NBER WORKING PAPER SERIES

PARTICIPATION IN HETEROGENEOUS COMMUNITIES

Alberto Alesina
Eliana La Ferrara

Working Paper 7155
http://www.nber.org/papers/w7155

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1999

We are grateful to Abhijit Banerjee, Francois Bourgignon, Ed Glaeser, Claudia Goldin, Luigi Guiso, Larry Katz, Caroline Minter-Hoxby, Daniele Paserman, Robert Putnam, Antonio Rangel, Howard Rosenthal, and to seminar participants at Harvard University (Saguaro Seminar, October 1998), Stanford University, and Ente Einuadi for very useful comments. We also thank Erzo Luttmer for sharing data and Spencer Glendon for help with data sources. This research is supported by an NSF grant to the NBER. We are grateful to both organizations for their support. Alesina also gratefully acknowledges financial support from the Weatherhead Center for International Affairs at Harvard University. All opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Participation in Heterogeneous Communities
Alberto Alesina and Eliana La Ferrara
NBER Working Paper No. 7155
June 1999

ABSTRACT

This paper studies both theoretically and empirically the determinants of group formation and of the degree of participation when the population is heterogeneous, both in terms of income and race or ethnicity. We are especially interested in whether and how much the degree of heterogeneity in communities influences the amount of participation in different types of groups. Using survey data on group membership and data on US localities, we find that, after controlling for many individual characteristics, participation in social activities is significantly lower in more unequal and in more racially or ethnically fragmented localities. We also find that those individuals who express views against racial mixing are less prone to participate in the groups the more racially heterogeneous their community is.

Alberto Alesina
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER
aalesina@harvard.edu

Eliana La Ferrara
IGIER, U. Bocconi
Via Salasco 5
20100 Milano, Italy
eliana.laferrara@uni-bocconi.it

1 Introduction

This paper studies the socio-economic determinants of participation in groups and social activities. This is an important issue for more than one reason. For example, political participation (voting and political action) has critical implications on policy choices.[1] If participation is low and the wealthy or more educated have a disproportionate propensity to vote and engage in political action, then public policies may be tilted in their favor. This may lead to vicious circles, in which disadvantaged minorities participate less, have less 'voice' and become even more disadvantaged, leading to a variety of social problems.[2] Second, participation has important economic effects. For instance, macroeconomists have argued that active interactions amongst individuals lead to transmission of knowledge, increases in aggregate human capital, and the development of 'trust', which improves the functioning of markets.[3] In addition, social interactions and networks affect many individual outcomes, from criminal activities, to fertility, to the labor supply.[4] Even though participation is typically associated with 'positive' socio-economic outcomes, social networks may also transmit 'negative' norms. For example, the so called 'culture of poverty and welfare' may find its roots in social networks propagating incentives to search for welfare rather than work.[5]

Most of the empirical literature studies how individual characteristics determine the individuals' choice to participate or not. We focus instead on how certain characteristics of communities influence the degree and nature of social interactions within them. We consider social activities like those occurring in religious groups, hobby clubs, youth groups, sport groups etc. Our interest in these social activities is motivated by Putnam (1993, 1995a,b) who suggests that this type of social interactions are particularly conducive to generating the beneficial effects of the so called *'social capital'*[6]. As for

[1]See Verba and Nie (1985) and Verba Scholtzman and Brady (1995).

[2]See Wilson (1987,1997) for a discussion of this type of vicious circles in american cities.

[3]On the effects of positive spillovers in the transmission of human capital see Romer (1986) Lucas (1988) and Benabou (1996). On trust, see La Porta et al. (1997).

[4]For theoretical work on the effects of tramsmission of informations in group and "informational cascades" see for instance Banerjee (1992) and Ellison and Fudemberg (1995). For empirical work on the importance of networks see Case and Katz (1992) and Betrand, Luttmer and Mullainhan (1998)

[5]See in particular Cutler and Glaeser (1997) and Betrand, Luttmer and Mullainathan (1998) for recent empirical work on this important question.

[6]See Coleman (1983) for an extensive discussion of the concept of 'social capial'.

the characteristic of the community, we study the role of heterogeneity in income, race and ethnicity. We want to know whether or not heterogeneity deters participation and social interaction.

Recent work has observed a positive bivariate correlation between inequality and social capital measures at the state level.[7] Our multivariate analysis, conducted with individual level data on participation and community level measures of income inequality, sheds light on this issue. As for race and ethnicity, much empirical research has studied the effects on public policy of ethnic and racial heterogeneity. For instance, Alesina, Baqir and Easterly (1999; 1998) show that the supply of 'core' productive public goods is lower and measures of patronage are higher in more racially fragmented localities. Glaeser, Scheinkman and Shleifer (1995), Cutler and Glaeser (1997), Poterba (1996), Luttmer (1997), and Goldin and Katz (1998) study the role of racial conflict as a determinant of education policies and several other characteristics of US cities.

In our model individuals prefer to interact with others who are similar to themselves in terms of income, race or ethnicity. If preferences are correlated to these characteristics, then our assumption is equivalent to saying that individual prefer to join groups composed of individuals with preferences similar to their own.[8] Given this setup, one may expect that diffuse preferences for homogeneity may decrease total participation in a mixed group if fragmentation increases. This is not necessarily the case. Going from a community which is, say, 90 percent white and 10 percent black, to one which is more fragmented, say 50-50, may or may not reduce participation in a mixed group. It depends on whether or not the loss of white participants is more than compensated by the increase in black participation. The implications of an increase in heterogeneity are even less clear if multiple groups can be formed. In this case, in fact, an increase in heterogeneity may lead to the creation of many homogenous small groups, leading to an increase or decrease of total participation. In our theoretical section, we develop conditions under which more heterogeneity lead to less (or more) participation.

Our empirical results on US localities suggest that income inequality and

[7] Robert Putnam, presentation at the Saguaro Seminar, October 1998, Harvard University.

[8] Theoretical results by Conley and Wooders (1996) are consistent with this assumption. They show that when agents can be crowded (positively or negatively) by the skills of other people in their jurisdictions, taste homogeneous jurisdictions are optimal. To the extent that tastes are correlated with income and race, our assumption follows.

racial and ethnic heterogeneity reduce the propensity to participate in a variety of social activities including recreational, religious, civic and educational groups.[9] Amongst the two, racial fragmentation seems to have a stronger negative effect on participation. Furthermore, and consistently with our model, these results are stronger for the groups in which direct contact amongst members is more important, like churches and youth clubs, while heterogeneity matters less or not at all in groups with low degree of interaction. Finally, our model predicts that individuals more averse to mixing with different types should be those more negatively influenced to heterogeneity in the community. To test this more stringent implication of the model, we exploit individual data on attitudes towards race relations (e.g., questions on mixed marriages, busing, etc.), and we find that the (negative) effect of racial heterogeneity on participation is significantly stronger for the individuals more averse to racial mixing.

This paper is organized as follows. Section 2 presents a model that generates predictions linking heterogeneity of the population and the level of participation in social activities. Section 3 describes our empirical strategy and data. Section 4 highlights some simple correlations at the state level between income inequality, racial and ethnic fragmentation, and measures of participation and social capital. Section 5 presents our econometric results. The last section concludes.

2 The model

Consider a community populated by two types, 'blacks' and 'whites', their size being $B > 0$ and $W > 0$, respectively. It is easier to think about 'race' (or ethnicity) as what identifies the types, for the moment. Below we discuss how to think about other discriminating characteristics, especially income.

Each individual has to decide whether or not to participate in a group. We assume that there exists a minimum size of the group equal to 1/2 of the total population, in order to ensure that at most one group will emerge in equilibrium. Below we briefly discuss the case of multiple groups. There are no congestion costs and no economies of scale: group size does not influence

[9] We define by 'race' the census classification of black, white, Asian, American Indian, and other. We define by ethnicity, the classification by ancestry, like Italian, Irish etc. Throughout this paper, we will use the terms 'white' and 'black' instead of Caucasian and African American, for the sake of brevity.

individual utility. Members of the group cannot exclude new members and entry in (exit from) the group is free and costless. The reservation utility from non participation is $\overline{u}$ for everybody. The reservation utility may of course be a function of individual characteristics. While this is important for the empirical analysis which follows, we will abstract from it in the theoretical discussion.[10]

Let P_W (P_B) be the proportion of whites (blacks) participating in the social activity. The utility from participation depends on individual characteristics and on the group's composition as follows:

$$u^i = u(\alpha_i^B, P_W) \text{ if } i \in B \quad (1)$$
$$u^i = u(\alpha_i^W, P_B) \text{ if } i \in W \quad (2)$$
$$u_\alpha(\cdot) < 0, \; u_P(\cdot) < 0$$

where $u(\cdot)$ is continuously differentiable in both arguments; $u_\alpha(\cdot)$ and $u_P(\cdot)$ represent the partial derivatives of $u(\cdot)$ with respect to α and to the proportion of the opposite type, respectively. Equation (1) represents the utility function of the generic member i of type B, and (2) is the analogous for types W. Underlying assumption $u_P(\cdot) < 0$ is the preference for participating in a social activity with members of one's own type. The parameter α_i, which is a taste indicator, captures the intensity of individual i's preference for participation in a group: higher values of α_i imply lower desire for participation. If we suppose that the cross derivative of $u(\cdot)$ with respect to α and P is negative (e.g., because α and P enter multiplicatively), higher values of α indicate more aversion to the opposite race. We do not need this assumption to develop the theory, but in the empirical part we discuss how 'aversion to the opposite race' strictly defined influences behavior. For brevity, we will identify α as the 'degree of intolerance' or 'aversion'.

An individual i of type B participates if and only if $u(\alpha_i^B, P_w) \geq \overline{u}$, or

$$\alpha_i^B \leq g(\overline{u}, P_w) \quad (3)$$

where $g(\cdot)$ is obtained inverting $u(\cdot)$. Our assumptions on $u(\cdot)$ imply that both the partial derivative of $g(\cdot)$ with respect to $\overline{u}$ and that with respect

[10]Note that since we only have at most one group formed, the alternative to belonging to the group is abstaining from participation altogether. This would not be the case if multiple groups were allowed.

to P are negative. In what follows we will indicate the latter derivative by $g'(\cdot) < 0$. Analogously, an individual of type W participates if and only if

$$\alpha_i^W \leq g(\overline{u}, P_B) \tag{4}$$

where again $g'(\cdot) < 0$. Note that in (3) and (4) the function $g(\cdot)$ is the same for the two types, since $u(\cdot)$ is the same in (1) and (2).

Assume α_i^W and α_i^B have the same differentiable cumulative distribution $F(\cdot)$. Then the total mass of individuals type B and type W who participate in the social activity is, respectively,

$$\widetilde{B} = \Pr\left\{\alpha_i^B \leq g(\overline{u}, P_w)\right\} \cdot B \equiv F\left[g(\overline{u}, P_w)\right] \cdot B \tag{5}$$

$$\widetilde{W} = \Pr\left\{\alpha_i^W \leq g(\overline{u}, P_B)\right\} \cdot W \equiv F\left[g(\overline{u}, P_B)\right] \cdot W \tag{6}$$

2.1 Equilibrium

Definition 1 *An* ***equilibrium*** *is a group composition* (P_B^*, P_W^*) *such that for both types none of the members wishes to leave the group and none of the non-members wishes to join.*

In equilibrium the proportion of individuals of type B in the group, P_B, must be equal to the ratio of the mass of the participants $\widetilde{B}$ to the total mass of participants $\widetilde{B} + \widetilde{W}$. The two conditions defining the equilibrium are therefore

$$P_B = \frac{F\left[g(\overline{u}, P_w)\right] B}{F\left[g(\overline{u}, P_w)\right] B + F\left[g(\overline{u}, P_B)\right] W} \tag{7}$$

$$P_W = 1 - P_B \tag{8}$$

which together give us the 'fixed point' equilibrium condition contained in the following.

Proposition 1 *There exists at least one equilibrium* $P_B^* \in [0, 1]$ *which solves*

$$P_B = \frac{F\left[g(\overline{u}, 1 - P_B)\right]}{F\left[g(\overline{u}, 1 - P_B)\right] + F\left[g(\overline{u}, P_B)\right](W/B)} \tag{9}$$

Proof: The proof of this proposition like all the other proofs is in Appendix A.

In what follows, for notational convenience we will suppress the term $\overline{u}$ from the arguments of $g(\cdot)$. We are interested in the properties of our equilibrium (equilibria).

Definition 2 *An equilibrium* $(P_B^*, 1 - P_B^*)$ *is* ***locally stable*** *if for given* W *and* B *a small perturbation, say to* $(P_B^* + \varepsilon, 1 - P_B^* - \varepsilon)$ *with* $\varepsilon \lesseqgtr 0$, *reverts to the original* $(P_B^*, 1 - P_B^*)$.

In other words, a group is 'stable' if when we add (remove) one member of either type, so that the composition of the group changes, this individual will choose to exit (re-enter) the group.

Remark 1 *A necessary and sufficient condition for* $(P_B^*, 1-P_B^*)$ *to be locally stable is that*

$$\left| \frac{F'[g(1-P_B^*)]\, g'(1-P_B^*)}{F[g(1-P_B^*)]} + \frac{F'[g(P_B^*)]\, g'(P_B^*)}{F[g(P_B^*)]} \right| \cdot \frac{W}{B} <$$

$$< \frac{F[g(1-P_B^*)]}{F[g(P_B^*)]} + \frac{F[g(P_B^*)]}{F[g(1-P_B^*)]} \left(\frac{W}{B}\right)^2 + 2\frac{W}{B}. \tag{10}$$

To gain some insights into the meaning of condition (10), it is useful to think in graphical terms. The equilibrium value(s) of P_B is (are) given by the intersection of the function in the right hand side of (9) with the 45° line. The right hand side of (9) represents the fraction of members type B in the group that is 'generated by the reactions' of both types to a given composition $(P_B, 1 - P_B)$. The intersection(s) with the identity line give(s) the value(s) of P_B at which both reactions are consistent with the actual proportions. Our stability condition requires that the slope of the above function at the point of intersection with the 45° line be less than one.

[Insert figure 1]

Figure 1 depicts various possible configurations of equilibria. In panel (a), there is a unique interior equilibrium, i.e. the group that forms is 'mixed', with a proportion $P_B^* \in (0, 1)$ of blacks and $(1 - P_B^*)$ of whites. This equilibrium is stable: in fact if we perturb P_B^* to the right or to the left, the movement along the curve brings us back to the initial equilibrium. This can be understood as follows. Suppose you add one or more black members to the group, so the composition becomes $P_B > P_B^*$. The shape of the curve in panel (a) suggests that this 'more favorable' composition for types B does not trigger enough participation of new B members, nor does it induce enough W members to exit, so the group goes back to the initial equilibrium. In panel (d) the opposite occurs, that is, any slight increase (reduction) in the fraction of B members from P_B^* will trigger an inflow (outflow) of B types and outflow (inflow) of W types, so that the composition of the group goes all the way to complete homogeneity with $P_B^* = 1$ ($P_W^* = 1$). While in panel (d) either $P_B^* = 0$ or $P_B^* = 1$ can in principle be stable equilibria[11], in panel (b) only $P_B^* = 0$ is. Finally, panel (c) illustrates the case of multiple equilibria. Of the five equilibria depicted in the figure, only P_{B1}^* and P_{B3}^* are stable, while $0, P_{B2}^*$, and 1 are unstable.

The issue of the selection among multiple equilibria will not be addressed in this paper. As a first approximation, we can think that initially all the population is in the group, or alternatively that the 'prior' of every individual on the group composition is equal to the composition of the population. This will ensure that the 'starting point' on the graph is $P_B^0 = B/(B+W)$ and that simple dynamics will bring the group to the stable equilibrium that is closest to P_B^0.

We are interested in two features of the equilibrium. The first is how the composition of the group relates to the composition of the total population; the second is *who* among the heterogeneous individuals of a given type will choose to participate and who will stay out. We answer these questions by establishing the following results.

Lemma 1 *Let (P_B^*, P_W^*) be a unique stable equilibrium. Then $B < W \iff P_B^* < P_W^*$.*

Corollary 1 *If $B \neq W$ and (P_B^*, P_W^*) is a unique stable equilibrium, then either $\frac{P_B^*}{P_W^*} < \frac{B}{W} < 1$ or $1 < \frac{B}{W} < \frac{P_B^*}{P_W^*}$.*

Corollary 2 *Let (P_B^*, P_W^*) be a unique stable equilibrium. The individuals type B and W who participate are those whose α_i's fall in the intervals*

[11] As a matter of fact, given our assumptions that $B < W$ and that the minimum size of the group is 1/2, only $P_W^* = 1$ can be a stable equilibrium.

$[0, \overline{\alpha}^B]$ *and* $[0, \overline{\alpha}^W]$, *respectively, where* $\overline{\alpha}^B \equiv g(\overline{u}, P_w^*)$ *and* $\overline{\alpha}^W \equiv g(\overline{u}, P_B^*)$. *In particular,* $B < W \Longleftrightarrow \overline{\alpha}^B < \overline{\alpha}^W$.

Lemma 1 says that the type who is a minority in the population will be a minority also in the group. Even more, corollary 1 shows that the unbalance between the types in the population is *magnified* within the group: not only is the minority type under-represented in the social activity, but it is less than proportionately represented compared to its weight in the population. Thus, the social group is more homogeneous than the whole population, and this result is consistent with the under representation of minorities which we observe in many social activities. Corollary 2 can help us understand this result. In general, among both types only the individuals more 'prone' to participating, or less averse to diversity (i.e. the low α_i's), will join the group. Suppose $B < W$. Although the two types have the same ex ante distribution of α, it is not possible to observe the same degree of 'aversion' for B and W in the equilibrium composition of the group. If this were true, the same α should be indifferent between participating as a majority or as a minority. Instead, the fact that B is under-represented in the population induces even some relatively 'moderate' B individuals (low α) to stay out, while W individuals manage to keep some relatively 'participation averse' (high α) members in the group. This coupled with the sheer unbalance in the numbers B and W produces the 'magnification' effect described above.

2.2 Heterogeneity and participation

We are finally ready to study how a change in the heterogeneity of the population influences the total amount of participants. First of all, we define precisely what we mean by 'heterogeneity'.

Definition 3 *The degree of* ***heterogeneity*** *is the probability that two randomly drawn individuals from the population belong to different types.*

We will use the same definition in our empirical analysis below. Obviously, in our case of two types, a fifty-fifty split has the maximum level of heterogeneity. Denote with w the fraction of 'whites' in the population, i.e.

$$w \equiv \frac{W}{W+B}.$$

Remark 2 *An increase in w represents a decrease in heterogeneity if $w \geq 1/2$ and an increase if $w < 1/2$.*

Definition 4 *The* ***aggregate level of participation*** *S is the share of the total population who belongs to the group:*

$$S \equiv \frac{\widetilde{B} + \widetilde{W}}{B + W} = F\left[g(1 - P_B^*)\right](1 - w) + F\left[g(P_B^*)\right]w \tag{11}$$

We are interested in how S changes with changes in the composition of the population. Under very mild sufficient conditions on $F(\cdot)$ and $g(\cdot)$,described in Appendix A, the following holds.

Proposition 2 *If a unique stable equilibrium exists, an increase in heterogeneity reduces total participation, i.e.*

$$\frac{dS}{dw} < 0 \Longleftrightarrow w < \frac{1}{2}$$

If more than one stable equilibrium exists, the previous result holds 'locally' around each of the equilibria.

The intuition underlying proposition 2 is simple. We have established that in equilibrium the minority type is under-represented in the social activity, e.g. if $B < W$ then $P_B^* < P_W^*$, and that the fraction of group members of a given type is increasing in the relative size of that type in the population ($\partial P_W^* / \partial w > 0$). Consider what happens when whites are the majority ($w > 1/2$) and the fraction of whites in the population decreases —an increase in heterogeneity. First of all, since the fraction of whites who participates is higher than that of blacks, i.e. $F\left[g(P_B^*)\right] > F\left[g(1 - P_B^*)\right]$, the fall in W creates an absolute loss in participants greater than the increase in B. Furthermore, the fractions of the two types who participate change with the new value of heterogeneity. The sufficient condition mentioned in the text ensures that the increase in the fraction of blacks is not so overwhelmingly larger than the decrease in the fraction of whites to overcome the first effect.

2.3 Two examples

Two simple examples highlight the critical features of our results.

2.3.1 Example 1

Suppose that α is distributed uniformly on $[0,1]$ and that the utility functions are:

$$\begin{aligned} u_i^B &= 1-\alpha_i\sqrt{P_W} \\ u_i^W &= 1-\alpha_i\sqrt{P_B} \end{aligned}$$

Therefore, $u^B_{P_w} < 0$ and $u^B_{P_W P_W} > 0$. A positive second derivative implies that increasing the proportion of whites decreases the marginal utility of blacks by more if there are very few whites. Suppose that a group is completely homogeneous; the first few participants of different types may require the adoption of different procedures, a different language etc. These costs would be declining as the minority becomes larger.

By definition:

$$\begin{aligned} \widetilde{B} &= \Pr\left\{\alpha_i^B \leq \frac{1-\overline{u}}{\sqrt{P_W}}\right\}\cdot B = \frac{1-\overline{u}}{\sqrt{P_W}}B \\ \widetilde{W} &= \Pr\left\{\alpha_i^W \leq \frac{1-\overline{u}}{\sqrt{P_B}}\right\}\cdot W = \frac{1-\overline{u}}{\sqrt{P_B}}W \end{aligned}$$

The fixed point equilibrium condition, i.e. the special case of equation (9), is:

$$P_B^* = \left(1+\sqrt{\frac{1-P_B^*}{P_B^*}\frac{W}{B}}\right)^{-1}$$

Note that the reservation utility $\overline{u}$ does not influence the composition of the group, but it does influence its size. Consider the following set of parameter values: $\frac{W}{B}=2$; $\overline{u}=0.6$. In this case we have a unique stable equilibrium in which $P_W^*=0.8$ and $P_B^*=0.2$, as shown in Figure 2a. Also, $\widetilde{B}\simeq .45B$ and $\widetilde{W}\simeq .89W$, so the aggregate participation rate is $S\simeq .74$.

[Insert figure 2]

As discussed in the general case, we have a magnification effect: the composition of the group is more tilted in favor of the majority than the distribution of types in the population. A ratio $B : W$ of $1 : 2$ in the population translates into $1 : 4$ within the group. Suppose now that $B = W$, i.e. the fragmentation is at its maximum value. In this case it is easy to check that $P_W^* = P_B^* = \frac{1}{2}$ and $\widetilde{W} = \widetilde{B} \simeq .57B$. The participation rate is now $S \simeq .57$, which is the minimum attainable given our parameter values.

2.3.2 Example 2

Consider now the following utility functions:

$$\begin{aligned} U_i^B &= 1 - \alpha_i (P_W)^2 \\ U_i^W &= 1 - \alpha_i (P_B)^2 \end{aligned}$$

In this case both the first and second derivatives of the utility functions relative to the P_j's ($j = B, W$) are negative. The interpretation of these utility functions is the opposite of that of example 1. For instance, we can think of groups where majority voting matters for certain decisions, so that the marginal utility of losing members of your own type may be increasing as you are approaching a half and half split. In this case the interior equilibrium is described by:

$$P_B^* = \left(1 + \left(\frac{1 - P_B^*}{P_B^*}\right)^2 \frac{W}{B}\right)^{-1}$$

Using the same parameter values of example 1, namely $\frac{W}{B} = 2$ and $\overline{u} = 0.6$, we get $P_B^* \simeq .67$ and $P_W^* \simeq .33$. Figure 2b plots the interior equilibrium. Notice that this equilibrium is unstable: starting for example at an initial value equal to the composition of the population, $P_B = 1/3$, the group would move towards the corner where only whites participate, i.e. $P_B^* = 0$ (indeed this is the only stable equilibrium feasible given our requirement that group size is at least half of the total population). Notice that in this equilibrium all the whites participate, so $S = 2/3$. If we increase fragmentation by decreasing W/B while still maintaining $W/B > 1$, the only stable equilibrium would still be $P_B^* = 0, P_W^* = 1$, and the participation rate would decrease (because the mass of participating whites is now smaller).

2.4 Extensions and discussion

The simple model outlined above is sufficient to generate the basic prediction that an increase in heterogeneity will be associated with less participation. There are, however, a number of issues from which we have abstracted and which we should discuss briefly.

2.4.1 Tolerance and population shares

The empirical literature on political participation suggests that, after controlling for socio-economic status, blacks have a higher propensity to participate in groups and to vote than whites.[12] In our empirical analysis below, we also find the same result. The explanation offered by political scientists is that blacks are more conscious of being a minority and have an extra incentive to engage in political action to preserve their identity and foster their political and civil rights. In our model this could be accommodated by assuming that $\alpha^i = \alpha^i(W/B)$, with $i = B, W$. That is, the propensity to participate is a function of the distribution of types in the population, with $\partial\alpha^B/\partial(W/B) > 0$ and $\partial\alpha^W/\partial(W/B) < 0$. Therefore if $B < W$, then we must have $\alpha^B > \alpha^W$. This implies that after controlling for all other determinants of participation, one should obtain the empirical finding that blacks participate more, since they are a minority.

2.4.2 Multiple groups

An important extension of our model would be to allow for the formation of multiple groups, by reducing their minimum feasible size. In this case an increase in heterogeneity might simply lead to the formation of multiple homogeneous groups with uncertain effects on total participation. However, to the extent that groups cannot be too small (either because there is a minimum size, or the benefits of participation are decreasing in size)[13] our basic insight generalize. That is, moving from a more to a less homogeneous community, the population would sort by 'pushing' on two margins: creating smaller groups (which are more costly or less beneficial), and increasing the

[12] See for instance Verba and Nie (1987)

[13] There could also be congestion costs, but this is less relevant for the issue under consideration; here we are concerned about the possibility of equilibria with many small homogeneous groups.

heterogeneity of the existing groups.[14]

More specifically, consider a population which is totally homogeneous, and suppose that the number of groups chosen is x: this is the number which optimizes the group size, or the trade off between congestion and the costs of being too small. Consider now an increase in heterogeneity, holding the total population constant. In the new equilibrium there will be more groups and, in general, they will be less homogeneous. This is indeed the result of reoptimizing along two margins: homogeneity and size. Since in the original equilibrium the group size was at its optimal level, the new groups will be smaller and, by definition of 'optimal size', they will provide lower utility to their members than the original ones. Thus, some of the previous members will drop out regardless of their preference for homogeneity. In addition, individuals who are less tolerant of heterogeneity will also drop out.

2.4.3 More than two types

The extension to more than two types is computationally quite cumbersome. The precise nature of the results would depend on the structure of preferences. For instance, one simple case would be that each individual prefers homogeneity and has no ranking of preferences amongst other types different from his or her own. Alternatively, individual preferences may be influenced by the mix of specific types which would be part of the group.[15]

2.4.4 Heterogeneity in income

Thus far we have focused on differences across types not based on income. In the empirical analysis we are also interested in the effect of an increase in income inequality on participation.

[14]Results obtained by Alesina and Spolaore (1997) on endogenous jurisdiction formation, economies of scale and preference for homogeneity are related to this point.

[15]Note that even in the case where an individual's utility is a function of the fraction of members different from his or her own type, the individual may not be 'indifferent' to the particular mix of types in the group if decision making rules depend upon relative shares in the group. For example, suppose there are three types: A, B, C and compare two possible equilibria with shares (.40, .30, .30) and (.40, .50, .10) respectively. Although individuals type A may not have any general 'distaste' for B versus C, they might well prefer the first type of equilibrium, in which they are likely to have more decision power, to the second. This type of analysis is certainly realistic but goes beyond the scope of the present work.

A vast literature in local public finance addresses the issue of group formation and income levels.[16] Particularly related to our analysis is the work by Fernandez and Rogerson (1996) and La Ferrara (1998), since both examine the effects of changes in income inequality. The first two authors show that the quality of publicly provided education (and, therefore, the benefit of acquiring it) is inversely related to income inequality. La Ferrara (1998) shows that in a sample drawn form Tanzania, the degree of participation in groups which provide economic benefits or informal insurance to their members is inversely related to income inequality in the community.

Our model is not a contribution game. Differences in income matter only to the extent that they are correlated with preferences and culture. In this case our formalization could be reinterpreted in terms of income rather than race, i.e. individuals would prefer to participate in social activities with people from their own income bracket. A complication, however, is that income is a 'continuous' variable, while race is much less so: in modelling income dispersion as a dispersion of 'types' we are therefore simplifying the analysis. To the extent that the issue of contributions would increase the preferences of rich individuals to exclude the poor, then income inequality would matter even more for participation. This suggests a test similar to that performed by La Ferrara (1998) namely that income inequality should matter more for participation in groups where the discrepancy in the net benefits perceived by the rich and the poor is larger.

3 Empirical strategy and data

We assume that at any point in time the 'latent variable' measuring the expected utility from participation in a group for individual i in community c can be modelled as:

$$Y_{ic}^{*} = X_{ic}\beta + H_c\gamma + S_c\delta + T\lambda + \varepsilon_{ic} \tag{12}$$

where X_{ic} is a vector of individual characteristics; H_c is a vector of community variables (including our variables of interest, namely heterogeneity of the population in terms of income, race, and ethnicity); S_c is a dummy for the state where the individual lives; T is a year dummy, and ε_{ic} is an error term normally distributed with mean 0 and variance σ_c. The vectors β, γ, δ, and

[16] See for instance Epple and Romer (1991) and the references cited therein.

λ are parameters. We do not observe the 'latent' variable Y_{ic}^* but only the choice made by the individual, which takes value 1 (participate in a group) if Y_{ic}^* is positive, and 0 (not participate) otherwise:

$$\begin{array}{l} P_{ic} = 1 \text{ if } Y_{ic}^* > 0 \\ P_{ic} = 0 \text{ if } Y_{ic}^* \leq 0 \end{array} \tag{13}$$

We estimate the Probit model (12)-(13) using individual level data and taking Metropolitan Sampling Areas (MSA) and Primary Metropolitan Sampling Areas (PMSA) as 'community' dimension. We are especially interested in the vector of coefficients γ, although many of the components of β will also be important to gain insights into the determinants of participation.

The main source of data for our regressions is the General Social Survey (from now on, GSS) for the years 1974-94. This survey interviews approximately 1,500 individuals every year from a nationally representative sample, and contains information on a variety of socio-political indicators, as well as on demographic and income characteristics of the respondents.[17] In particular, the questionnaire prompts the respondent to answer questions regarding his or her membership in organizations such as political groups, religious groups, unions, school associations, service groups, fraternities, sports and hobby clubs, etc. We use the answers to these questions to construct our dependent variables. Another interesting feature of the GSS is that it contains information about individual attitudes towards race relations and racial mixing. This will prove useful for us in that we will draw on these questions to construct proxies for our parameter of 'participation aversion' (α) in order to test the implications of our model as strictly as we can.

Among the explanatory variables we include individual controls taken from the GSS, as well as community variables capturing heterogeneity in race, ethnicity, and income in the place where the individual lives. The individual controls, as well as all other variables, are described in Appendix B. The remainder of this section illustrates our procedure for constructing community level variables.

It is possible to match most respondents from the GSS 1972-1994 with the

[17]Note that the survey was not conducted in 1979, 1981, and 1992. Moreover, in 1982 and 1987 black individuals were oversampled, therefore in our regressions we will use the weights provided by the cumulative GSS file 1972-1994 to correct for this oversampling. For more detailed information about the GSS, the reader is referred to Davis and Smith (1994).

MSA/PMSA where they live. We have therefore used Census data to build community level variables, adopting the MSA/PMSA as our geographic notion of 'community'. Our measure of income inequality is the Gini coefficient for the MSA/PMSA computed using family income figures from the 1970, 1980, and 1990 Censuses. The values of *Gini* for the remaining years were obtained by linear interpolation and extrapolation. Our results are not overly sensitive to the interpolation procedure. Moreover, we computed Gini coefficients at the state level from the Current Population Survey (CPS) every year between 1974 and 1994, and the correlation between the CPS and the Census interpolated Gini's was extremely high. The state level Gini's from the CPS are the ones we use in figures 4, 7, and 8.

Our racial fragmentation index (*Race*) is constructed from the Census 1990 according to the following formula:

$$Race_i = 1 - \sum_k s_{ki}^2 \qquad (14)$$

where i represents a given MSA/PMSA and k the following races: i) White; ii) Black; iii) American Indian, Eskimo, Aleutian; iv) Asian, Pacific Islander; v) other. Each term s_{ki} is the share of race k in the population of MSA/PMSA i. The index (14) measures the probability that two randomly drawn individuals in area i belong to different races. Therefore, higher values of the index represent more racial fragmentation.

The Census did not identify 'Hispanic' as a separate racial category. However, Alesina Baqir and Easterly (1999), who use the same measure of racial fragmentation, note that the category 'Hispanic' (which they obtain from a different source) has a correlation of more than 0.9 with the category 'other' in the Census data. Our interpretation is that individual of Hispanic origin checked the 'other' category in the Census survey, since they did not feel well represented by the four other possibilities. Thus, for all practical purposes, the category 'other' in the Census is virtually a measure of the Hispanic population.

The ethnic fragmentation index (*Ethnic*) is computed by an analogous formula to (14), using *ancestry* instead of race. In other words, s_{ki} in the formula now represents the fraction of people in area i whose first ancestry is type k. The original ancestries reported by the 1990 Census (35 categories) have been aggregated into 10 different groups on the basis of common language, culture, and geographic proximity (see Appendix B for a precise definition). We have chosen to aggregate these data in order not to give the same 'weight' in the definition of *Ethnic* to very similar countries of origin, say Norway and Sweden, and two very different ones, say India and Ireland. Our results are not unduly sensitive to reasonable changes in our aggregation rules.

Note that we use the values of *Race* and *Ethnic* in 1990 for the whole sample. Our reasons for not interpolating are twofold. First, we believe that racial and ethnic fragmentation within MSAs are sensibly more stable over time than, say, income inequality. Second, and most importantly, in order to get variation over time, we should have resorted to the 1970 and 1980 Censuses, which contained fewer categories. For example, all Censuses before 1990 distinguished only three races: white, black, and other. Relying on years earlier than 1990 would thus have meant sacrificing the precision of our heterogeneity measures to a considerable extent. We felt that the loss in explanatory power due to this oversimplification outweighed the potential gain from time variation of the above indexes, hence we chose to adopt the 1990 measures as our best proxies for racial and ethnic fragmentation.

4 Descriptive results

We begin by presenting summary statistics and a few simple correlations among membership rates and our measures of heterogeneity. Summary statistics and definitions of the full set of variables can be found in Appendix B.

[Insert Table 1]

The top panel of Table 1 shows the sample characteristics of some of our data. Participation rates are on average very high: overall, 72 percent of the respondents are members of at least one group, the average number of group memberships being 1.8 per person. Also, there is considerable variation in participation rates, both across individuals and across groups: the standard deviation of our basic membership variable is 0.45. The fraction of participants in the various groups ranges from 0.02 for farmers' associations to 0.35 for religious groups. Sport groups are the second most popular category, with a participation rate of 0.21, followed by professional associations (0.17), unions (0.15), and school service groups (0.14). Literary groups, hobby clubs, fraternities, and service groups (Rotary, Lions, etc.) have participation rates of 0.10-0.11. Most notable is the low enrollment in political associations: only 5 percent of the respondents are members of a political group. Nationality groups, which we will consider in more detail below, are joined by about 4 percent of the respondents.

The last three variables in panel A of the table are measures of heterogeneity in income, race, and ethnicity in the MSA's where the respondents live. The mean of the Gini coefficient is 0.41, with a standard deviation of 0.03. Our racial fragmentation index has a mean of 0.36 (standard deviation 0.14), while ethnic fragmentation is higher at 0.67 (standard deviation 0.07). The correlations among these three measures of heterogeneity is quite high, as shown in the bottom panel of Table 1: *Gini* is correlated 0.43 with *Race* and 0.16 with *Ethnic*; the correlation between *Race* and *Ethnic* is 0.56. The high correlation between these variables should be remembered for the discussion of our regressions.[18] On the other hand the simple correlation between average membership in the MSA and the Gini index is -0.23, while that between membership and racial or ethnic fragmentation is not significant. This will no longer be true when we turn to multivariate analysis.

Figures 3, 4, 5 and 6 illustrate the geographic distribution of our variables of interest by reporting sample averages at the State level.

[Insert figures 3,4,5,6]

Figure 3 shows the distribution of participation rates from the GSS data set, that is the percentage of respondents in each state who are members of at least one group (average from 1974 to 1994). As one can see this percentage is highest in states of the north central and northwest regions, and lowest in the south and south east. Figures 4, 5, and 6 show the distribution of the Gini coefficient (average 1974-94 from CPS data), of Racial fragmentation, and of Ethnic fragmentation (both measured in 1990), all calculated at the State level. These maps show a rather striking pattern, when compared with Figure 3: racial and ethnic fragmentation, as well as income inequality, are highest in the south east and lowest in the north east, i.e. those regions where participation is, respectively, lowest and highest.

[18] We also explored the correlation among our heterogeneity variables and the measures of racial or income segregation of Cutler, Glaeser and Vigdor (1999). It turns out that segregation by income is positively correlated with income inequality, and that the various measures of racial segregation are positively correlated with racial and ethnic fragmentation. This may suggest that segregation is 'valued' relatively highly in places where heterogeneity is higher, an observation consistent with the argument that individuals prefer contact with people similar to themselves.

[Insert figures 7,8]

Similar implications can be gathered from the three panels of Figure 7, which plot state level participation rates against the Gini index, Race fragmentation, and Ethnic fragmentation. In all three panels a negative correlation between membership in groups and heterogeneity clearly emerges. We tried similar plots for other dimensions of 'social capital' and 'civic culture' derived from the GSS 1972-94, namely for *Trust* —defined as the percentage of respondents in each state who say that "most people can be trusted"— and for *Vote* —the percentage of respondents who voted in the last presidential election— and the pattern was strikingly similar. A synthetic view of these results can be gathered by figure 8, which plots *Gini*, *Race*, and *Ethnic* against an *aggregate index of social capital*. This index is constructed by extracting the principal components from the three variables above which we computed from the GSS, i.e. *Member*, *Trust*, and *Vote*.[19] Again, these simple correlations suggest a strong negative correlation between various forms of heterogeneity and what we generally think of as 'social capital'.

5 The Econometric evidence

5.1 Basic Regressions

Table 2 displays our basic probit regression using the GSS data set and including only individual controls. The dependent variable takes the value 1 if the respondent belongs to at least one group, and 0 otherwise. The regressors include a set of individual characteristics which, in our model, may influence either the individual's reservation utility if not participating, $\overline{u}$, or the preference for participation for given group composition (captured by the parameter α). The political science literature has generally looked at these individual determinants of participation in isolation, i.e. correlating one or two variables at a time with membership rates.[20]

[19] Principal component analysis has been typically used by the related literature to construct a single index of social capital (see, among the others, Putnam and Yonish (1998)).

[20] See for instance Verba and Nie (1987), and Verba, Schlozman and Brady (1995). A multivariate analysis with demographic controls similar to those we include is in Glaeser and Glendon (1997), but their dependent variable is church attendance rather than group participation.

[Insert table 2]

The estimates in the first column of Table 2 are marginal probit coefficients evaluated at the means; in the second column we report heteroskedasticity corrected standard errors adjusted for intra-MSA clustering of the residuals. First of all, the cohort variable suggests a decline in participation by younger cohorts. Secondly, the age distribution variables show a dip in participation for individual in their thirties. Child raising activities reduce the time available for participation: in fact the coefficient on the variable that captures whether the respondent has children below the age of 5 is negative and significant. Note that the dummy for age group 30-39 and that for children below age 5 are highly correlated. More generally, both variables capture a period of individual lifetime which is particularly 'busy' because of marriage, having children, setting up new households, etc. There is some weak evidence that older people participate more, probably because they have more time if they are retired, although health considerations may work the other way. This result, together with the cohort effect, accounts for the notion of 'older civic generation' emphasized by Putnam (1995a,b).

Years of schooling are positively associated with participation: high school dropouts participate significantly less, while college graduates significantly more. The coefficients on the education variables remain highly significant and stable throughout all specifications. Among the possible explanations for this strong association, Verba and Nie (1987) suggest that more education is generally combined with a higher evaluation of one's own ability to influence socio-political outcomes, and with a higher level of social interaction. In Table 2 we also see that women participate significantly less than men. Our interpretation is that they often carry the weight of a job plus a preponderant share of household chores and child raising activities. This heavy load leaves women with less time for leisure and participation. This is corroborated by the fact that women do *not* participate less in voting, an act of participation which does not require a significant amount of time[21]. Participation is increasing in family income of the respondent, suggesting that participation is a 'normal good'.

Consider now the time spent at work. The omitted category captures people who are not working, including housekeepers, retirees, students and unemployed. After controlling for the level of income, the effect of time

[21] Results on this point using the GSS data are available from the authors.

spent at work could be twofold. On the one hand, a constraint on time may decrease participation; on the other hand, socialization in the workplace may increase social interaction, incentives, and ability to participate. Our results on this point are consistent with basic economic principles. The coefficient on part time workers is larger (and significantly so) than the one on full time workers. This suggests that, even though socialization in the workplace helps (in fact full time workers participate more than those out of the labor force), the time constraint is binding for people who work full time. It is less binding for part time workers who, on the other hand, still get the benefits of social interactions in the workplace.[22] When we control for all these variables, marital status does not seem to affect participation (contrary to the common notion in sociological and political analyses based on partial correlations, which indicate that married people participate significantly more).

For our purposes, a particularly interesting variable is 'black'. As we can see from Table 2, *ceteris paribus* members of this racial group participate significantly more, a result which emerges clearly when we control for the other individual determinants of participation but which is obscured if we only look at partial correlations. Note that this result is not driven by the higher church attendance of blacks in the south: in fact it survives if churches are left out of the definition of groups and if the south is omitted from the regression. As discussed above, our model could be extended to incorporate a feature of group consciousness, in which the minority type participates more to preserve identity and to defend its role in the community. More importantly, since blacks are a minority in virtually all MSA's, the percentage of black residents is positively associated with racial fragmentation. The result on black propensity to participate implies that if we find that participation is *lower* in more racially fragmented communities, this result is not due to the positive correlation between percentage of blacks and racial fragmentation.

Regressors not shown include year dummies and state dummies. The pattern of year dummies is broadly consistent with the declining trend in participation rates, already partly captured by the variable cohort. Many of the state dummies are statistically significant, indicating a need to include them.

[22] This interpretation is confirmed by additional sensitivity analysis which explores alternative specifications and distinguishes between the participation levels of various components of the non workers. Interestingly, the unemployed participate less even after controlling for income. All these results are available from the authors.

The coefficients on individual controls are very stable and robust to different specifications.[23] Therefore, to economize in space from now on we will not report them, although it should be intended that they have always been included in the regressions, together with the state and year dummies. We next extend our analysis by incorporating variables which capture the characteristics of the community where the respondent lives.

[Insert Table 3]

In Table 3 we include the size of the place where the individual lives, the median income level in the MSA, and its squared (all in logs), together with our measures of heterogeneity. Size has a negative but not significant coefficient, while the coefficients on the income variables indicate that richer communities participate more but at a decreasing rate. Finally, we move to the characteristics of communities which are the focus of the present paper. The first measure of heterogeneity included in column 1 is income inequality. The coefficient on *Gini* is negative and significant at the 1 percent level, indicating that *people living in more unequal communities are less likely to participate in groups.* Column 2 includes our measure of racial fragmentation, which also has a negative and significant coefficient: *individuals living in more racially fragmented areas participate less.* In column 3, 'racial' fragmentation is replaced by 'ethnic' fragmentation as measured through the ancestry data. Again, the negative and significant coefficient on this variable suggests that *participation is lower in more ethnically fragmented communities.*

In the last three columns of Table 3 we introduce in the same regression both inequality and our measures of racial and/or ethnic heterogeneity. *Gini* and racial fragmentation remain significant when introduced jointly; however, the absolute values of their coefficients fall due to the positive correlation among the two variables, highlighted in Table 1. Ethnic fragmentation remains negative and significant in column 5, but looses significance in the last column when all three measures of heterogeneity are included. This result may suggest that among the three kinds of heterogeneity we consider, *Ethnic* is the one that is less strongly associated with participation, a hypothesis confirmed by the results on the group by group analysis of the next

[23]We do not include homeownership among our regressors because this would restrict our sample to 3101 observations only. When included in the regression, homeownership has a positive and significant coefficient, consistently with the findings of DiPasquale and Glaeser (1998).

section. In any case, given the high degree of correlation between the three indexes of heterogeneity, in what follows we present results from regressions where the above measures have been introduced one at a time.[24]

5.2 Sensitivity analysis and causality

In Table 4 we conduct a sensitivity analysis by controlling for influential observations whose presence would sensibly bias our estimators. We do this by calculating the DFbetas from each original regression, and dropping those observations that lead to major changes in the coefficients of our heterogeneity measures.[25]

[Insert Table 4]

The results in Table 4 are even stronger than those in table 3, in the sense that the coefficients on *Gini*, *Race* and *Ethnic* are larger in absolute value. Using the estimated coefficient in column 1, we calculate that, starting from the sample mean, an increase in *Gini* by one standard deviation leads to a reduction in the probability of participation of 24 percentage points. This is quite a sizeable effect, if we compare it to the impact of other significant determinants of participation. Take for example education. Ceteris paribus, going from the status of high school dropout of high school graduate or higher increases the probability of being member of a group by about 13 percentage points. From the above estimate, moving in a community which is one standard deviation below the mean for inequality increases this probability by almost twice as much. A similar ratio of magnitudes holds if we compare the impact of inequality and of college education. Having more than 16 years of education raises the probability of participating by about 14 percentage points. Again, this is slightly more than half the impact of a decrease in the Gini coefficient by one standard deviation.

Racial and ethnic fragmentation also have sizeable, though lower, coefficients. An increase of one standard deviation in racial fragmentation (column 2) implies a reduction in the propensity to participate of about 8 percentage

[24] When we introduced them jointly, the results we obtained were broadly consistent but in some instances some of the three would lose statistical significance at standard levels of confidence.

[25] Specifically, we dropped those observations for which $abs(DFbeta) > 2/\sqrt{\#obs}$ (see e.g., Belsley et al. (1980), p.28).

points. A similar result (6 percentage points) holds for the ethnic fragmentation. This is not too high compared to the impact of education, but is still definitely relevant compared to other variables. For instance, we have seen that with children below 5 years in age participate significantly less in social activities, and in particular that, ceteris paribus, their probability of being members of a group is 3 percentage points lower than that of someone who does not have a small child. Well, living in a community that is one standard deviation above the mean in racial fragmentation reduces the probability of participating by almost three time as much. Similarly for ethnic fragmentation (two times as much).

Finally, it is worth noting that one of the categories included in our dependent variable is 'nationality groups'. One should expect that for this type of groups racial and especially ethnic fragmentation should not have a negative effect, but rather a *positive* effect (see the next section). We ran the same regressions of Table 3 excluding nationality groups, and the results on the effects of racial and ethnic fragmentation were in fact stronger. The estimated coefficients on *Race* and *Ethnic* when the dependent variable is membership in any group other than a nationality group are reported in columns 4 and 5 of Table 4. Compared to the estimates of Table 3 (columns 2 and 3, respectively), the impact of racial and ethnic heterogeneity is now quantitatively more important, as expected.[26]

We next consider the issue of potential endogeneity. A high degree of participation may reduce income inequality by increasing availability of information, options and opportunities. Also communities prone to social activities may be more favorable to redistributive policies. These problems are much less important for measure of ethnic or racial fragmentation, therefore in Table 5 we concentrate on instrumenting for Gini.

We consider two instruments: the number of municipal and township governments in 1962, and the share of the labor force employed in manufacturing. The number of governments in 1962 can be safely considered exogenous to participation in 1974-94 and it can have influenced the degree of income inequality in the MSA. Within a metropolitan area that was fragmented in many smaller jurisdictions it is more likely that significant differences in policies, local public good provision, and income levels persist among those jurisdictions. The share of manufacturing is not exogenous to

[26] Analogous results were obtained on the sample purged of influential observations (i.e. on the sample comparable to columns 2 and 3 of Table 4).

union participation, so when we use this instrument we check our results by excluding unions.

[Insert Table 5]

In column 1 of Table 5 we reproduce our OLS regression for comparison sake. In column 2 and 3 we instrument Gini with the number of governments in 1962, and with the same variable plus the share of manufacturing. Gini remains highly significant in both cases. Since union participation is certainly not exogenous to share of manufacturing, in columns 4 and 5 we exclude unions from our membership dependent variable. Once again in the IV specification Gini remains significant and the coefficient is higher in absolute value in the IV specification relative to the OLS one.

We should pause to analyze the fact that the coefficient on GINI in the IV regressions is larger in absolute value that in the OLS ones. Suppose that individuals who are more likely to participate in social activities are more favorable to redistribution, thus reducing inequality. This would imply an upward bias (in absolute value) of the estimated coefficient on Gini in the participation regression. An alternative argument is that individuals who are more prone to participate are those who are less averse to mixing with individuals with a different income level. This is in fact the basic idea of our model. But then individuals less averse to income heterogeneity may also be more prone to live in communities with more income heterogeneity. This would imply a downward bias (in absolute value) of the OLS coefficient on Gini. The patterns of the coefficient in Table 5 seems to support more the second interpretation, even though in the regression without unions the comparisons between coefficients and the Hausman test suggests that perhaps there is no bias up or down.

Note that a similar argument applies to Racial and Ethnic fragmentation. Suppose that individuals less averse to racial mixing move to more heterogenous communities. These individuals are also those who are more prone to participate in mixed groups. Thus, our OLS estimates imply a bias against finding an effect of racial fragmentation on participation and implies that the values of the OLS coefficients on Race and Ethnic may be biased downward, in absolute value.

5.3 Types of groups

The groups included in the GSS questionnaire are quite diverse, ranging from unions to literary clubs to church groups. It is therefore instructive to analyze participation in each of them separately to see if heterogeneity plays a different role in different types of groups. This is done in table 6.

[Insert table 6]

We have run the same regressions of Table 3, using as dependent variable individual membership in a given type of group. Each cell in the table refers to a separate regression and shows the marginal probit coefficient on the variable listed by column (*Gini*, *Race*, or *Ethnic*) for the type of group listed by row.[27] For each group we have excluded from the sample those respondents that for one reasons or another *cannot* be members of a particular group. For instance, individuals below a certain age cannot be member of a veteran group, non farmers cannot be a member of a farmer group, retirees cannot be members of a union, etc. The exact exclusion rules from the regression for each group are reported at the bottom of the table (the qualitative nature of our results is robust to the specification of these exclusions).

What patterns are we looking for? According to the basic ideas underlying our model, measures of heterogeneity should be less important for groups with a high degree of excludability and/or a low degree of close interaction amongst members than for groups with the opposite characteristics.

The results of Table 6 are broadly supportive of these hypotheses. Church groups are those with the strongest effect of all three types of heterogeneity. These are groups with virtually no excludability and a high degree of interaction. It is worth remembering also that Church groups are the largest ones in terms of number of participants.

At the opposite extreme we have professional associations and farmers groups, which have a very low level of personal interaction; the coefficients on *Gini*, *Race* and *Ethnic* are in fact insignificant in these regressions. Service groups, hobby clubs, sports clubs and youth clubs have a high degree of interaction and a less than perfect degree of excludability. Participation in

[27] All regressions include the individual and community controls listed in previous tables. We have also run the group by group regressions using the DFbeta method employed in Table 4. Results are qualitatively similar to those reported in Table 6 and are available upon request.

these groups tends to be influenced by some form of heterogeneity, even though not all the coefficients are significant in every regression.

Interestingly, and consistently with what one would expect, Ethnic fragmentation is positively associated with nationality groups. This observation suggests that in fragmented communities, individuals may feel more of a need to preserve and actively promote their own cultural identity and values, an observation broadly consistent with the spirit of our model.

Finally, notice that, aside from church groups, school service groups are the ones for which income inequality has the strongest negative impact. This is not surprising, given that high degrees of inequality are likely to be associated with marked heterogeneity in preferences for the type of education and services that schools should provide.

5.4 Individual preferences and participation

Our model implies that in heterogeneous and non segregated communities those who choose not to participate should be individuals who are more strongly averse to mixing with different types. This implication can be investigated, since the GSS asks several questions aimed at directly identifying individual preferences and attitudes toward racial mixing.

[Insert Table 7]

In Table 7 the dependent variable is membership in a subset of groups for which heterogeneity is important. More specifically, we choose all the groups for which at least one of the three measures of heterogeneity is negative and significant at the five per cent level in Table 6. These groups are: church, hobby, sport, youth, service, political, and school service groups.[28] We split the sample amongst individuals who can be identified as strongly averse to racial mixing on the basis of their answer to a given question, and individuals who are indifferent or favorable to racial mixing. Our model implies that, if we run the participation regression separately for the two sub-samples, *the coefficient on Race should be larger in absolute value for the sub-sample of individuals relatively more averse to racial mixing.*

The GSS includes many questions concerning attitudes towards race relations. Some of them are 'yes' or 'no' questions, others range on a scale of 1

[28]The results of table 7 are not unduly sensitive to the choice of groups . They are robust to focusing on a smaller or larger subset of groups.

to 3 or 1 to 4. In all cases, we have created a binary variable distinguishing 'averse' from 'non averse' people, and we have reported a rough description of the criterion (or question) in rows [1] to [7] of Table 7. Details on the definition of each binary variable, as well as the exact wording of the question in the GSS, are provided in appendix B. Out of the many questions on race relations contained in the GSS, we have chosen those for which we had a sufficiently high number of respondents on both sides.[29] Table 7 reports the marginal probit coefficients on the racial fragmentation variable from a participation regression whose sample is split among individuals who answered 'yes' to the question described in each row and individuals who answered 'no' (first and second column, respectively). Each cell of the table therefore refers to a separate regression which includes all our standard individual and community controls. The last column in the table simply reports the fraction of respondents on the 'yes' side of the question.

The estimates in Table 7 provide considerable support for this implication of our model. For six out of seven questions concerning attitudes toward race, the effect of racial heterogeneity is significantly stronger for the individuals more averse to racial mixing ad racial cooperation.

We start from what would seem the strongest measure of aversion to racial mixing, namely whether one "thinks there should be laws against marriages between blacks and whites" or not (line [1]). Twenty percent of almost 6,000 respondents answered 'yes' to this question. The negative effect of racial fragmentation in the community for these people is significantly stronger than for those who answered 'no' or 'I don't know'.

A particularly interesting question is the one about whether one has had a black person home for dinner in the past few years.[30] In fact, this is a question concerning individuals' actual *behavior*, as opposed to a test of a generic *attitude* toward race in an abstract sense. Therefore it is a better

[29]In practice this means that we have excluded several questions for which the yes answer had only about 200 observations. For these cases, the sizes of the coefficients on *Race* were consistent with our hypothesis, even though the scarcity of observations made the estimation unreliable.

[30]Due to the framing of the question, we restricted the sample to non-blacks only. However, in examining the GSS data we noted that a considerable number of black respondents had been ansked this question and had answered 'no'. We feel that black respondents may have interpreted this question as saying : "have you had a person of the opposite race home for dinner?". For this reason we also tried including black respondents in our regression, and our results were basically unchanged both in terms of significance and of quantitative values or the estimated coefficients.

measure of individuals' true attitude toward racial mixing and interracial direct contacts, which is the essence of our model. In fact, recent results by Glaeser and Laibson (1999) are quite supportive of this interpretation. They find that when individuals are asked in abstract they claim to have a large amount of 'trust' towards others, but their behavior in actual experiments is much less prone to trusting others. Our results based upon this behavioral question are very strongly in favor of our hypothesis, as shown in line [2]. The coefficient on *Race* is negative and highly significant in the sample of those who did not have a person of the opposite race for dinner and totally insignificant in the other sample.

We also get particularly strong results for the question concerning busing, (line [7]), a topic which has been hotly debated and has generated much racial conflict. The effect of racial fragmentation on the likelihood of participating in a group is negative and significant for those individuals who oppose busing, while it is insignificant and close to zero for those who do not. Analogous results are obtained when we consider respondents who would not vote for a black president (line [6]), who would object sending their children to a school where half of the children are of the opposite race, or who think that "blacks should not push themselves where they are not wanted".

In only one case the results are opposite to our hypothesis: the case of the question about the right to teach for people who "believe that blacks are genetically inferior". Note, however, that this question is more related to one's view about civil liberties than to personal interactions with members of the opposite race.[31]

Unfortunately we cannot perform similar tests on our two other measures of heterogeneity. The GSS does not include questions which can proxy for attitudes toward ethnicity. As for income heterogeneity, the survey ask a few questions on attitudes toward redistributive fiscal policy (e.g., whether the government should actively help the poor, or whether the federal income tax is too high), but nothing that may capture attitudes toward *interactions* with individuals of a different income or social level.

An additional set of variables which we considered are the measures of racial and income segregation constructed by Cutler, Glaeser, and Vigdor (1999). Our model does not deal directly with segregation, but one may ar-

[31] The GSS also includes a question concerning the right to *speak* for racist individuals. The answers to this question are very highly correlated with those to the question about the right to teach for racist individuals.

gue that in more segregated communities racial and income mixing is lower, and therefore even people relatively averse to heterogeneity may be willing to participate in groups. If this interpretation is correct, and if heterogeneity and segregation were positively correlated (as indeed they seem to be), omitting segregation from our regressions would bias our results *against* finding an effect of heterogeneity on participation. We have tested the effect of segregation and found results which are only weakly supportive of this interpretation. In particular, the various measures of segregation do not appear to have always significant coefficients, neither alone nor interacted with the heterogeneity variables.[32] A possible explanation is that several of the groups in the GSS questionnaire are generally city based, thus segregation may not be sufficient to prevent mixing, particularly in small cities.

6 Conclusions

Participation in social activities is positively associated with several valuable phenomena, like trust and human capital externalities. The propensity to participate is of course influenced to a large extent by individual characteristics, but it also depends on the composition and degree of heterogeneity of the community. In the theoretical part of this paper we show under which conditions more heterogeneity of the population leads to less social interaction. We then explore the evidence on US cities and find that income inequality and racial fragmentation are strongly inversely related to participation. Ethnic fragmentation also influences negatively participation, but perhaps less than racial fragmentation. The groups that are more affected by heterogeneity are those in which members directly interact to a significant extent, and in which excludability is low. Also, in accordance with our model, we find that the individuals who choose to participate less in racially mixed communities are those who most vocally oppose racial mixing.

This paper leaves open many avenues for research. For example, the theory for the formation of multiple groups in the same community is only sketched. A more fully developed model would lead to empirical predictions concerning how many groups of the same type one should find in the same community, as a function of the type of group and of the community characteristics. Also, one could focus on one specific group and analyze in more

[32] All these regressions are available upon request.

detail its actual evolution over time, composition of membership and internal functioning and relate that to the evolution of income inequality and racial/ethnic fragmentation in communities. This experiment would require data which go well beyond what is included in the GSS, such as historical data collected by specific group organizations.

References

[1] Alesina, A., R. Baqir and W. Easterly (1999), "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics,* forthcoming.

[2] Alesina, A., R. Baqir and W. Easterly (1997), "Redistributive Public Employment," NBER Working Paper.

[3] Alesina, A. and E. Spolaore (1997), "On the number and Size of Nations", *Quarterly Journal of Economics*, November.

[4] Banerjee A. (1992):"A simple model of herd behavior" *Quarterly Journal of Economics, 107, 797-818.*

[5] Belsley, D.A., E. Kuh, and R.E. Welsch (1980), *Regression Diagnostics*, New York: John Wiley and Sons.

[6] Benabou R. (1996):" Heterogeneity, Stratification and Growth" *American Economic Review*, 86, 584-609

[7] Bertrand M., E. Luttmer and S. Mullainathan (1998) "Network Effects and Welfare Cultures" NBER Working Paper n.6832 December

[8] Borjas, G. (1992). "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 107, 123-150.

[9] Borjas, G. (1995). "Ethnicity, Neighborhoods and Human Capital Externalities," *American Economic Review*, 85, 365-390.

[10] Coleman, J. (1990), *Foundations of Social Theory*, Cambridge, MA: Harvard University Press.

[11] Conley, J.P. and M. Wooders (1996),"Taste-Homogeneity of Optimal Jurisdictions in a Tiebout Economy with Crowding Types and Endogenous Educational Investment Choices", *Ricerche Economiche*, 50(4), 367-87.

[12] Cutler, D. and E. Glaeser (1997), "Are Ghettos Good or Bad?", *Quarterly Journal of Economics*, 112(3), 827-72.

[13] Cutler D. E. Glaeser and J. Vigdor (1999), "The rise and Decline of the American Ghetto", *Journal of Political Economy*, forthcoming.

[14] Davis, J.A. and T.W. Smith (1994), *General Social Surveys, 1972-1994*, Machine-readable data file, Chicago: National Opinion Research Center (producer), Storrs, CT: The Roper Center for Public Opinion Research (distributor).

[15] Di Pasquale, D. and E. Glaeser (1998), "Incentives and Social Capital: Are Homeowners Better Citizens?", NBER Working Paper n. 6363, January.

[16] Ellison D. and D. Fudenberg (1995), "Words of Mouth Communication and Social Learning", *Quarterly Journal of Economics,* 110(1), 93-126.

[17] Fernandez R. and R. Rogerson (1996), "Income Distribution, Communities, and the Quality of Public Education", *Quarterly Journal of Economics*, 111(1), 135-64.

[18] Glaeser, E., J. Scheinkman and A. Shleifer (1995). "Economic Growth in a Cross Section of Cities," *Journal of Monetary Economics*, 36, 117-143.

[19] Glaeser, E. and S. Glendon (1997). "The Demand for Religion," unpublished.

[20] Glaeser, E. and D. Laibson (1998). "What is Social Capital?", unpublished.

[21] Goldin, C. and L. Katz (1998). "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910 to 1940," NBER Working Paper 6439.

[22] La Ferrara, E. (1998). "Inequality and Participation: Theory and Evidence from Rural Tanzania," unpublished.

[23] La Porta, R., F Lopez de Silanes, A. Shleifer and R. Vishny (1997),"Trust in Large Organizations", *American Economic Review Papers and Proceedings*, 333-8.

[24] Lucas R. (1996): " On the mechanics of Economic Development" *Journal of Monetary Economics* 22, 3-42.

[25] Luttmer, E. (1997). "Group Loyalty and the Taste for Redistribution", unpublished.

[26] Poterba, J. (1996), "Demographic Structure and the Political Economy of Public Education", NBER Working Paper n. 5677, July.

[27] Putnam R. (1993) *Making democracy work* Princeton University Press, Princeton NJ

[28] Putnam, R. (1995a). "Bowling Alone: America's Declining Social Capital," *Journal of Democracy,* 6 (January), 65-78.

[29] Putnam, R. (1995b). "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America," *P.S.: Political Science and Politics,* 28 (December), 1-20.

[30] Putnam, R. and S. Yonish (1998), "New Evidence on Trends in American Social Capital and Civic Engagement: Are we Really Bowling Alone?", unpublished.

[31] Romer P. (1986) "Increasing returns and long run growth" *Journal of Political Economy* 94, 1002-37

[32] Uslaner, E. (1997). "Volunteering and Social Capital: How Trust and Religion Shape Civic Participation in the United States," unpublished.

[33] Verba, S. and N.H. Nie (1987), *Participation in America. Political Democracy and Social Equality,* Chicago: The University of Chicago Press.

[34] Verba, S., Schlozman, K. and H. Brady (1995). *Voice and Equality: Civic Volunteerism in American Politics,* Cambridge, MA: Harvard University Press.

[35] Wilson, W. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy,* Chicago: University of Chicago Press.

[36] Wilson, W. (1996). *When Work Disappears: The World of the New Urban Poor.* New York: Knopf (distributed by Random House, Inc.).

Appendix A

Proof of Proposition 1

Proposition 1 can be proved by applying Brouwer's fixed point theorem to the function $f(P_B)$ defined by the right-hand side of (9).[33] Notice first of all that $f(P_B)$ maps the interval $[0,1]$ into itself, and that $[0,1]$ is clearly a non empty, compact, and convex set. All we need to show is therefore that the right-hand side of (9) is a function, and that it is continuous. This follows from our assumptions that $F(\cdot)$ is differentiable, $u(\cdot)$ is well-behaved (hence $g(\cdot)$ is continuously differentiable in P_B), and $B, W > 0$. $\square$

Proof of Remark 1

The derivative of the right hand side of (9) with respect to P_B is equal to

$$-\frac{F'[g(1-P_B)]g'(1-P_B)F[g(P_B)] + F'[g(P_B)]g'(P_B)F[g(1-P_B)]}{\{F[g(1-P_B)] + F[g(P_B)](W/B)\}^2}\frac{W}{B} \tag{15}$$

Notice that since $F'(\cdot) > 0$ and $g'(\cdot) < 0$, expression (15) is unambiguously positive. The stability condition (10) requires that (15) is less than 1.

Proof of Lemma 1

The proof of Lemma 1 can be divided in two parts.

(a) $B < W \Longrightarrow P_B^* < P_W^*$.

Apply the implicit function theorem to (9) to get

$$\frac{\partial P_B^*}{\partial(W/B)} = \left\{ \begin{array}{c} \left[-\frac{F'[g(1-P_B^*)]g'(1-P_B^*)}{F[g(1-P_B^*)]} - \frac{F'[g(P_B^*)]g'(P_B^*)}{F[g(P_B^*)]}\right] \cdot \frac{W}{B} - \\ - \left[\frac{F[g(1-P_B^*)]}{F[g(P_B^*)]} + \frac{F[g(P_B^*)]}{F[g(1-P_B^*)]}\left(\frac{W}{B}\right)^2 + 2\frac{W}{B}\right] \end{array} \right\}^{-1} \tag{a.1}$$

Under the stability condition (10) this derivative is unambiguously negative. Note that by symmetry when $B = W$ the unique equilibrium must be $P_B^* = \frac{1}{2} = P_W^*$. Condition (a) follows from these two facts.

[33] Brouwer's fixed point theorem says that if $S \subset \Re^N$ is a nonempty, compact, convex set, and $f : S \to S$ is a continuous function from S into itself, then $f(\cdot)$ has a fixed point.

(b) $P_B^* < P_W^* \Longrightarrow B < W$.

By contradiction. Suppose that $B > W$. Then by the same arguments as in part (a) we should have $P_B^* > P_W^*$, which contradicts the hypothesis. □

Proof of Corollary 1

Let us start by showing that $B < W \Rightarrow \frac{P_B^*}{P_W^*} < \frac{B}{W}$. From (7) we can write the ratio of the two proportions as

$$\frac{P_B^*}{P_W^*} = \frac{F\left[g(P_W^*)\right]}{F\left[g(P_B^*)\right]}\frac{B}{W}.$$

>From Lemma 1, $B < W$ implies $P_W^* > P_B^*$. Given that $F'(\cdot) > 0$ and $g'(\cdot) < 0$, this in turn implies $F\left[g(P_W^*)\right] < .F\left[g(P_B^*)\right]$, which proves the first part of the corollary. The second part, namely $B > W \Rightarrow \frac{P_B^*}{P_W^*} > \frac{B}{W}$, can be proved with the same arguments. □

Proof of Corollary 2

Follows from Lemma 1 together with definitions (5) and (6), recalling that $g'(\cdot) < 0$. □

Proof of Proposition 2

>From (11) we obtain

$$\begin{aligned} \frac{dS}{dw} &= \{F\left[g(P_B^*)\right] - F\left[g(1-P_B^*)\right]\} + \frac{\partial P_B^*}{\partial w}\{wF'\left[g(P_B^*)\right]g'(P_B^*) - \\ &\quad -(1-w)F'\left[g(1-P_B^*)\right]g'(1-P_B^*)\} \end{aligned} \tag{a.2}$$

We need to find the conditions under which (a.2) is negative. Let us start by showing that $w < 1/2$ implies $dS/dw < 0$. As proven in lemma 1, when $w < 1/2$ we have $P_B^* > P_W^*$. Given that $F'(\cdot) > 0$ and $g'(\cdot) < 0$, the expression in the first curly brackets in (a.2) is thus negative. As for the second part of the derivative, we know that in a stable equilibrium $\partial P_B^*/\partial w < 0$, so it is sufficient (but not necessary) to show that the expression in the second curly brackets is positive to prove our result. This amounts to requiring that

$$\frac{w}{1-w} < \frac{F'\left[g(1-P_B^*)\right]g'(1-P_B^*)}{F'\left[g(P_B^*)\right]g'(P_B^*)} \tag{a.3}$$

Notice that the left hand side of (a.3) is less than one by assumption. Therefore a sufficient, though not necessary, condition for the above inequality to hold is that $F''(\cdot) \geq 0$ and $g''(\cdot) \geq 0$. Intuitively, $F''(\cdot) \geq 0$ says that the negative effect of ethnic fragmentation on participation will likely be observed when the distribution of the α's is uniform or skewed to the right, i.e. when there is a significant fraction of the population who dislikes interaction with the opposite race. On the other hand, $g''(\cdot) \geq 0$ describes a situation in which the fraction of people whose utility exceeds the reservation level decreases relatively more at low levels of the proportion of the opposite type in the group.[34] Note, however, that even when $F''(\cdot) < 0$ and/or $g''(\cdot) < 0$, i.e. when most individuals have mild preferences on racial relations, it is still possible that $dS/dw < 0$ because all other effects work in this direction. The conditions $F''(\cdot) \geq 0$ and $g''(\cdot) \geq 0$ are in fact 'twice sufficient': the first time because they are sufficient but not necessary for (a.3) to hold; the second time because (a.3) is sufficient but not necessary for (a.2) to be satisfied.

The second half of the proposition, namely the fact that $dS/dw < 0$ implies $w < 1/2$ under the above sufficient conditions, can be proved by contradiction taking into account the first part of the proof. □

Appendix B

Variable definition

The following is a list of the variables we use and of their sources, followed by summary statistics. The data sources are abbreviated as follows: GSS stands for 'General Social Survey, cumulative file 1972-94'; CensusCD90 refers to the CDrom "CensusCD+Maps" by GeoLytics, Inc. (1996-98) which contains data from the Summary Tape Files 3F of the 1990 Census. In all cases from variables constructed from the GSS, 'no answer' and 'not applicable' were coded as missing values.

Member: dummy equal to 1 if respondent is member of at least one group. [Source: authors' calculation on GSS]

Member (excl. nationality): dummy equal to 1 if respondent is member of at least one group other than a nationality group [Source: authors'

[34] This condition is much less restrictive than it looks. In fact it is *always* satisfied when α and P enter multiplicatively in the utility function, whatever the exact functional form. See examples 1 and 2 in the text for an illustration.

calculation on GSS]

Member (excl. unions): dummy equal to 1 if respondent is member of at least one group other than a union [Source: authors' calculation on GSS]

Cohort: year of birth of the respondent [Source: GSS]

Age<30: dummy equal to 1 if respondent is less than 30 years old [Source: authors' calculation on GSS]

Age30-39: dummy equal to 1 if respondent is between 30 and 39 years old [Source: authors' calculation on GSS]

Age50-59: dummy equal to 1 if respondent is between 50 and 59 years old [Source: authors' calculation on GSS]

Age≥60: dummy equal to 1 if respondent is 60 years old or more [Source: authors' calculation on GSS]

Married: dummy equal to 1 if respondent is married [Source: authors' calculation on GSS]

Female: dummy equal to 1 if respondent is female [Source: authors' calculation on GSS]

Black: dummy equal to 1 if respondent is African American [Source: authors' calculation on GSS]

: dummy equal to 1 if respondent is [Source: authors' calculation on GSS]

Educ<12 yrs: dummy equal to 1 if respondent has less than 12 years of education [Source: authors' calculation on GSS]

Educ>16 yrs: dummy equal to 1 if respondent has more than 16 years of education [Source: authors' calculation on GSS]

Children≤5 yrs: dummy equal to 1 if respondent has children age 5 or less [Source: authors' calculation on GSS]

Children 6-12: dummy equal to 1 if respondent has children age 6 to 12 [Source: authors' calculation on GSS]

Children13-17: dummy equal to 1 if respondent has children age 13 to 17 [Source: authors' calculation on GSS]

ln(real income): logarithm of respondent's family income (constant 1986 US$) [Source: authors' calculation on GSS]

Full-time: dummy equal to 1 if respondent works full time [Source: authors' calculation on GSS]

Partime: dummy equal to 1 if respondent works part time [Source: authors' calculation on GSS]

Size of place: logarithm of the size of place where respondent lives (thousands of people) [Source: authors' calculation on GSS]

Med HH income: logarithm of median household income in MSA/PMSA where respondent lives [Source: authors' calculation on CensusCD90]

Med HH income: square of the logarithm of median household income in MSA/PMSA where respondent lives [Source: authors' calculation on CensusCD90]

Gini: Gini coefficient on family income in MSA/PMSA where respondent lives. Actual Gini coefficients were computed for the years 1970, 1980, 1990. The values for the remaining years in the sample were obtained by linear interpolation (and extrapolation for 1991-94). [Source: authors' calculation on IPUMS 1%, Census 1970, 1980, 1990]

Race: racial fragmentation index in MSA/PMSA where respondent lives, defined in expression (14) in the text. The five categories used for the shares are the original Census categories: i) white; ii) black; iii) American Indian, Eskimo, Aleutian; iv) Asian, Pacific Islander; v) other. [Source: authors' calculation on CensusCD90]

Ethnic: ethnic fragmentation index in MSA/PMSA where respondent lives, defined in expression (14) in the text. The 10 categories used for the shares are obtained aggregating the original 'first ancestries' from the Census as follows: (1) Arab; (2) Subsaharan African; (3) West Indian; (4) Race or Hispanic origin; (5) Canadian, United States or American; (6) Austrian, Belgian, Dutch, English, French Canadian, German, Irish, Scotch-Irish, Scottish, Swiss, Welsh; (7) Czech, Hungarian, Lithuanian, Polish, Romanian, Russian, Slovak, Ukrainian, Yugoslavian; (8) French, Greek, Italian, Portuguese; (9) Danish, Finnish, Norwegian, Swedish; (10) other. Each share is computed a share of people in that category over the total population in the MSA/PMSA (excluding people with 'ancestry unclassified' and 'ancestry not reported'). [Source: authors' calculation on CensusCD90]

NGOV62: number of municipal and township governments in the MSA/PMSA in 1962 [Source: Cutler, Glaeser and Vigdor (1999)]

MANSHR: share of the labor force employed in manufacturing in the MSA/PMSA in 1990 [Source: Cutler, Glaeser and Vigdor (1999)]

NOMIXMARRIAGE: dummy equal to 1 if respondent is against mixed marriages. Original GSS survey question: "Do you think there should be laws against marriages between blacks and whites?". Prompted answers coded in the GSS variable 'RACMAR': 1=Yes; 2=No; 3=Don't know. Our variable takes value 1 if RACMAR=1 and zero otherwise. [Source: authors' calculation on GSS]

NOBLKDINNER: dummy equal to 1 if respondent has not had a black person home for dinner in past few years. Original GSS survey question: "During the last few years, has anyone in your family brought a friend who was a black home for dinner?". Prompted answers coded in the GSS variable 'RACHOME': 1=Yes; 2=No; 8=Don't know; 9=No answer. Our variable takes value 1 if RACHOME=2 and zero otherwise. [Source: authors' calculation on GSS]

BLKNOPUSH: dummy equal to 1 if respondent thinks that blacks should not push. Original GSS survey question: "Here are some opinions other people have expressed in connection with black-white relations. Which statement on the card comes closest to how you, yourself, feel? The first one is: Blacks shouldn't push themselves where they're not wanted". Prompted answers coded in the GSS variable 'RACPUSH': 1=Agree strongly; 2=Agree slightly; 3=Disagree slightly; 4=Disagree strongly; 8=No opinion; 9=No answer. Our variable takes value 1 if RACPUSH=1 and zero otherwise. [Source: authors' calculation on GSS]

NOHALFSCHOOL: dummy equal to 1 if respondent would not send children to school with half opposite race. Original GSS survey question: "Would you yourself have any objection to sending your children to a school where half of the children are Whites / Blacks?". Prompted answers coded in the GSS variable 'RACHAF': 1=Yes; 2=No; 3=Don't know. Our variable takes value 1 if RACMAR=1 and zero otherwise. [Source: authors' calculation on GSS]

RACTEACH: dummy equal to 1 if respondent thinks that racists should be allowed to teach. Original GSS survey question: "Consider a person who believes that blacks are genetically inferior. Should such a person be allowed to teach in a college or university, or not?". Prompted answers coded in the GSS variable 'COLRAC': 4=Yes, allowed; 5=Not allowed; 8=Don't know; 9=No answer. Our variable takes value 1 if COLRAC=4 and zero otherwise. [Source: authors' calculation on GSS]

NOBLKPRESID: dummy equal to 1 if respondent would not vote for black president. Original GSS survey question: "If your party nominated a Black for President, would you vote for him if he were qualified for the job?". Prompted answers coded in the GSS variable 'RACPRES': 1=Yes; 2=No; 8=Don't know; 9=No answer. Our variable takes value 1 if RACPRES=2 and zero otherwise. [Source: authors' calculation on GSS]

NOBUSING: dummy equal to 1 if respondent opposes busing. Original GSS survey question: "In general, do you favor or oppose the busing of black and white school children from one school district to another?". Prompted answers coded in the GSS variable 'BUSING': 1=Favor; 2=Oppose; 8=Don't know; 9=No answer. Our variable takes value 1 if BUSING=2 and zero otherwise. [Source: authors' calculation on GSS]

Table B1: Summary statistics

	Mean	*Std. Dev.*	*No. obs.*
Member	.72	.45	10031
Member (excl. nationality)	.71	.45	9922
Member (excl. unions)			9922
Cohort	1939.85	17.59	10031
Age<30	.25	.43	10031
Age30-39	.24	.43	10031
Age50-59	.13	.34	10031
Age≥60	.20	.40	10031
Married	.56	.50	10031
Female	.56	.50	10031
Black	.16	.37	10031
Educ<12 yrs	.24	.43	10031
Educ>16 yrs	.22	.41	10031
Children≤5 yrs	.19	.39	10031
Children 6-12	.22	.41	10031
Children 13-17	.18	.38	10031
ln(real income)	10.03	.93	10031
Fulltime	.52	.50	10031
Partime	.10	.30	10031
Size of place	4.28	2.18	10031
Med HH income	10.38	.14	10031
Med HH inc^2	107.74	3.0	10031
Gini	.41	.03	10031
Race	.36	.14	10031
Ethnic	.67	.07	10031
NGOV62	89.84	90.78	10031
MANSHR	.17	.05	10031
NOMIXMARRIAGE	.19	.40	5901
NOBLKDINNER	.68	.47	4493
BLKNOPUSH	.32	.47	2885
NOHALFSCHOOL	.19	.39	5717
RACTEACH	.44	.50	4661
NOBLKPRESID	.12	.33	5872
NOBUSING	.73	.44	6303

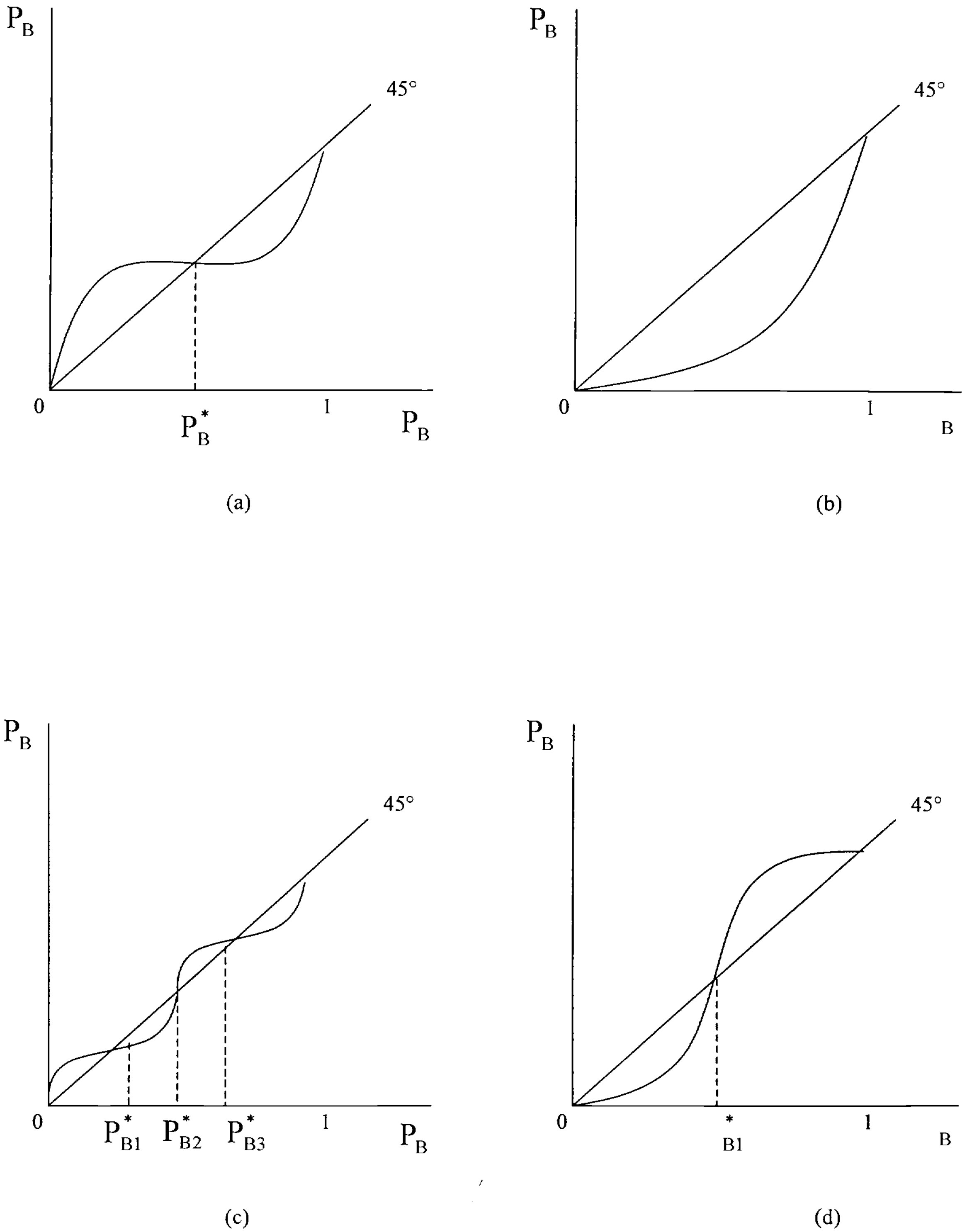

Figure 1: Equilibrium configurations

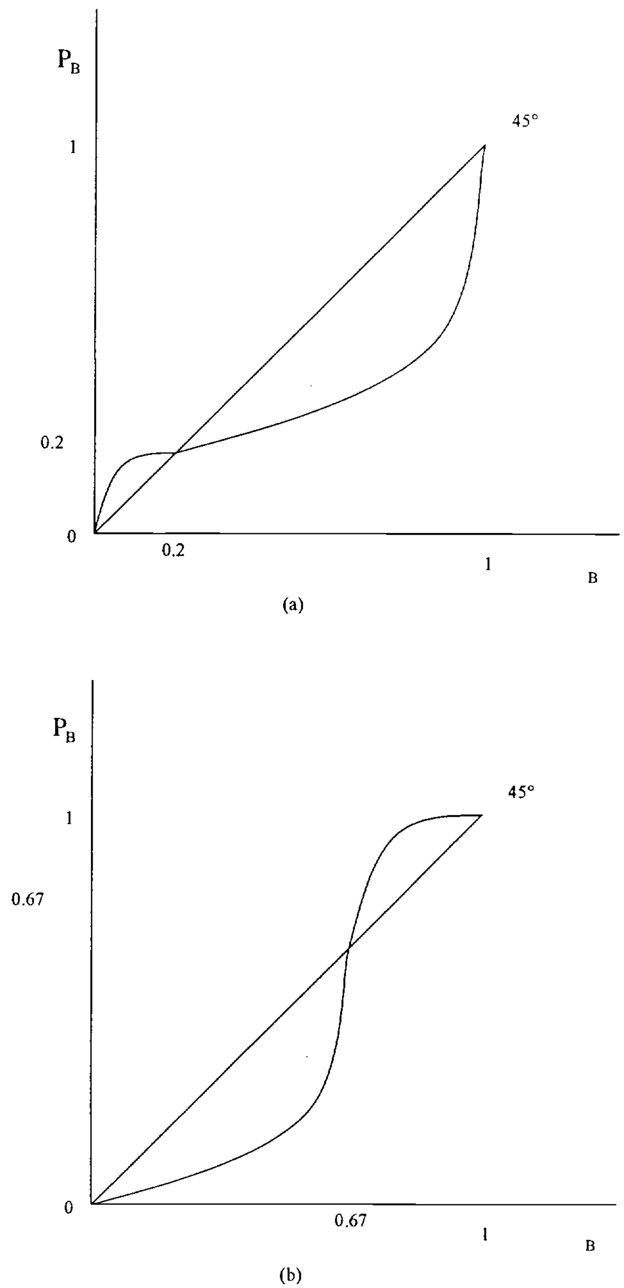

Figure 2: Examples

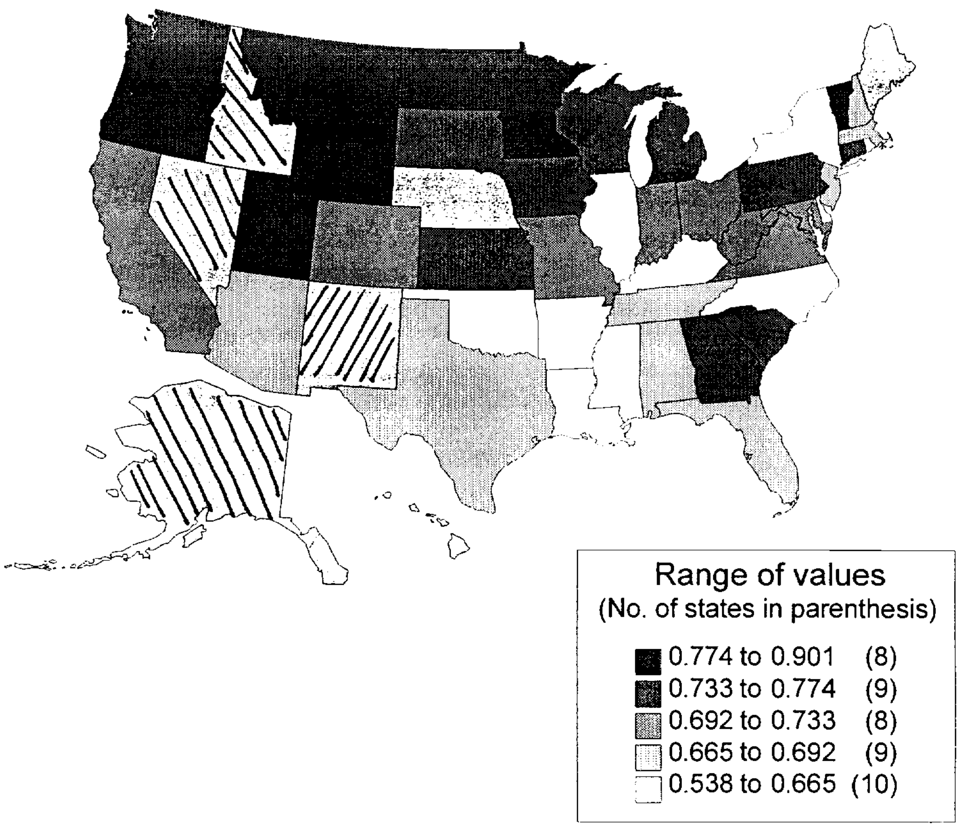

Figure 3: Average membership rate, 1973-94

//// = missing

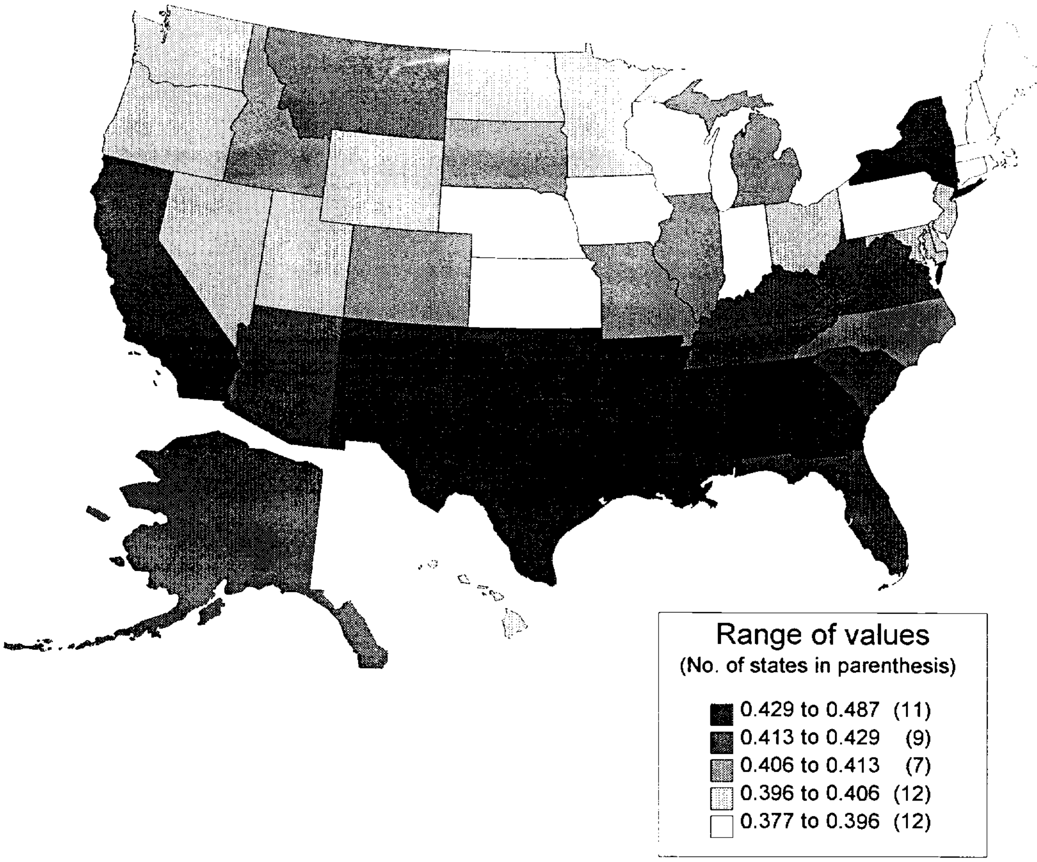

Figure 4: Gini coefficient, 1972-94

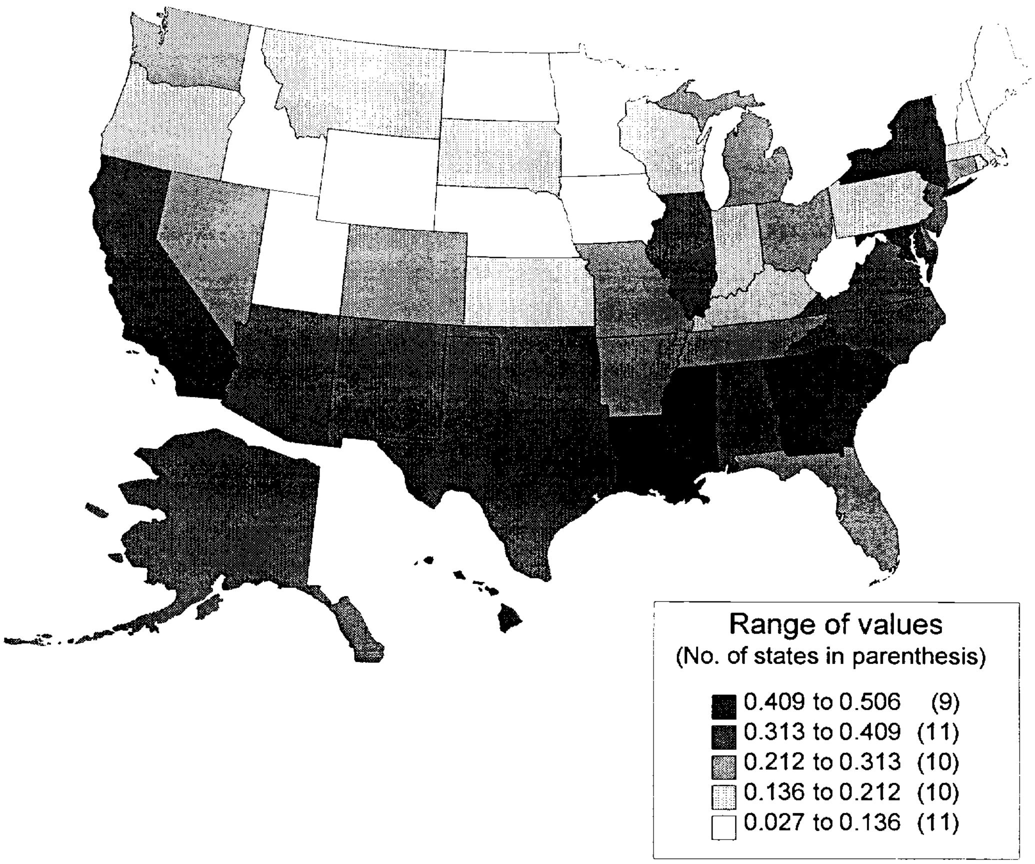

Figure 5: Racial fragmentation, 1990

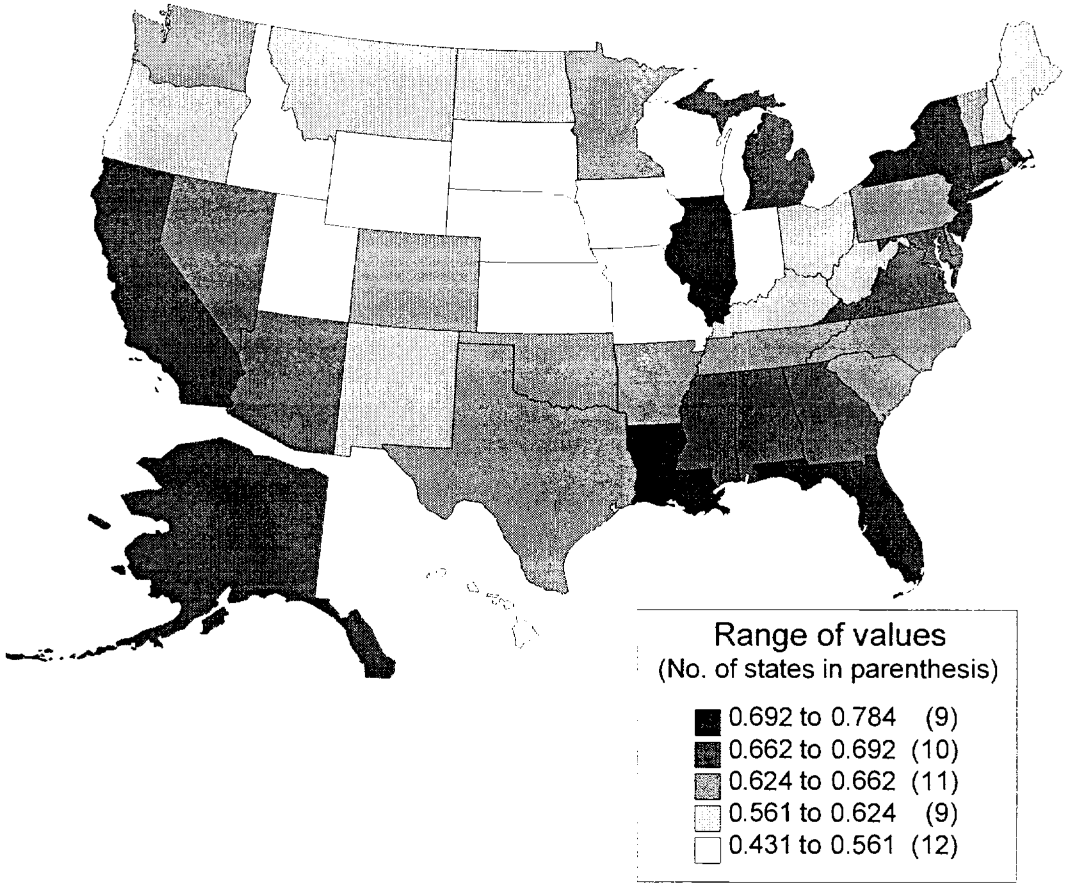

Figure 6: Ethnic fragmentation, 1990

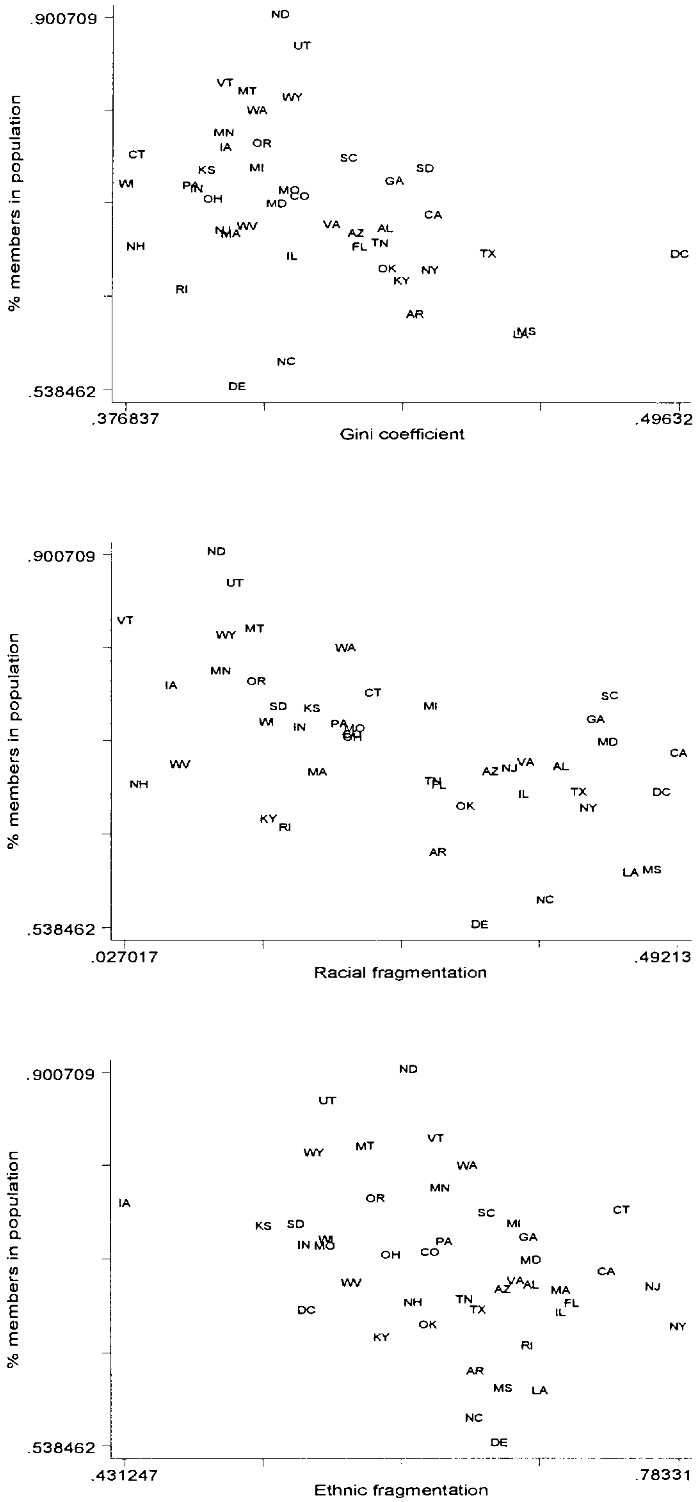

Figure 7: Heterogeneity and participation in groups

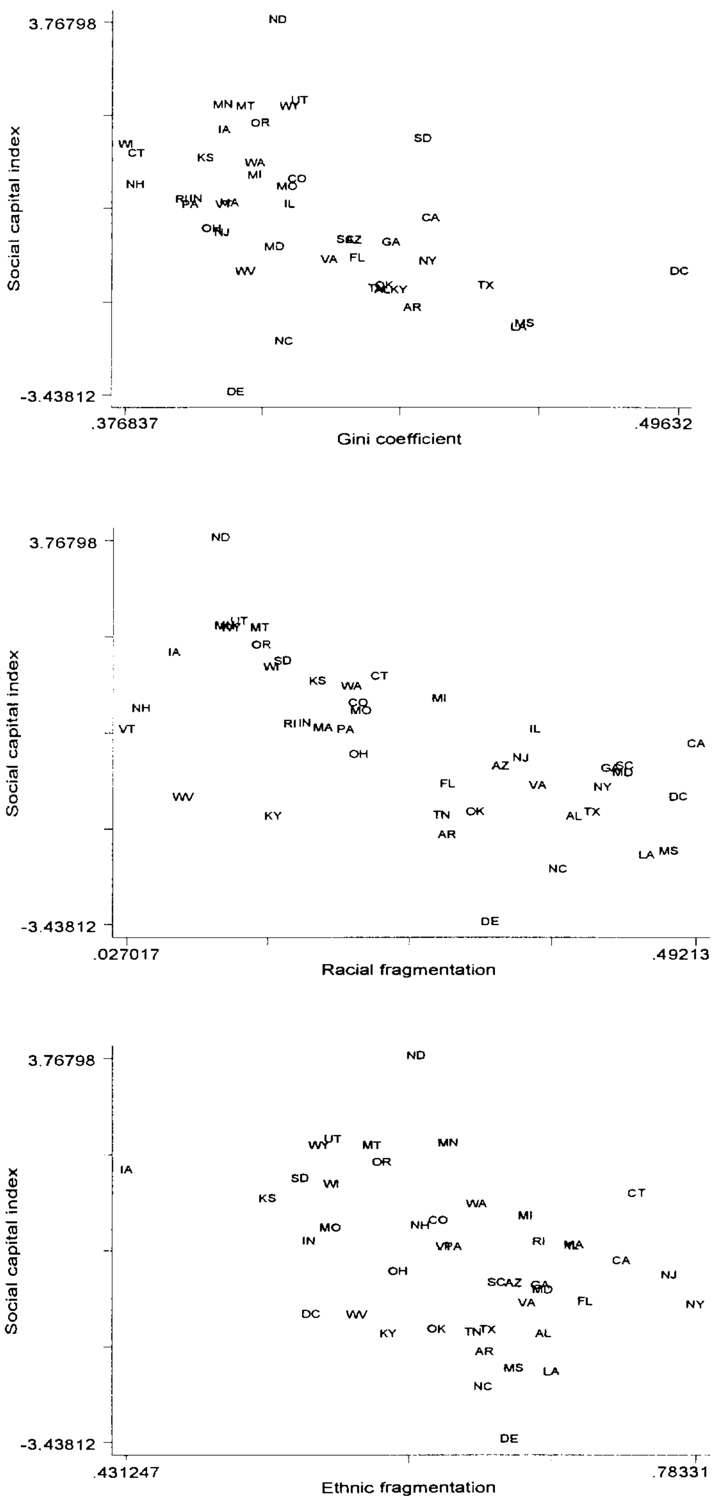

Figure 8: Heterogeneity and social capital

Table 1: Descriptive statistics

Panel A: Summary statistics

	Mean	*Std. Dev.*
Member of any group	.72	.45
Member of church group	.35	.48
Member of fraternity	.10	.29
Member of service group	.10	.30
Member of hobby club	.10	.30
Member of sport club	.21	.41
Member of youth group	.10	.30
Member of literary group	.11	.31
Member of school service group	.14	.35
Member of school fraternity	.06	.23
Member of veterans' group	.06	.24
Member of political group	.05	.21
Member of nationality group	.04	.20
Member of union	.15	.36
Member of professional association	.17	.37
Member of farmers' group	.02	.14
Member of other group	.11	.31
Gini	.41	.03
Racial fragmentation ('Race')	.36	.14
Ethnic fragmentation ('Ethnic')	.67	.07

Panel B: Correlations

	Membership[a]	Gini	Race
Gini	-.23*		
Race	-.03	.43**	
Ethnic	.06	.13	.56**

Notes:

* denotes significance at the 5 percent level, ** at the 1 percent level.

(a) Average membership rate in the MSA.

Table 2: Individual determinants of participation

	Marg. Probit coeff.$^{(a)}$	*Std. error*^{b}
Cohort	-.002*	(.001)
Age<30	-.032	(.028)
Age30-39	-.028*	(.017)
Age50-59	.010	(.016)
Age≥60	.041	(.028)
Married	-.001	(.011)
Female	-.042**	(.011)
Black	.044**	(.011)
Educ<12 yrs	-.125**	(.013)
Educ>16 yrs	.145**	(.011)
Children ≤5 yrs	-.033**	(.014)
Children 6-12	.074**	(.012)
Children 13-17	.000	(.015)
ln(real income)	.075**	(.006)
Fulltime	.027**	(.012)
Partime	.062**	(.017)
STATES	Yes	
YEARS	Yes	
No. obs.	10031	
Pseudo Rsq	.09	
Observed P	.72	
Predicted P	.74	

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

(a) Marginal probit coefficients calculated at the means.

(b) Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

Table 3: Heterogeneity and participation

	[1]	[2]	[3]	[4]	[5]	[6]
Size of place	-.002	-.001	-.003	-.001	-.002	-.001
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Med HH income	3.441**	5.002**	3.861**	4.404**	3.401*	4.195**
	(1.668)	(1.858)	(1.858)	(1.848)	(1.816)	(1.915)
Med HH inc. ^2	-.176**	-.248**	-.193**	-.221**	-.173**	-.211**
	(.079)	(.088)	(.089)	(.088)	(.087)	(.091)
Gini	-1.017**			-.503**	-.870**	-.533**
	(.247)			(.254)	(.231)	(.268)
Race		-.220**		-.152**		-.122*
		(.058)		(.069)		(.072)
Ethnic			-.281**		-.178*	-.087
			(.118)		(.101)	(.096)
INDIV CONTROLS$^{(a)}$	Yes	Yes	Yes	Yes	Yes	Yes
STATES	Yes	Yes	Yes	Yes	Yes	Yes
YEARS	Yes	Yes	Yes	Yes	Yes	Yes
No. obs.	10031	10031	10031	10031	10031	10031
Pseudo Rsq	.09	.09	.09	.09	.09	.09
Observed P	.72	.72	.72	.72	.72	.72
Predicted P	.74	.74	.74	.74	.74	.74

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

Marginal probit coefficients calculated at the means. Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

(a) Individual controls: all those listed in Table 2.

Table 4: Sensitivity analysis

	Dependent variable: Member *Excl. influential observations*$^{(b)}$			Dependent variable: Member (excl. nationality groups)	
	[1]	[2]	[3]	[4]	[5]
Gini	-2.257**				
	(.813)				
Race		-.514**		-.233**	
		(.156)		(.057)	
Ethnic			-.806**		-.321**
			.308		(.117)
INDIV CONTROLS$^{(a)}$	Yes	Yes	Yes	Yes	Yes
STATES	Yes	Yes	Yes	Yes	Yes
YEARS	Yes	Yes	Yes	Yes	Yes
No. obs.	9464	9414	9503	9922	9922
Pseudo Rsq	.11	.13	.12	.09	.09
Observed P	.74	.75	.74	.72	.72
Predicted P	.77	.79	.77	.74	.74

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

Marginal probit coefficients calculated at the means. Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

(a) Individual controls: all those listed in Table 2.

(b) Influential observations identified by predicting DFbetas for the relevant variable from the full sample regression and then dropping those observations for which $abs(DFbeta) > 2/\sqrt{\#obs}$.

Table 5: instrumenting Gini

	Dependent variable: Member			Dependent variable: Member (excl. unions)	
	OLS	*2SLS*		*OLS*	*2SLS*
		Instrument set:			*Instrument set:*
		NGOV62	NGOV62 MANSHR		NGOV62 MANSHR
	[1]	[2]	[3]	[4]	[5]
Gini	-.977**	-2.954**	-3.114**	-1.127**	-1.886**
	(.245)	(1.194)	(1.122)	(.241)	(1.083)
INDIV CONTROLS[(a)]	Yes	Yes	Yes	Yes	Yes
STATES	Yes	Yes	Yes	Yes	Yes
YEARS	Yes	Yes	Yes	Yes	Yes
No. obs.	10031	10031	10031	9941	9941
R sq.	.11	.10	.10	.11	.11
Hausman (p-value)		.09	.01		.48
Sargan (p-value)			.81		.75

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

Marginal probit coefficients calculated at the means. Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

(a) Individual controls: all those listed in Table 2.

Table 6: Participation by type of group[a]

Dependent variable is Membership in	*Marginal Probit coefficients on:* Gini	Race	Ethnic
Church groups	-1.174**	-.185**	-.346**
	(.221)	(.044)	(.103)
Fraternities	-.324*	-.046	-.041
	(.188)	(.031)	(.087)
Service groups[b]	-.363**	-.049*	-.046
	(.139)	(.028)	(.054)
Hobby clubs	-.542**	-.079**	.018
	(.135)	(.030)	(.059)
Sport clubs	-.826**	-.126*	-.060
	(.310)	(.067)	(.087)
Youth groups[c]	-.530**	-.062	-.023
	(.251)	(.044)	(.064)
Literary groups[b]	-.071	-.025	-.039
	(.214)	(.051)	(.079)
School service groups[d]	-1.019**	-.101	-.150
	(.410)	(.069)	(.141)
School fraternities[b]	-.202*	-.007	-.043
	(.119)	(.031)	(.049)
Veterans' groups[e]	-.134	-.027	-.045
	(.106)	(.024)	(.045)
Political groups	-.313**	-.030	-.018
	(.122)	(.025)	(.054)
Nationality groups	.047	.017	.071*
	(.092)	(.016)	(.039)
Unions[f]	-.076	-.021	.078
	(.247)	(.050)	(.082)
Professional associations[g]	.225	.156	.331*
	(.429)	(.103)	(.174)
Farmers' groups[h]	-1.685	.365	.184
	(5.770)	(.547)	(1.139)
Other groups	.061	-.044	-.050
	(.230)	(.039)	(.073)

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

Marginal probit coefficients calculated at the means. Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

Each cell reports the marginal probit coefficient on the variable listed in the column heading from a regression in which the dependent variable is membership in the type of group described in the row heading. All regressions include the individual controls listed in Table 2, state, and year dummies.

(a) The sample for each regression is restricted to those individuals who can potentially be members of that particular group.

(b) Sample includes individuals with at least 12 years of education.

(c) Sample includes individuals younger than 50.

(d) Sample includes individuals with children age 6 to 17.

(e) Sample includes cohorts 1920 to 1955.

(f) Sample includes production, clerical, sales, and service workers.

(g) Sample includes professional workers.

(h) Sample includes only workers whose occupation code corresponds to Agriculture. Due to the small size of the sample, state and year dummies are omitted from these regressions.

Table 7: **Participation and aversion to racial mixing**

Dependent variable: Member		*Probit coeff. on Race for those who answer* Yes	*No*	Fraction of Yes
[1]	Would favor a law against mixed marriages (NOMIXMARRIAGE)	-.575** (.226)	-.401** (.074)	.20
[2]	Have *not* had black home for dinner[(b)] in last few years (NOBLKDINNER)	-.545** (.078)	-.127 (.152)	.68
[3]	Think that blacks should not push[(a)] (BLKNOPUSH)	-.458* (.257)	-.248** (.094)	.32
[4]	Oppose your children going to school with half opposite race (NOHALFSCHOOL)	-.460** (.148)	-.310** (.099)	.19
[5]	Racist has right to teach (RACTEACH)	-.254** (.101)	-.377** (.102)	.44
[6]	Would oppose black president[(b)] (NOBLKPRESID)	-.451** (.224)	-.358** (.096)	.14
[7]	Oppose busing (NOBUSING)	-.433** (.086)	-.148 (.181)	.73

Notes:

* denotes significance at the 10 percent level, ** at the 5 percent level.

Marginal probit coefficients calculated at the means. Standard errors corrected for heteroskedasticity and clustering of the residuals at the MSA level.

All regressions include the individual controls listed in Table 2, state, and year dummies.

(a) The results do not change significantly if the sample is restricted to whites only.

(b) Sample includes non-blacks only.