Luxembourg Income Study Working Paper Series

Working Paper No. 383

Fractionalization and the Size of Government

Jo Thori Lind

June 2004

Luxembourg Income Study (LIS), asbl

Fractionalization and the size of government*

Jo Thori Lind†

June 18, 2004

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. All these results hold even if a poor group forms a majority. Using a panel data set constructed from US micro data, I find support for the hypothesis that within race inequality increases redistribution while between race inequality decreases redistribution.

*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, and Klaas Staal as well as participants at the 2003 Econometric Society European Winter Meeting, 2003 EEA conference, 3rd Norwegian-German Seminar on Public Economics, 2002 meetings of the European Public Choice Society, the Norwegian Economic Society and seminars at the University of Oslo for help and valuable comments.

†Department of Economics, University of Oslo, PB 1095 Blindern, 0317 OSLO. E-mail: j.t.lind@econ.uio.no

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 will 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. However, Milanovic (2000) claims this is mainly due to data problems, and using an improved data set, he finds support for the theory.[1]

A separate literature has recently 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

[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, Glaeser, and Sacerdote (2001).

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 these influence support 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 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

[2]Based on such factors as differentiated tastes (Alesina, Baqir, and Easterly 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 2003). 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.

[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, Baqir, and Easterly 1999, Alesina, Glaeser, and Sacerdote 2001, Alesina and La Ferrara 2000). This is corroborated by similar findings in Pakistan (Khwaja 2002) and Kenya (Miguel and Gugerty 2003). Furthermore, comparing Kenya, where ethic conflicts are important, to Tanzania, where there is less ethnic conflict, Miguel (2003) finds that ethnic fractionalization is important in Kenya but insignificant in Tanzania.

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 the validity of the key insights of the model, I use a panel of US states observed in six 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. I find empirical support for the model. 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 (2003), 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. Vigdor (2001) 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 middleclass 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 $\int_\Omega u(x)\, dF(x)$.

There is a government that redistributes resources before production 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. I could of course model this 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. For a given mean income (tax base) $\bar{x}$, social welfare is given by

$$S(t, F) = \int_\Omega u[(1-t)x + T(t)\bar{x}]\, dF(x). \quad (1)$$

[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.

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

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

where α is a coefficient of social conscience. Throughout the paper, I assume $\alpha \in [0, 1]$.[5] 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

[5] 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.

weight the agent puts on persons from different income groups. 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[6]. 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

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

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[7] within the groups are described by F_A and F_B which are both assumed to have support Ω. Hence $F = qF_A + (1-q) F_B$. I will say that one group is richer than the other if the two groups' income distributions

[6]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.

[7]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.

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

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

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 $\beta = 0$.[8] 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.[9] 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.[10] They have a common preferred tax rate, but in general their endowments differ. Call the endowment of the A median voter x_A^m and that of the B median voter x_B^m. Then the tax rate t chosen by the median voters satisfies the

[8] 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.

[9] 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).

[10] 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.

system

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

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, Fong, and Gintis 2001; 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 gets 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.

In the current model, the only objective of the government is to transfer income between individuals. However, 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). Notice that even if agents have the same degree of social conscience for both their own group and other groups, a segmented labour market, in the sense that new incomes are drawn from different distributions for different groups, is sufficient to make it appear as if the agent had a group biased social conscience. Hence we could reinterpret the whole model as an analysis of the consequences of a segregated labour market.

3 The size of government

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 $i \in \{A, B\}$ and $x \in \Omega$. Implicit now is that the two distributions have common support. Some deviations from these assumptions are discussed in an appendix available upon request. Differentiation of the system (8) yields

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

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

$$q f_A \left(x_A^m\right) dx_A^m + (1-q) f_B \left(x_B^m\right) dx_B^m = 0 \tag{9c}$$

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

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

$$\text{and} \tag{10}$$

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

Then the implicit function theorem yields

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

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 median voter from the two groups both are the median income in society x^m and

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

This expression is negative if $f_A(x^m) > f_B(x^m)$. When β rises marginally from $\beta = 0$, the A-median voter cares less about group B, and consequently prefers a lower tax rate whereas the B-median voter now cares less about group A and therefore prefers a higher tax rate. Consequently, as β increases, the median voters will be an A-agent with endowment $x_A^m < x^m$ and a B-agent with endowment $x_B^m > x^m$. Notice that this change in preferences is very similar to the one discussed by Persson and Tabellini (1994: 168f). 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-median voter will be high relative to x^m, the income of the former median 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, i.e. the median voters. If tax preferences change a lot within a group, this decreases that group's power in the political struggle as their median voter is quickly swapped with a new median 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 $s_A f_A(x_A^m) > s_B f_B(x_B^m)$.

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 median voters. At $\beta = 0$, both median voters have the same endowment x^m. However, since the As are richer than the Bs, the median 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 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 iff $q > 1/2$.

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

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

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 (13) is negative. This is the case if agents are not too risk averse.[11] 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.

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.*

We see from equation (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. There are also indirect effects of changing q on S^i_{tt} and through changes in x^m_A and x^m_B, but these are of uncertain sign, and are unlikely to dominate. This easily extends to the case of multiple groups. Hence we have the following result:

Proposition 2 *When Assumption 1 holds, then conditional on the group income distributions and the degree of group antagonism, increased fractionalization increases the effect of group antagonism on taxes.*

However, as the overall income distribution F and hence the mean income $\bar{x}$ will depend on q, it is difficult to properly study the effect of changing q without changing any other parameters. Nevertheless, it seems likely that when Assumption 1 holds, a society with positive group antagonism will ceteris paribus have lower taxes the higher the degree of fractionalization is. This extend Alesina, Baqir, and Easterly's (1999) and Miguel and Gugerty's (2003) results on public good provision to redistribution.

I showed above that the effect of fractionalization and group antagonism, given by equations (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 fractionalization when $\beta = 0$ (no group antagonism) will depend

[11] 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 $\tau = [T(t) + T'(t)]\bar{x}$.

on the difference at the median income of the population as a whole. An example is depicted in Figure 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, reenforcing the effect.

Figure 1 about here

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 B. 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 predicts 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 median voter is pushed to the right, increasing the density and vice versa for the White median voter.

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 median voter from group A is $S_t\left(t, (1-\alpha) D_{x_A^m} + \alpha\beta F_A + \alpha(1-\beta) F\right) = 0$, which we may rewrite

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

where

$$\Psi(t) = (1-\alpha) S_t\left(t, D_{x_A^m} - F\right) + \alpha\beta(1-q) S_t(t, F_A - F_B), \tag{14}$$

and of course a similar expression holds for a median voter from group B. Ψ is the deviation from social welfare in the agent's maximand.[12] The first term of Ψ is the effect of the median voter caring more about herself than other individuals in society and the last term stems from the median 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 $\alpha = 1$,

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

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.[13] If $S_t(t, D_{x^m} - F) > 0$, i.e. the original median 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 median voter prefers a tax rate that is too low, then antagonism is detrimental to efficiency.

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 belong. 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. This is illustrated in Figure 2. If group antagonism is too strong, however, it may push the tax rate below the social optimum.

Figure 2 about here

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 median 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 median voter before the shift is determined as the solution to $S_t(t, F_i) = 0$.

[13] This is done by differentiation (14) with regard to β and inserting from (9a) and (12).

$\frac{\partial}{\partial t} u \left[(1-t)\, x + T\,(t)\, \bar{x}\right]$, the integrand of S_t, is decreasing and convex in x when $u'' < 0$ and $u''' > 0$.[14] Hence when F_i^0 second order stochastically dominates F_i^1, $S_t\left(t, F_i^0\right) < S_t\left(t, F_i^1\right)$, a converse to the well-known result on second order stochastic dominance in the theory of choice under uncertainty (Rothchild and Stiglitz 1970). As S_t is decreasing in t, this implies that the median 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 iff 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 median voter in group i also prefers a higher tax rate in cases with less then full group antagonism. These results are very similar to those found in the ordinary Romer-Roberts-Meltzer-Richard model.

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 $\tilde{F}_{\gamma i} := \gamma F_i + (1-\gamma)\, F$. 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:

$$S_t\left(t, (1-\alpha)\, D_{x_A^m} + \gamma\alpha F_A + (1-\gamma)\, \alpha F\right) = 0 \tag{15a}$$

$$S_t\left(t, (1-\alpha)\, D_{x_B^m} + \gamma\alpha F_B + (1-\gamma)\, \alpha F\right) = 0 \tag{15b}$$

$$q\tilde{F}_{\gamma A}\left(x_A^m\right) + (1-q)\, \tilde{F}_{\gamma B}\left(x_B^m\right) = \frac{1}{2}. \tag{15c}$$

As γ enters (15c), the analysis of this system is slightly more involved than of (8). However, the results are almost identical. In Appendix A I prove the following result:

[14] 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.

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 median 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 Testing the model

6.1 The data

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 support for redistribution. However, at the time being I don't know any method to measure β, so I will limit the test to the following somewhat simpler predictions:

1. For a given level of group antagonism, a higher degree of fractionalization leads to less redistribution (Proposition 2)

2. Within group inequality should increase the support for redistribution (Proposition 3)

3. Between group inequality should reduce the support for redistribution (Proposition 3)

To perform the tests, I employ a panel of US states with six observations per state.[15] The main reason for using a single country is that the definition of groups and the collection of data on groups are more homogeneous. We need measures of inequality both between and within groups. As such data are not readily available, I had to construct the measures from micro data. Income data are taken from March Current Population Survey, made available through the Luxembourg Income Study (LIS).[16] 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 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

$$I(\mathbf{y}) = \sum_{g=1}^{G} w_g I(\mathbf{y}_g) + I(\bar{\mathbf{y}}_1, \ldots, \bar{\mathbf{y}}_G) \tag{16}$$

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

$$I_{GE}^{\kappa} = \frac{1}{\kappa(\kappa-1)} \int \left[\left(\frac{x}{\mu} \right)^{\kappa} - 1 \right] dF(x), \tag{17}$$

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 $I_{GE}^{0} = -\int \ln\left(\frac{x}{\mu}\right) dF(x)$, the mean logarithmic deviation.

I use two measures of redistribution. The first is the average share of transfers received by households as a share of disposable income, calculated from the LIS data. The second

[15] 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.

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

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.[17] This first 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. The second measure is probably rather narrow. Ideally, we should have a measure less broad than the first, but more broad than the first. However, such data do not exist. Consequently, I use both, and it seems plausible that the truth should be 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. The groups are African American, white, and other in 1969, African American, Spanish, white, and other in 1974 and 1986, and African American, American Indian/Aleut/Eskimo, Asian/Pacific Islander, Hispanic white, non-hispanic white, and other thereafter. 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 1990 census, I get an overall correlation of .87, ranging from .67 in 1969 to .98 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 1 about here

Figure 3 about here

Table 1 gives basic descriptive statistics of the data and Figure 3 shows the geographical distribution of fractionalization, within- and between group inequality, and average transfers as a share of disposable income. For the figure, all numbers are measured in 2000. We notice that the degree of fractionalization follow quite similar patterns with high values in the South and South-West. Within group inequality is uncorrelated with between group inequality (the correlation coefficient is -.05) and does not seem to follow

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

any strong geographical patterns. Finally, transfers are generally high in the Midwest and the North East.

6.2 Empirical results

Table 2 shows the main empirical results. In column (1) to (7) the dependent variable is 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.[18] A one standard deviation increase in inequality increases the fraction of transfers by .016 or about half a standard deviation, which should be judged a quite large effect. This result is also strongly significant in the fixed effects panel data model reported in column (2). Fractionalization seems to have a negative effect on transfers. A one standard deviation increase in fractionalization reduces transfers by about 0.005 or about 0.15 of a standard deviation. Hence the magnitude of this effect is far smaller. This effect does not seem to be robust to the introduction of state fixed effects. As fractionalization changes little over time, this is not surprising. Furthermore, the positive coefficient on fractionalization in column (2) is mainly due to a few outliers, most importantly Idaho 1974. Non-reported robust regressions also confirm this.

Table 2 about here

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. We see that the estimates conform to the expectations from my theoretical model, although the coefficient on between group inequality is not significantly different from zero. However, the two parameters are significantly different from each other at the 5% level of confidence, so we can reject the Romer-Roberts-Meltzer-Richard conjecture of all inequality having the same effect. We also notice that the coefficient on within group inequality when we control for between group inequality is

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

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. Introducing state fixed effects give almost identical results.

One may worry that the results are driven by a few outliers. To check this, I rerun some of the results using median regressions 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 estimator for median regression has not yet been developed, I introduce eight Division dummies to partially pick up state fixed effects. Now, between group inequality gets a positive effect on transfers, but still smaller than within inequality. However, the difference is no longer significantly different.

The measure of transfers also contains federal transfers, so it may be argued that it is too broad. Hence I repeated the estimations using the fraction of state welfare expenditure in state personal income as dependent variable. This measure is arguably to limited, but inequality and fractionalization should still have the predicted effects upon it. However, the results are somewhat less appealing. Fractionalization still has a negative effect on transfers, but the effect is hardly significant in any of the specifications. However, total inequality has a positive and strongly significant effect. A one standard deviation in inequality increases welfare expenditure per capita by .0024 or about a quarter of a standard deviation. Although this effect is smaller than for the first measure of transfers, the effect is still important. When we compare between and within inequality, there appear to be little difference between the two. In the state fixed effects specification, between group inequality even has a stronger positive effect than within inequality. However, is seems that this may be driven by outliers. A quite large fraction of the observations have large DFITS. It seems that the District of Columbia is the most important outlier with DFITS above .5 in all years. If we remove it from the sample, we find that between group inequality has a significantly lower effect than within group in equation (10) and the difference is insignificant in equation (11). When we use median regression instead of least squares, the effect of between group inequality is essentially zero, and significantly lower than the coefficient on within group inequality at the 10% level.

To see whether my particular choice of inequality measure may be driving the results, I rerun the basic regressions in columns (1) and (3) using different values for the parameter

κ. The results are reported in Table 3. It is seen that the results are essentially the same: Inequality has a significantly positive effects on transfers, and when we decompose into between and within group inequality, within has a somewhat stronger effect whereas the effect of between is about zero. The effect is less strong for $\kappa \neq 0$. However, we also see that the fit of the model as measured by R^2 is highest at $\kappa = 0$, so it may seem that this is the most suitable measure of inequality to explain redistribution.

The table also reports results where I use the Gini coefficient rather than the generalized entropy measure. As the Gini coefficient is not decomposable, column (10) reports results from a regression with total inequality and between group inequality rather than between and within as before. The results are similar to the ones found above, so it does not seem that the results are an artefact of the particular choice of inequality measure.

Table 3 about here

The estimates are also almost the same if I run separate regressions for each year or use other robust estimators than median regression. 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. However, it is 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 use lagged values of the inequality measures. As the time between observations differ, I interact the lagged inequality measures with time dummies to allow for the effect declining over time. Table 4 report results from these estimations. In the models without individual effects, all measures of inequality are seen to have negative effects on redistribution. However, these coefficients are not significantly different from zero. When we introduced state fixed effects, results comparable to those in Table 2 reappear. Again, within inequality has a positive and significant effect on redistribution whereas between group inequality has a negative effect. The latter is not significantly different from zero, and is now not significantly different from the within effect either. However, we may conclude that controlling for endogeneity does not appear to have large effects on the conclusions.

Table 4 about here

To conclude, the first set of regressions using the share of transfers in household disposable income give strong support for the predictions of the model. When we turn to the fraction of state welfare expenditure, the conclusions are weaker. However, this may to some extent be due to the measure being too limited to capture the total picture of state redistributive efforts.

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 tend 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 Spolaore 1997, Goyal and Staal 2003).

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.

A Proof of Proposition 3

Application of the implicit function theorem on the system (15) yields

$$\frac{dt}{d\gamma} = \Xi \Bigg\{ \frac{(1-\alpha)\,\Gamma}{s_A q \tilde{f}_{\gamma A}(x_A^m) + s_B (1-q) \tilde{f}_{\gamma B}(x_B^m)} + \alpha q (1-q) \left[s_A \tilde{f}_{\gamma A}(x_A^m) - s_B \tilde{f}_{\gamma B}(x_B^m) \right] S_t (F_A - F_B) \Bigg\}. \tag{18}$$

where

$$\Xi = -\frac{s_A q \tilde{f}_{\gamma A}(x_A^m) + s_B (1-q) \tilde{f}_{\gamma B}(x_B^m)}{s_A q \tilde{f}_{\gamma A}(x_A^m) S_{tt}^A + s_B (1-q) \tilde{f}_{\gamma B}(x_B^m) S_{tt}^B} > 0, \tag{19}$$

$$\Gamma = q (F_A - F)(x_A^m) + (1-q)(F_B - F)(x_B^m), \tag{20}$$

and s_i is given by (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 $\tilde{f}_A (x_A^m) > \tilde{f}_B (x_B^m)$, so the square brackets in the second term in (18) is positive, and the second term in (18) is negative.

Now we need to show that $\Gamma \leq 0$, to establish that $dt/d\gamma < 0$. Furthermore, differentiation of (15c) 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 (15a) and (15b), the term in square brackets may be rewritten

$$\frac{-1}{1-\alpha}\left[qs_Af_AS_{tt}^A+(1-q)\,s_Bf_BS_{tt}^B\right]\frac{dt}{d\gamma}$$
$$-\frac{\alpha}{1-\alpha}q\,(1-q)\,(s_Af_A-s_Bf_B)\,S_t\,(t,F_A-F_B)\,.$$

From (18) it follows that if $\Gamma<0$, then $dt/d\gamma<0$, so the first term in this equation is negative if $\Gamma<0$. Since $s_Af_A>s_Bf_B$ by assumption, the second term is also negative. Consequently, if at some level γ we have $\Gamma<0$, $d\Gamma/d\gamma$ 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 /18) is negative.

B 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 A-1). 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.

Table A1 here.

References

Alesina, A., R. Baqir, and W. Easterly (1999): "Public goods and ethnic division." *Quarterly Journal of Economics* 114: 1243-84.

Alesina, A., E. Glaeser, and B. Sacerdote (2001): "Why doesn't the US have a European-style welfare system." *Brookings Papers on Economic Activity* 187-278.

Alesina, A., and E. La Ferrara (2000): "Participation in Heterogeneous Communities." *Quarterly Journal of Economics* 115: 847-904.

Alesina, A. and E. Spolaore (1997): "On the number and size of nations." *Quarterly Journal of Economics* 112: 1027-56.

Austen-Smith, D., and M. Wallerstein (2003): "Redistribution in a divided society." Unpublished paper.

Barro, R. (1974): "Are government bonds net wealth?" *Journal of Political Economy* 82: 1095-117.

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

Bénabou, R. (2000): "Unequal societies: Income distribution and the social contract." *American Economic Review* 90: 96-129.

Benhabib, J. and A. Rusticini (1996): "Social conflict and growth." *Journal of Economic Growth* 1:125-42.

Bjorvatn, K. and A. W. Cappelen (2002): "Inequality, segregation, and redistribution." *Journal of Public Economics* 87: 1657-79.

Bobo, L. and J. R. Kluegel (1993): "Opposition to race-targeting: Self-interest, stratification ideology, or racial attitudes?" *American Sociological Review* 58: 443-64.

Bourguignon, F. (1979): "Decomposable income inequality measures." *Econometrica* 47: 901-20.

Bowles, S., C. Fong, and H. Gintis (2001): "Reciprocity, redistribution, and the welfare state." Forthcoming in L.-A. Gerard-Varet, S.-C. Kolm, and J. Mercier Ythier: *Handbook on the Economics of Giving, Reciprocity, and Altruism*, Amsterdam: Elsevier.

Bowles, S. and H. Gintis (2000): "Reciprocity, self-interest, and the welfare state." *Nordic Journal of Political Economy* 26: 33-53.

Charness, G., and M. Rabin (2002): "Understanding social preferences with simple tests." *Quarterly Journal of Economics* 117: 817-69.

Collier, P. (2000): "Ethnicity, politics and economic performance." *Economics and Politics* 12: 225-45.

Collier, P. (2001): "Implications of ethnic diversity." *Economic Policy* 16:129-166.

Cowell, F. A. (2000): "Measurement of inequality." Ch. 2 in A. B. Atkinson and F. Bourguignon (eds.): *Handbook of Income Distribution*, Amsterdam: Elsevier.

Downs, A. (1957): *An Economic Theory of Democracy.* New York: Harper&Row.

Esteban, J.-M. and D. Ray (1994): "On the Measurement of Polarization." *Econometrica* 62: 819-51.

Galasso, V. (2003): "Redistribution and fairness: A note." *European Journal of Political Economy* 19: 885-92.

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

Goyal, S. and K. Staal (2003): "The political economy of regionalism." Forthcoming, *European Economic Review.*

Keefer, P. and S. Knack (2002): "Polarization, politics and property rights. Links between inequality and growth." *Public Choice* 111: 127-54.

Khwaja, A. I. (2002): "Can good projects succeed in bad communities? Collective Action in Public Good Provision." Unpublished paper.

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

Knack, S. (1992): "Civic norms, social sanctions, and voter turnout." *Rationality and Society* 4: 133-156.

Knack, S. and P. Keefer (1997): "Does social capital have an economic payoff? A cross-country investigation." *Quarterly Journal of Economics* 1252-1288.

Luttmer, E. F. P. (2001): "Group loyalty and the taste for redistribution." *Journal of Political Economy* 109: 500-28.

Meltzer, A. and S. Richard (1981): "A rational theory of the size of government." *Journal of Political Economy* 89: 914-27.

Miguel, E. (2003): "Tribe or Nation? Nation-Building and Public Goods in Kenya versus Tanzania." Unpublished paper.

Miguel, E. and M. K. Gugerty (2003): "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." Unpublished paper.

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. and M. Wallerstein (2001): "Inequality, social insurance, and redistribution." *American Political Science Review* 95: 859-874.

Moffitt, R., D. Ribar, and M. Wilhelm (1998): "The decline of welfare benefits in the U.S.: The role of wage inequality ." *Journal of Public Economics* 68: 421-52 .

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

Persson, T. and G. Tabellini (1994): "Does centralization increase the size of government?" *European Economic Review* 38: 765-73.

Persson, T. and G. Tabellini (2000): *Political economics. Explaining economic policy.* Cambridge, MA: MIT Press.

Roberts, K. W. S. (1977): "Voting over income tax schedules." *Journal of Public Economics* 8:329-40.

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: 163-85.

Rotschild, M., and J. E. Stiglitz (1970): "Increasing risk: I. A definition." *Journal of Economic Theory* 2: 225-43.

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: 361-71.

Shorrocks, A. F. (1981): "The class of additively decomposable inequality measures." *Econometrica* 48: 613-26.

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

US Department of Commerce (various years): *Government Finances*. Washington, DC: US Government Printing Office.

Vigdor, J. L. (2001): "Community Heterogeneity and Collective Action: Analyzing Initial Mail Response to the 2000 Census." Forthcoming, *Review of Economics and Statistics*.

Wilson, W. J. (1978): *The Declining Significance of Race. Blacks and Changing American Institutions*. 2nd ed., Chicago: University of Chicago Press.

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

Table 1: Descriptive statistics

Variable		Totoal obs	Periods	States	Mean	Std. dev.	Between std. dev.	Within std. dev
Total inequality	κ=-1	357	7	51	31.72	33.20	8.73	32.05
	κ=0	357	7	51	.570	.134	.047	.125
	κ=1	357	7	51	.353	.076	.029	.070
	κ=2	357	7	51	.290	.103	.039	.096
Between group inequality	κ=-1	357	7	51	.077	.547	.202	.509
	κ=0	357	7	51	.015	.026	.013	.023
	κ=1	357	7	51	.012	.015	.010	.011
	κ=2	357	7	51	.007	.009	.007	.006
Within group inequality	κ=-1	357	7	51	31.64	33.22	8.77	32.06
	κ=0	357	7	51	.555	.133	.049	.123
	κ=1	357	7	51	.341	.073	.027	.068
	κ=2	357	7	51	.282	.101	.037	.094
		357	7	51	.250	.162	.147	.072
Fraction above 65		357	7	51	.094	.027	.018	.020
Log per capita income		357	7	51	9.51	.752	.148	.738
Average share of transfers to disp. income		306	6	51	.146	.033	.023	.023
Fraction expenditure on welfare		357	7	51	.024	.010	.007	.008

Inequality is measured by the generalized entropy measure with coefficient κ. Between standard deviations are standard deviations of the state averages and within the average within state standard deviation.

Table 2: Inequality and redistribution

Dependent variable	Average fraction of transfers in disposable income							Fraction expenditure on welfare in per capita personal income						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Fraction above 65	0.675***	0.719***	0.656***	0.698***	0.619***	0.574***	0.635***	-0.032*	-0.022	-0.033	-0.012	-0.026	-0.040**	-0.029
	(0.045)	(0.060)	(0.046)	(0.059)	(0.045)	(0.046)	(0.032)	(0.019)	(0.015)	(0.020)	(0.015)	(0.021)	(0.017)	(0.022)
Log per capita income	-0.064***	-0.063***	-0.065***	-0.060***	-0.057***	-0.058***	-0.075***	0.005**	-0.019	0.005*	-0.020***	-0.002	-0.002	-0.019***
	(0.006)	(0.023)	(0.006)	(0.023)	(0.006)	(0.006)	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)	(0.003)	(0.002)	(0.004)
Fractionalization	-0.028***	0.017***	-0.020***	0.028	-0.032***	-0.027***	-0.034***	-0.004	-0.017**	-0.004	-0.019***	-0.005	-0.003	-0.002
	(0.006)	(0.026)	(0.007)	(0.026)	(0.006)	(0.007)	(0.006)	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.002)	(0.004)
Total inequality	0.121***	0.104***			0.130***			0.018***	0.006**			0.018***		
	(0.012)	(0.013)			(0.012)			(0.005)	(0.003)			(0.005)		
Within group inequality			0.128***	0.112***		0.148***	0.105***			0.018***	0.003		0.022***	0.006
			(0.012)	(0.013)		(0.012)	(0.008)			(0.005)	(0.003)		(0.004)	(0.006)
Between group inequality			-0.039	-0.140		-0.018	0.063***			0.018	0.030***		0.002	0.016
			(0.074)	(0.085)		(0.075)	(0.049)			(0.017)	(0.011)		(0.010)	(0.013)
Intercept	0.666***	0.649***	0.669***	0.618	0.589***	0.593***	0.806***	-0.033	0.227**	-0.033	0.241***	0.041	0.044*	0.220***
	(0.065)	(0.236)	(0.065)	(0.232)	(0.066)	(0.065)	(0.059)	(0.028)	(0.059)	(0.029)	(0.059)	(0.032)	(0.023)	(0.042)
Different			-2.190**	-2.900***		-2.150**	-0.830			-0.010	2.330**		-1.800*	0.740
			[0.029]	[0.004]		[0.032]	[0.410]			[0.994]	[0.020]		[0.073]	[0.462]
R^2	0.762	0.716	0.766	0.719	0.538	0.576	0.576	0.425	0.275	0.425	0.266	0.295	0.297	0.400
Observations	306	306	306	306	306	306	306	357	357	357	357	357	357	357
Ind. effects		States		States			Divisions		States		States			Divisions
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	LS	LS	LS	LS	Med	Med	Med	LS	LS	LS	LS	Med	Med	Med

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 t-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 division-dummies.
Standard errors in parenthesis. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence. p-values in square brackets.

Table 3: Robustness to the parameter κ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
κ	-1	0	1	2	-1	0	1	2
Total inequality	.0001***	0.121***	0.110***	0.020				
	(0.000)	(0.012)	(0.024)	(0.016)				
Within group inequality					.0001***	0.128***	0.123***	0.023
					(0.000)	(0.012)	(0.025)	(0.016)
Between group inequality					-.0002	-0.039	-0.041	-0.074
					(0.002)	(0.074)	(0.098)	(0.146)
Different					-0.160	-2.190**	-1.590	-0.650
					[0.869]	[0.029]	[0.113]	[0.516]
R^2	0.689	0.762	0.693	0.682	0.689	0.766	0.704	0.682
Observations	306	306	306	306	306	306	306	306
Individual effects	No	No	No	No	No	No	No	No
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Estimator	LS	LS	LS	LS	LS	LS	LS	LS

Dependent variable is the average fraction of transfers in disposable income. All measures of inequality are generalized entropy measures with parameter κ. Control variables are the fraction of the population above 65, log of per capita income, fractionalization, and year dummies. Different is the t-test of the parameters on between and within group inequality being different. Estimation is by ordinary least squares. Standard errors are in parenthesis, p-values in square brackets. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence.

Table 4: Instrumental variables estimation

	(1)	(2)	(3)	(4)
Fraction above 65	1.413*	1.195	0.604***	0.586***
	(0.690)	(0.768)	(0.126)	(0.109)
Log per capita income	-0.090**	-0.096**	-0.069**	-0.064**
	(0.028)	(0.034)	(0.025)	(0.024)
Fractionalization	-0.001	0.086	0.014	0.032
	(0.029)	(0.081)	(0.028)	(0.029)
Total inequality	-0.303		0.167**	
	(0.393)		(0.062)	
Within group inequality		-0.225		0.170***
		(0.441)		(0.051)
Between group inequality		-2.134		-0.242
		(1.576)		(0.247)
Intercept	1.121*	1.158*	0.682**	0.625*
	(0.447)	(0.523)	(0.248)	(0.245)
		-1.32		-0.98
		[0.189]		[0.327]
Observations	306	306	306	306
Individual effects	No	States	No	States
Year effects	Yes	Yes	Yes	Yes
Estimator	2SLS	2SLS	2SLS	2SLS

Dependent variable is the average fraction of transfers in disposable income. All inequalities refer to the generalized entropy measure with parameter 0. Estimation is by 2SLS instrumenting inequality measures with lagged values interacted with year dummies. Different is the t-test of the parameters on between and within group inequality being different. R^2 is omitted as it was negative for all estimations. Omitted category for year-dummies is 2000.

Standard errors in parenthesis. Significantly different than zero at 90% (*), 95%(**), and 99% (***) confidence. p-values in square brackets

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

Year	Median income	Density x100	
		Blacks	Whites
2001	42228	1.03	1.05
2000	43162	1.03	1.04
1999	43355	0.96	1.04
1998	42173	0.98	1.07
1997	40699	1.05	1.11
1996	39869	1.04	1.12
1995	39306	1.06	1.16
1994	38119	1.03	1.18
1993	37688	1.06	1.20
1992	37880	1.06	1.20
1991	38183	1.08	1.22
1990	39324	1.06	1.24
1989	39850	1.04	1.18
1988	39144	0.98	1.19
1987	38835	1.04	1.21
1986	38365	1.04	1.23
1985	37059	1.12	1.28
1984	36343	1.12	1.31
1983	35438	1.13	1.35
1982	35423	1.16	1.36
1981	35478	1.11	1.34
1980	36035	1.15	1.35
1979	37192	1.12	1.31
1978	37234	1.17	1.32
1977	34989	1.23	1.39
1976	34792	1.26	1.42
1975	34219	1.32	1.45
1974	35159	1.29	1.46
1973	36278	1.24	1.41
1972	35560	1.25	1.46
1971	34126	1.38	1.53
1970	34481	1.34	1.54
1969	34714	1.39	1.57
1968	33436	1.41	1.65
1967	32081	1.43	1.68

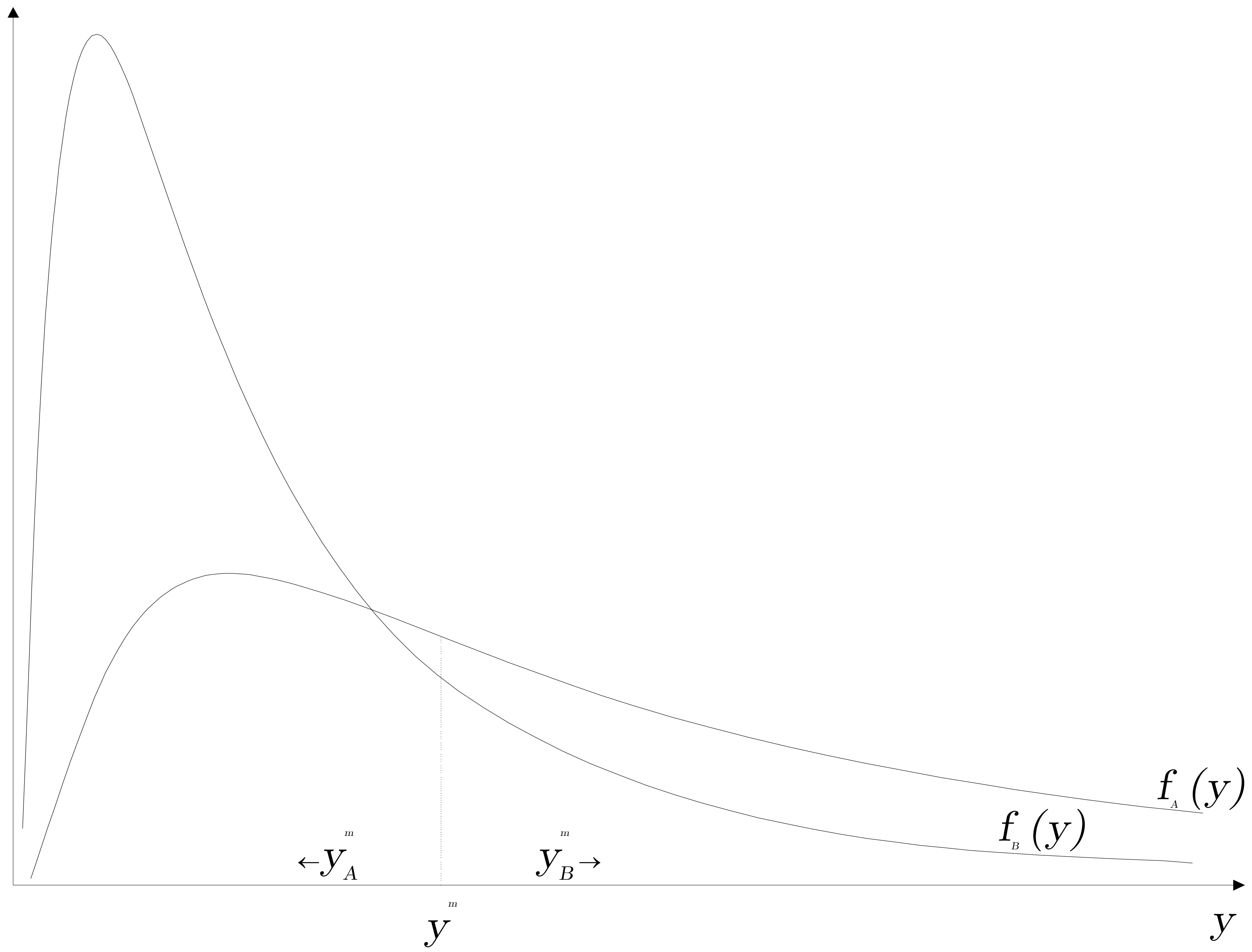

Figure 1: Example of densities at the mean

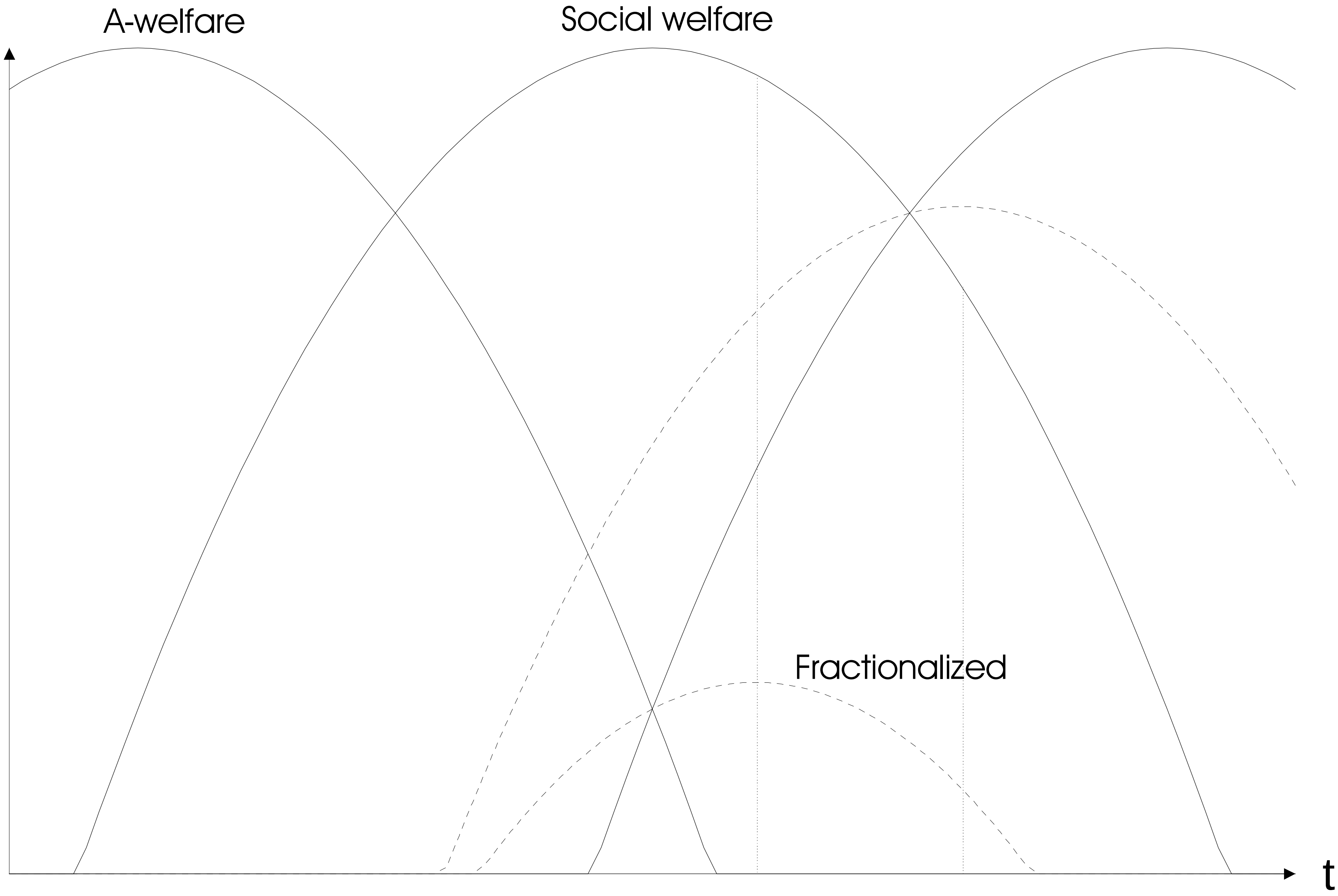

Figure 2: The effect of fractionalization on social welfare

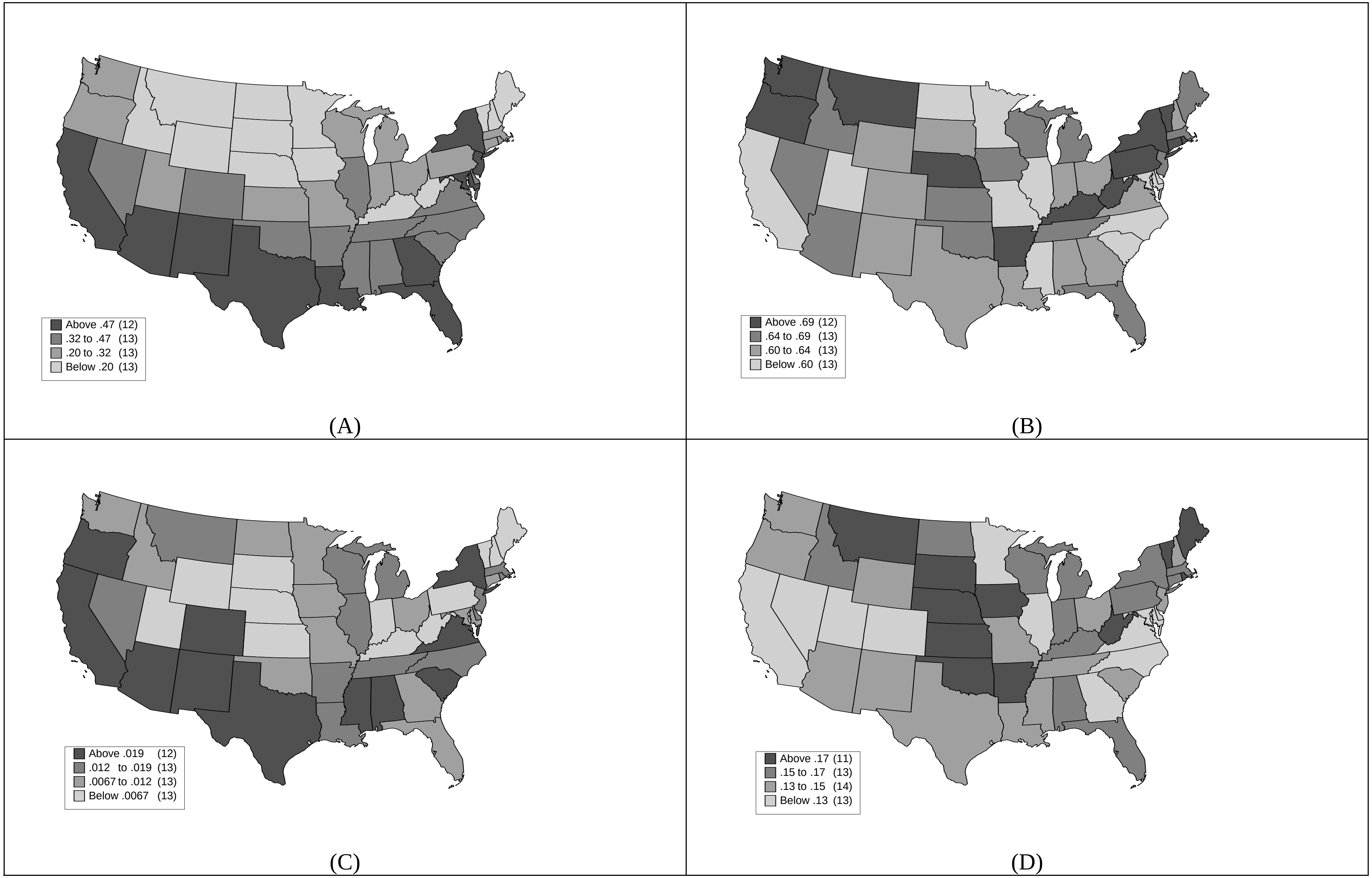

Figure 3: (A) Fractionalization, (B) Within group inequality, (C) Between group inequality, and (D) Average transfers received as share of household disposable income. All measured in 2000, inequality measure is generalized entropy measure with parameter 0.