also reject the Romer–Roberts–Meltzer–Richard conjecture of all inequality having the same effect. Furthermore, we notice that the coefficient on within group inequality when we control for between group inequality is numerically larger than the coefficient on overall inequality. Hence aggregating between and within inequality tends to hide some of the effect of within group inequality on redistribution. Again, introducing state fixed effects reduces these effects, so they seem to be mostly driven by between state variation.

One may worry that the results are driven by a few outliers. To check this, I rerun some of the results using median regression instead of least squares, reported in columns (5) to (7). The changes in the estimates are not large, and the overall conclusions are the same. As a fixed effects

[23] Another explanation for this fining could be that the level of group antagonism β differ among people so some are more group biased than others. Then the low-β agents would push the outcome in the direction of a standard Romer–Roberts–Meltzer–Richard model whereas the high-β agents would push the outcome in the direction of Proposition 3. The sum could then be that both within and between group inequality leads to more redistribution, but the former having a stronger effect than the latter.

Table 4
Inequality and redistribution measured by average fraction of transfers in disposable income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction above 65	0.498***	0.407***	0.485***	0.404***	0.479***	0.470***	0.361***
	(0.037)	(0.045)	(0.038)	(0.045)	(0.054)	(0.051)	(0.033)
Log per capita income	−0.088***	−0.121***	−0.091***	−0.121***	−0.080***	−0.083***	−0.106***
	(0.006)	(0.011)	(0.006)	(0.011)	(0.008)	(0.008)	(0.006)
Fractionalization	−0.010	−0.022	−0.003	−0.022	−0.014	−0.011	−0.001
	(0.007)	(0.028)	(0.008)	(0.028)	(0.010)	(0.011)	(0.009)
Total inequality	0.127***	0.103***			0.128***		
	(0.011)	(0.011)			(0.016)		
Within group inequality			0.130***	0.106***		0.129***	0.115***
			(0.011)	(0.012)		(0.015)	(0.010)
Between group inequality			−0.046	−0.004		0.032	−0.105
			(0.106)	(0.104)		(0.142)	(0.080)
Constant	0.936***	1.308***	0.958***	1.302***	0.853***	0.883***	1.150***
	(0.059)	(0.112)	(0.061)	(0.112)	(0.085)	(0.081)	(0.059)
Different			2.68	1.07		0.45	7.30
			[0.10]	[0.30]		[0.50]	[0.01]
Observations	300	300	300	300	300	300	300
R-squared	0.80	0.74	0.81	0.74	0.58	0.58	0.64
Ind. effects		States		States			Regions
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	LS	LS	LS	LS	Median	Median	Median

All inequalities refer to the generalized entropy measure with parameter 0. Estimator is either least squares (LS) or least absolute deviations (Med). Different is the F-test of the parameters on between and within group inequality being different. R^2 is overall R^2 for fixed effects models and pseudo-R^2 for median regressions. Omitted categories are 2000 for year-dummies and East North Central for regional dummies. District of Columbia not included.
Standard errors in parenthesis. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence. p-values in square brackets.

estimator for median regression has not yet been developed, I introduce eight Region dummies to partially pick up state fixed effects. Now, between group inequality gets a positive effect on transfers, but still smaller than within inequality.

The fraction of state expenditure on welfare is a quite narrow measure of redistribution. Hence I have rerun the analysis using a broader measure of redistribution, the average share of transfers to disposable income. These estimates are reported in Table 4. These effects are stronger than the ones for the more restricted measure of redistribution used in Table 3. Quite generally, the results are more significant and numerically larger. Furthermore, they now quite generally remain significant when we control for state fixed effects. Hence the support for both parts of Proposition 3 is stronger when using this broader measure of redistribution.

To see whether my particular choice of inequality measure may be driving the results, I have rerun the basic regressions in columns (1) and (3) using different values for the parameter κ as well as by using the Gini coefficient. The results are not reported, but available upon request. They are essentially the same, but the effects are somewhat less strong for $\kappa \neq 0$. The fit of the model as measured by R^2 is highest at $\kappa=0$, so this seems to be the most suitable measure of inequality to explain redistribution.

A final worry may be that inequality and transfers are jointly determined so inequality is an endogenous regressor. The obvious solution is instrumental variables estimation, but it is

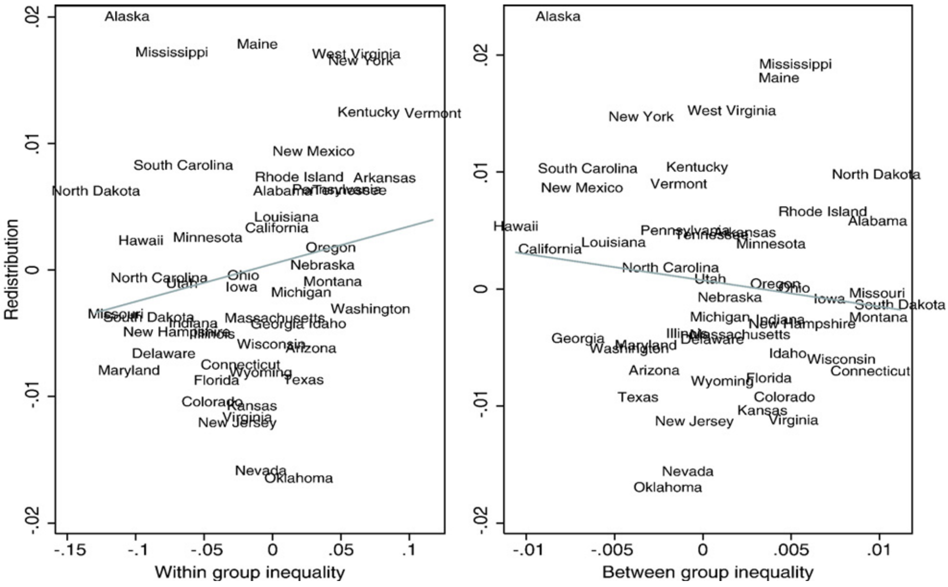

Fig. 3. Scatter plot of redistribution measured by fraction expenditure on welfare in per capita personal income against within and between group inequality, partialing out the other variables in Table 3. Only year 2000 shown. All inequalities refer to the generalized entropy measure with parameter 0.

notoriously hard to find good instruments for inequality as almost everything that affects inequality also will affect transfers. To get some idea of the impact of potential endogeneity, I used lagged values of the inequality measures. This is not a perfect instrument, but it may still throw some light on the magnitude of a potential endogeneity problem. The results are not reported, but correcting for endogeneity in this way does not have large effects on the conclusions.

To conclude, we have quite good support for the model. In the first set of regressions, using the fraction of state expenditure on welfare, the effects are not fully robust. This may, however, to some extent be due to this measure of redistribution being too narrow to capture the total picture of state redistributive efforts as the second set of regressions, using the share of transfers in household disposable income, give strong support for the predictions of the model.

7. Conclusion

Fractionalization in general, and racial divide in particular, has a major impact on politics. I have provided a theoretical basis showing that it tends to reduce the amount of redistribution in democratic polities. Furthermore, when a society is fractionalized, inequality between and within groups have opposite effects on the support for redistribution. The former will reduce the support and the latter increase it. These predictions also have reasonably good empirical support.

This may also be an explanation for the fact that many very unequal societies have small governments. The reason is twofold. In the first place, fractionalized countries tend to have a more uneven distribution of income than less fractionalized countries. As fractionalization reduces the support for redistribution, this implies a negative correlation between inequality and the size of government. Furthermore, inter-group inequality tends to reduce the support for redistribution in fractionalized societies. Hence if both inter- and intra-group inequality is increasing, this might lead to less support for public redistribution. Although most of the analysis was performed within a relatively simple model of policy determination, it seems plausible that most of the main conclusions also hold in richer models. It also supports the view that fragmentation along racial lines is a barrier to policies that benefits the poor in racially divided countries like the US, a view emphasized by e.g. Wilson (1978, 1999).

Observe that if the groups are geographically segmented, it is quite probable that redistribution takes place locally so most of the tax levied from one agent is transferred to her fellow group members. This may to some extent limit the consequences of high fractionalization but excludes possibly beneficial redistribution between groups. One could imagine an extension of the model in this direction, which is closely related to the literature on the optimal size of nations (Alesina and Spolare, 1997; Goyal and Staal, 2004). Another interesting extension would be to study the effect of polarization between groups in the spirit of Esteban and Ray (1994).

The theory also has implications for the development of a welfare state in democratizing states. In countries with heavy fractionalization and intense groups' conflicts, it will usually be difficult to obtain democratic support for a large welfare state. Then one has the choice between two paths: On the one hand, one could opt for a small government and little redistribution through central budgets. On the other hand, it may be possible to go through a nation building process where the tension between the groups is reduced and a European style welfare state becomes politically feasible. However, in the long run the degrees of social conscience and group antagonism may also change. A conjecture is that high inequality will tend to reduce social conscience and between group inequality increase group antagonism due to segregation and polarization.

Appendix A. Detailed data on the densities at the median

The table underneath gives details of the density of the income distribution for an income equal to the overall median income for the Blacks and Whites since 1967. Median incomes are given in 2001 dollars. The data are taken from US Census Bureau (2001: Table A1). The cumulative density function of the income distribution is then approximated by a cubic spline and densities are found by numerical differentiation. Micro data for 2000 from the Luxembourg Income Study give virtually identical results.

Appendix B. Proof of Proposition 3

Application of the implicit function theorem on the system ((17a) (17b) (17c)) yields

$$\frac{\mathrm{d}t}{\mathrm{d}\gamma} = \Xi\left\{\frac{(1-\alpha)\Gamma}{s_A q\tilde{f}_{\gamma A}(x_A^m) + s_B(1-q)\tilde{f}_{\gamma B}(x_B^m)} + \alpha q(1-q)\left[s_A\tilde{f}_{\gamma A}(x_A^m) - s_B\tilde{f}_{\gamma B}(x_B^m)\right]S_t(F_A - F_B)\right\}. \tag{20}$$

where

$$\Xi = -\frac{s_A q\tilde{f}_{\gamma A}(x_A^m) + s_B(1-q)\tilde{f}_{\gamma B}(x_B^m)}{s_A q\tilde{f}_{\gamma A}(x_A^m)S_{tt}^A + s_B(1-q)\tilde{f}_{\gamma B}(x_B^m)S_{tt}^B} > 0, \tag{21}$$

$$\Gamma = q(F_A - F)(x_A^m) + (1-q)(F_B - F)(x_B^m), \tag{22}$$

and s_i is given by Eq. (10). When group A is richer than B in the sense of first order stochastic dominance we have $S_t\,[t, F_A - F_B] < 0$. We assume that $s_A f_A\,(x_A^m) > s_B f_B\,(x_B^m)$. Hence $\tilde{f}_A(x_A^m) > \tilde{f}_B(x_B^m)$,

Table A1
Density at the median of the US income distribution for Blacks and Whites

Year	Median income	Density ×100	
		Blacks	Whites
2001	42,228	1.03	1.05
2000	43,162	1.03	1.04
1999	43,355	0.96	1.04
1998	42,173	0.98	1.07
1997	40,699	1.05	1.11
1996	39,869	1.04	1.12
1995	39,306	1.06	1.16
1994	38,119	1.03	1.18
1993	37,688	1.06	1.20
1992	37,880	1.06	1.20
1991	38,183	1.08	1.22
1990	39,324	1.06	1.24
1989	39,850	1.04	1.18
1988	39,144	0.98	1.19
1987	38,835	1.04	1.21
1986	38,365	1.04	1.23
1985	37,059	1.12	1.28
1984	36,343	1.12	1.31
1983	35,438	1.13	1.35
1982	35,423	1.16	1.36
1981	35,478	1.11	1.34
1980	36,035	1.15	1.35
1979	37,192	1.12	1.31
1978	37,234	1.17	1.32
1977	34,989	1.23	1.39
1976	34,792	1.26	1.42
1975	34,219	1.32	1.45
1974	35,159	1.29	1.46
1973	36,278	1.24	1.41
1972	35,560	1.25	1.46
1971	34,126	1.38	1.53
1970	34,481	1.34	1.54
1969	34,714	1.39	1.57
1968	33,436	1.41	1.65
1967	32,081	1.43	1.68

so the square brackets in the second term in Eq. (20) is positive, and the second term in Eq. (20) is negative.

Now we need to show that $\Gamma \leq 0$, to establish that $\mathrm{d}t/\mathrm{d}\gamma < 0$. Differentiation of Eq. (17c) and rearranging yields

$$\gamma \frac{\partial \Gamma}{\partial \gamma} = -\Gamma - \left[q f_A(x_A^m) \frac{\partial x_A^m}{\partial \gamma} + (1-q) f_B(x_B^m) \frac{\partial x_B^m}{\partial \gamma} \right].$$

The term $-\Gamma$ is equilibrating and tends to keep Γ close to zero. Inserting from Eqs. (17a) and (17b), the term in square brackets may be rewritten

$$\frac{-1}{1-\alpha} \left[q s_A f_A S_{tt}^A + (1-q) s_B f_B S_{tt}^B \right] \frac{\mathrm{d}t}{\mathrm{d}\gamma} - \frac{\alpha}{1-\alpha} q(1-q)(s_A f_A - s_B f_B) S_t(t, F_A - F_B).$$

From Eq. (20) it follows that if $\Gamma<0$, then $\mathrm{d}t/\mathrm{d}\gamma<0$, so the first term in this equation is negative if $\Gamma<0$. Since $s_A f_A > s_B f_B$ by assumption, the second term is also negative. Consequently, if at some level γ we have $\Gamma<0$, $\mathrm{d}\Gamma/\mathrm{d}\gamma$ is the sum of an equilibrating term and a negative term. Hence Γ will remain below zero.

At $\gamma=0$, we have $\Gamma=0$ and

$$\left.\frac{\partial \Gamma}{\partial \gamma}\right|_{\gamma=0} = q(1-q)(f_A-f_B)(x^m)\left(\left.\frac{\partial x_A^m}{\partial \gamma}\right|_{\gamma=0} - \left.\frac{\partial x_B^m}{\partial \gamma}\right|_{\gamma=0}\right),$$

which is negative when $(f_A-f_B)(x^m)>0$, Assumption 1 applied at $\gamma=0$. Hence for small values of γ, $\Gamma\leq 0$. As a rise in γ will keep Γ below zero is $\Gamma<0$ we have established that for all γ, $\Gamma<0$. Then it follows from Assumption 1 that Eq. (20) is negative.

References

Alesina, A., La Ferrara, E., 2000. Participation in heterogeneous communities. Quarterly Journal of Economics 115, 847–904.

Alesina, A., La Ferrara, E., 2005. Ethnic diversity and economic performance. Journal of Economic Literature 43, 721–761.

Alesina, A., Spolaore, E., 1997. On the number and size of nations. Quarterly Journal of Economics 112, 1027–1056.

Alesina, A., Baqir, R., Easterly, W., 1999. Public goods and ethnic division. Quarterly Journal of Economics 114, 1243–1284.

Alesina, A., Glaeser, E., Sacerdote, B., 2001. Why doesn't the US have a European-style welfare system. Brookings Papers on Economic Activity, pp. 187–278.

Austen-Smith, D., Wallerstein, M., 2006. Redistribution and affirmative action. Journal of Public Economics 90, 1789–1823.

Barro, R., 1974. Are government bonds net wealth? Journal of Political Economy 82, 1095–1117.

Bénabou, R., 1996. Inequality and growth. In: Bernanke, B.S., Rotemberg, J.J. (Eds.), Macroeconomics Annual. MIT Press, Cambridge, MA.

Bénabou, R., 2000. Unequal societies: income distribution and the social contract. American Economic Review 90, 96–129.

Benhabib, J., Rusticini, A., 1996. Social conflict and growth. Journal of Economic Growth 1, 125–142.

Bjorvatn, K., Cappelen, A.W., 2002. Inequality, segregation, and redistribution. Journal of Public Economics 87, 1657–1679.

Bobo, L., Kluegel, J.R., 1993. Opposition to race-targeting: Self-interest, stratification ideology, or racial attitudes? American Sociological Review 58, 443–464.

Bourguignon, F., 1979. Decomposable income inequality measures. Econometrica 47, 901–920.

Bowles, S., Gintis, H., 2000. Reciprocity, self-interest, and the welfare state. Nordic Journal of Political Economy 26, 33–53.

Bowles, S., Fong, C., Gintis, H., 2006. Strong reciprocity and the welfare state. In: Kolm, S.-C., Mercier Ythier, J. (Eds.), Handbook on the Economics of Giving, Altruism, and Reciprocity, vol. 2. North-Holland, Amsterdam.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117, 817–869.

Collier, P., 2000. Ethnicity, politics and economic performance. Economics and Politics 12, 225–245.

Collier, P., 2001. Implications of ethnic diversity. Economic Policy 16, 129–166.

Cowell, F.A., 2000. Measurement of inequality. In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. North-Holland, Amsterdam.

Downs, A., 1957. An Economic Theory of Democracy. Harper & Row, New York.

Esteban, J.-M., Ray, D., 1994. On the Measurement of Polarization. Econometrica 62, 819–851.

Galasso, V., 2003. Redistribution and fairness: a note. European Journal of Political Economy 19, 885–892.

Gilens, M., 1999. Why Americans Hate Welfare. Race, Media, and the Politics of Antipoverty Policy. The University of Chicago Press, Chicago.

Goyal, S., Staal, K., 2004. The political economy of regionalism. European Economic Review 48, 563–593.

Keefer, P., Knack, S., 2002. Polarization, politics and property rights. Links between inequality and growth. Public Choice 111, 127–154.

Khwaja, A.I., Can good projects succeed in bad communities? Collective Action in Public Good Provision. Mimeo, Harvard University; 2002.

Kinder, D.R., Sanders, L.M., 1996. Divided by Color. Racial Politics and Democratic Ideals. University of Chicago Press, Chicago.

Knack, S., 1992. Civic norms, social sanctions, and voter turnout. Rationality and Society 4, 133–156.

Knack, S., Keefer, P., 1997. Does social capital have an economic payoff? A cross-country investigation. Quarterly Journal of Economics 112, 1252–1288.

Luttmer, E.F.P., 2001. Group loyalty and the taste for redistribution. Journal of Political Economy 109, 500–528.

Meltzer, A.H., Richard, S., 1981. A rational theory of the size of government. Journal of Political Economy 89, 914–927.

Miguel, E., 2004. Tribe or Nation? Nation-building and public goods in Kenya versus Tanzania. World Politics 56, 327–362.

Miguel, E., Gugerty, M.K., 2005. Ethnic diversity, social sanctions, and public goods in Kenya. Journal of Public Economics 89, 2325–2368.

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.

Moene, K.O., Wallerstein, M., 2001. Inequality, social insurance, and redistribution. American Political Science Review 95, 859–874.

Moffitt, R., Ribar, D., Wilhelm, M., 1998. The decline of welfare benefits in the U.S.: the role of wage inequality. Journal of Public Economics 68, 421–452.

Mueller, D.C., 1987. The voting paradox. In: Rowley, C.K. (Ed.), Democracy and Public Choice. Basil Blackwell, New York.

Persson, T., Tabellini, G., 1994. Does centralization increase the size of government? European Economic Review 38, 765–773.

Persson, T., Tabellini, G., 2000. Political Economics. Explaining Economic Policy. MIT Press, Cambridge, MA.

Roberts, K.W.S., 1977. Voting over income tax schedules. Journal of Public Economics 8, 329–340.

Rodrik, D., 1999. Where did all the growth go? External shocks, social conflict, and growth collapses. Journal of Economic Growth 4, 385–412.

Roemer, J.E., 1998. Why the poor do not expropriate the rich: an old argument in new garb. Journal of Public Economics 70, 399–424.

Roemer, J.E., 1999. The democratic political economy of progressive income taxation. Econometrica 67, 1–19.

Romer, T., 1975. Individual welfare, majority voting and the properties of a linear income tax. Journal of Public Economics 4, 163–185.

Rotschild, M., Stiglitz, J.E., 1970. Increasing risk I: A definition. Journal of Economic Theory 2, 225–243.

Saint-Paul, G., 2001. The dynamics of exclusion and fiscal conservatism. Review of Economic Dynamics 4, 275–302.

Sen, A.K., 1965. Labour allocation in a cooperative enterprise. Review of Economic Studies 33, 361–371.

Sethi, R., Somanathan, R., 2004. Inequality and segregation. Journal of Political Economy 112, 1296–1321.

Shorrocks, A.F., 1981. The class of additively decomposable inequality measures. Econometrica 48, 613–626.

US Census Bureau, 2001. Money income in the United States: 2000. U.S. Government Printing Office, Washington, D.C. Available online at http://www.census.gov/prod/2001pubs/p60–213.pdf.

US Department of Commerce. Government Finances. US Government Printing Office: Washington, DC; Various years.

Vigdor, J.L., 2004. Community heterogeneity and collective action: analyzing initial mail response to the 2000 census. Review of Economics and Statistics 86, 303–312.

Wilson, W.J., 1978. The Declining Significance of Race, Blacks and Changing American Institutions, second ed. University of Chicago Press, Chicago.

Wilson, W.J., 1999. The Bridge over the Racial Divide. University of California Press, Berkeley.