CENTRO STUDI LUCA D'AGLIANO

www.dagliano.uni-bocconi.it

DEVELOPMENT STUDIES WORKING PAPERS

Inequality and Group Participation: Theory and Evidence from Rural Tanzania

Eliana La Ferrara

N. 138 June 2000

http://www2.qeh.ox.ac.uk/

Inequality and Group Participation: Theory and Evidence from Rural Tanzania

Eliana La Ferrara[1]

Bocconi University and IGIER

First version: April 1998
This version: December 1999

[1]I am grateful to Abhijit Banerjee and Alberto Alesina for insightful discussions and suggestions. Useful comments were also provided by Philippe Aghion, Oriana Bandiera, Tim Besley, Francois Bourguignon, Micael Castanheira, Jean Drèze, Emanuela Galasso, Caroline Minter Hoxby, Massimiliano Marcellino, Eric Maskin, Jonathan Morduch, Debraj Ray, Guido Tabellini, and by seminar participants at Harvard University, 1998 NEUDC Conference at Yale, Bocconi University, University of Namur, LSE, and The World Bank. I thank Deepa Narayan and Lant Pritchett for providing the Tanzania Social Capital and Poverty Survey data. The usual disclaimer applies. Correspondence: IGIER, Universita' Bocconi. Via Salasco 5, 20136 Milano, Italy. E-mail: eliana.laferrara@uni-bocconi.it

Abstract

This paper investigates the determinants of group membership, and in particular the effect of income inequality on individual incentives to join economic groups. Drawing on a simple model, we show that an increase in inequality has an ambiguous effect and that the type of access rule (open versus restricted access) is key in determining what income categories are represented in the group. Furthermore, the shape of the income distribution can be crucial to determine whether increased inequality leads to more or less group participation. Using survey data from rural Tanzania we find that inequality at the village level has a negative impact on the likelihood that the respondents are members of any group. This effect is particularly significant for relatively wealthier people, both when relative wealth is 'objectively' measured, and when it is 'subjectively' defined. However, when we disaggregate groups by type of access rule, we find that inequality can have a positive impact on participation, depending on the shape of the income distribution. Finally, we assess the impact of inequality on various dimensions of group functioning.

1 Introduction

This paper investigates the determinants of participation in groups, and in particular the effect of heterogeneity on individual incentives to join groups that provide some shared economic benefits. The particular form of heterogeneity considered is income inequality. We address such questions as: Is group participation higher or lower in more unequal communities? When inequality increases, is it the 'poor' or the 'rich' who drop out of the groups? Does this depend on the type of access rule to the group? How are individual incentives affected by the *shape* of the distribution of income (i.e. by the fact that heterogeneity may be concentrated in the lower or in the upper quintiles of the distribution)?

These questions are relevant for a variety of reasons. First, there has been growing attention on group formation and participation, as the latter is recognized as an essential component of 'social capital'. To the extent that social capital has nonnegligible effects on economic performance[1], it becomes important to understand the *determinants* of participation in groups, and the role of members' characteristics in shaping economic incentives within a group. Surprisingly, no extensive work has been done on this topic. Empirical investigations on the determinants of participation in socio-political groups have been conducted on a sample of US cities by Di Pasquale and Glaeser (1998), and by Alesina and La Ferrara (1999). Di Pasquale and Glaeser find that social capital is positively associated with home ownership, while Alesina and La Ferrara find that individuals are less likely to join groups in unequal and racially fragmented communities because they prefer to interact with people similar to themselves. This paper differs from the above studies in that it includes 'economic' groups and focuses on the asymmetry in benefits and contributions between people with different levels of wealth.

A second reason underlying the relevance of this topic is that it nicely complements the literature on inequality and collective action. Starting with the seminal contributions by Olson (1965) and Bergstrom et al. (1986), who predicted, respectively, a positive and a neutral effect of income inequality on public good provision, recent studies have emphasized the possibility that the impact of increased inequality on collective action may indeed be *negative*. Baland and Platteau (1997) have shown that increased inequality may lead

[1] On the impact of 'social capital' and participation on economic performance, see among others Wade (1988), Ostrom (1990), Putnam et al. (1993), Isham, Narayan and Pritchett (1995), Isham (1998), Narayan and Pritchett (1999).

to less collective action when the free rider problem gets worse for the poor and the set of contributors shrinks substantially. Baland and Ray (1997) study the problem under different specifications for the production function of the public good and find that if there is high complementarity among the inputs of rich and poor members, higher disparities in income may lead to less efficiency. Finally, Bardhan and Ghatak (1999) show that when there are market imperfections in inputs that are complementary to the collective goods, the relationship between inequality and efficiency will typically be inverse U-shaped. Our paper does not study public goods, but something that is closer to the notion of 'club goods'.[2] However, it shares with the above literature the interest in the effects and the workings of inequality.

Finally, there is very little empirical work on the determinants of group formation. This paper exploits an unusually rich dataset from Tanzania containing information on individual membership in a variety of groups that differ for type of economic service provided, size, and access rule. Such information is often difficult to come across even for industrialized countries. Furthermore, the relevance of issues like group composition and functioning is particularly high for developing economies because in those contexts many individuals rely on informal groups to get employment and production opportunities, as opposed to considering group participation as a side 'social' activity. We can thus be relatively confident that the type of economic incentives analyzed by the theory are at work in the settings we consider. This is confirmed by a recent study (Abraham et al., 1998) that has investigated the patterns of participation in economic groups in the informal settlements of Nairobi. However, the analysis by Abraham et al. (1998) is conducted only at a descriptive level, and the role of inequality is not addressed.

The main contributions of this paper can be summarized as follows. In the theoretical section we present a simple model in which heterogeneous individuals can choose to join a group which provides an excludable good to its members, and we derive predictions on the equilibrium composition of the group and on its size under two alternative access rules. The first is one of 'open access', by which anyone can join provided he or she pays the cost. The second rule instead allows the members of the group to exclude

[2] In contrast with club models, where an exclusion mechanism is usually assumed, we will study group composition both when there is excludability and when there is unrestricted access to the group. For a model of jurisdiction formation in which the choice of the admission rule is endogenous, see Jehiel and Scotchmer (1997). For a general review of club theory, see Cornes and Sandler (1996).

someone by majority vote. We show that an increase in income inequality has an ambiguous effect both on group composition and on aggregate levels of participation, and that the type of access rule is key in determining what income categories are represented in the group. In particular, open access groups will be formed by relatively poor individuals, while the composition of restricted access groups will be unbalanced in favor of the relatively rich. The impact of increased inequality on participation is ambiguous and depends on the access rule and on the shape of the income distribution. In particular, aggregate membership decreases under open access when heterogeneity increases in the lower part of the distribution, while participation can actually *increase* under restricted access if the upper part of the distribution is sufficiently skewed.

We test the predictions of the theory using survey data on rural households from Tanzania. Our main empirical results are the following.

(*i*) Higher inequality in assets at the village level has a negative impact on the likelihood that the respondents are members of some group. This result holds when controlling for other kinds of heterogeneity and for the possible endogeneity of inequality.

(*ii*) Inequality acts differentially on rich and poor people: when inequality increases, it is the relatively richer who drop out of groups, possibly because they have less to gain. The motives behind the decision of the rich to withdraw from groups are explored using both objective measures of relative wealth and discrepancies in the subjective rankings given by the respondent and by evaluation teams. We find that, for given 'objective' wealth, those individuals who overestimate their relative rank in the village participate less when inequality increases.

(*iii*) The impact of inequality on participation depends on the shape of the distribution of wealth and on the access rule to the group. In particular, it is negative for open access groups when there are wide disparities at the bottom of the distribution, while it is positive for restricted access groups when the disparities are around the middle and top part of the distribution.

(*iv*) Finally, group functioning in more unequal communities displays the following features: decisions are less likely to be taken by vote; members tend to sort into homogeneous income and ethnic groups; they more often report poor group performance and misuse of funds; they interact less frequently, and in general they feel less encouraged to participate. These effects are estimated separately for different categories of groups.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 briefly describes the data set and comments on the trends in participation and group composition. Section 4 illustrates our empirical strategy. Section 5 contains the main results on the effect of inequality — and of heterogeneity in general— on individuals' decisions to join groups. In section 6, an attempt is made to estimate the impact of inequality on aggregate outcomes at the group level, e.g. on group functioning, decision-making processes, quality of information sharing, frequency of interaction, etc. The results are presented separately for different categories of groups, such as economic, religious, and political. Finally, section 7 concludes.

2 The model

In this section we present a simple model of participation in groups in which the net benefits from participation vary with individual wealth. The aim of the model is to highlight one of the possible channels through which income distribution and participation can be related. It should by no means be regarded as an exhaustive theory, but rather as a framework to interpret our empirical results.

The model studies the impact of increased inequality on individual and aggregate levels of participation, comparing three possible mechanisms of group functioning. First, the social planner solution, where some external authority can assign each individual to some (or no) group, and the individual is obliged to join. Second, the 'open access' case, in which whoever wants to join a group is free to do it, provided he or she pays the dues. Third, the 'restricted access' case, in which the admission of a new individual into a group must be agreed upon by a majority of the members.

We study the case in which *only one group can be formed*. Though certainly restrictive, this assumption is made because we are interested in understanding the impact of inequality in settings in which individuals with different income cannot perfectly 'segregate' themselves into homogeneous income groups. On the other hand, it is not too irrealistic to think that this assumption may apply to some communities (e.g., villages) whose small scale does not allow for the coexistence of many groups of the same type. The extension of our results to the case of multiple groups is discussed in section 2.3.

Unless otherwise stated, all proofs are in appendix A.

2.1 Setup

Consider a continuum of individuals uniformly distributed on the interval $[0, 1]$ and ranked according to their income (or wealth) y_i so that the poorest individual has index 0 and the richest 1.[3] We denote the initial income distribution by $\{y_i\}$. Each individual can choose whether or not to participate in a group which provides an excludable good to its members. This good is consumed in the same (fixed) quantity by all and only the members of the group. Let H be the set of individuals who join the group. The total cost of providing the good is assumed to be fixed and equal to C. This cost is covered through a proportional tax on members' income:[4]

$$C = \int_{i \in H} t y_i di.$$

The tax rate is assumed to be exogenous and, from the above expression, is decreasing in the size of the group and in the wealth of the participants:[5]

$$t = \frac{C}{\int_{i \in H} y_i di}. \tag{1}$$

In order to abstract from the free-rider problem, we assume that individual wealth is observable and that whoever does not pay the fee ty_i cannot have access to the services provided by the group. Also, in what follows we exclude the possibility of side payments among individuals.

Everyone has the same preferences, represented by the continuous utility function $u(\cdot)$. We assume that the quality of the good provided by the group is the same for every member and is decreasing in the size of the group, say because of congestion. Let H indicate the size of the group, with a slight abuse of notation. We assume that the 'reduced form' utility function $u(H)$ is continuously differentiable and that

[3] In the theoretical section, we will interchangeably refer to income and wealth; when turning to the empirics, however, the distiction will be made clear.

[4] All our qualitative results hold under a more general tax schedule $t(i)$, as long as $t(\cdot)$ is strictly increasing in i. While particularly convenient from an analytical point of view, the assumption of a proportional tax can be motivated, for example, by the existence of informational contraints which prevent the use of more complex tax schemes. In section 2.3 we will discuss the consequences of the adoption of a lump-sum tax instead of a proportional tax.

[5] For a model in which the tax rate is endogenously determined by voting, see Roberts (1977).

(A1) $u'(H) < 0$; $u''(H) \leq 0$.

Assumption (A1) says that utility decreases at an increasing rate with group size, i.e. for a small group the addition of one more member has a negative but small impact on the quality of the services provided; when the group becomes very large, however, each additional member induces a larger and larger decrease in quality, hence in utility.[6]

Any individual who does not join the group gets a reservation utility $\overline{u}$.

Both the assumption that $u(H)$ is independent of i and the assumption that the reservation utility $\overline{u}$ is the same for everyone may seem unrealistic. For the sake of parsimony, we choose to concentrate all wealth effects in the cost of participation, ty_i, but we could obtain analogous results by letting the gross benefits from participation and/or the outside option vary across individuals with different wealth. All we really need in the following analysis is some monotone relationship between the benefit from participation *in excess* of individual contribution and reservation utility on the one hand, and individual wealth on the other.

The net utility from participation for an individual j belonging to the set of participants H is defined as

$$U(j) \equiv u(H) - \frac{C}{\int_{i \in H} y_i di} y_j \tag{2}$$

2.1.1 The social planner solution

As a benchmark, we start by considering what group composition and size would be chosen by a benevolent social planner who maximizes aggregate welfare of both members and non-members. The planner solves the problem:

$$\max_{0 \leq H \leq 1} \left\{ \int_{i \in H} \left[u(H) - \frac{C}{\int_{i \in H} y_i di} y_i \right] di + \int_{i \notin H} \overline{u} di \right\}$$

Under our assumption that the distribution of the i's is uniform on $[0, 1]$, the above problem simplifies to:

$$\max_{0 \leq H \leq 1} \{u(H) H + \overline{u}(1 - H)\}$$

[6] In general, we can think that initially there may be utility gains, rather than losses, from having more people in the group, and that congestion would take place only beyond a certain group size. We choose to ignore this possibility and to incorporate all positive effects if increased group size through the cost reduction channel.

Notice that the social planner does not care about *who* participates in the group. Any 'unbalance' in the relative wealth of the members translates into a redistribution of the costs from some members to others, which is a wash in the objective function of the planner. Proposition 1 characterizes the solution to the optimization problem. We first introduce two assumptions.[7]

(A2) $u(0) > \overline{u}$

(A3) $u'(1) + u(1) < \overline{u}$

Proposition 1 *Under assumptions (A2)-(A3), the social planner will select enough individuals to form a group of size* $0 < H^* < 1$, *where* H^* *solves*

$$-u'(H^*)H^* = u(H^*) - \overline{u}. \tag{3}$$

Proof. Straightforward from continuous differentiability of $u(\cdot)$, given (A2)-(A3). □

Intuitively, the social planner will add people to the group until the marginal loss from increased congestion (left hand side of (3)) equals the marginal benefit from participation (right hand side).

2.1.2 Open access group

Consider next the case where the participation decision is decentralized and anyone is allowed to join the group, with or without the consent of the other members. We will call this the 'open access' group. In this case, the *equilibrium* size of the group must be *self-enforcing*, in the sense that all and only the individuals who obtain a positive net benefit from participating are in the group. This is captured in our definition of an open access (OA) equilibrium:

Definition 1 *A group* H *is an* ***OA equilibrium*** *if the following two conditions hold:*

(i) no $j \in H$ *wishes to leave* H*;*

(ii) no $j \notin H$ *wishes to join* H*.*

[7] Assumptions (A2) and (A3) are necessary and sufficient conditions for the existence of an *interior* solution to the social planner's problem, which is the only type of solution of interest to us.

The above definition amounts to saying that group H will be formed by all and only the individuals j for whom the following participation constraint is satisfied:

$$u(H) - \frac{C}{\int_{i \in H} y_i di} y_j \geq \overline{u} \tag{4}$$

Proposition 2 *In an OA equilibrium the group is formed by the poorest* $[0, h^*]$ *individuals, where* h^* *is the largest root of the equation:*

$$u(h) - \frac{C}{\int_0^h y_i di} y_h = \overline{u} \tag{5}$$

In what follows we will often refer to h^* as the 'marginal' member or participant. Proposition 2 is a consequence of the fact that when individual contributions increase with wealth, as under a proportional income tax, *the net utility from participation is higher for the poorer members than for the richer ones*. In fact the benefit from a given level of congestion is the same for everyone (first addendum on the left hand side of (4)), while the cost is higher for richer people (second addendum). There is a threshold level h^* beyond which the burden of financing group activities would be too heavy for the richest individuals to bear, and they prefer to stay out of the group. All we need to know to predict group size is then the identity of the last richest member, which is the h^* solution of (5).

Notice that while the optimal size for the social planner, H^*, did not depend on the distribution of income, the self-enforcing equilibrium size h^* does. This point will be addressed in the next section. For the time being, we turn to a third concept of equilibrium.

2.1.3 Restricted access group

Consider now the case where the members can regulate access to the group, effectively preventing someone from joining or expelling a member they no longer want. We will maintain the assumption that the act of joining must be voluntary, i.e. the group can *decrease* its size without the consent of the involved party, but cannot *increase* it unless the new member agrees to join.

Decisions within the group are taken by majority rule. Note that since for given group size individual preferences are strictly monotone in individual

income, the median income member is decisive. The extensive form of the game can be described as follows.

1. An initial group H_0 is given.[8]

2. The individual who has the median income among the members of the group proposes a new group H_1.

3. A majority vote is taken by the members of H_0 on the proposal H_1.

4a. If a majority votes against H_1, the original group H_0 remains in place.

4b. If a majority votes in favor of H_1, then:

4b_1. If all the new members of H_1 accept to stay in the group, H_1 replaces H_0.

4b_2. If any member of H_1 refuses to join, the original H_0 is reinstated.

5. Taxes are levied on the members of the final group (be it H_0 or H_1), the good is provided to them, and the game ends.

We next define our equilibrium concept under this restricted access (RA) regime.

Definition 2 *A group H_0 is a **RA equilibrium** if the following two conditions hold:*

(i) no $j \in H_0$ wishes to leave H_0;

(ii) $H_1 \equiv H_0$, i.e. given the initial size and composition of the group, the median income member does not propose to alter it.

The following consideration can help understand our definition of a RA equilibrium. Note that any median member would always like to expel the bottom half income members of an initial group H_0 and include richer people in their place, i.e. shift the composition of the group to the right: this way congestion could be maintained the same and taxes decreased. In this effort, he would always be supported by the majority (the top half members). An equilibrium is reached when the size and composition of the group is such that: all the members get a higher utility from staying in the group than from leaving (condition (i) in the definition); and no richer individual beyond the current top income member would agree to join the group, so it is not profitable for the median member to propose a change (condition (ii) in the definition). Alternatively, an RA equilibrium can be seen as the group chosen

[8] We are not interested in how the original group is formed. One can think for example of a random draw by nature, both in terms of size and of composition of the group.

by a median member who maximizes his/her utility across the set of groups all of which will have him/her as a member.

Notice that in this framework the median member is given *one chance* to alter the composition of the group and then the game ends. In other words, the median member does not have to worry about the fact that, if group composition shifts right, in the following period he/she will no longer be 'median' and the new decisive individual may choose to move even more to the right, so that eventually the utility of the original median voter may be lower than under the initial configuration. While more satisfactory from a dynamic point of view, the latter approach would yield a larger set of equilibria. We choose therefore to adopt the seemingly more 'myopic' game structure described in the text in order to get sharper predictions on the equilibrium outcome. This way we obtain a 'fixed point' type of equilibrium, in which group size and composition are optimal for the median member even if he/she were free to alter it without risking 'unraveling' in the future. In the terminology of Roberts (1999), who develops a fully dynamic framework for voting equilibria in clubs, we are restricting our attention to an 'extrinsic' steady state.

Lemma 1 *Any group that is a RA equilibrium must be a single interval.*

Given the above lemma, the equilibrium can be found by choosing the boundaries of a generic group $[h_1, h_2]$ so that the utility of the median voter $(h_1 + h_2)/2$ within this group is maximized, provided that the participation constraint holds for all the individuals in $[h_1, h_2]$:[9]

$$\max_{0 \leq h_1 < h_2 \leq 1} u(h_2 - h_1) - \frac{C}{\int_{h_1}^{h_2} y_i di} y_{(h_1+h_2)/2} \tag{6}$$

$$\text{s.t. } u(h_2 - h_1) - \frac{C}{\int_{h_1}^{h_2} y_i di} y_{h_2} \geq \overline{u} \tag{7}$$

The solution to the problem is described in the following proposition.

Proposition 3 *Depending on the functional form of $u(\cdot)$ and on the income distribution $\{y_i\}$, in equilibrium the group can take one of two forms:*

[9] Notice that in (7) we only write *one* constraint, namely that for the richest member h_2. In fact given that $U(j)$ is strictly decreasing in y_j for give group size, if the participation constraint holds for h_2, it will necessarily hold for any $j \in [h_1, h_2]$.

Case A: $[h_1^*, 1]$, *where* h_1^* *is the solution to*

$$-u'(1-h_1) - \frac{C}{\int_{h_1}^{1} y_i di}\left[\frac{\partial y_{(h_1+1)/2}}{\partial h_1} + \frac{y_{h_1} y_{(h_1+1)/2}}{\int_{h_1}^{1} y_i di}\right] = 0 \tag{8}$$

Case B: $[h_1^{**}, h_2^{**}]$, *where* h_1^{**}, h_2^{**} *and* λ *solve*

$$-u'(h_2-h_1)(1+\lambda) - \frac{C}{\int_{h_1}^{h_2} y_i di}\left[\frac{\partial y_{(h_1+h_2)/2}}{\partial h_1} + \frac{y_{h_2}\left(y_{(h_1+h_2)/2} + \lambda y_{h_2}\right)}{\int_{h_1}^{h_2} y_i di}\right] = 0 \tag{9}$$

$$\begin{aligned} - \quad & u'(h_2-h_1)(1+\lambda) - \frac{C}{\int_{h_1}^{h_2} y_i di} * \\ * \quad & \left[\frac{\partial y_{(h_1+h_2)/2}}{\partial h_2} - \frac{y_{h_2} y_{(h_1+h_2)/2}}{\int_{h_1}^{h_2} y_i di} + \lambda\left(\frac{\partial y_{h_2}}{\partial h_2} - \frac{y_{h_2}^2}{\int_{h_1}^{h_2} y_i di}\right)\right] = 0 \end{aligned} \tag{10}$$

$$u(h_2 - h_1) - \frac{C}{\int_{h_1}^{h_2} y_i di} y_{h_2} = \overline{u} \tag{11}$$

where $\lambda \geq 0$ *is the multiplier on the participation constraint (7).*

Case A describes an equilibrium in which the group is formed by the richest segment of the population and the poorest $[0, h_1^*)$ individuals are left out. In case B, instead, both the low and the top end of the income scale are out of the group, the former because they are not allowed to join (though they would like to), and the latter because they choose not to (the tax burden is relatively too high for them). The group is thus formed by an intermediate segment $[h_1^{**}, h_2^{**}]$ of the population.

Compared to the OA equilibrium, the possibility to exclude people from the group generates a *composition unbalanced in favor of the rich*. A poor member is always less desirable than a rich one because it generates the same congestion and contributes less resources. A group that votes on its composition will always try to incorporate as many rich individuals as possible, that is, it will try to push membership as 'far to the right' as possible, compatibly with individual participation constraints.

2.2 Income distribution, group composition and size

Starting from the initial income distribution $\{y_i\}$, we now consider the effects of an increase in inequality taking the form of a redistribution of wealth from the poor to the rich. In particular, we assume that the new income distribution $\{\widetilde{y}_i\}$ has the following characteristics:

(A4) $\int_0^1 \widetilde{y}_i di = \int_0^1 y_i di$
(A5) $\forall i < i'$, $y_i < y_{i'} \Longrightarrow \widetilde{y}_i < \widetilde{y}_{i'}$
(A6) $\exists \widehat{i}$ s.t. $\widetilde{y}_{\widehat{i}} = y_{\widehat{i}}$, and $\widetilde{y}_i < (>) y_i$ for all $i < (>) \widehat{i}$

Assumption (A4) says that aggregate wealth is unchanged. Under (A5), each individual maintains the same wealth *ranking*, i.e. if i was poorer than i' under the old distribution, the same will be true under $\{\widetilde{y}_i\}$. Finally, (A6) says that there is some individual $\widehat{i}$ whose wealth is the same as under the original income distribution, all the people in $[0, \widehat{i})$ being poorer than before, and all those in $(\widehat{i}, 1]$ being richer.[10]

In the present context, in which only one group is formed, the aggregate level of participation coincides with group size. Will participation, as measured by group size, be higher or lower under the more unequal distribution of income? The answer to this question is straightforward if group formation is in the hands of the social planner. In fact we have seen that the group size chosen by the planner, H^*, is independent of the distribution of wealth. An increase in inequality, therefore, will leave H^* unaffected. In this section, we concentrate on the effect of inequality in the two regimes of open access and restricted access.

2.2.1 Inequality under open access

When anyone is free to join the group, the impact of increased inequality is described by the following proposition.

Proposition 4 *Let h^* be the solution to (5). An increase in inequality taking*

[10] Of these assumptions, only (A4) is really needed: (A5) and (A6) are introduced to ease exposition, but could be dispensed with. Suppose in fact that redistribution were non-monotonic, in the sense that individual ranking under $\{\widetilde{y}_i\}$ changed compared to that under $\{y_i\}$. In this case, the arguments provided below could be applied in the same way by simply *re-sorting* individuals in ascending wealth order and studying whether the individual that under the new distribution occupies the 'marginal position' h (or h_1^*, h_1^{**}, h_2^{**}) is richer or poorer than before.

the form described in (A4)-(A6) will decrease participation if and only if

$$\frac{\widetilde{y}_{h^*}}{y_{h^*}} > \frac{\int_0^{h^*} \widetilde{y}_i di}{\int_0^{h^*} y_i di} \tag{12}$$

Corollary 1 *If (12) holds, the group that forms under* $\{\widetilde{y}_i\}$ *is constituted by the poorest* $[0, h']$ *individuals, where* $h' < h^*$.

Intuitively, the condition expressed in proposition 4 requires the relative increase (decrease) in the wealth of the marginal member h^* to be larger (smaller) than the cumulative change in the wealth of the poorer members. We have seen, in fact, that each individual's incentive to join the group depends on his/her relative wealth vis-a-vis that of the other participants. Consider the case in which individual h^* is richer under the new income distribution, i.e. $\widetilde{y}_{h^*} > y_{h^*}$. In this case, (12) holds unambiguously because the right-hand side is less than unity by construction. Individual h^* is now 'too rich', and the other members 'too poor', for h^* to benefit from participation.

If, on the other hand, $\widetilde{y}_{h^*} < y_{h^*}$, it can be profitable for h^* to belong to the group as long as he/she has become sufficiently poor and $\widetilde{y}_{h^*}$ is close to that of the rest of the individuals in $[0, h^*)$, as depicted for example in figure 1a. In this case group size may actually *increase* with inequality, because the many people who have become poor get comparable benefits from participation and can gain by extending membership to decrease per-capita costs. Conversely, if $\widetilde{y}_{h^*} < y_{h^*}$ but (12) holds, as depicted for example in figure 1b, participation will be lower under the new distribution of income.

[Insert figure 1]

To sum up, under open access an increase in income inequality has an ambiguous effect on group size, hence on participation. Roughly speaking, the effect is negative if the redistribution leads to more inequality in the bottom part of the distribution; it can be positive if the redistribution benefits few rich people at the expense of a large mass of poor people so that there is relatively more equality at the bottom of the distribution. In all instances, if participation does decrease, it is the relatively richer members who choose to drop out. In the empirical section of this work, we will test all these implications.

2.2.2 Inequality under restricted access

When access to the group can be regulated by its members, the effect of inequality on participation is even more ambiguous. In what follows we discuss the results qualitatively.[11] The formal analysis for case A is relegated to Appendix A.[12]

Let us start by recalling the mechanisms underlying the choice of group size under the initial distribution of income. An important element of the analysis in section 2.1.3 was that, since the net benefit from participation is lower for richer members, *ceteris paribus the 'ideal' group size is increasing in individual wealth.* In fact any increase in size brings the same congestion to all the members, but generates greater savings for the richer ones through lower tax rates. The equilibrium size chosen by the median voter is a compromise between the relatively large group sizes desired by the top half members and the relatively small ones preferred by the bottom half members. The key factor for assessing the impact of increased inequality on equilibrium group size will therefore be the new relative income of the median voter vis-a-vis the rest of the group.

Consider first case A, where the group under the initial income distribution was constituted by $[h_1^*, 1]$. By assumption (A4), under $\{\widetilde{y}_i\}$ the aggregate wealth of the individuals in $[h_1^*, 1]$ is higher. This means that, if group size stayed the same, the tax rate required to cover the cost C would now be lower. Loosely speaking[13], if the median member $(h_1^* + 1)/2$ had the same (or lower) wealth than before, as in figure 2b, he would certainly be better off by reducing the level of congestion and choosing a smaller group size.

[11] If there are multiple equilibria, the conditions we derive on the impact of inequality on participation must be intended to hold *locally* around each of the equilibria. In appendix A, however, we show that, under mild sufficient conditions on the shape of the income distribution, the RA equilibrium type A is unique, hence for that case our conclusions hold globally.

[12] A formal treatment of case B is not reported due to the large number of possibilities arising from the fact that in case B nothing can be said a priori on the aggregate wealth of group members under the new versus the old distribution of income. For this case, we therefore choose to give a qualitative account of the possible ambiguities.

[13] Notice that in general the individual who was the median member under the initial income distribution will no longer be the median member under $\{\widetilde{y}_i\}$. All we say in what follows should be read as saying that if the same individual were still median, he or she would no longer be optimizing by choosing the same group size, hence group size must change, in one direction or the other.

The initial group size could no longer be an equilibrium, and the new group should be $[h_1', 1]$, with $h_1' > h_1^*$. If, on the other hand, the median voter had become so rich to actually benefit from *further* decreases in tax rates (see for example figure 2a), participation may actually increase and the new lower bound h_1' may be smaller than h_1^*.

[Insert figure 2]

Consider now case B. Under the original income distribution, the group was joined by the individuals in $[h_1^{**}, h_2^{**}]$, defined in proposition 3. Notice that in the initial equilibrium the richest member, h_2^{**}, was held at the reservation utility $\overline{u}$. Take the case where the original 'median member' $(h_1^{**} + h_2^{**})/2$ is richer under $\{\widetilde{y}_i\}$, and thus wants to extend membership. If the increase in h_2^{**}'s individual wealth is particularly high and the decrease in the tax rate is not big enough, h_2^{**} may actually drop out of the group. More interestingly, contrary to case A where the total wealth of the participants is necessarily higher under the new income distribution, in case B it is possible that $\int_{h_1^{**}}^{h_2^{**}} \widetilde{y}_i di < \int_{h_1^{**}}^{h_2^{**}} y_i di$. This implies that if the group remained as before, the tax rate should be higher and, unless h_2^{**} had become substantially poor and 'close' to the other members in wealth, his or her net utility would fall below $\overline{u}$. Under such circumstances, we would expect the relatively richer members to drop out and group composition to shift left.

To sum up, in the RA equilibrium an increase in inequality may either increase or decrease participation. Furthermore, in contrast with the OA case in which any decrease in group size happens with the richer members dropping out, in the RA case it can be the rich or the poor who leave the group, depending on the shape of the income distribution.

2.3 Discussion and extensions

The analysis so far has assumed that only one group can be formed. A first question is whether our results generalize to the case where *multiple groups* can coexist. In this section we briefly address this question at the qualitative level.[14] Given the irrelevance of income distribution for the social

[14] An analytical treatment of the multiple group model goes beyond the scope of this paper, also because for some of the groups considered in the empirical analysis (e.g., burial societies, church groups, political groups) it is realistic to think that only one group can form in a village. For the other groups, the theory will serve as a benchmark to analyize the mechanisms underlying the impact of inequality on participation, with the caveat that when multiple groups can form the results will generally be more ambiguous.

planner solution, a change in inequality will not affect the number and size of groups even when this number can exceed one. The interesting cases are the open access and the restricted access regimes. What makes the multiple groups setting interesting in these cases is the possibility that more inequality actually leads to an *increase* in participation by increasing the equilibrium number of groups. In fact from the local public finance - multicommunity literature we know that when heterogeneous agents can choose among multiple 'jurisdictions' they will stratify into homogeneous income groups (e.g., Epple and Romer, 1991; Fernandez and Rogerson, 1996). While we can expect the results here obtained for a single group to hold *within* each of the multiple groups, it is possible to get contradicting results in the aggregate. Take for example the open access case, in which N groups will form starting at 0 until some wealth level $\overline{h}$ beyond which no individual is part of any group. In this case an increase in inequality will shift the composition of each group to the left, and this may 'free up' some of the richer members so they can form an $N+1$th group including some $j > \overline{h}$. Aggregate participation may in this case be higher, rather than lower, even in the open access case.

Another assumption in the above analysis was that members paid a *proportional tax* on income. This had two implications. The first was that anyone could sustain the costs of joining a group, because by definition only a fraction of their income was taken. The exclusion of poor people from a group then crucially depended on the availability of a 'restricted access' technology, rather than on self-restraint. The second implication was that for the whole group a richer individual was preferable to a poor one, because he or she contributed more. This generated our 'sorting' result.

Suppose instead that group activities were financed through an exogenous lump-sum tax, T. In this case a new constraint should be added to the problem, namely that the individual can afford paying the membership dues. For every member i of a group of size H we must have $y_i \geq T$, i.e. $y_i \geq C/H$. In the open access - single group case our result that the group is formed by the poorest $[0, h^*]$ individuals would no longer hold. First of all, the very poorest individuals may not be able to afford the costs of participation if their income is lower than T. Second, within the set of individuals whose income exceeds T no prediction could be made on who exactly would join the group and who would not. All the model could determine is an equilibrium group size at which the net utility for any member exactly equals the reservation level $\overline{u}$. In this case the presence of heterogeneity in preferences and/or in

individual outside options would deliver predictions on *who* should have an incentive to join. For example, if individual outside options $\overline{u}_i$ were increasing in i, we could still predict that the group will be formed by an interval $[h_0, h^*]$, where the lower bound h_0 is defined by the budget constraint $y_{h_0} = C/(h^* - h_0)$ and the upper bound h^* is defined by the indifference condition $u(h^* - h_0) - C/(h^* - h_0) = \overline{u}_{h^*}$. In any case, it will still be true that if participation decreases when inequality increases, it will be the relatively richer people who stay out of the group.

The introduction of a lump-sum tax modifies the analysis of the RA group in a more substantial way. If fact, ceteris paribus, it is no longer true that the median voter would be made better off by a one time shift in the composition of the group to the right. In fact, as long as an individual is able to pay membership dues, everyone looks the same to the other group members. Enriching the model with the possibility of idiosyncratic income shocks, a limited liability argument may push in the direction that a rich individual is relatively more 'desirable' than a poor one because he/she has a higher probability of being able to pay the fee in the event of an adverse shock.

A strong assumption of the theoretical framework presented in this paper is the semi-linearity of the utility function, and in particular the fact that income does not enter $u(\cdot)$. Removing this assumption would certainly affect our results and bring risk-sharing into the picture. The impact of heterogeneity on group size and composition would then depend on the relative costs and benefits of the insurance function of the group versus the local public good provision. The analysis of group composition when individuals have heterogeneous risks in the context of group lending has been conducted, among the others, by Ghatak (2000) and Sadoulet (1999). While the former predicts an 'assortative matching' result whereby groups are formed by individuals with identical risks, the latter finds conditions under which heterogeneous groups form in which the less risky members of the group receive transfers from the riskier ones.

Regarding the decision process in the restricted access model, we have made the assumption 'one head one vote'. Realistically, in many groups the relatively richer individuals may have greater influence (direct or indirect) on the decisions. If this feature were incorporated in our analysis of RA equilibrium, it would not undermine our conclusion, but simply *reinforce* the possibility that an increase in inequality leads to an increase of participation,

thus accentuating the difference between the OA and the RA equilibria.

Finally, we choose to disregard the possibility that altruism may affect individual decisions to join groups (or to admit people into a group). Any assumption in this direction would of course affect our results.

Our model has highlighted some aspects of the link between inequality and participation in groups. We may summarize our results in four main points:

i) inequality has an ambiguous impact on participation;

ii) ceteris paribus, when inequality is higher we can expect the relatively richer individuals to stay out of the groups, if they have relatively less to gain;

iii) under open access, the impact of inequality depends on the shape of the income distribution at the *bottom*;

iv) under restricted access, the impact of inequality depends on the shape of the income distribution around the *mean* and the top.

In what follows we turn to the empirical analysis to try and shed light on some of the ambiguities present in the theory and to test the predictions of the model.

3 The data

The data we will be using come from two recent surveys conducted by the World Bank: the Tanzania Social Capital and Poverty Survey (SCPS) of 1995 and the Tanzania Human Resource Development Survey (HRDS) of 1993. The SCPS sampled 1376 rural households, asking a broad set of questions on group membership, group composition and performance, and also on the values and degree of social ties in the community. For approximately half of the households in the sample, data on demographic characteristics and expenditure was also collected. An alternative source of information on village-level aggregates is the HRDS: although the surveyed households may be different, the clusters of the social capital survey are all comprised in this latter, broader survey. Where useful to remedy the incomplete coverage of the SCPS, this coincidence will be exploited. In this section, some descriptive statistics from the SCPS will be presented to gain some insight into the

diffusion and characteristics of associational activity in our sample.[15]

[Insert table 1]

According to the SCPS, 72 percent of the individuals in the sample were members of some group, the average number of groups per participant being 1.6. Table 1 lists the main types of groups, reporting the percentage of the total sample who belong to the given group(s). We will classify the various groups under three broad categories: (i) *religious*, such as churches (joined by 15% of the respondents), Muslim groups and mosques (6%); (ii) *political* (12%); and (iii) *economic*, including burial societies (12%), women's groups (5%), farmers' groups (5%), cooperatives (2%), rotating credit associations (1%), and dairy/cattle rearing groups (1%). The characteristics of the different groups can be described with the help of table 2.[16]

[Insert table 2]

Panel A of the table contains information on the organizational rules of the groups, while panel B summarizes the degree of heterogeneity of the members. Starting from the top panel, for all categories of groups more than half of the respondents report that they joined voluntarily. A significant proportion of memberships for religious groups occur by birth (36%); in the rest of the cases members of political and economic groups are either required to join (13% and 16%, respectively) or pay a fee (35% and 19%, respectively). Approximately 39% of the respondents report that their groups are supported by an external agency: in most cases, the agency is an NGO (39% for religious, 40% for political, and 54% for economic groups), other relevant funding sources being the government for political associations (53%) and the church for religious ones (41%). Enforcement mechanisms vary among the different categories: when there is a fee and members do not pay it, chances of being expelled are high for political and economic groups (about 40%), while almost 80% of the members of religious associations report that nothing happens.

[15] For a detailed description of the Tanzania Social Capital and Poverty Survey, the reader is referred to Narayan (1997).

[16] The questions on group organizational structure listed in table 2 were mainly asked to people who said they were members of the respective group.

In panel B we turn to the degree of heterogeneity of the members in terms of ethnic identity, type of economic activity, and income level. This data differs from the measures of heterogeneity we will use in the multivariate analysis under two respects: (*i*) the figures in table 2 refer to heterogeneity *within* the group, while in our regressions we will link participation to heterogeneity in the whole village (as requested by the theory); (*ii*) the data in table 2 comes from individual responses to questions directly aimed at assessing homogeneity among group members. In the multivariate analysis, on the contrary, we will construct measures of heterogeneity starting from the primitives, i.e. from raw data on individual ethnicity, education, economic activity, income and assets.

Regarding the ethnic composition of the groups, in the vast majority of the cases, group membership is open to anyone: only for economic groups, 7% of the respondents say that the members all belong to the same clan or tribe.[17] Group members are also diverse in the type of activity on which they earn their living. Approximately 78 to 87 percent of the respondents consider the groups 'mixed' under this respect. Again, only for economic groups the fraction reporting that 'all' ('most') members make a living in the same way is significant, namely 7% (14%). In a separate question, the same individuals were asked whether the *leaders* of the groups made a living in a different way from the rest of the members: in this case, 32 to 34 percent of the people answered 'yes'. Finally, all groups seem to aggregate people with different income levels: only 1% of the respondents reported that the members were all rich or all poor, and 15 to 29 percent —depending on the type of group— said that there was little income diversity among the participants.

Overall, the evidence reported in panel B seems to suggest that the individuals in our sample do *not* sort into highly homogeneous groups. At least in principle, therefore, there is scope for differential incentive mechanisms among relatively rich and relatively poor people, as stressed by the theory.

4 Empirical strategy

The first model we will estimate relates the individual decision to participate in a group to individual attributes as well as village-level variables, among

[17] The relative importance of ethnic affiliation for this category of groups compared to the others is not surprising, given that tribal links can enlarge the scope for reciprocity and enforcement mechanisms in economic transactions (La Ferrara, 1997).

which inequality. The structural equations underlying our estimates can be presented as follows. The expected net benefit from participation for individual i in village v is:

$$B^*_{iv} = \beta' X_{iv} + \gamma' H_v + \delta G_v + \varepsilon_{iv} \tag{13}$$

where X_{iv} is a vector of individual characteristics, such as age, sex, education, and wealth; H_v is a vector of village characteristics, such as average wealth, heterogeneity in education , in economics activity, in tribe, and income inequality; G_v measures income or assets inequality in village v, and ε_{iv} is an error term normally distributed with mean 0 and variance σ_v. The vectors β, γ and δ are parameters. We do not observe the 'latent' variable B^*_{iv}, but only the choice made by the individual, which takes value 1 (participate) if the expected net benefit is positive, and 0 (not participate) otherwise:

$$P_{iv} = 1 \text{ if } B^*_{iv} > 0 \tag{14}$$

$$P_{iv} = 0 \text{ if } B^*_{iv} \leq 0 \tag{15}$$

The probit model (13)-(15) will be estimated, correcting for heteroskedasticity and clustering of the residuals at the village level.

A potential concern with the above strategy is that inequality is measured with error or that it is endogenous, in that higher participation may create opportunities for the poor to advance and decrease inequality. To account for this possibility, we estimate a linear probability model and instrument inequality using geographical variables. The two-stage-least-squares model estimated is in this case:

$$P_{iv} = \beta' X_{iv} + \gamma' H_v + \delta G_v + \varepsilon_{iv} \tag{16}$$

$$G_v = \mu' X_{iv} + \psi' H_v + \lambda Z_v + \eta_v \tag{17}$$

where P_{iv} is equal to 1 if individual i in village v is member of some group and 0 otherwise; X_{iv}, H_v and G_v are as above; Z_v is a vector of instruments; finally, $\beta, \gamma, \delta, \mu, \psi, \lambda$ are vectors of parameters. The key assumption for the two stage estimation of this linear probability model is that Z_v is correlated with G_v but uncorrelated with ε_{iv}.

After estimating the basic reduced form we will turn to a more specific prediction of the model, namely that when inequality is higher, participation

should be relatively less attractive for relatively richer people. To test this prediction we will replace expression (13) with the augmented form

$$B^*_{iv} = \beta' X_{iv} + \gamma' H_v + \delta G_v + \zeta A_{iv} + \xi G_v A_{iv} + \varepsilon_{iv} \tag{18}$$

where A_{iv} is the difference between individual i's assets and the average in village v. A positive value of A_{iv} characterizes a richer than average individual. Our conjecture is that the sign on the interaction term between A_{iv} and the inequality measure G_v should be negative. Since relative wealth may capture some unobservable individual characteristics which influence participation regardless of the 'relative benefits' explanation proposed by the theory, we shall try and sharpen our test by exploiting the difference between 'objective' and 'subjective' relative wealth. The latter test will be illustrated below.

Regarding the role of the shape of the income distribution under different access rules, we will test the predictions of our model by estimating the probit model (13)-(15) separately for OA and RA groups, and using inequality measures with varying degrees of sensitivity to the lower vs. upper part of the income distribution. According to our theoretical framework, the negative impact of inequality in open access groups should be best captured by inequality indexes that put a lot of weight on heterogeneity in the lower part of the distribution. The potential positive role of inequality in RA groups, instead, should be best captured by measures that emphasize heterogeneity in the middle and upper part of the distribution.

Finally, the last part of our empirical analysis will explore some issues not directly linked to the model presented in this paper, but which we find interesting and worth investigating given the richness of the data set. In particular, we will examine how inequality relates to relevant group characteristics, namely group composition and functioning. Variable definitions, as well as summary statistics, are reported in appendix B.

5 Inequality and group membership

5.1 Basic regressions

Table 3 presents our estimates of the basic model. The dependent variable is a dummy taking value 1 if the respondent –generally the head of the household– is member of at least one group. The individual controls include:

age, sex, and education of the respondent; household size, and a dummy equal to one if the respondent has been living in the same village at least for the past ten years. Individual wealth is measured through an 'asset index' constructed in the survey by assigning predetermined weights to the ownership of various durable consumption goods.[18] The average of this index at the village level, as well as the 'size' of the village in terms of population, are included among the regressors to control for potential differences in the 'demand' for participation across communities. Regional dummies are also included.

The explanatory variables in which we are most interested are measures of *heterogeneity* among residents in the village, and in particular:

i) inequality, measured with the Gini coefficient built from the assets index of all respondents from the village;

ii) tribal fragmentation, measured through the following 'fractionalization' index:

$$F_v = 1 - \sum s_{kv}^2 \qquad k = 1, ..., K_v \tag{19}$$

where s_{kv} is the share of respondents in village v who belong to tribe k, and in each village there is a number K_v of different tribes. This index represents the probability that two randomly drawn individuals in the same village belong to different tribes;

iii) heterogeneity in economic activity, measured with an index analogous to (19), where k denotes an individual's economic activity rather than tribe;[19]

iv) heterogeneity in education, measured as the standard deviation of the highest grade in school obtained by household heads interviewed in the village.

Before commenting on our estimates, we should pause to discuss our choice of the inequality variable. Self-reported household expenditure is also available from the SCPS and HRDS and can be used to build income inequality measures. Similarly, the HRDS contains data on the number and size of *shambas* (or gardens) owned by a household, and can be used to con-

[18] The durable consumption goods and weights used were the following: van or truck, car, motorcycle, bicycle, sewing machine = 16, radio = 8, table = 6, bed, clock or watch = 4, chair = 3, lamp = 2. If the asset was not working, these weights were approximately halved.

[19] The most common activities are: working on own farm, government employee, self-employed, and retired.

struct an index of land inequality. We choose to employ asset inequality for a variety of reasons. First of all, compared to income inequality, inequality in assets is likely to be less endogenous, and a better measure of *permanent* inequality in the village. Second, while land inequality may not suffer from the same endogeneity problem as income inequality, it is not clear that it captures differences in relative wealth: for example, an individual who solely relies on farm labor and cultivates two *shambas* should not be considered richer than a government employee who may not own a *shamba*. Finally, there is much more variation in the Gini coefficient built from assets (min = .16, max = .69, std. dev. = .11) than in the Gini built from expenditure (min = .13, max = .44, std. dev. = .06) or from land (min = .24, max = .68, std. dev. = .10). We have estimated all regressions using income and land inequality in place of assets inequality, and our results were generally similar in qualitative terms, though the coefficients on inequality were not always significant in statistical terms.[20]

[Insert table 3]

The first column of table 3 presents the baseline specification of our probit model, including individual and village controls but not heterogeneity measures. The age and sex of the respondent do not seem to matter very much, while education has a positive and significant impact on the probability of being member of a group.[21] Household size and the length of residence in the village do not have a significant impact on participation. On the other hand, the size of the community in terms of population has a negative effect on participation rates, though not always statistically significant at conventional levels. Average wealth in the village is positively associated with group membership, suggesting that participation may be a 'normal good'.[22]

In column 2 we introduce various measures of heterogeneity at the village level. Neither tribal fractionalization nor heterogeneity in education turn out to be significant in our regression. The diversity of economic activities –which might be relevant for the 'consumption' interpretation– is also not

[20] Results are available from the authors.

[21] As will be clear from columns 3 and 4, this effect is partly due to the fact that education here proxies for individual wealth, not included among the regressors.

[22] We must be careful in interpreting this link in a causal sense, given that higher wealth in the community might be a *consequence* (rather than a cause) of higher levels of social capital (see Narayan and Pritchett, 1999).

significant: one possible reason is that there is not much variation in this variable: almost 84 percent of the respondents belong to the same category, i.e. 'farmers'. On the contrary, the estimated coefficient on 'Gini' is negative and significant at the 1 percent level: *increased inequality in one's village has a negative effect on the likelihood that somebody will join a group.* Based on the marginal coefficients of the probit model in table 3, an increase in 'Gini' by one standard deviation decreases the probability of participating in a group by about 4 percentage points. This is quite a sizeable effect if compared to the other determinants of participation: for example, it is almost three times the effect of one more year of education.

In the last two columns we consider non linear effects of individual income on participation. This is partly to verify that our results on inequality are not a statistical artifact. In fact, if the relationship between individual income and group participation were concave and we omitted nonlinear income terms from the specification, we may find a negative and significant coefficient on Gini even if inequality per se were not a determinant of participation.[23] In this case, adding nonlinear income terms to the regression should wipe out the effect of inequality. We can see from column 3 that, when introduce individual assets and its square, Gini retains a negative and significant coefficient. The coefficients on individual assets suggest that the relationship between individual wealth and participation is increasing and concave. Finally, in column 4 we further explore nonlinearities by introducing two dummies that indicate whether the respondent belongs to the first or the last wealth quartile. We find that they have, respectively, a negative and a positive coefficient, though not highly significant. On the other hand, the coefficient on inequality is quite robust and fairly stable in all these specifications.

5.2 Instrumenting Gini

Before proceeding with more specific tests of our theory, we need to deal with the potential endogeneity of the variable Gini and with the possibility of measurement error. In fact, we may conjecture that a high degree of participation may reduce inequality by generating opportunities of advancement for the poor in terms of education, diffusion of information, public good pro-

[23] This point has been discussed in the health economics literature when assessing the impact of income inequality on health outcomes (Gravelle, 1998).

vision, etc. Also, villages prone to social activities may have higher levels of risk-sharing, which would attenuate the differences in expenditure among their members (though this would be less of a concern if we had a reliable measure of permanent wealth). Finally, we want to account for the possibility of measurement error for the variable Gini.

To address the above problems, we estimate a linear probability model using two-stage-least-squares, and instrumenting Gini with demographic controls and with the following *geographic* variables:

(i) the average precipitation in the 'critical month(s)', i.e. in those months of the growing season in which scarcity of rain can jeopardize the harvest the most.[24] The critical months were identified by looking at the patterns of rain in the different regions, e.g. unimodal versus bimodal (Gommes and Houssiau, 1982), and by checking the calendars for the relevant crops. Given that agriculture is the main activity in rural Tanzania, we expect to find a negative relationship between this measure of rain and inequality in the village;

(ii) rainfall variability, i.e. the standard deviation of the monthly precipitation during the year. The higher this variability, the more vulnerable to shocks should be people who rely on farming to earn a living, hence the higher inequality;

(iii) average temperature during the year (in C degrees);

(iv) temperature variability, i.e. the standard deviation of monthly temperature.

(v) median distance of village households from the livestock market;

(vi) median distance of village households from the market for farm products;

(vii) a dummy for villages located in touristic areas, e.g. Kilimanjaro or other national parks, which plausibly have prospects for higher earnings by a few people.

The coefficients of the first stage regression of Gini on the above variables are presented in table 4.

[Insert table 4]

[24] The only year for which we could find monthly precipitation data minute by minute is 1987. Despite the lag between the year of the social capital survey and 1987, the latter is recognized as a 'typical' year in terms of precipitation (Gommes and Petrassi, 1994), and should therefore be adequate to relate to 'permanent' wealth as measured by asset ownership.

All the geographic variables have the expected sign, and many of them are statistically significant at conventional levels. The adjusted R-square of the regression is .40.

When the predicted value of inequality is used in the second stage (third column of table 4), the coefficient on Gini remains negative and significant, and larger in absolute value than that of the linear probability model without instrumentation (first column). This seems to suggest that our concerns about reverse causation are not warranted empirically and suggests the presence of attenuation bias to some degree, as it would be produced by an OLS model in which Gini were measured with error. However, the Hausman test fails to reject the joint null of weak exogeneity and no measurement error in Gini. Also, notice from table 4 that according to the Sargan overidentification test our instruments are valid.

5.3 Differential impact of inequality for the rich and the poor

An important question is whether the negative effect of inequality on participation holds across households with different wealth levels, or rather if there is a systematic tendency for richer (or poorer) individuals to drop out of groups when inequality increases. This question is addressed in table 5.

[Insert table 5]

In the first column table 5 the difference between the asset index of the respondent's household and the village average is included among the regressors. This variable measures how much richer (or poorer, if the variable takes a negative value) than the average neighbors the respondent is. An interaction term between this difference and the Gini coefficient is also included, to see if increased inequality has a differential impact on the participation of those people who are above (or below) the average.

The coefficient on 'Assets above average' shows that, *ceteris paribus*, individuals from richer households are more likely to be members of some group.[25] The negative coefficient on the interaction term, however, shows

[25] Due to the potential endogeneity problem, it is questionable that this link can be interpreted in a causal sense. Two considerations can be made to reduce the importance of the reverse causality argument here. First, the benefits of participation may be 'social' rather than 'individual' (Narayan and Pritchett, 1999). Second, our assets index refers to the entire *household*, and not just to the individual who participates.

that *it is precisely the 'richer than average' who are less likely to join or stay in groups when inequality increases.* Our interpretation for this result is that the economic gains from participation are asymmetric and that when low-income households become poorer and high-income ones become richer, the latter have less to gain from joining a group. This is quite plausible for groups such as, for example, burial societies or women's groups, where the gains from participation are not proportional to individual wealth and where the rich have less to gain from participating relative to the poor.[26]

One may argue that relative wealth captures some unobservable individual characteristics that affect participation independently of the 'relative benefits' that can be obtained from the group. In the next two columns we address this concern by exploiting the nature of 'Participatory Poverty Assessment' of the SCPS. In addition to collecting information on asset ownership, this survey asked each respondent to rank his or her household on a scale from 1 to 5, 1 being 'very poor', 2 'poor', 3 'average', 4 'rich', and 5 'very rich'. Separately, a mixed group of people from the village was asked to agree on a definition for the five categories and then rank themselves and the households included in the survey sample according to those categories.[27] Using this information, we build a measure of 'subjective overestimation' by subtracting the score assigned by the evaluation group from the score self-assigned by the respondent. A positive value of this variable indicates that the respondent perceives him or herself richer than the community does, and vice versa for a negative value.

We therefore estimate the same specification as in column 1, splitting the sample between people who underestimate or correctly rank themselves (column 2), and people who overestimate themselves (column 3). The coefficients on the individual controls and on average village assets are similar between the two sub-samples; those on inequality, instead, are remarkably different. In particular, *the negative effect of inequality on participation only holds for those people who overestimate their rank.* The coefficients on 'Assets above average' and on the interaction term with 'Gini' do not significantly differ from zero for the sub-sample of 'underestimators or neutral', while they

[26] This evidence is consistent with other empirical studies of the determinants of collective action (e.g., Gaspart et al., 1997), which show that participation —in the form of labor contribution to the provision of a public good— is higher for those individuals who can expect to benefit relatively more from the public good.

[27] The group evaluation process is described in Narayan (1997), ch.2. For an empirical analysis of the determinants of subjective welfare, see Ravallion and Lokshin (1998).

do for the sub-sample of 'overestimators'. In other words, people whose objective wealth is higher than average do not respond to increased inequality differently from poorer individuals if they underestimate or correctly estimate their relative wealth. On the contrary, people whose objective wealth is higher than average *and* who believe to be even richer than they are, respond to increased inequality by participating less.

Our interpretation of these results is the following. Each individual's decision to be part of a group depends on the expected net benefit from membership. Controlling for all other relevant factors, two individuals with the same wealth should have the same *objective* benefits and costs from participating. If one of them believes to be richer than he or she actually is relative to the rest of the village, and the other believes to be poorer, the divergence in their reaction to increased inequality could be due to the conviction that richer people have less to gain from participating in groups with poorer people. Both individuals in our thought experiment have the same wealth. The one who overestimates his or her rank believes that when inequality increases he or she will be relatively richer than the other potential members of the group and will have less to gain, hence stays out. The other one does the opposite.[28] The plausibility of this interpretation will be reinforced in section 6, where the analysis is conducted separately for different kinds of groups from which rich and poor people have different benefits.

5.4 Skewedness of the income distribution and access rules

We next investigate the relationship between the shape of the distribution of wealth and the impact of inequality under different access rules. In the top panel of table 6, our dependent variable is a dummy equal to 1 if the respondent belongs to and open access group (church, Muslim, and political groups) and 0 otherwise. In the bottom panel, the dependent variable takes the value 1 if the respondent is a member of a restricted access group (burial

[28] Another possible interpretation is that the selection into either sub-sample is endogenous, e.g. it is *because* they are not part of any group that the wealth of individuals belonging to the second sub-sample is underestimated by the evaluation team. In order for this objection to be compelling, one should argue that the bias in the team's evaluation of those who are not members of any group is *systematically* in the direction of underestimating their wealth. However, it is very likely that such a bias, if it exists, would go in the opposite direction.

society, women's group, farmers', dairy/cattle rearing group, coop, rosca) and 0 otherwise. Explanatory variables include individual and village controls, plus various measures of inequality. The idea is to exploit the fact that different measures are sensitive to different parts of the income distribution, to link our results to the predictions of the theory.

[Insert table 6]

The first column of table 6 reports the marginal probit coefficient on asset inequality when the inequality index used is the Generalized Entropy with parameter equal to 2. This measure is very sensitive to wealth differences at the top of the distribution but not at the bottom. As we can see, in the bottom panel of the table it has a positive and significant coefficient, while in the top panel it is negative and not significant. A similar pattern obtains when the Gini coefficient is used (column two). Gini is in fact most sensitive to differences around the median. These findings are consistent with the analysis of section 2.2.2, where we showed that inequality can increase participation in restricted access groups if there are significant income disparities in the middle and top part of the distribution (see figure 2a).

In the third and fourth column of table 6, instead, we use inequality measures very sensitive to the bottom part of the distribution, namely the Atkinson index (with parameter 2) and the standard deviation of the logarithm of assets. The results are now quite different: inequality is no longer significant in RA groups, while it has a negative and significant impact on participation in open access groups. This is consistent with the analysis in section 2.2.1, where we showed that increased inequality lowers participation if there are sufficient disparities in the lower part of the distribution (see figure 1b).

Another way to read these results is to look at table 6 by row. In OA groups (top panel) we fail to capture the negative impact of inequality if we use measures that do not discriminate sufficiently between situations where 'the poor are all poor' and situations where relevant differences exist within the bottom quintiles. We capture instead this impact when we use measures like Atkinson and the Standard Deviation of Logs. In RA groups (bottom panel), on the other hand, we capture the positive effect of inequality on participation when we use inequality measures which are sensitive to differences around the middle and top parts of the distribution, but we don't get significant results with the other measures. Having an accurate measure of

the skewedness of the income distribution seems to matter to understand the differential impact of inequality on participation by the rich and the poor.[29]

6 Inequality and group characteristics

Finally, we address the question of whether inequality is systematically associated with certain characteristics of the groups, both in terms of composition and in terms of functioning. Table 7 presents the probit marginal coefficients from a set of regressions of various group characteristics on assets inequality, controlling for average assets in the village, heterogeneity in schooling and in economic activity, and tribal fractionalization. Each row in the table refers to a different dependent variable (all dummies, listed on the left), and each column to a different type of group (listed on top). Each cell therefore reports the coefficient on 'Gini' from a separate regression.[30]

The types of groups considered are: religious (church, mosque, or Muslim group), political, and five of the most widespread 'economic' groups: burial societies, women's groups, farmers' associations, cooperatives and rotating saving and credit associations (ROSCAs). Burial societies are essentially a means of pooling resources to organize and pay for unexpected expenses such as funerals. There is in general only one burial society in a village, so that if rich and poor people want to participate, they will be members of the same society. While there are no fees, all members are supposed to pay and provide labor when someone dies. For this reason, we can think that poor people have relatively more to gain than rich people from being members in a burial society. Women's groups can serve a variety of functions. Some of them are essentially political organizations, others serve religious or social purposes, others still serve economic functions. Among these are microenterprise activities such as tree planting, beer brewing, and credit

[29] A similar pattern emerges when a single index is used (e.g. the Generalized Entropy or the Atkinson index) varying the values of the parameters. Results are available from the authors.

[30] In general, the questions on group composition and functioning of the different groups were asked only to those individuals who reported that they were members of those particular groups. Compared to the sample used for the membership regressions, those individuals who do not belong to any group are left out, while those who belong to more than a group are asked the same questions more than once. Each data point used in the regressions is therefore the assessment of one member of the relevant type of group in the village.

provision.[31] Again, the possibility to take part in this kind of activities is relatively less appealing for people at the top end of the income scale. Farmers' associations deal with agricultural production and fertilizers, and as such can comprise both rich and poor members. Finally, cooperatives and Rotating Savings and Credit Associations (ROSCAs) serve the usual functions described in the literature.

[Insert table 7]

The top half of table 7 shows the impact of inequality on group composition. In villages with higher inequality group members are generally more likely to belong to the same clan or tribe (row 1). They are also more likely to make a living in the same way (row 2) and less likely to be from a mixed income group (row 4), suggesting that *when inequality increases people tend to sort into more homogeneous groups*. It should be noted that these effects are significant in particular for burial societies, women's groups, cooperatives and ROSCAs, i.e. those groups where the 'rich' have less to gain if the rest of the members become 'poorer'. For these groups the likelihood that the "members are all poor" is in fact higher the higher the inequality in the village (row 5). These findings are consistent with the interpretation on the differential impact of inequality for rich and poor people given above.

In the bottom half of table 7 group composition is taken as given and the question asked is: does higher inequality harm the functioning of a group in any way? Various aspects of group functioning are considered. First of all, in more unequal communities people are less likely to respond that decisions are taken by vote (row 6). There seems to be a tendency towards hierarchic decision-making, especially in those groups —political and farmers' associations— where both rich and poor members coexist. This is of particular interest when evaluating the effect of inequality on 'participation' because, although this effect may not show up as a decrease in raw membership numbers, the nature of the groups may still be not very 'participatory'. Also, when inequality is higher members feel less "encouraged to participate" (row 7), again especially in religious and political groups, where members with different levels of wealth coexist.

When asked to evaluate the functioning of their groups, people living in villages with higher inequality are less likely to report that it is "good"

[31] The source of the information on burial societies and women's groups is Narayan (1997), p.56.

or "excellent" (row 8), although this relationship is statistically significant only for religious and political groups. Members of political groups tend to report that the disadvantage from participating in the group is that they are "misinformed" (row 9) when inequality is higher, consistently with the less democratic decision process noted above. On the other hand, for members of burial societies and women's group the main disadvantage seems to be bad economic management (row 10), e.g. misappropriation of funds by some member or unprofitable activities.

In villages with more income disparities, the likelihood that membership has increased in the past four years (row 11) is lower, which may be seen as an implicit assessment of bad performance. Finally, in more unequal areas groups themselves interact less frequently (row 12). The fact that the negative impact of inequality on many aspects of group functioning is especially significant for burial societies and women's groups is of particular concern because it reveals a potentially perverse effect of inequality on groups that are already comprised of low-income individuals.

Similar results obtain when one looks at other dimensions of 'social capital', such as trust and conflict. In particular, ceteris paribus individuals living in more unequal villages are less likely to trust others and more likely to report that there is a high degree of conflict in the community.[32]

7 Conclusions

The impact of group participation on economic performance has received increasing attention in the literature. This paper has attempted to understand the *determinants* of group membership and how groups function by looking at the role of heterogeneity, and in particular of income inequality. Using household level data from rural Tanzania, we have addressed four categories of questions.

First, does more inequality lead to less membership in groups? The answer we have suggested is 'yes'. *Ceteris paribus*, individuals living in villages with higher inequality in assets are less likely to be members of groups, and this result is robust to instrumentation of inequality with geographical variables.

Second, does inequality have a different impact on the decision to participate for poor and rich people? We have shown that when inequality

[32] These results are available from the authors.

increases, the decline in membership is higher for people who are wealthier than their village average. Furthermore, holding individual wealth constant, this effect is particularly significant for those people who overestimate their relative income rank.

Third, we ask if the shape of the income distribution matters depending on the access rules to the group. We provide suggestive evidence that inequality has a negative impact on participation in open access groups when there are wide disparities in the bottom part of the distribution, while it can actually *increase* participation in restricted access groups when the disparities are wider around the median (and the mean) of the distribution.

Fourth, we explore what group characteristics are associated with higher inequality. In terms of group composition, we find that individuals living in more unequal villages tend to sort into homogeneous income groups. In terms of decision making and group functioning, we find that when inequality increases: decisions are less likely to be taken by vote; members feel less encouraged to participate, and they are more likely to report poor group performance and misuse of funds.

Though far from definitive, the evidence presented seems certainly suggestive and calls for a deeper investigation of the mechanisms through which heterogeneity and inequality affect individual incentives to participate in groups.

References

[1] Alesina, A. and E. La Ferrara (1999), "Participation in heterogeneous communities", NBER Working Paper n. 7155, June.

[2] Abraham, A., J-M. Baland and J-P. Platteau (1998), "Groupes informels de solidarité dans un bidonville du Tiers-Monde: le cas de Kibera, Nairobi", *Non-Marchand*, 2, 29-52.

[3] Baland, J.-M., and J.-P. Platteau (1997), "Wealth inequality and efficiency in the commons", CRED Cahier de Recherche No. 193, University of Namur, Belgium.

[4] Baland, J.-M. and D. Ray (1997), "Technical complementarities and the impact of inequality on collective action: An application to the commons", mimeo.

[5] Bardhan, P. and M. Ghatak (1999), "Inequality and Collective Action Problems: Was Olson Right?", mimeo.

[6] Bergstrom, T., L. Blume and H. Varian (1986), "On the private provision of public goods", *Journal of Public Economics*, 29, 25-49.

[7] Cornes, R. and T. Sandler (1996), *The Theory of Externalities, Public Goods, and Club Goods*, Cambridge University Press.

[8] Di Pasquale, D. and E. Glaeser (1998), "Incentives and social capital: Are homeowners better citizens?", NBER Working Paper 6363.

[9] Epple, D. and T. Romer (1991), "Mobility and Redistribution", *Journal of Political Economy*, 99, 828-58.

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

[11] Gaspart, F., M. Jabbar, C. Melard, and J. Platteau (1997), "Participation in the construction of a local public good: A case study of watershed management in the Ethiopian highlands", CRED, University of Namur, Cahier de Recherche No. 181.

[12] Ghatak, M. (2000), "Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect", *Economic Journal*, forthcoming.

[13] Gommes, R. and Houssiau (1984), "Rainfall variability, types of growing seasons and cereal yields in Tanzania", Proceedings of the Technical Conference on Climate in Africa, No. 596, World Meteorological Organization.

[14] Gommes, R. and F. Petrassi (1994), "Rainfall variability and drought in Sub-Saharan Africa", Agrometeorology Series Working Paper No. 9, Rome: FAO.

[15] Gravelle H. (1998) "How much of the relation between population mortality and unequal distribution of income is a statistical artifact?" *British Medical Journal* 16:382-85.

[16] Hausman, J. (1975), "An instrumental variable approach to full-information estimators for linear and certain nonlinear models", *Econometrica*, 43, 727-38.

[17] Isham, J., D. Narayan and L. Pritchett (1995), "Does participation improve performance? Establishing causality with subjective data", *World Bank Economic Review*, 9(2):175-200.

[18] Isham, J. (1998), "How is information diffused among agricultural households?", mimeo, University of Maryland.

[19] Jehiel, P. and S. Scotchmer (1997), "Constitutional Rules of Exclusion in Jurisdiction Formation", Working Paper No. 231, Graduate School of Public Policy, UC Berkeley.

[20] La Ferrara, E. (1997), "Ethnicity and reciprocity: A model of credit transactions in Ghana", mimeo, Harvard University.

[21] Narayan, D. (1997),"Voices of the poor: Poverty and Social Capital in Tanzania", ESSD Studies and Monograph Series No. 20, The World Bank, Washington DC.

[22] ____ and L. Pritchett (1999), "Cents and Sociability: Household Income and Social Capital in Rural Tanzania", *Economic Development and Cultural Change*.

[23] Olson, M. (1965), *The logic of collective action,* Cambridge: Harvard University Press.

[24] Ostrom, E. (1990), *Governing the commons: The evolution of institutions for collective action*, Cambridge University Press.

[25] Putnam, R., with R. Leonardi and R. Nanetti (1993), *Making democracy work: Civic traditions in modern Italy*, Princeton, N.J.: Princeton University Press.

[26] Ravallion, M. and M. Lokshin (1998), "Subjective Economic Welfare", mimeo, The World Bank.

[27] Roberts, K. (1977), "Voting over Income Tax Schedules", *Journal of Public Economics*, 8, 329-40.

[28] Roberts, K. (1999), "Dynamic Voting in Clubs", STICERD Discussion Paper No. TE/99/367.

[29] Sadoulet, L. (1999), "Non-monotone Matching in Group Lending", mimeo.

[30] Sargan, J. (1958), "The estimation of economic relationships using instrumental variables", *Econometrica*, 26, 393-415.

[31] Wade, R. (1988), *Village Republics*, Cambridge University Press.

Appendix A

Proof of proposition 2

The first part of the proposition says that the group is formed by an interval of people, and in particular by the poorest $[0, h]$ individuals. This follows from the fact that, for given group size, the left-hand side of (4) is decreasing in y_j, hence in j. Therefore if (4) is satisfied for some $j \in (0, 1)$, then it must be satisfied for all $j' < j$.

The second part of the proposition gives the equilibrium size of the group, i.e. the 'identity' of the richest individual in the group. The equilibrium value of h must be such that the participation constraint is binding, thus it must solve (5). Whether the solution to (5) is *unique* depends on the functional form of $u(\cdot)$, as well as on the distribution of y_i. We can gain some insights by looking at how $U(h)$ changes with h:

$$\frac{\partial U(h)}{\partial h} = u'(h) - \frac{C}{\int_0^h y_i di}\frac{\partial y_h}{\partial h} + y_h^2 \frac{C}{\left(\int_0^h y_i di\right)^2} \tag{A.1}$$

An increase in h has three effects on the net utility of the hth individual. First of all, the quality of the good provided by the group is lower because of increased congestion, hence utility is lower (first addendum on the right-hand side of (A.1)). Second, for given tax rate, the cost paid by the hth individual is higher because his or her wealth is higher (second term on the right-hand side). Finally, for given income level, the tax rate is lower because the fixed cost C is divided among more members (third addendum in (A.1)). The sign of (A.1) will be positive if and only if this last 'tax rate effect' (positive) is stronger than the 'congestion' and 'taxable income' effects (negative).

It is easy to show that for $h \to 0$ the tax rate effect prevails, hence it is optimal to increase the size of the group beyond 0. For $h \to 1$, on the other hand, two possibilities arise. The first, relatively uninteresting, is that the tax rate effect still prevails over the others so that $U(h)$ is still increasing in h and the only stable group is the whole population. The second possibility is that for large group sizes the congestion and taxable income effects prevail, so that (A.1) evaluated at 1 is negative. In this case there will be two solutions to (5), h_A and h_B, where $h_A < h_B$. However, only h_B is *stable*. Due to $U'(h_A) > 0$, in fact, if one increased infinitesimally the size of the group to $h_A + \varepsilon$ the participation constraint would still be met for the $(h_A + \varepsilon)$th individual. The latter would be strictly better off by joining the group than by staying outside, hence group size would increase. This holds as long as h's net utility is above $\overline{u}$, i.e. until the richest member is the second root of (5), h_B, where $U'(h_B) < 0$. In the text we concentrate only on the stable equilibrium and refer to h_B simply as h^*. $\square$

Proof of Lemma 1

Suppose not, e.g. H is the union of two intervals $[h_1, h_2] \cup [h_3, h_4]$, where $0 \leq h_1 < h_2 < h_3 < h_4 \leq 1$. Note that since (4) is satisfied for the individuals in $[h_3, h_4]$, then it must be satisfied for those in (h_2, h_3), i.e. any individual in this interval would accept to join. Let $\widehat{j}$ be the median income member of group H. If $\widehat{j} \in [h_3, h_4]$, then all individuals in $[\widehat{j}, h_4]$ —a majority— would be better off by keeping group size the same and exchanging some low income members $[h_1, \widehat{h})$ for members in $[\widehat{h}, h_3)$, where $h_1 < \widehat{h} < h_3$. In fact congestion would be the same but per capita taxes would be lower. If $\widehat{j} \in [h_1, h_2]$, then all individuals in $[\widehat{j}, h_2] \cup [h_3, h_4]$ —a majority— would be better off by keeping group size the same and exchanging some low income members $[h_1, \widehat{h})$ for members in $[h_2, h_3)$, where $h_1 < \widehat{h} < h_2$ and $\widehat{h} - h_1 = h_3 - h_2$. Therefore H cannot be a RA equilibrium. $\square$

Proof of proposition 3

Let $\lambda \geq 0$ indicate the multiplier on the constraint (7), and $\mu \geq 0$ the multiplier on the constraint $h_2 \leq 1$. The first order conditions to the problem are given by

$$-u'(h_2-h_1)(1+\lambda)-\frac{C}{\int_{h_1}^{h_2} y_i di}\left[\frac{\partial y_{(h_1+h_2)/2}}{\partial h_1}+\frac{y_{h_1}}{\int_{h_1}^{h_2} y_i di}\left(y_{(h_1+h_2)/2}+\lambda y_{h_2}\right)\right]=0 \tag{A.2}$$

$$u'(h_2-h_1)(1+\lambda)-\mu-\frac{C}{\int_{h_1}^{h_2} y_i di}\ast$$

$$\ast\quad\left[\frac{\partial y_{(h_1+h_2)/2}}{\partial h_2}-\frac{y_{h_2}y_{(h_1+h_2)/2}}{\int_{h_1}^{h_2} y_i di}+\lambda\left(\frac{\partial y_{h_2}}{\partial h_2}-\frac{y_{h_2}^2}{\int_{h_1}^{h_2} y_i di}\right)\right]=0 \tag{A.3}$$

$$\lambda\left[u(h_2-h_1)-\frac{C}{\int_{h_1}^{h_2} y_i di}y_{h_2}-\overline{u}\right] \;=\; 0,\ \lambda\geq 0,\ [\cdot]\geq 0, \tag{A.4}$$

with complementary slackness.

$$\mu(1-h_2) \;=\; 0,\ \mu\geq 0,\ (1-h_2)\geq 0, \tag{A.5}$$

with complementary slackness.

Claim 1: Either $h_2 = 1$, or $\lambda > 0$, or both.

Proof of Claim 1. By contradiction. Suppose $h_2 < 1$ and $\lambda = 0$, i.e. $U(h_2) > \overline{u}$. Then the median member would be strictly better off in the group $[h_1+\varepsilon, h_2+\varepsilon]$, $\varepsilon \to 0^+$. In other words, the median voter can maintain the same congestion and save on taxes by excluding the poorest member and extending membership to the individual immediately richer than h_2. By continuity of $U(\cdot)$, the latter will have a net benefit no lower than $\overline{u}$, hence will accept to join the group. Therefore the initial group was not an equilibrium. □

Case A in the text is obtained by substituting in the first order conditions respectively, $h_2 = 1$ and $\lambda = 0$. Case B is obtained by setting $U(h_2) = \overline{u}$ and $\mu = 0$. The two cases are used as extreme exemplifications, but are not necessarily exclusive.

Notice that, while in general it is difficult to establish *uniqueness* in case B, for case A there is a simple sufficient condition. The partial derivative of

(8) with respect to h_1 is in fact unambiguously negative when $\frac{\partial^2 y_{(h_1+1)/2}}{\partial h_1^2} \geq 0$, i.e. when the distribution of income is sufficiently skewed to the right. Under this condition, if an equilibrium h_1^* exists, it is unique.

Proof of proposition 4

The mass of members will be smaller than $[0, h^*]$ if and only if the individual who was the marginal member under $\{y_i\}$ —i.e. the h^* defined by (5)— will drop out when the distribution becomes $\{\widetilde{y}_i\}$. In fact assumption (A5) and monotonicity of $U(j)$ guarantee that if h^* does not participate, no $j > h^*$ will. Consider the difference between $U(h^*)$ under the new and under the old distribution:

$$u\left(h^*\right) - \frac{C}{\int_0^{h^*} \widetilde{y}_i di}\widetilde{y}_{h^*} - \left[u(h^*) - \frac{C}{\int_0^{h^*} y_i di} y_{h^*}\right] = C\left(\frac{y_{h^*}}{\int_0^{h^*} y_i di} - \frac{\widetilde{y}_{h^*}}{\int_0^{h^*} \widetilde{y}_i di}\right)$$

Given that individual h^* was held at the reservation utility $\overline{u}$ under the initial distribution, for $y_i > 0$, $\forall i$, a necessary and sufficient condition for h^* to stay out of the group is that the above expression is negative, i.e. that (12) holds. □

The impact of inequality under restricted access

For case A, where the initial group composition before the redistribution of income is $[h_1^*, 1]$, we can establish the following.

Claim: An increase in inequality taking the form described in (A4)-(A6) will decrease participation if and only if

$$\frac{\int_{h_1^*}^1 \widetilde{y}_i di}{\int_{h_1^*}^1 y_i di} > \frac{\frac{\partial \widetilde{y}_{(h_1+1)/2}}{\partial h_1} + \widetilde{y}_{h_1^*}\widetilde{y}_{(h_1^*+1)/2}\left(\int_{h_1^*}^1 \widetilde{y}_i di\right)^{-1}}{\frac{\partial y_{(h_1+1)/2}}{\partial h_1} + y_{h_1^*} y_{(h_1^*+1)/2}\left(\int_{h_1^*}^1 y_i di\right)^{-1}} \tag{A.6}$$

Proof. Follows from (8), observing that $-u'(1-h_1)$ is monotone decreasing in h_1.

Notice that under mild sufficient conditions on the shape of the new income distribution, participation unambiguously decreases when $\widetilde{y}_{(h_1^*+1)/2} <$

$y_{(h_1^*+1)/2}$, i.e. when the original median income member is poorer under the new distribution.[33] That is the situation represented in Figure 2b.

Appendix B

B.1: Variable definitions

Age: Age of respondent (years). Source: SCPS

Assets above avg: Indiv assets - Avg assets in village. Source: SCPS

Atkinson (2): Atkinson inequality measure, parameter=2, computed from Indiv assets. Source: SCPS

Avg assets in village: mean of Indiv assets in village. Source: SCPS

Decision by vote: dummy = 1 if decisions in the group taken by vote. Source: SCPS

Disadvantage misinformed: dummy = 1 if respondent says that disadvntage from being in the group is that he/she is misinformed.

Distance farm mkt: distance of village from closest farm market. Source: HRDS

Distance livestock mkt: distance of village from closest livestock marjet . Source: HRDS

Education: highest grade attained by respondent. Source: SCPS

Gen. Entr. (2): Generalized Entropy inequality measure, parameter=2, computed from Indiv assets. Source: SCPS

Gini: Gini coefficient, computed from Indiv assets. Source: SCPS

Good or excellent functioning: dummy = 1 if respondent says that group has good or excellent functioning. Source: SCPS

Groups frequently interact: dummy = 1 if respondent says that groups interact frequently. Source: SCPS

Heterog activity: probability that two randomly drawn individuals from the same village have different economic activities. Source: HRDS

Heterog education: standard deviation of the highest grade in school obtained by household heads interviewed in the village. Source: HRDS

HH size: number of household members. Source: SCPS

[33] The sufficient condition is that, when evaluated at h_1^*, $\partial \widetilde{y}_{(h_1+1)/2}/\partial h_1 < \partial y_{(h_1+1)/2}/\partial h_1$. In fact when $\widetilde{y}_{(h_1+1)/2} < y_{(h_1+1)/2}$ this guarantees that the right hand side of (A.6) is less than 1, while the left hand side is greater than one by construction. Notice, though, that even if $\partial \widetilde{y}_{(h_1+1)/2}/\partial h_1 > \partial y_{(h_1+1)/2}/\partial h_1$, condition (A.6) may still be satisfied.

Indiv assets: index of individual assets computed by summing scores of each item owned, according to the following values: van or truck, car, motorcycle, bicycle, sewing machine = 16, radio = 8, table = 6, bed, clock or watch = 4, chair = 3, lamp = 2. If the asset was not working, these weights were approximately halved. Source: SCPS

Leaders different living: dummy = 1 if respondent says that leaders of group earn a living in a different way from the rest of members. Source: SCPS

ln(Pop) in village: logarithm of village population. Source: HRDS

Members same living: dummy= 1 if respondent says that members of group earn a living in the same way. Source: SCPS

Members all poor: dummy= 1 if respondent says that members of group are all poor. Source: SCPS

Members mixed income group: dummy= 1 if respondent says that members of group are from mixed income group. Source: SCPS

Members same clan: dummy= 1 if respondent says that members of group are from same clan or tribe. Source: SCPS

Membership increased (4yrs): dummy= 1 if respondent says that membership has increased in past 4 years. Source: SCPS

Rainfall crit. month avg: average precipitation (in mm.) in the critical months of the growing season in 1987. The geographic area of reference is the cluster (identified by latitude and longitude), with a precision of minute by minute. Sources: HRDS documentation; Tanzania Gazzetteer; Gommes and Houssiau (1984); data collected at Harvard Institute for International Development.

Rainfall crit. month std. dev: standard deviation of Rainfall in critical month (see above).

Regions: dummies for the following six areas: coast, arid, quasi-arid, plateaux, south west plains, northern highlands. Source: SCPS

Resid in village $\geq$ 10 yrs: dummy= 1 if respondent has been living in same village for at least 10 years. Source: SCPS

Sex(F): dummy= 1 if respondent is female. Source: SCPS

StdDevLogs: standard deviation of ln(Indiv assets). Source: SCPS

Temperature avg: average temperature (in C degrees) in 1987. The geographic area of reference is the cluster (identified by latitude and longitude), with a precision of minute by minute. Sources: HRDS documentation; Tanzania Gazzetteer; data collected at Harvard Institute for International Development.

Temperature std. dev: standard deviation of Temperature (see above).

Tourism: dummy = 1 if village is in a touristic area. Sources: HRDS documentation; Tanzania Gazzetteer.

Tribal fragmentation: probability that two randomly drawn individuals from the same village belong to different tribes. Source: HRDS

Very encouraged to participate: dummy= 1 if respondent feels very encouraged to participate. Source: SCPS

B.2: Summary statistics

	Mean	*Std Dev*
Age	45.14	14.96
Assets above avg	0	39.59
Atkinson	.42	.16
Avg assets in village	36.54	19.12
Decision by vote	.70	.46
Disadvantage-misinformed	.11	.31
Distance farm mkt	1.16	1.96
Distance livestock mkt	4.73	5.53
Education	4.51	3.87
Gen Entr (2)	.36	.47
Gini	.36	.11
Good or excellent functioning	.66	.47
Groups frequently interact	.72	.45
Heterog activity	1.94	1.64
Heterog education	3.62	.98
HH size	6.31	3.31
Leaders different living	.33	.47
ln(Pop) in village	12.49	.35
Members same living	.18	.38
Members all poor	.07	.25
Members mixed income group	.92	.26
Members same clan	.04	.20
Membership increased (4yrs)	.56	.50
Rainfall crit. month avg	141.8	61.8
Rainfall crit month std dev	2.54	.94
Regions		
Coast	.05	.22
Arid	.12	.33
Quasi-arid	.25	.43
Plateaux	.28	.45
South West Plain	.20	.40
Northern Highlands	.09	.28
Resid in village $\geq$ 10 yrs	.86	.35
Sex(F)	.12	.33
StdDevLogs	.78	.22
Temperature avg	26.52	1.82
Temperature std dev	3.39	1.27
Tourism	.06	.25
Tribal fragmentation	.33	.25
Very encouraged to participate	.66	.47

Table 1: Membership

	Mean
Member of any group	.72
Member of political group	.12
Member of church group	.15
Member of mosque/muslim group	.06
Member of burial society	.12
Member of women's group	.05
Member of farmers' group	.05
Member of cattle rearing group	.01
Member of coop	.02
Member of Rosca	.01

Source: author's calculations on the SCPS.

Table 2: Group characteristics

Panel A: Organizational rules

How become member?

	Born	Required	Fee	Voluntary
Religious	.36	.04	.07	.55
Political	.02	.13	.35	.51
Economic	.02	.16	.19	.63

Supported by agency? (Yes = 39%)

	Government	NGO	Foreign	Church
Religious	.03	.39	.07	.41
Political	.53	.40	.01	.02
Economic	.13	.54	.03	.07

Penalty if not pay fee?

	Expelled	Pay late	Nothing
Religious	.10	.11	.79
Political	.38	.29	.32
Economic	.39	.16	.45

Panel B: Members' heterogeneity

Ethnicity

	Same clan or tribe	Diff. tribes	Anyone
Religious	.01	.30	.69
Political	.01	.24	.75
Economic	.07	.23	.70

Make a living in the same way?

	All same	Most same	Mixed	Leaders differ
Religious	.05	.12	.83	.34
Political	.05	.08	.87	.32
Economic	.07	.14	.78	.34

Income diversity

	All rich or all poor	Little income diversity	Very mixed income group
Religious	.01	.23	.76
Political	.01	.15	.84
Economic	.01	.29	.71

Source: author's calculations on the SCPS.

Table 3: Basic regressions (marginal Probit coefficients)
(Dependent variable =1 if member of a group)

	[1]	[2]	[3]	[4]
Age	-.0004	-.0001	-.0004	-.0003
	(.002)	(.002)	(.002)	(.002)
Sex (F)	.067	.089*	.070	.097*
	(.047)	(.045)	(.050)	(.045)
Education	.014**	.014**	.008	.011*
	(.006)	(.006)	(.006)	(.006)
HH size	.01	.009	.004	.006
	(.006)	(.006)	(.007)	(.006)
Resid. in village≥10 yrs	.046	.034	.040	.05
	(.061)	(.057)	(.053)	(.059)
Avg assets in village	.003**	.005**	.003**	.005**
	(.001)	(.001)	(.001)	(.001)
ln(Pop) in village	-.031	-.09*	-.073	-.09*
	(.05)	(.052)	(.052)	(.055)
Heterog education		.003	.004	.008
		(.020)	(.023)	(.022)
Heterog activity		.044	.017	.046
		(.102)	(.108)	(.106)
Tribal fragmentation		-.112	-.121	-.114
		(.084)	(.091)	(.085)
Gini		-.436**	-.378*	-.46**
		(.181)	(.218)	(.193)
Indiv assets			.002**	
			(.001)	
Indiv assets^2			-5.3e-06**	
			(2.3e-06)	
1st assets quartile				-.090*
				(.050)
4th assets quartile				.053
				(.049)
REGIONS	Yes	Yes	Yes	Yes
No. obs.	624	624	581	581
Pseudo Rsq	.06	.07	.07	.07
Observed P	.73	.73	.73	.73
Predicted P	.75	.75	.75	.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 village level.

Table 4: instrumenting Gini

Dependent variable = 1 if member				
	Linear Prob	*Std error*	*2SLS*	*Std error*
Gini	-.450**	(.169)	-.696*	(.401)
INDIV & VILLAGE CONTROLS[(a)]	Yes		Yes	
REGIONS	Yes		Yes	
No. obs.	573		573	
R sq.	.07		.07	
1st stage: Dependent variable = Gini				
Rainfall crit. month avg.			-.0001	(.0001)
Rainfall crit. month std.dev.			.009	(.006)
Temperature avg.			-.014**	(.004)
Temperature std. dev.			-.0005	(.004)
Distance livestock mkt			.003**	(.001)
Distance farm mkt			-.009**	(.002)
Tourism			.229**	(.036)
INDIV & VILLAGE CONTROLS[(a)]			Yes	
REGIONS			Yes	
No. obs.			573	
R sq.			.40	
Hausman (p-value)			.54	
Sargan (p-value)			.16	

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 village level.

(a) Individual & village controls: all those listed in Table 3, column 1.

Table 5: Inequality and relative wealth
(Dependent variable =1 if member of a group)

	Full sample [1]	*Underestimators or neutral* [2]	*Overestimators* [3]
Assets above avg.	.004*	.001	.01**
	(.002)	(.002)	(.004)
Gini	-.488**	-.290	-.908**
	(.219)	(.289)	(.331)
Assets above avg * Gini	-.006*	-.001	-.016**
	(.003)	(.005)	(.007)
INDIV & VILLAGE CONTROLS[(a)]	Yes	Yes	Yes
REGIONS	Yes	Yes	Yes
No. obs.	581	398	178
Pseudo Rsq	.07	.07	.11
Observed P	.73	.73	.74
Predicted P	.75	.74	.77

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 village level.

(a) Individual & village controls: all those listed in Table 3, column 1.

Table 6: Access rules and skewedness of income distribution

	Inequality measure used:			
	Gen.Entr.(2)	*Gini*	*Atkinson(2)*	*StdDevLogs*
OA groups				
Ineq.	-.068	-.403*	-.255*	-.195**
	(.063)	(.246)	(.134)	(.097)
[+1 std.dev.].	[-.03]	[-.03]	[-.04]	[-.03]
No. obs.	581	581	581	581
R sq.	.10	.10	.10	.10
RA groups				
Ineq.	.153**	.547**	.206	.145
	(.052)	(.256)	(.137)	(.099)
[+1 std.dev.]	[.07]	[.04]	[.03]	[.02]
No. obs.	581	581	581	581
R sq.	.08	.08	.08	.08

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 village level.

All regressions include individual, village, and region controls as in column 2 of table 3.

Table 7: Group characteristics, by type

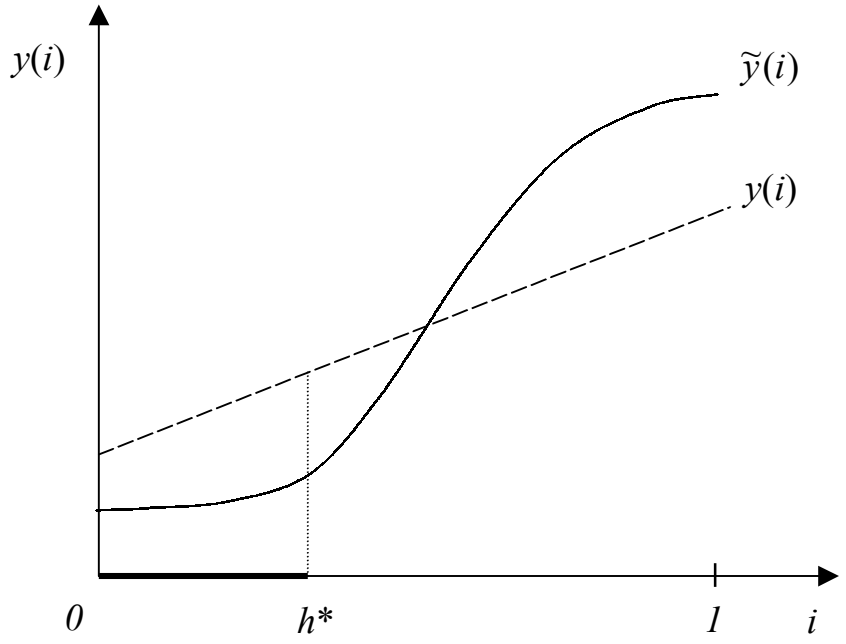
		Type of group:					
		Religious groups	Political groups	Burial societes	Women's groups	Farmer's groups	Coops & RO
	Dependent variable:						
	GROUP COMPOSITION:						
[1]	Members same clan	0.054	-0.031*	0.282*	0.124	0.293**	-0.002
		(1.39)	(-1.75)	(1.86)	(1.23)	(2.50)	(-0.44)
[2]	Members same living	0.176	-0.323	0.415*	0.643*	0.544	0.284
		(0.85)	(-1.41)	(1.89)	(1.82)	(1.21)	(0.84)
[3]	Leaders different living	0.069	-0.454*	-0.249	0.417	0.159	0.327
		(0.34)	(-1.67)	(-0.85)	(0.79)	(0.38)	(0.56)
[4]	Members mixed income group	-0.305**	-0.045	-0.469**	-0.483*	-0.315	-0.014**
		(-2.83)	(-0.40)	(-4.65)	(-1.67)	(-0.86)	(-2.51)
[5]	Members all poor	0.234**	-0.071	0.469**	0.433	0.315	—
		(2.39)	(-0.57)	(4.65)	(1.55)	(0.86)	—
	GROUP FUNCTIONING:						
[6]	Decision by vote	-0.166	-0.823**	-0.214	-0.059	-0.714*	0.119
		(-0.65)	(-2.72)	(-1.17)	(-0.12)	(-1.70)	(0.24)
[7]	Very encouraged to partecipate	-0.351	-0.642*	0.129	-0.349	0.011	-0.301
		(-1.56)	(-1.85)	(0.78)	(-1.05)	(0.03)	(-0.55)
[8]	Good or excellent functioning	-0.711**	-0.788*	-0.168	0.135	-0.373	-0.376
		(-3.45)	(-1.93)	(-0.95)	(0.29)	(-0.69)	(-0.57)
[9]	Disadvantage: misinformed	-0.122	0.162**	—	—	—	0.171
		(-0.65)	(2.60)	—	—	—	(0.36)
[10]	Disadvantage: econ management	0.122	0.071	0.315	1.37**	—	0.479
		(0.47)	(0.14)	(1.66)	(1.98)	—	(0.68)
[11]	Membership increased (4yrs)	-0.309	-0.715	-0.297	-0.819**	-0.312	-0.997
		(-1.08)	(-1.55)	(-0.83)	(-1.97)	(-0.73)	(-1.48)
[12]	Groups frequently interact	-0.337	-0.465	0.435*	-0.523*	0.139	-0.455
		(-1.24)	(-1.55)	(1.93)	(-1.80)	(0.37)	(-0.85)

Notes:

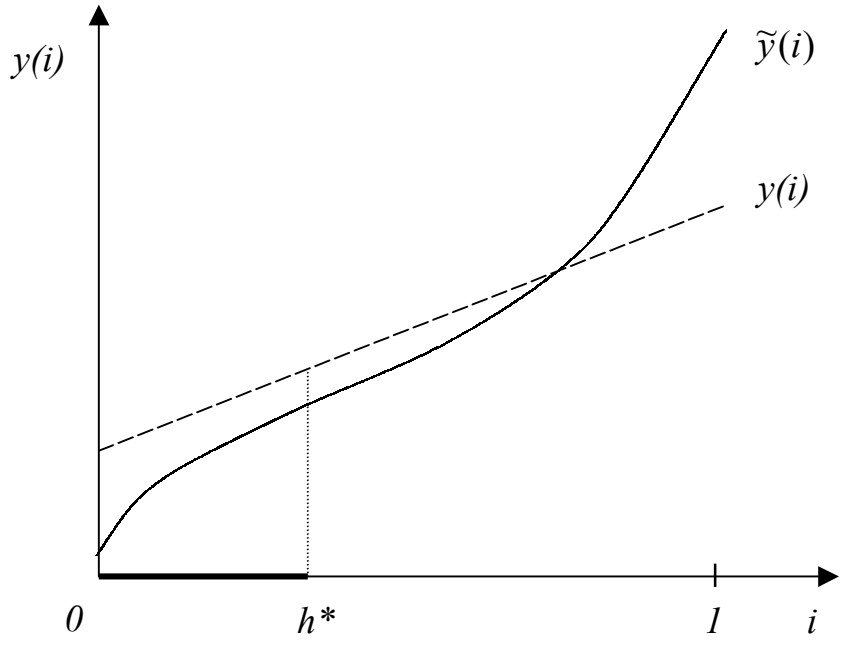
* denotes significance at the 10 percent level, ** at the 5 percent level.

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

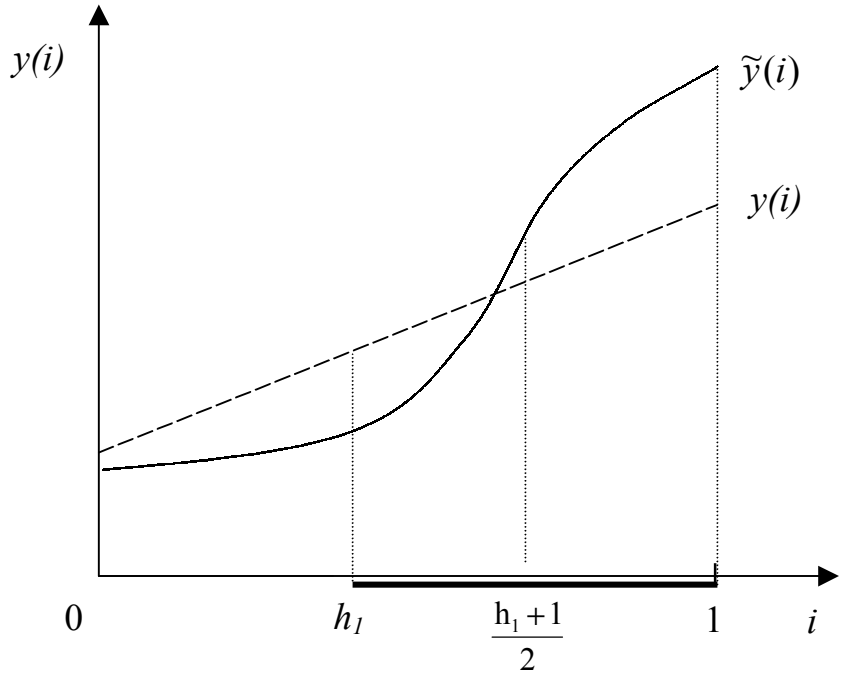
(a)



(b)



(a)



(b)