www.dagliano.unimi.it

CENTRO STUDI LUCA D'AGLIANO

DEVELOPMENT STUDIES WORKING PAPERS

N. 193

December 2004

Ethnic Diversity and Economic Performance

*Alberto Alesina**
*Eliana La Ferrara***

*Harvard University, NBER and CEPR
**Bocconi University and IGIER

Ethnic Diversity and Economic Performance

Alberto Alesina, Harvard University, NBER and CEPR
Eliana La Ferrara, Bocconi University and IGIER*

Abstract

We survey and assess the literature on the positive and negative e.ects of ethnic diversity on economic policies and outcomes. Our focus is on communities of di.erent size and organizational structure, such as countries, cities in developed countries, villages and groups in developing countries. We also consider the endogenous formation of political jurisdictions and we highlight several open issues in need of further research, in particular the endogenous formation of ethnic identity and the measurement of ethnic diversity.

**We thank Roger Gordon, David Laitin, John McMillan and two anonymous referees for very useful comments. Angelo Mele provided excellent research assistance. Alesina is grateful to the NSF for financial support through a grant to the NBER. This paper is part of the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme. This article reflects only the authors' views and the Community is not liable for any use that may be made of the information contained therein.*

1 Introduction

New York and Los Angeles are amongst the two most troubled American cities in terms of racial relations; at the same time they are constant producers of innovation in the arts and business. The United States itself is an economically successful melting pot, but many of its social problems are related to racial and ethnic cleavages. The "tragedy of Africa" is, according to many, largely a result of ethnic conflict, which is indeed pervasive in many parts of the developing word. So, what are the pros and cons of "diversity", being that racial, ethnic, religious or linguistic?

The potential costs of diversity are fairly evident. Conflict of preferences, racism, prejudices often lead to policies which are at the same time odious and counter productive for society as a whole. The oppression of minorities may lead to political unrest or even civil wars. But a diverse ethnic mix also brings about variety in abilities, experiences, cultures which may be productive and may lead to innovation and creativity. In what follows we try to highlight the trade off between the benefits of "diversity" and the costs of heterogeneity of preferences in a diverse multi-ethnic society.

In order to bring more evidence to bear on this question we plan to examine jointly two strands of the literature that have proceeded in a parallel way: one on cross country comparisons, and one on local communities. The latter is itself split into two sub areas with little communication between the two, namely the public and urban economics literature on US cities on the one hand, and the development literature which focuses on groups and local communities on the other. Within both strands of the literature, one approach takes the size and number of jurisdictions (countries or localities) as given, and studies the effects of different degrees of diversity on quality of government, economic policies, growth, unrest, crime, civil wars etc. A second and less developed approach focuses on the fact that the number, size and shape of jurisdictions (countries or localities) is itself endogenous, namely determined by the interaction of individuals and groups.

So, is diversity "good" or "bad"? Fragmented societies are often more prone to poor policy management and pose more politico economic challenges than homogenous ones; it is easy to find rather voluminous evidence on this point. However, to the extent that not all diverse societies are a failure but in fact some work much better than others, and in fact rather well, it is important to understand why and how. We propose a simple theoretical framework in which the skills of individuals from different ethnic groups are complementary in the production process for a private good, implying that more diversity translates into increased productivity. On the other hand, individual utility also depends on the consumption of a shared public good, and since different ethnic groups may have different preferences on the type of public good to provide, increased diversity lowers the utility from public good consumption. The size of the public sector and the number of ethnic groups are determined by the trade off between these two forces. We verify the consistency of this theory using repeated cross sectional data on countries and localities in the US and we find that, while ceteris paribus increases in ethnic diversity

are associated with lower growth rates, the interaction between diversity and the income level of the community under study is positive. This suggests that ethnic diversity can be beneficial (or at least less detrimental) at higher levels of development. One potential explanation for this effect is that the productivity benefits of skill complementarities are realized only when the production process is sufficiently diversified, as in advanced economies. Another -possibly complementary- explanation, is that richer societies have developed institutional features that allow them to better cope with the conflict element intrinsic in diversity, and isolate or moderate its negative effects. From the micro to the macro level, in fact, the importance of adequate "rules of the game" to manage diversity is stressed by all disciplines.[1]

Before going any further we need to clarify three points. First, we are aware that American cities are very different from African villages, but we believe that highlighting similarities and differences in the findings may shed some light on the question at hand, for instance how different levels of development and different types of racial, linguistic or religious conflict play out in the political economy of various parts of the world. Also, comparing microeconomic effects at the team level with macroeconomic effects at the country or city level is useful to get an idea on the mechanisms underlying the economic effects of diversity.

Second we need to clarify what we, and the literature which we review, mean by various terms like diversity, fractionalization, ethnicity, race etc. The empirical literature on cross country studies has typically used various measure of ethno-linguistic fractionalization. An "ethno-linguistic group" (often referred for brevity as "ethnic group") is identified by a language only in some cases and in other cases by language and skin color or other physically attributes; a variety of indexes have been suggested and we will discuss below similarity and differences. In the context of the literature on American cities, racial groups are identified with the Census definition of "race", based on five categories: (i) white, (ii) black, (iii) American Indian, Eskimo, Aleutian, (iv) Asian, Pacific Islander, and (v) other (including Hispanic). Some studies also look at "ancestry" or ethnic origin, most often defined in this context as the country of birth of the American individual (for in stance, Western European, Eastern European, Indian etc.). In the development literature on village communities, diversity is measured with reference to language/ethnic group, and more seldom through membership in different clans or tribes. We will use the terms "fractionalization" and "diversity" when we want to be generic and not refer to any particular type of identifying characteristics of the groups; we will use ethnic, racial, religious fragmentation and diversity when we want

[1]For a series of case studies on work teams supporting the notion that diverse groups function better when inter-group dynamics and allocation rules have been targeted to minimize the conflict between minority and majority members, see Susan Jackson and Marian Ruderman (1996) and Thomas Kochan, Katerina Bezrukova, Robin Ely, Susan Jackson, Aparna Joshi, Karen Jehn, Johnatan Leonard, David Levine, and David Thomas (2002). For a more macroeconomic survey on the interplay between ethnicity and institutions in the rational choice literature, see Andreas Kyriacou (2004).

to be more specific. With the term "diverse society" (or city or country) we mean a non homogenous place. The term fractionalization, on the other hand, will be directly related to a specific measure of number and size of groups: specifically, a more fractionalized place is one in which the probability the two randomly drawn individuals belong to the same group is lower. In surveying the existing literature we do not touch on the question of what identifies an ethnic group, and we take the classifications adopted by the authors as given. We briefly discuss the issue of how ethnic identity is defined in section 5, although we are aware that this topic would deserve a much more in depth treatment (which however goes beyond the scope of the present article).

Finally, and this is the third point, when reviewing a broad strand of the literature at the intersection of economics, political science, sociology and history it is important to set some boundaries. We limit ourselves to "direct" economic effects of diversity; we leave aside indirect effects that may go through civil wars, crime, revolutions etc. To put it differently we do not review in any detail the literature linking fractionalization to civil wars nor the literature on the effects of wars (civil or otherwise) on the economy. However boundaries are never neatly defined, and judgement calls are often necessary about which paper or issues fall on which side of the boundary. We tried our best but we readily plead guilty of making judgement calls that may reflect our uneven knowledge of the literature.

We proceed in the following way. In section 2 we discuss the theoretical underpinnings of the relationship between ethnic diversity and economic performance. We also sketch a simple model, which has no pretence of being innovative but illustrates clearly the pros and cons of ethnic fragmentation and sets the stage for the discussion of the literature (mostly empirical) that follows. Section 3 discusses the effects of ethnic and racial fragmentation in various types of communities holding the number and size of communities as exogenous. We examine evidence collected on three types of communities: social groups, localities and nations. Section 4 discusses the question of endogenous formation of groups, localities and nations. Section 5 concludes by discussing several open questions in this area of research. The last section draws some tentative conclusions and policy implications.

2 Theories on diversity

The goal of this section is to briefly highlight some basic economic forces underlying the relationship between ethnic diversity and economic performance. We begin by trying to offer a coherent picture of the microfoundations for this relationship. We then move to an analysis of the impact of diversity on public policies and productivity through a simple model.

2.1 Some "microfoundations"

First diversity can affect economic choices by directly entering individual *preferences*. Early work on social identity theory has established that patterns of intergroup behavior can be understood considering that individuals may attribute positive utility to the well being of members of their own group, and negative utility to that of members of other groups (see e.g., Henri Tajfel, Michael Billig, Robert Bundy and Claude Flament (1971)). A recent formalization of this concept is the analysis of group participation by Alberto Alesina and Eliana La Ferrara (2000), where the population is heterogeneous and individual utility from joining a group depends positively on the share of group members of one's own type and negatively on the share of different types.[2]

Second, diversity can affect economic outcomes by influencing the *strategies* of individuals. Even when individuals have no taste for or against homogeneity, it may be optimal from an efficiency point of view to transact preferentially with members of one's own type if there are market imperfections. For example, Avner Greif (1993) argues that traders in Medieval times formed coalitions along ethnic lines in order to monitor agents by exchanging information on their opportunistic behavior. Ethnic affiliation helped sustain a reputation mechanism in the presence of asymmetric information. But strategies can be conditional on one's ethnic identity also in the presence of perfect information. Eliana La Ferrara (2003a) shows that when contracts cannot be legally enforced (and therefore have to be "self-enforcing"), membership in ethnic groups allows an enlargement of the set of cooperative strategies that can be supported. The reason is that both punishment and reciprocity can be directed not only at the individual but to other members of his/her group. A similar reasoning is proposed by James Fearon and David Laitin (1996) to explain inter-ethnic cooperation. Using a social matching model, they show that cooperation amongst different ethnic groups can arise through either of two channels. The first, which they label "spiral" equilibrium, occurs when conflict between individuals is expected to spiral to the whole groups, and the fear of this induces cooperation on the equilibrium path. The second, labelled "group-policing" equilibrium, is one in which deviations by members of other ethnic groups are ignored, and each group sanctions deviations by its own members. Finally, an interesting application of the "strategic" role of diversity concerns the incentives to innovate through individual initiative. Two recent studies shed light on this point. Eli Berman (2000) uses a club good model with social interactions to argue that small communities can ensure the loyalty of their members by "taxing" activities outside the club (e.g., innovations). Tanguy Bernard, Alain de Janvry and Elisabeth Sadoulet (2004) also study a context in which local communities try to restrain innovations by sub-groups, but once enough diversity exists within a local community, "differentiating organizations" may actually

[2] A "business counterpart" of the preference element in diversity may be seen in the theories of "customer discrimination". According to these theories, businesses whose employees reflect the ethnic mix of the communities in which they operate perform better than those who do not, as customer satisfaction increases from interacting with service providers similar to themselves.

emerge.[3]

Finally, diversity may enter the *production* function. People differ in their productive skills and, more fundamentally, in the way they interpret problems and use their cognitive abilities to solve them. This can be considered the origin of the relationship between individual heterogeneity and innovation or productivity. An elegant formalization of this concept is provided by Lu Hong and Scott Page (1998), who prove two key results on this point. First, a group of "cognitively diverse" problem solvers can find optimal solutions to difficult problems; second, under certain conditions a more diverse group of people with limited abilities can outperform a more homogeneous group of high-ability problem solvers. The intuition is that an individual's likelihood of improving decisions depends more on her having a different perspective from other group members than on her own high expected score. At a more applied level, Alberto Alesina, Enrico Spolaore and Romain Wacziarg (2000) employ a Dixit-Stiglitz production structure where more variety of "intermediate inputs," that can be interpreted as more variety of individual skills, increases total output. Andrea Prat (2002) raises related points in the context of team theory. In teams where jobs are complementary homogeneity has positive effects and the other way around. Gianmarco Ottaviano and Giovanni Peri (2003) also investigate the pros of diversity in production. Diversity and related amenities affect the value of land, which enters the production function. These models, however, do not identify a trade off in the production function since more heterogeneity is always better than less. The costs of heterogeneity are outside the production function. Edward Lazear (1999 a, b) also discusses how different skills in a production unit may increase overall productivity. He identifies a trade off between the productive benefits of diversity and the possible costs that may arise due to difficult communication between people with different languages, culture etc. Purely from the point of view of maximization of productivity there is an optimal degree of heterogeneity that is identified by the optimal point of this trade off given also the nature of the production unit and its technology. This trade-off also emerges from a number of recent studies on organization performance, surveyed among others by Jackson and Ruderman (1996), Katherine Williams and Charles O'Reilly (1998) and Orlando Richard, Thomas Kochan and Amy Mcmillan-Capehart (2002). The majority of these studies rely on laboratory experiments to test the link between diversity and performance, and generally find a positive effect of racial and gender diversity on creativity and task completion. For example, Charles O'Reilly, Katherine Williams and Sigal Barsade (1997) analyze 32 project teams and find that more diversity leads to more conflict and less communication, but controlling for the latter it also leads to higher productivity. The few existing studies carried on real organizations offer a more complex picture. Summarizing the findings of their re-

[3]Note that all these studies indicate a role for ethnic diversity in facilitating the enforcement of informal contracts, but this role becomes less and less important as more transactions are mediated through the market, i.e. as the level of economic development of the society increases. We return to this issue below.

cent research projects, Kochan et al. (2002) argue that no significant direct relationship between diversity and team performance emerges unless the specific organizational context and policies are accounted for. The importance of the organizational context to minimize conflict within the group is also highlighted in a model by Michelle Garfinkel (2004). She considers a setting in which individuals compete in a winner-take-all market. The formation of groups reduces conflict over the resource compared to the case in which all individuals compete with each other, but it also adds another dimension of rivalry within the groups, as members need to agree on how to distribute the resource between themselves. The availability of institutional mechanisms to solve the "within group" allocation more or less easily than the "between groups" one is a crucial factor in understanding the role of diversity in teams.

2.2 Costs and benefits of diversity: a simple model

2.2.1 Private goods, public goods and diversity

We provide here a simple model that helps to clarify the pros and cons of ethnic diversity and offers a useful perspective for a review of the empirical literature. Consider a community, say a country, with K different types of individuals, for a total population of N individuals. For simplicity, every group has the same size $s = N/K$. Output produced in the country is given by:

$$Y = Nf(x; K) \tag{1}$$

where x is the fixed amount of input, say labor, equal for every person and type. We assume that $f_x > 0$, $f_{xx} < 0$, where subscripts denote partial derivatives. If variety in production is "good" then we have $f_K > 0$, $f_{KK} < 0$. This is the simplest possible way of capturing a benefit from variety in production, since per capita income is increasing in the number of different types in the population. We also assume complementarity, i.e. $f_{xK} > 0$. [4]

Output can be either consumed privately or used to produce a public good, g. Individual utility is separable in the private and public good and is given by:

$$U^i = u(c_i) + v(g, K) \tag{2}$$

where $u_c > 0$, $u_{cc} < 0$, $v_g > 0$, $v_{gg} < 0$. We also assume $v_K < 0$, $v_{KK} \leq 0$ and $v_{gK} < 0$, implying that the enjoyment of the public good is decreasing with the number of types in the population. These preferences can be rationalized in two ways. One is that sharing a public good implies contacts between people, and contacts across types produce negative

[4] This can be considered a reduced form simplification of a production function with a variety of inputs a la Dixit-Stiglitz as used by Alesina, Spolaore and Wacziarg (2000). While we could derive our results using the fully specified Dixit-Stliglitz productioin function, we find it more useful for the reader to offer a simpler work horse model.

utility, as in Alesina and La Ferrara (2000). A different rationalization follows Alberto Alesina and Enrico Spolaore (1997). They distinguish between different kinds of public goods in a context where the public good chosen is the one preferred by the median voter. The larger the number of types in the population the larger the average distance between each type and the median one that chooses the public good.[5]

The budget constraint implies:

$$g = tNf(x, K) \tag{3}$$

where t is the income tax rate. Suppose that a benevolent government can choose the tax rate, for given number of types. The problem is:

$$\begin{aligned} \max \quad & N\left[u(c) + v(g, K)\right] \\ \text{s.t.} \quad & Nc + g = Nf(x, K) \\ & g = tNf(x, K) \end{aligned}$$

The first order condition that defines an interior solution for this problem is:[6]

$$Nv_g\left(tNf(x, K); K\right) = u_c\left((1-t)f(x, K)\right). \tag{4}$$

This equation implies that the marginal benefit of taxation in terms of production of public good (LHS) has to be equal to the marginal cost of taxation in terms of reduction of private consumption (RHS). Distortionary taxes on, say, the labor supply would not change the basic message. Standard applications of the implicit function theorem lead to the following result:

$$sign\ \{dt/dK\} = sign\ \left\{tN^2 v_{gg} f_K + Nv_{gK} - (1-t)u_{cc} f_K\right\}. \tag{5}$$

Note that we are holding N constant to isolate the effects of more fragmentation without changing total population size. While the sign of (5) is generally uncertain, $dt/dK < 0$ as long as v_{gK} is large enough in absolute value. The intuition for this condition is clear: as long as the marginal benefit of public consumption goes down substantially with an increase in ethnic fragmentation, then a larger K means that the social planner will choose a smaller size of the public good in favor of more private good. The only force working against this effect is the decreasing marginal utility of the private good. In what follows we refer to the case where $dt/dK < 0$ as our "benchmark" case. This benchmark implies that, as a country becomes more ethnically fragmented, it may become more productive but it will choose to have a smaller size of government (remember that $t = g/Y$, thus t represents the size of government). More generally

[5] In Alesina and Spolaore (1997) there are multiple kinds of public goods to be supplied with fixed quantities. More generally, both the type and the quantity of public goods could change.

[6] Note that the second order condition is always satisfied under our assumptions.

private consumption will increase but public consumption will decrease. This is an empirical implication which we shall test below.[7]

Another application of the implicit function theorem leads to the following result:

$$sign\ \{dt/dx\} = sign\ \left\{tN^2 v_{gg} - (1-t)u_{cc}\right\}. \tag{6}$$

Note that if $dt/dx < 0$, then, a fortiori, $dt/dK < 0$ in expression (5). However one could have $dt/dx > 0$ and $dt/dK < 0$, i.e., it is perfectly possible that the size of government is increasing with the level of individual productivity x, and thus in GDP, but decreasing in fragmentation.

We next allow the social planner to choose not only the level of taxation but also the optimal number of types, K, again holding the size of the country constant. The first order condition for an interior solution with respect to K is:

$$u_c(\cdot)(1-t)f_K + v_g(\cdot)tNf_K = -v_K(\cdot) \tag{7}$$

and the second order conditions are satisfied. Condition (7) equalizes the marginal benefit of letting in an additional group in terms of increased productivity and tax revenues (LHS) to the marginal cost of having one more groups to share the public good with (RHS).[8]

An interesting comparative statics exercise regards the effect of an increase in x (individual level of input/productivity) on the optimal number of groups. Straightforward algebraic computations applying the implicit function and the envelope theorem establish, under fairly general conditions, the following:[9]

Remark 1 *If f_{xK} is positive and sufficiently large, then $dK/dx > 0$.*

A higher level of per capita input raises the benefits of variety and increases the optimal number of groups if the cross partial f_{xK} is large enough. In this case, as the level of individual output increases, the productivity gains from variety go up as well, so the benefit from more ethnic fragmentation are increasing with the level of per capita output. This is an empirically plausible implication: the benefits of skill differentiation are likely to be more relevant in more advanced and complex societies.

[7]Note of course that if $f_k < 0$, then income per capita would go down as fragmentation increases and the allocation of this lower total output between private and public consumption would depend on the marginal benefits of the two.

[8]Note that if there were no benefits in production from variety ($f_K \leq 0$), then the solution would be at a corner with the minimum number of groups, possibly 1, that is, a fully homogeneous society. The first order condition for the choice of t is of course unchanged.

[9]Intuitively, these conditions require that the indirect effects of a a change in t caused by a change in K do not override the direct effect of a change in x on K. Details are provided in a theoretical appendix available from the authors upon request.

2.2.2 On the number of jurisdictions

The same theoretical framework can be extended to analyze the optimal number of jurisdictions, along the lines of Alberto Alesina and Enrico Spolaore (1997, 2003).[10] We can think of the optimal size of a jurisdiction (say a country) as emerging from the trade off between the benefits from variety and the costs of heterogeneity. In the language of our model above we could think of a social planner choosing the size of the country with the goal of maximizing total welfare subject to the trade off between benefits and costs of heterogeneity. The larger the (positive) effect of variety in production and the lower the utility costs of heterogeneity, the larger the size of the jurisdiction chosen by the social planner. More specifically, imagine that each place specializes in the production of an intermediate good and assume that crossing the boundaries of political jurisdictions is costly. Then the more beneficial is variety in production the more efficient is to employ many different intermediate goods and the larger the total border crossing costs. Should we then expect larger countries to be more productive because they have more variety? The answer depends on the structure of international trade. Trade flows are generally hampered by country borders even when the trade regime is free and without explicit policy-induced barriers.[11] With severe trade restrictions, the size of a country would be very important for productivity; on the other hand with free trade countries can be small, enjoy the benefit of homogeneity as far as public goods provision is concerned but take advantage of diversity in production (and consumption) by means of international trade. One implication of this is that the effects of the size of countries on economic success is mediated by the extent of freedom of trade, a result empirically supported by Peter Katzenstein (1985), Ronald Rogowski (1987), Alberto Ades and Edward Glaeser (1995), Alesina, Spolaore and Wacziarg (2000) and Francisco Alcala and Antonio Ciccone((2004) amongst others. Note that some diversity in a country may favor trade as well. For instance a certain ethnic minority in country A can be a "link" with a country B where that ethnic group is a majority, therefore facilitating trade between A and B. The extent to which ethnic and cultural relations facilitate trade and more generally economic integration is well established; see for instance Samuel Huntington (1998) for an informal discussion and James Rauch (2001), and Allesandra Casella and James Rauch (2001, 2003) for models and empirical evidence. An important question is under what conditions the optimal solution would be reproduced by the "market" without a social planner, a question explored in depth by Alesina and Spolaore

[10]For an earlier contribution on endogenous jurisdictions which is however not as focused on diveristy of prefences see David Friedman (1977). Also, the literature on the size of nations has some connection with "club theory" (see in particular James Buchanan and Roger Faith (1987)). However, there are two key differences. One is that club theory generally does not consider the geographic distribution of club members; second, club theory emphasizes the issue of congestion which is absent from the size of nations literature.

[11]A large literature has documented this fact; see Alberto Alesina, Robert Barro and Silvana Tenreyro (2003) for a review and some new results.

(2003). In general the answer is no and the equilibrium size of jurisdictions varies as a function of the type of political institutions and rules available to change borders, a set of issues that we do not pursue here.[12]

The same ideas about endogenous border formation can be applied to sub-national governments; this in fact the goal of Alberto Alesina, Raza Baqir and Caroline Hoxby (2004) who extend the Alesina-Spolaore framework and adapt it to localities with a special reference to the US. They produce a model in which the number of localities depends on the distribution of different groups and their density in various geographical areas; once again the key is the trade off between economies of scale, heterogeneity of preferences, and in this case mediated by density of the population. The traditional literature on jurisdiction formation in the US is overwhelmingly focused on differences in income, with the rich trying to isolate themselves from the poor.[13] We do not review this literature, and we focus instead on the much more limited empirical literature that concentrates on jurisdiction formation based upon racial and ethnic cleavages.

2.2.3 Summing up the implications of the theory

The potential benefits of heterogeneity come from variety in production. The costs come from the inability to agree on common public goods and public policies. One testable implication is that more heterogenous societies may exhibit higher productivity in private goods but lower taxation and lower provision of public goods (in relative terms). The benefits in production from variety in skills are more likely to be relevant for more advanced societies. While in poor economies ethnic diversity may not be beneficial from the point of view of productivity, it may be so in rich ones. The more unwilling to share public goods or resources are the different groups, the smaller the size of jurisdictions. [14] The larger the benefits in production from variety, the larger the size. If variety in production can be achieved without sharing public goods, different groups will want to create smaller jurisdictions to take advantage of homogeneity in the enjoyment of the public good.

2.2.4 What is not in the model

Many important aspects are missing in this model, but a few in particular are worth pausing upon. First, we have not modelled explicitly the possible benefits of information

[12]Under certain conditions, Alesina and Spolaore (2003) show that when borders are decided by majority voting, in equilibrium countries may be "too small" because the benefits of size are not fully internalized by the median voter.

[13]Important references include Dennis Epple and Thomas Romer (1991) and Stephen Calabrese, Glenn Cassidy and Dennis Epple (2002) who also provide a good review of the previous literature.

[14]In principle various ethnic groups could segregate within the same jurisdiction and use different public goods. However segregation is often imperfect, may entail other costs and some public goods are by nature jurisdiction wide.

diffusion and enforcement of contracts related to ethnic diversity. These mechanisms are probably most likely to be important for developing countries or the poorer regions of middle income countries with very rudimentary forms of market institutions and legal enforcement. Fafchamps (2004) discusses a variety of models of the pros and cons of ethnically based economic networks as way to substitute for "western style" markets, and reviews much evidence with specific reference to several Sub-Saharan African countries in Africa. We will return on this point below.

Second, the model is institution free so we cannot address the question of the interaction between political institutions and diversity. Certain types of institutions may be more conducive to harmony than others. Paul Collier (2000, 2001) for instance argues that ethnic fragmentation is less disruptive in democracies because minorities feel represented. Alberto Alesina and Edward Glaeser (2004) discuss at length how in the United States proportional representation both at the federal level and at the local level was opposed in the nineteenth and early twentieth century precisely on the ground that it would allow more representation of the blacks, something that would not please the white majority. More generally the type of political institutions is endogenous to the nature of inter groups conflict, a point made in theory and also tested by Philippe Aghion, Alberto Alesina and Francesco Trebbi (2004a). First of all if a group is politically dominant it may impose a type of government that restricts freedom of the minority. On the other hand a more fractionalized society in which no group is dominant may end up with a constitution especially careful to defend the rights of minorities.

The third missing aspect in the model is that while pure public goods may be lower in more fragmented communities, the amount of publicly provided "private" goods – especially those that can be targeted to specific groups– may be larger. We can then have a positive correlation between fragmentation and ethnically based patronage.[15]

Finally in the model an increase in diversity would simply lead to smaller jurisdictions. In practice this process may be peaceful or not, leading to violent civil wars. This is an important topic that we do not investigate directly here; we refer the reader to James Fearon and David Laitin (2003), James Fearon (2002) and the references cited therein.

3 The consequences of fragmentation

In this section we review the main contributions that have linked ethnic fragmentation to economic outcomes, going from the more aggregate level (country level fragmentation and performance) to the more micro level, i.e. local jurisdictions (cities, districts) down to the level of small groups (schools, associations, cooperatives).

[15]See Alesina, Baqir and Easterly (2000) for some evidence on this point on US cities.

3.1 Countries

3.1.1 Effects on productivity and income level

Economists have started to pay attention to the effects of racial fragmentation across countries at least since a paper by William Easterly and Ross Levine (1997) who argued that, ceteris paribus, more racially fragmented countries grow less and that this factor is a major determinant of Africa's poor economic performance.[16] Several subsequent papers confirmed these results in the context of cross country growth regressions. In their overview of Africa's problems Paul Collier and Jan Gunning (1999) also place much emphasis on ethno- linguistic fractionalization (coupled with low political rights) as a major explanation for the lack of social capital, productive public goods and other growth enhancing policies.

Easterly and Levine's paper, as much of the literature that followed, used as a measure of fragmentation the probability that two randomly drawn individuals from the unit of observation (say, country) belong to two different groups. Their ethno-linguistic fractionalization (ELF) measure is an Herfindahl-based index defined as follows:

$$ELF = 1 - \sum_i s_i^2 \tag{8}$$

where s_i is the share of group i over the total of the population. This index represents the probability that two randomly drawn individuals from the population belong to different ethnic groups. The source used by Easterly and Levine to construct the ethno-linguistic groups is the Atlas Narodov Mira, originally compiled by Soviet researchers. Apart from issues of measurement (to which we return below), the robustness of Easterly and Levine's results has been called into question by Jean-Louis Arcand, Patrick Guillaumont and Sylviane Jeanneney (2000) due to problems of data missingness.[17] Despite the criticisms, subsequent estimates have taken Easterly and Levine's results as a benchmark.

Using the updated data set of Alesina et al. (2003), we now test whether the negative correlation between ethnic fragmentation and growth holds irrespective of the level of economic development or, as our model suggests, is mitigated when the benefits of heterogeneity for productivity are taken into account. Measurement and data issues are discussed below in section 5.2. A brief description of the data is contained in the Appendix. For now it is enough to note that Alberto Alesina, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg (2003) construct two indices

[16] An early and never published paper by David Canning and Marianne Fay (1993) used ethnic fractionalization as an instrument for growth.

[17] Arcand et al. (2000) note that African countries constitute only 27 of the 172 observations in Easterly and Levine's main regression, and highlight the potential sample selection bias generated by the fact that the data is missing precisely for those countries (in Africa) that have experienced slower growth.

with the same structure as above but using two different (although closely related) characterizations of groups. One which is more comprehensive and is labelled ELF and extends the Easterly and Levine index by differentiating groups that may speak the same language but have different ethnicity based upon certain physical characteristics. A striking example would be blacks and whites in the US, or various ethnic groups in Latin America all speaking the same language, often that of a former colonizer.[18] The second index relies exclusively on language spoken.

[Insert Table 1]

Table 1 shows some standard growth regressions adopting the baseline specification of Alesina et al. (2003). The dependent variable is the growth rate of GDP per capita from 1960 to 2000 and we use a SUR method in four 10-years periods. The first two columns use the more comprehensive index of fractionalization (which we label ELF), while columns 3 and 4 use the one based solely on language. Columns 1 and 3 show a baseline regression with very few controls: regional dummies, initial income and schooling. Columns 2 and 4 include additional controls, such as measures of political stability and quality of policy. One may argue (and in fact we explore this point below) that the effect of fractionalization on growth may go through exactly these variables; therefore by controlling for these variables one may underestimate the effects of fractionalization on growth. Overall table 1 shows considerable support for the negative effects of fractionalization on growth.[19] In terms of magnitude, the estimates in column 1 suggest that ceteris paribus going from perfect homogeneity to maximum heterogeneity (i.e., increasing ELF from 0 to 1) would reduce a country's growth rate by 2 percentage points per year. Increasing ethnic fractionalization by one standard deviation would reduce growth by 0.6 percentage points per year. These are quite sizeable effects. All the other controls have signs consistent with the vast literature on growth.

[Insert Table 2]

An important question is whether or not these negative effects from ethnic fractionalization on growth depend on the level of income or other features of society. In the model of section 2 we showed that under reasonable conditions on technology, fractionalization may have positive (or less negative) effects on output at higher levels of development. Table 2 adds to all the regressions of Table 1 an interaction term between fractionalization and GDP per capita. In all four regressions the interaction of initial

[18] In fact several countries in Latin America appear as more fractionalized compared to Easterly and Levine's classification using this more comprehensive index. See Alesina et al. (2003) for more details.

[19] These results are very similar to those reported by Alesina et al. (2003). The only difference is that they use both a linear and a quadratic term for initial per capita income. We use only the linear one because below we explore interactions of the initial level of income with other variables and we want to keep a simpler specification. In any case results with a quadratic term for initial income are very similar for our variables of interest.

GDP per capita and fractionalization has the expected (positive) sign, suggesting that indeed fractionalization has more negative effects at lower levels of income. In two out of four regressions, this effect is strongly statistically significant.

[Insert Table 3]

Collier (2000) argues that fractionalization has negative effects on growth and productivity only in non democratic regimes, while democracies manage to cope better with ethnic diversity. This is an important result worth exploring further. It is well known that per capita GDP and democracy are positively correlated: richer countries are more democratic. From a statistical point of view, this high correlation makes it quite difficult to disentangle the effects of democracy from the effects of the level of income on any dependent variable that might be affected by either one or both.[20] Table 3 considers the effects of the interaction of ethnic and language fractionalization with the Gastil index of democracy. This index is *decreasing* in the level of democracy so the expected sign on the interaction with fractionalization is negative. The estimates in table 3 are consistent with Collier's findings that fractionalization has less negative effects in democracies.

[Insert Table 4]

Table 4 uses the two basic specifications to try and disentangle the effects of income and democracy. Since we are adding several variables with interactions we use the simpler specification. Overall the effect of income seems more robust and more precisely estimated than the effect of democracy. However these results have to be taken cautiously given the high correlation between democracy and GDP per capita.

The punch line is that rich democracies are more capable of "handling" productively ethnic diversity. Note, however, that as argued above the variable "democracy" may be endogenous to ethnic diversity. It may be the case that racially fragmented societies that choose democratic institutions are also those in which ethnic cleavages are less deep and/or the power distribution of groups is such that none can impose a non democratic rule.

Related to the issue of how democracy interacts with ethnic conflict and with the level of development is the role played by institutions in general. William Easterly (2001) constructs an index of institutional quality aggregating Stephen Knack and Philip Keefer's (1995) data on contract repudiation, expropriation, rule of law and bureaucratic quality. He finds that the negative effect of ethnic diversity is significantly mitigated by the presence of "good" institutions, and the marginal effect of ethnic diversity at the maximum level of institutional development is actually zero. Again, the institutional variables used as explanatory factors are likely not exogenous, and more work needs to

[20]This is a well known and common stumbling bloc for anybody who has tried to estimate empirically the costs and benefit of democracy on economic variables, a vast literature that we do not review here; see Jose Tavares and Romain Wacziarg (2001) for one of the most recent and careful contributions.

be done to assess the marginal impact of institutional arrangements. Nonetheless, it seems important to take into account that, whatever the mechanisms relating ethnic diversity to economic growth, channelling diversity towards productive uses may require a particular set of "rules of the game".

3.1.2 Effects on public policies

An important prediction of the model sketched in section 2 is that the propensity to supply true public goods should be lower in more ethnically fragmented societies. The empirical literature has focused more on the "quality" than on the "quantity" of public goods partly because of data availability. In order to carefully test the implication of the model on the *quantity* of public goods provision one would need aggregate measures of the various components of the government budget for a relatively large group of countries. These data are notoriously of poor quality and not disaggregated enough. Therefore results in this area have to be taken cautiously.

Rafael La Porta, Florencio Lopez de Silanes, Andrei Shleifer and Robert Vishny (1999) and Alesina et al. (2003) show that ethnic fragmentation is negatively correlated with measures of infrastructure quality, literacy and school attainment and positively correlated with infant mortality. These correlations are very strong in regressions without income per capita (that may be endogenous to ethnic fragmentation). They lose some of their significance in regressions where on the right hand side one controls for GDP per capita.[21] In any case, neither of these studies argues that ethnic fragmentation is the only cause of "poor quality of government": La Porta et al. (1999), for instance, argue that legal origins are at least as important.

An interesting related question regards the size of public transfers rather than public goods. For a large sample of countries, Alberto Alesina, Edward Glaeser and Bruce Sacerdote (2001) show an inverse relationship between the size of government social spending and transfers relative to GDP on the one hand, and ethnic fractionalization on the other. One explanation is that altruism does not travel well across ethnic lines. Relating this point to the model above, one can view redistributive policies as a "public good" in a society that values equality as a public benefit. On this point a comparison between US and Europe seems especially suggestive. In the US welfare spending and redistributive policies are much smaller than in Europe, consistent with the fact that the US is much more racially and ethnically diverse than most countries in Continental Europe, a point explored in much detail by Alesina and Glaeser (2004). One implication of this analysis is that, to the extent that Western European countries will become more ethnically fragmented, their welfare systems may be under stress.

[21] Another variable that is correlated with racial fragmentation is "latitude" and this high correlation makes it sometimes difficult to disentangle the two effects separately, although it is unclear why latitude per se (leaving aside its possible effects on GDP per capita) should affect public policies. Often both variables (latitude and fragmentation) used together are insignificant while they are significant if used separately.

3.2 American localities: counties and cities

3.2.1 Effects on productivity and income level

American localities are an ideal setting to study the effects of ethnic fragmentation because we have many observations and excellent data, compared, say, to cross country data. Edward Glaeser, Jose Scheinkman and Andrei Shleifer (1995) have examined the growth of US cities using a similar structure to cross country growth regressions. They argue that the most appropriate measure of growth to use in this case is population growth. They note that income growth is a natural measure for cross country growth regressions because labor is relatively immobile across countries. Instead within the US the high mobility of individuals suggests that population growth is the correct measure to use to capture areas and cities that are becoming increasingly more attractive economically and as a place to live in. As Olivier Blanchard and Lawrence Katz (1992) have noted, migration within the US responds strongly and relatively quickly to income opportunities.

Glaeser, Scheinkman and Shleifer (1995) do not find any effect of racial fragmentation on the growth of cities in the sample 1960 to 1990. Their only finding concerning racial composition is that population growth is positively correlated with racial segregation in cities with large non-white communities. This result suggests that growth is higher when racial interaction is lower because of segregation. A suggestive interpretation of this result that goes back to our model may be that racial fractionalization with segregation may allow for diversity in production and lower interaction in public good consumption and social activity. Jordan Rappaport (1999) also studies population growth in cities and counties in the US. He controls for many more determinants of counties' characteristics and amenities and he finds that more racially fragmented counties grow less in terms of population.

[Insert Tables 5, 6]

In Table 5 and 6 we present some results on population growth in counties that are in line with our cross country results. For the reason discussed above we follow the literature in using population growth as our dependent variable. Table 5 reproduces for counties instead of cities and for a different sample (1970-2000) the basic specification of Glaeser, Scheinkman and Shleifer (1995). We do not find any effect of fractionalization on population growth. In Table 6 we add an interaction of initial per capita income level and fractionalization and we experiment with different samples, noting that we do not have data on income per capita before 1970. Broadly speaking, the results are consistent with the cross country results: we find that fractionalization has a negative effect on population growth in initially poor counties and a less negative (or even positive) effect for initially richer counties. This result significantly corroborates the cross country evidence in a setting where institutional and political differences should be definitely lower than across countries.

Two recent papers have looked directly at the productivity enhancing effects of diversity in American cities. Ottaviano and Peri (2003) use data on rents and wages in US cities and find that US born individuals living in more "culturally diverse" cities in their terminology (i.e., cities with a larger share of foreign born people) pay higher rents than those living in more homogeneous cities. In other words, diversity seems to have positive "amenity effects" on production and consumption. Their findings are robust to instrumenting the share of foreign born people with the distance from the closest "port of entry" into the US. Along similar lines Richard Florida (2002a,b) argues that amenities and diversity in US cities attracts human capital. He constructs indices of heterogeneity of a place that are not directly related to ethnicity but involve proportions of gay households, diversity of night life, etc. and finds that places that score higher in these indices have also higher human capital. The direction of causality is however unclear. Further work, possibly using firm and plant level data, would be useful in this area.

In subsequent work Gianmarco Ottaviano and Giovanni Peri (2004) find that the wage of white individuals, after controlling for various other determinants are higher in more diverse cities where diversity is measured with index based on main language spoken at home. They interpret this results as an indication of higher productivity with diversity. A possible criticism of this finding is that cities in which the language spoken at home is not English may be cities of recent immigration and immigrants may be attracted to faster growing cities. While they try to correct for this problem of reverse causation, it is not clear that there is way of completely disregarding this possibility.

3.2.2 Effects on public policies

A very large literature in political science and sociology examines the role of race in the history of public policies in American cities (for a recent contribution, see Nancy Burns (1994) and the references cited therein). Several papers within the economics literature have argued that public good provision is lower and/or less efficient in more racially fragmented American cities, results which are consistent with those obtained in cross country samples and in many ways follow similar procedures that involve cross cities (rather than cross country) regressions. Alesina, Baqir and Easterly (1999, 2000) show that in more fragmented cities the provision of "productive" public goods (roads, hospitals, schools etc.) is lower while the types of expenditures that more closely resembles transfers targeted to ethnic and racial groups are larger. In particular spending on roads, schools and other public goods are smaller in racially fragmented cities. Preferences over roads are likely to be different if different racial groups are located in different parts of the city and preferences over schools are certainly ethnically sensitive. Instead spending on public employment is larger in more ethnically fragmented cities, a result consistent with a use of public jobs with ethnically or racially motivated patronage. Interestingly they find that racial divisions have stronger effects than ethnic ones (where ethnicity is

identified with country of origin), a result consistent with evidence discussed in the next section on the endogenous formation of localities in the US.

A particularly important type of local public good is public education. James Poterba (1997) finds that in US states government per child spending on K-12 education decreases with the fraction of the population aged 65 and above, and that this effect is strengthened when the difference between the fraction of nonwhite population aged 5-17 and the fraction nonwhite aged 65+ is included among the controls. This suggests an interplay of demographic and racial composition effects, as if older citizens were less inclined to spend on public goods that benefit younger generations when these generations belong disproportionately to a different race. Using historical data on US states, Claudia Goldin and Lawrence Katz (1999) find a similar role for heterogeneity, be it ethnic, racial, religious or economic. Eliana La Ferrara and Angelo Mele (2003) investigate the relationship between racial segregation and spending on public education across US cities. Jacob Vigdor (2004) finds that the greater a community's racial heterogeneity, the lower its rate of response to the 2000 Census form. Response is interpreted as a local public good in that the amount of federal funds allocated to the community depend on its response rate. Alberto Alesina and Eliana La Ferrara (2002) also show how redistributive policies are deeply affected by racial politics. In more racially fragmented communities people are less willing to redistribute income because the white majority feels that redistributive flows would favor a racial minority. Survey evidence suggests that those respondents who express attitudes less favorable to racial integration are also more averse to government intervention on redistributive matters. Erzo Luttmer (2001) finds that respondents' in surveys show more support for welfare policies as the share of population of the same race of the respondent increases in his/her community.

One of the reasons why public policies in racially fragmented communities are worse is that social capital is lower. Two key aspects of social capital are participation in social activities or groups and trust.[22] Using data from the General Social Survey (GSS), Alesina and La Ferrara (2000) provide evidence that in American cities individuals of different races are less willing to participate in social activities in racially mixed communities. There are two non mutually exclusive explanations. One is that members of different racially identified groups have different preferences on what a group should do or how it should be run, and the other is that there is a cost in sharing a group with different races simply because of aversion to racial mixing. Alesina and La Ferrara (2002) show that in American cities individuals living in more racially fragmented communities have a lower propensity to trust other people, while they do not exhibit lower levels of trust towards institutions. Similar results were later obtained by Costa and Khan (2003b). Interestingly, all these authors also show that income inequality reduces participation and social capital but the effect of racial conflict seems stronger. Experimental evidence on trust and participation included in Edward Glaeser, David Laibson,

[22]For survey on the relationship between ethnic diversity and social capital, see Dora Costa and Matthew Kahn (2003a).

Jose Scheinkman and Christine Soutter (2000) is also consistent with these results: even in experimental settings and amongst a relatively homogeneous group of individuals (in terms of education), trust does not travel well across racial lines.

In summary, looking at US cities, there seem to be two empirically relevant types of diversity. One is ethnic fractionalization, where ethnic groups are defined according to the five Census categories (White; Black; American Indian - Eskimo - Aleutian; Asian - Pacific Islander; and Other). This type of fractionalization seems to be associated with poor public policies, low trust and low city growth. On the other hand, there are measures of cultural diversity based on language, life style and attitudes. These measures seem to be associated with a positive evaluation of amenities and higher productivity. The two sets of results are not contradictory because the two indices of diversity are very imperfectly correlated. Most notably, an index based on language does not distinguish between whites and blacks and in this sense may underestimate diversity in racially heterogeneous cities. But it is also possible that, to the extent that a totally white city is very diverse in terms of language spoken at home, the language-based diversity index can overestimate heterogeneity compared to the race-based one. The choice of the notion of cultural diversity to be used is ultimately a matter of context, and of the particular problem to be analyzed.

3.3 Village communities in developing countries

3.3.1 Effects on productivity

A particularly relevant setting in which to study the productivity effects of ethnic diversity is that of developing economies. The reason is that a large share of economic transactions occur outside the boundaries of the formal sector, and need to be supported by enforcement schemes similar to those described in section 2.1. Although direct empirical evidence is seldom available, a number of recent studies on developing countries allow to draw preliminary inference on the impact of diversity on productivity and economic performance at the micro level.

Manufacturing firms in Africa have been studied by several authors. Arne Bigsten, Peter Kimuyu and Karl Lundvall (2000) use a data set on Kenyan firms in the food, wood, textile and metal industries, and examine what factors account for the choice of going formal, and for the degree of economic efficiency. They find that kinship and community ties among entrepreneurs of Asian origin reduce the barriers to entry in the formal sector, so that even after accounting for differences in education, "African" firms are much more likely to be informal at start-up.[23] In addition to the advantage that the "formal status" gives to Asian-managed firms (e.g., in terms of access to formal credit),

[23]The distinction between the Asian and the African business community in Kenya goes back to the colonial period, when the British organization rested upon a three-tier society in which traders and businessmen of Indian origin occupied an intermediate position.

Asian firms are shown to be more capital intensive and more productive.[24]

Marcel Fafchamps (2000) focuses on the relationship between ethnicity and access to credit for manufacturing firms in Kenya and Zimbabwe and finds that, after controlling for observable firm characteristics (e.g., size) African firms are *not* discriminated against in the allocation of bank overdraft and formal loans. The way in which ethnicity seems to make a difference is by offering network relationships that improve access to supplier credit. This in turn affects productivity and allows firms to remain in the market in the presence of negative shocks, as a common way to absorb cash flow variations is to delay payments to suppliers. The relationship between trade credit and productivity is further explored by Raymond Fisman (1999, 2003), who shows that entrepreneurs of Asian and European origin are more likely to obtain supplier credit, and that firms that do not have access to supplier credit have a higher probability of facing inventory shortages and have lower rates of capacity utilization.

Marcel Fafchamps (2004) reviews more broadly the evidence on the effects of ethnic networks on economic performance and on the allocation of credit in several African countries. He identifies two effects, one positive and one negative. On the one hand, ethnically based networks can provide insurance, facilitate transactions, increase trust and simply substitute for rudimental market institutions. On the other hand, the same networks may lead to bias against various groups. For instance, Fafchamps finds "the presence of an ethnic bias in trade credit usage amongst manufacturing firms in sub Saharan Africa. The direction of the bias is generally detrimental to entrepreneurs of African descent" (p. 368). This author, going beyond some of his previous work, tries to identify whether this amounts to active discrimination due to the lack of connection with appropriate networks. At least for Zimbabwe and Kenya, he finds that this is indeed the case for black and female entrepreneurs. Interestingly, the same author finds virtually no effects of ethnically based networks on agricultural trade (chapter 19), a result that leads to be cautious about blanket generalizations across all Africa and all trades on the role of ethnic networks.

Despite their focus on employer-level ethnicity as opposed to ethnic *fractionalization*, the above studies potentially bear interesting implications for the relationship between ethnic diversity at the community level and firm performance. In fact, for a given level of credit supply, the greater the number of ethnic groups in the business community, the lower the chances that supplier credit is allocated efficiently if the criterion is purely ethnic affiliation, which can ultimately harm economic productivity.[25]

An explicit focus on ethnic heterogeneity and economic performance is in the study by Eliana La Ferrara (2002b). She uses an original data set on production cooperatives in the informal settlements of Nairobi, and has information on all members of the surveyed

[24]Further evidence on the relationship between ethnic networks and access to credit in the Kenyan manufacturing sector is provided by Tyler Biggs, Mayank Raturi, and Pradeep Srivastavaet al. (2002).

[25]This obviously depends on the way in which network structure endogenously responds to the ethnic composition of the community, a point we address in section 4.3.

groups, which allows to construct exact measures of group composition in terms of income, education, age and ethnicity. She finds that ethnicity matters for gaining access to group resources, especially in the form of cheap loans: members who share the same ethnicity as the chairperson are 20 to 25 percentage points more likely to borrow from the group or from other members. Ethnic heterogeneity also seems to influence the organization of production: members of more ethnically heterogeneous groups are less likely to specialize in different tasks and more likely to all do the same job. Prima facie, this seems in contrast with the assumption of positive complementarities in production among different ethnic groups, which we made in our theoretical framework. However, the difficulty of allocating different tasks across different groups in the particular context of La Ferrara's study stems from the governance structure of the groups, and on the lack of transparent allocation rules that characterizes most of the groups under study. For example, ethnically fragmented groups more often adopt remuneration schemes in which every worker gets the same fixed amount, rather than being paid on the basis of the amount of work put in. These choices on division of labor and wage structure may be due to the relative difficulty of reaching consensus on "unequal" task allocations and remuneration schemes in ethnically heterogeneous groups. In this case, the consequences of ethnic diversity on differential access to inputs get reinforced by its impact on within-firm organization of production. Viewed in this light, our hypothesis that the beneficial effects of diversity are stronger in more economically advanced societies is actually fully compatible with these findings.

A recent application to lending groups is provided by Dean Karlan (2003). He uses data on members of a Peruvian micro finance organization, and exploits the random selection of people into groups to estimate the effect of group composition on repayment performance. He finds that members of more "homogeneous" groups, both in terms of geographical proximity and of cultural affiliation, are more likely to save and to repay their loans. Interestingly, "cultural" homogeneity is measured through a score attributed by enumerators to each respondent on the basis of his/her language, dress and hair style. These findings suggest that monitoring and enforcement within groups are easier the greater the social affinity among their members, as argued in section 2.1.

Finally, although very limited evidence exists on the subject, ethnic diversity can have an impact also on agricultural productivity in developing countries. A recent study by Karen Macours (2003) suggests that informal enforcement of property rights in the land market creates incentives for rental transactions to remain within ethnic groups. In turn, in a highly fragmented environment, the exclusion of minority groups leads to ethnic conflict, further weakening property rights and reinforcing segmentation.

3.3.2 Effects on public policies

Most of the literature on group heterogeneity and collective action in small communities has focused on the role of income inequality as opposed to ethnic or racial heterogeneity.

The focus of that literature, exemplified by Mancur Olson's (1965) seminal contribution, has been on the relationship between inequality in the shares of the benefits from the commonly provided good that accrue to different types and their incentives to contribute.[26] However, the extent to which this literature can be generalized to the impact of ethnic diversity is limited, and requires a context in which types (e.g., ethnic groups) matter for public good provision only through their contributing capacity, and a mapping between inequality and ethnic fractionalization can be traced.

A recent literature, however, has looked specifically at the role of ethnic heterogeneity with a particular emphasis on public good provision in poor communities. An example is the study by Cagla Okten and Una Okonkwo-Osili (2004), who use micro level data from Indonesia to estimate how ethnic diversity affects monetary and time contributions to community organizations. They propose three ways in which ethnic diversity can affect voluntary contributions: (i) diverse communities may have more difficulty in defining common policies as their preferences diverge; (ii) transaction costs are higher in more diverse communities, e.g. because of lower effectiveness of informal enforcement mechanisms; and (iii) an altruistic orientation to contribute to one's own ethnic group. Their empirical results suggest that increased ethnic heterogeneity decreases both contributions to local community organizations, and the prevalence itself of such organizations. Adi Brender (2004) finds similar results for Israel.[27]

Edward Miguel and Mary Gugerty (2004) also investigate the mechanisms through which ethnic heterogeneity may harm public good provision. In particular, they focus on the role played by social sanctions. As we argued in section 2, in environments with weak legal enforcement most informal transactions rely on the availability of "self-enforcing" mechanisms related to repeated interaction and reputation, as well as on the imposition of social sanctions. Miguel and Gugerty assume that such sanctions are more effective if imposed *within* ethnic groups than *between* groups. They test this hypothesis using data on 337 primary schools in rural Kenya. In addition to information on students and teachers, their data contains school committee records which report the threat or application of sanctions and the fund raising activities of the school. They find that local ethnic diversity is negatively correlated to school funding and to the quality of school facilities. According to their estimates, moving from complete homogeneity to complete heterogeneity would reduce average local funding by about 20 percent.

An insight into the motivations underlying the failures of collective action in heterogeneous communities is offered by the recent work of Abigail Barr (2003). She conducted field experiments in Zimbabwe exploiting the resettlement policies promoted by the government, which generated a set of socially and ethnically heterogeneous villages (treatment) to be compared with non-resettled communities (control). From the results

[26]For a survey of this topic with a specific emphasis on economic inequality see Eliana La Ferrara (2003b).

[27]This author also discusses the effect of ethnic minority's control of municipalities and finds that it is associated for poorer tax collection.

of her trust game, Barr concludes that the lower propensity to trust of resettled villagers is due not to differences in altruism or in socially transmitted norms, but to the lower density of kinship ties. Again, this is consistent with the hypothesis that informal enforcement does not travel well across kinship (and *a fortiori* ethnic) lines.

Finally, Jeff Dayton-Johnson (2000) and Asim Khwaja (2000) look at more production-oriented collective activities. The former uses data on Mexican irrigation projects and finds that canal maintenance is worse the more unequal is the distribution of land and the higher is "social" heterogeneity, proxied by the number of different farming communities represented in the same maintenance unit. Khwaja (2000) uses data on community-maintained infrastructure projects in Northern Pakistan, the complexity of which ranges from simple irrigation channels to sophisticated electricity units. Again, he finds that "social" heterogeneity - measured as the fragmentation into different clans, political and religious groups - is negatively associated with project maintenance. An interesting difference between Khwaja's work and the other existing studies is its focus on institutional design. In fact he finds that good task design can potentially compensate for fragmentation in allowing heterogeneous communities to succeed in collective action. This parallels the findings of the literature on organization behavior described in section 2.1.

4 Ethnic fragmentation and endogenous community formation

In this section we discuss how fragmentation affects not just the economic policies and performance of given communities, but the formation and composition of the relevant communities. In other words, what happens when community size and/or composition can be simultaneously determined with the policies?

4.1 Countries

A line of research by Alesina and Spolaore (1997, 2003), Alesina, Spolaore and Wacziarg (2000), Enrico Spolaore and Romain Wacziarg (2002) emphasizes the role of racial conflict as a determinant of the number and size of countries. The argument is as follows. The size of a country emerges from a trade off between the benefits of scale (broadly defined) and the cost of heterogeneity of preferences in the population. Benefits of size include economies of scale in the production of some public goods, internalization of policy externalities, the size of the market, defense and protection from foreign aggression, regional insurance schemes. The costs of heterogeneity arise because in large and diverse countries individuals with different preferences have to share common policies so the average utility of these policies is decreasing with heterogeneity. Empirically, racial fragmentation is often associated with differences in preferences, so racial cleavages are

a major determinant of the determination of borders, secessions and various centrifugal forces.[28]

A potentially testable implication of this approach is that as the benefits of size diminish, then it becomes more likely that countries can split into more homogenous smaller political entities. One building bloc of this argument is of course that openness to trade is particularly beneficial for small countries. Results by Ades and Glaeser (1999), Alesina, Spolaore and Wacziarg (2000), and Alcala and Ciccone (2004) suggest that as freedom of trade increases, the benefit of size for economic growth diminishes. In a completely autarkic world the political size of a country also determines its economic size; in a world of free trade they become more disjoint. That is, from an "economic" point of view (our production of private goods in the simple model above) trade makes economic size "larger". On the other hand since countries can retain their independence while trading they do not have to share common public policies on which there are differences of opinions. In ethnically diverse societies, then, increased economic integration should make it more likely that conflicts are resolved with break down of countries. Some insights on this issue can be gathered from the political science literature on partition as a solution to ethnic civil war, supported among others by Chaim Kaufmann (1996, 1998). A critical assessment of the view that separation is the best solution for civil wars generated by ethnic conflict is provided by Nicholas Sambanis (2000), who uses a cross sectional data set of all civil wars since 1944 and estimates the probability of partition as a function of the type of civil war (ethnic/religious as opposed to ideological) and of several socioeconomic factors, among which ethnic heterogeneity of the population.

The relationship between ethnic heterogeneity and the likelihood of country breakdowns is also mediated by the role of natural resources, and this is a particularly relevant issue for developing countries. Natural resource discoveries tend to be located in remote areas at the periphery of a country, as resources more centrally located have likely been discovered already. It is often the case that people living in peripheral areas have ethnic identities that do not coincide with the majority of the country as a whole. The availability of new natural resources makes these regions more economically viable on their own and therefore increase pressure for separation or autonomy.[29]

In addition to economies of scale, another benefit of country size is defense and protection from aggressions, so as the world becomes more peaceful one should observe centrifugal forces. Alesina and Spolaore (2003) discuss historical evidence arguing that this implication is consistent with the data concerning the evolution of country size, international trade and threats of conflicts. Recently, the collapse of the Soviet Union by reducing the threat of and East West conflict has certainly facilitated political separatism in Eastern Europe. Huntington (1998) notes how the end of the Cold War allowed the realignment of peoples into countries that better reflected homogenous "civilizations".

[28]Patrick Bolton and Gerard Roland (1997) explore how income differences and redistribution may lead to break down of countries.

[29]We are grateful to a referee for suggesting this point.

In most cases this movement meant breakdown of countries and in a few cases movement toward reunification.

Finally, an important issue is the relationship between ethnic heterogeneity, country formation and democracy. Alesina and Spolaore (2003) discuss the effect of authoritarian systems on measured racial, linguistic or religious fragmentation and country size. Dictators prefer large countries for several reasons. One is that they can extract rents from larger populations, another one is that they can support with size their bellicose attitudes. Historically, one of the main problems of dictators has been to repress ethnic conflict in an attempt to create artificially homogeneous countries – an issue to which we return below when we discuss the endogeneity of the notion of fragmentation. In fact often dictators use racial hatred to create support for the dominance of one group over others, a result consistent with models and empirical evidence by Edward Glaeser (2002). One of the implications of this artificial repression of diversity is that centrifugal forces typically explode when dictators falls, as happened for example in the Soviet Union, Spain, Yugoslavia and Iraq. James Fearon (1998) provides an insightful game theoretic model of civil wars that follow the collapse of dictators.

4.2 Cities

A very large literature based on the celebrated Charles Tiebout (1956) model has discussed the formation and organization of jurisdictions based upon a very simple but powerful idea. The rich want to isolate themselves from the poor to escape from redistributive policies and the poor want to be close to the rich to gain from redistribution. Until recently, virtually all the economic literature on jurisdiction formation in urban economics was based on this income conflict. That is, if the wealthy want to segregate away from the poor, the number of communities should increase as income inequality increases.[30]

On the other hand, a vast body of sociological literature has emphasized the importance of racial divides in the formation and organization of American cities. Alesina, Baqir and Hoxby (2004) provide a model of formation of political jurisdictions which expands upon the models of country formation described above. Again, the formation of local jurisdictions emerges from a trade off between the benefits of scale and the costs of racial heterogeneity. These authors look both at recent evidence and at historical evidence on the formation and break down of school districts, special districts and cities. In particular they consider the Great Migration of African Americans from the South to some areas of the North to support the war industries during the two world wars. They examine how the pattern of jurisdiction formation differs in counties where the immigration of blacks occurred and in those in which it did not, confirming the result that the desire for racial homogeneity was the driving force in the formation of localities.

[30]For an excellent recent contribution in this line, which also summarizes much of the earlier work, see Calabrese, Cassidy and Epple (2002).

The trade-off between economies of scale and racial heterogeneity tends to be larger in magnitude and more robust empirically than the trade-off between economies of scale and income heterogeneity.

An important issue is how different dimensions of heterogeneity *interact* to determine jurisdiction formation. In a recent paper, Sethi and Somanathan (2004) propose a model in which individuals care both about the racial composition of their communities and about its wealth, and in which races differ in income. They show that it is crucial to consider the *interplay* between preferences on inter-racial interactions and income differentials between races in order to understand patterns of residential location (i.e., segregation). An application of their framework to jurisdiction formation would enrich existing theories in interesting ways.

Heterogeneity can also affect jurisdiction formation through the choice of the "admission rule" into the jurisdiction. Phillipe Jehiel and Suzanne Scotchmer (2001) provide a model in which agents are heterogeneous in their taste for a public good, and the choice of the admission rule into the jurisdiction is endogenous. They consider different possible admissions rules (free mobility, majority vote, unanimity, and conditional on demand) and ask which partition is stable for each given rule. While not directly applied to the issue of ethnic heterogeneity, their theoretical framework seems useful for a research agenda in which changes in ethnic diversity do not automatically translate into break down or consolidation of jurisdictions, but can be mediated through an endogenous choice of specific rules of the game. This seems a promising avenue of research for the future.

4.3 Groups and clubs

Compared to the literature on country size and jurisdiction formation, the economics literature on the role of ethnic diversity in the endogenous formation of social groups is significantly smaller. The traditional approach within economics has been one in which groups were seen as "clubs" into which individuals could self-select, and within which they would interact. The benefits from consumption depended on the characteristics of other members of the club (hence on group composition), mostly in terms of income. More recently, a game-theoretic literature has developed on the formation of clubs, networks and coalitions that provide economic benefits to their members (see the volume by Gabrielle Demange and Myrna Wooders (2005) for a survey and a collection of results). Within this literature, it is possible to isolate the role of players' heterogeneity as a determinant of group formation. For example, Bryan Ellickson, Birgit Grodal, Suzanne Scotchmer and William Zame (1999) propose a general equilibrium framework in which agents are heterogeneous and can sort into different clubs, and the equilibrium number and composition of the groups is determined endogenously to clear the market for club memberships. Another formulation by Igal Milchtaich and Eyal Winter (2002) focuses specifically on the effects of endogenous group formation for the composition of

the groups, and in particular on the conditions under which the resulting equilibrium is one with segregation. Starting from the assumption that individuals prefer to associate with others similar to themselves, Milchtaich and Winter study the "stable" partitions of society into groups and show one crucial element is whether people minimize the "average distance" from other group members or the "distance from the average" group member.[31]

In recent work, Debraj Ray and Rajiv Vohra (1999, 2002) study coalition formation and public good provision in a setting where agents have complete information and can write binding agreements. They focus on coalition formation as a potential source of inefficiency and show that only in some cases full cooperation (efficient public good provision) emerges as the equilibrium; in other cases, several coalitions coexist in equilibirium and the level of public good provided is inefficient.

Alesina and La Ferrara (2000) focus on the role of ethnic heterogeneity in a setting in which individuals can choose whether to join groups or not, and derive the equilibrium composition of the groups as a function of the degree of ethnic heterogeneity in the society as a whole. In particular, they study under what conditions increased heterogeneity in the population leads to less aggregate participation in groups, even when individuals can sort into multiple homogeneous groups. Using survey data for the US, they find that participation in socioeconomic groups is negatively affected by local indexes of racial fractionalization and heterogeneity in ethnic origin. Thanks to the availability of direct individual responses on questions regarding racial mixing, they can test the effects on different sub-groups of the population and the find that the negative effect of racial fragmentation on participation only holds for people relatively averse to racial mixing.

A similar question is addressed by Eliana La Ferrara (2002a) in the context of developing countries. The model here focuses on the relationship between heterogeneity and group participation in the presence of different admission rules. Under one rule, labelled "open access", anyone can join the group provided he or she pays the cost; another rule instead allows the members of the group to exclude someone by majority vote. La Ferrara shows that an increase in heterogeneity 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 categories are represented in the group. Empirical findings from informal groups in rural Tanzania are consistent with the predictions of the theory. A more general treatment of group formation and decision rules, but without a focus on ethnic diversity, can be found in the survey by Gabrielle Demange (2005). This author also places the argument in the context of a trade off between the benefits of size (increasing returns) and the costs of increased preference diversity, much in the spirit of the literature on country and jurisdiction formation surveyed above.

[31]Note that while in order to calculate the "average distance" one needs to know the identity of each and every member of the group, a "distance from the average" approach only requires knowledge of the representative individual within the group. Which assumption is more realistic depends on the situation under study.

5 Open questions

In this section we highlight the main questions that in our opinion need to be addressed to get a better understanding of how much and why ethnic diversity matters for economic outcomes. The first is the endogeneity of ethnic diversity; the second is how diversity should be measured.

5.1 The endogeneity of diversity

All the work surveyed above shares the assumption that ethnic groups are "objective categories" into which individuals can be classified, and that such classification is commonly shared and exogenous. However, the validity of this assumption can be called into question on several grounds. First, people may not agree on what are the relevant ethnic groups into which they are supposed to "classify" others, i.e., the boundaries of these groups may not be objectively known to all. Secondly, even under the most conventional definition of cultural category (ethnic, linguistic or religious), the latter may not be determined independently of economic and policy choices at a given point in time. This can occur both because political leaders may actively pursue policies that influence (historically, often reduce) ethnic diversity, and because citizens may "choose" their identity differently in response to political and economic conditions.

5.1.1 What makes ethnicity identifiable?

Underlying most research undertaken so far is the assumption that people's ethnicity is easily identifiable and can be used to construct categories of "homogeneous" individuals. Indeed, the supposedly "objective" nature and visibility of ethnic identity is often advocated as the primary reason why economic or political conflict may organize around ethnic lines even when the underlying preferences are not intrinsically about ethnicity. For example, Francesco Caselli and Wilbur Coleman (2002) state that "ethnicity allows groups fighting over resources to enforce membership in the respective coalitions. Without the distinguishing marks of ethnicity, these coalitions would be porous and subject to infiltration".[32] James Fearon (1999) argues that using ethnicity as a criterion for the allocation of "pork" is a way for those who win elections to prevent losers from entering the winning coalition. Several recent contributions, however, have started to challenge this assumption.

First of all, individuals' ability to correctly classify *others* into a given ethnic category may not be taken for granted. Donald Horowitz (2001) and Macartan Humphreys, Daniel Posner and Jeremy Weinstein (2002) report evidence from case studies in Sri Lanka, Burundi and Ethiopia, where identifying members from different ethnic groups was at times difficult despite the fact that local conflicts were revolving around ethnic

[32] Caselli and Coleman (2002), p.1.

roots. In those cases, the possibility to fake one's accent or to dress in a particular way made it impossible to recognize people's ethnic origin even for their local counterparts. In a recent paper, Macartan Humphreys and Habaye Mohamed (2002) compare the experiences of Mali and Senegal in terms of the ability to identify specific ethnic groups leading separatist movements. They argue that the fact that the Tuaregs and Maures in Mali were relatively "white" compared to the rest of the population led to a polarization of forces and to escalating communal violence. On the other hand, ethnic violence towards the Diola minority group has been limited by the difficulty of identifying them. Finally, James Habyarimana, Macartan Humphreys, Daniel Posner and Jeremy Weinstein (2004) conducted an experiment on undergraduate students in the US recruited from seven different ethnic groups, and found that their subjects managed to "pass" as members of other groups 45 percent of the time.[33] Among the factors influencing a person's ability to correctly identify others were his or her own exposure to other ethnic groups and the level of information about those groups.

Secondly, individuals' choice of *their own* ethnic identity may not be unresponsive to the economic environment.[34] In India a well known phenomenon, known as "Sanskritization" since the early work of Mysore Srinivas (1966), denotes the efforts of lower caste members to raise their social status by adopting the practices and language of upper castes. Using a simple model, Francis Bloch and Vijavendra Rao (2001) show that in societies where the minority group suffers from statistical discrimination, social assimilation can occur as minority members adopt the behavior of the dominant group to signal high productivity to potential employers. An alternative formulation of a similar phenomenon is provided by David Laitin (1998) using a Schelling tipping model. He argues that as countries become richer, there may be a tendency for lower income ethnic groups to mimic and assimilate with higher income groups. This "ethnic mimicking" by lower status ethnic groups decreases the social costs of heterogeneity in rich economies, and thus contributes to explain why ethnic diversity appears to be more costly in poorer countries (see our econometric results above).

In the context of data collection, self reported racial classifications may be partly endogenous to government policies. Users of Census data know how sometimes questions about ethnic affiliation can be a politically charged issue. For example, if the government is known to favor (or hinder) a given ethnic group, people may have an incentive to report (or not report) themselves as part of that group.[35] How empirically important

[33]The ethnic categories in their sample were: African American, Arab, Asian, Caucasian, Indian, Persian/Iranian, and Latino.

[34]Among earlier contributions highlighting the responsiveness of ethnic identities to political and economic incentives, see Anderson (1983) and Donald Horowitz (1985).

[35]Steven Wilkinson (2002) discusses two interesting examples in this respect. One is from a Bohemian town where about a third of the respondents who had declared to be "Germans" in the 1910 Census switched to "Czech" in 1921 to avoid discrimination. The second is from the Indian state of Punjab, where in the 1961 Census the fraction of Punjabi speakers dropped by over 20 percentage points because many Hindu Punjabi speakers who wanted to block the attempts of a Sikh movement to partition the

this "tyranny of the Census" is remains to be seen.

While the notion of endogenous ethnic identity is becoming increasingly popular among social scientists, to our knowledge the only attempt at formalizing it in the context of an economic model is the recent work by Caselli and Coleman (2002). In their model, resources are allocated based on the ethnic composition of the society, and individuals can choose their identity strategically, i.e. can switch ethnicity by paying a cost. The greater the "physical" or cultural distance among the groups, the greater this cost. As we shall see below, this formalization also has important implications for building relevant measures of ethnic diversity.

5.1.2 Why are some ethnic differences perceived as "salient"?

While ethnic diversity is often associated with poor politico- economic outcomes, as discussed above, it is not always the case. Also while in some cases ethnic conflict explodes in violent civil wars in many other cases it does not. [36] Why do ethnic or cultural differences matter in some cases and not in others?

Daniel Posner (2004b) offers an interesting "natural experiment" originated from the arbitrary drawing of the border between Zambia and Malawi. When the border between the two countries was drawn, two ethnic groups –the Chewas and the Tumbukas– were partitioned so that approximately two thirds of each group remained in Malawi, and the rest in Zambia. Coming from an identical cultural background, the evolution over time of the relationship between the two ethnic groups in each country can be presumed to be the result of the difference in economic and political institutions. In particular, since their division, the Chewas and the Tumbukas have been political allies in Zambia and adversaries in Malawi. Posner suggests that the explanation for this difference lies in the relative size of each group compared to the relevant country's population. While in Malawi both groups represent a large fraction of the country's population, hence they can compete for power at the national level, in Zambia they are a minority compared to other ethnic groups and they often ally as an "Eastern" coalition against the remaining political forces. This example powerfully suggests that there is nothing intrinsic to physical differences or to the content of cultural traditions that should make a given ethnic divide "salient" or not: rather, it is the structure of domestic political and economic competition that shapes potential ethnic divisions into meaningful realities.

As a matter of fact, even within a given institutional structure the salience of ethnic divisions can change over time as a response to politico-economic incentives. Alesina et al. (2003) discuss the example of Somalia, which until the onset of the 1991 civil war was considered an ethnically homogeneous country because 85 percent of the population was Somali. The war shifted the relevant dimension of ethnic cleavage to that of "clans", and

state declared themselves as speaking Hindi.

[36] For example, James Fearon and David Laitin (1996) argues that compared to the degree of ethnic fragmentation in the African continent, the actual occurrence of conflicts is small.

individual self-identification to groups changed in a way that made the country more "ethnically" fragmented. James Fearon (2003) argues that the only way to really measure ethnic fragmentation is to get the salient issue "right", that is to identify correctly for every country what the salient divisions are.

Finally, individual socioeconomic background can be an important factor in determining ethnic identification. Using the Afrobarometer surveys collected in the early 2000's in nine Sub-Saharan African democracies, Alicia Bannon, Edward Miguel and Daniel Posner (2004) estimate the likelihood that an individual will identify him/herself primarily in ethnic terms. According to their estimates, this likelihood increases with the individual's education, with the occupation in non-traditional sectors, and with exposure to political mobilization. in other words, ethnic identification is not the result of backwardness. These authors also find a negative relationship between ethnic diversity and ethnic salience, which they interpret as evidence against the maintained assumption that ethnic divisions are more salient in more diverse societies. While this is certainly a reasonable interpretation, it is also possible that respondents to surveys tend to underreport the importance of the ethnic factor in context where ethnicity is a particularly sensitive issue.[37]

5.1.3 Mobility and ethnic diversity

Even if one were ready to accept the definition of ethnic groups as objective categories with exogenous borders, we should worry about the potential endogeneity of ethnic diversity measures as a result of individual mobility. Consider for example US cities. Changes over time in the economic growth of different metropolitan areas have induced massive flows of migration that have sensibly altered some cities' ethnic composition. Local economic policies have also played a role: the structure of public policies such as education spending is such that the racial or ethnic composition of a given area can also shift over time as a result of policy changes. An empirical solution to this issue is provided for example by Alesina, Baqir and Hoxby (2004), who use historical evidence on the pattern of South-North migration to develop the war industry in the early XX century as an instance of pre-determined local ethnic composition. In a cross country setting, endogeneity of ethnic differences due to geographic mobility is less likely to be relevant, except possibly as a result of diasporas following civil wars.

The nature of the geographic landscape may also influence mobility and the ethnic composition of various countries. However note that borders are only in part a predetermined geographic feature, as in general they are chosen by a combination of political forces in motion. For instance after the First World War the superpowers of Britain, France and the US met in Versailles and redrew the world borders in ways that only partially reflected the goal of ethnic homogeneity; they were much more interested in

[37] A similar problem, for example, has been noted in the reliability of survey-based perceptions of corruption.

grabbing for their empires and their allies as much territory as possible. The failures of the border arrangement in Versailles are still responsible for many of today's conflicts.[38]

Finally, a more general concern related to the geographical coverage of ethnic diversity measures relates to what statistical geographers call the "modifiable area unit problem". This term, introduced by Stan Openshaw and Peter Taylor (1979), is associated with the distortions that may arise when individual level data are aggregated into somewhat arbitrary geographical units (e.g., census tracts).[39] These authors showed that the sign of the correlation between two variables could change with the spatial definition used, which potentially undermines the validity of empirical analyses in which the unit of aggregation does not have a clear social or political justification. It is conceivable, however, that the severity of this problem will diminish as Geographical Information Systems software and computational tools become increasingly available.

5.2 Measuring ethnic diversity

5.2.1 What dimension of heterogeneity?

How to classify ethnic groups is a difficult and politically charged issue. While for the US the Census Bureau provides a classification in five major groups which is fairly broadly accepted, similar classifications for other countries are more problematic. Individuals differ in skin color, language, origin of birth, religion: in some countries language is the key dividing line, in others it is skin color. What dimension should one use? Can ethnicity be measured in a multidimensional way?

The raw data originally used by Easterly and Levine (1997) come from the *Atlas Narodov Mira*, a compilation of ethno-linguistic groups present in 1960 based on historical linguistic origin. A first weakness of this data is that linguistic heterogeneity does not necessarily coincide with ethnic heterogeneity.[40] For instance, most Latin American countries are relatively homogenous in terms of language but less so in terms of "ethnicity" or "race". Fearon (2003) and Alesina et al. (2003) have compiled various measures of ethnic heterogeneity which try to tackle the fact that the difference amongst groups manifests itself in different ways in different places. The two classifications are constructed differently. Alesina et al. (2003) do not take a stand on what characteristics (ethnicity, language or religion) are more salient, and adopt the country breakdown suggested by original sources, mainly the Encyclopedia Britannica (see the Appendix for more details). Fearon (2003) instead tries to construct a list of "relevant" ethnic groups which "depends on what people in the country identify as the most socially relevant ethnic groupings" (page 198). This approach has the advantage of being closer to

[38] For an insightful discussion of these issues and more generally of the Treaty of Versailles see Margaret McMillan (2003).

[39] For a survey of the literature on this problem, see Geoff Dudley (1991).

[40] Another issue is that multilingualism make linguistic fractionalization less salient, a point emphasized by David Laitin (2000).

what the theory would want and the disadvantage of having to make judgement calls (or adopt others' judgement calls) about what the "relevant list" is. The sources used by Fearon (2003) are carefully described in his paper, but an especially useful one to identify "salient" cases of ethnic conflict is Ted Gurr (1996) who classifies minorities at risk in many countries around the world.

Alesina et al. (2003) identify language groups as well as ethnic groups that are defined by other characteristics, such as skin color.[41] The correlation between their more comprehensive ELF index and the one based purely on language is between 0.6 and 0.7, depending on the period and sample of countries used. An interesting example of the differences between the two indices is Latin America. In this region the language index shows more homogeneity because the language of the former colonizers (Spanish, Portuguese, English) is often spoken by most, but the index based on skin color or ethnic origin (say black, mulattos, white, mestizos, Indian, etc.) shows more heterogeneity. The correlation between the Alesina et al. (2003) measure of ethnic fragmentation and Fearon's (2003) is about 0.76. It is therefore quite high, considering the different criteria of construction.

A second weakness of the *Atlas Narodov Mira* has to do with the way in which the various groups were formed. Daniel Posner (2004a) argues that the *Atlas* data suffers from a "grouping problem" at two different levels. On the one hand, many groups are aggregated into a single category while they are distinct political actors –even enemies– at the national level. The most striking example of this concerns the Tutsis and the Hutus in Rwanda, which are aggregated into a single category "Banyrwanda". At the opposite extreme stand a number of groups that are listed as separate linguistic categories, but whose distinction has no political or economic relevance. Posner thus proposes a classification based on "politically relevant ethnic groups" (PREG), defined as groups that can influence economic policy decisions either directly or indirectly (e.g., by threatening to remove politicians from power). However, it is difficult to argue that the realized structure of power at a given point in time is exogenous and can be used as an underlying determinant of the definition of ethnic groups. To date, it is still unclear how to integrate linguistic or "ethnic" differences with other dimensions that make the latter politically or economically salient.

5.2.2 What index?

Most of the existing literature on ethnic diversity and economic performance focuses on the "fractionalization index" defined by expression (8) in section 3.1. This index

[41] An innovative approach to the use of information such as skin color can be found in recent work by Hugo Nopo, Jaime Saavedra and Maximo Torero (2002). They use survey data in which every respondent is assigned a score from 1 to 10 for each of the four main racial groups in Peru: White, Indigenous, Black, and Asian. This way heterogeneity can be measured through a multidimensional index of "racial intensity". It should be explored to what extent it is feasible, and profitable, to move in this direction.

captures the probability that two individuals randomly drawn from the population belong to different groups, and reaches a theoretical maximum of 1 when every individual belongs to a different group. This measure implies that a country composed of say 100 equally sized groups is more fractionalized than a country with two equally sized groups. However, an argument that goes back as far as the Founding Fathers is that a country composed of many small groups may actually be more stable than one composed by two equally sized ones, which are more likely to be in direct conflict with each other.[42] Based upon the theoretical results of Joan-Maria Esteban and Debraj Ray (1994), Jose Montalvo and Martha Reynal-Querol (2002) propose the following "polarization index":

$$RQ = 1 - \sum_{i=1}^{N} \left(\frac{1/2 - s_i}{1/2} \right)^2 s_i. \tag{9}$$

where s_i is the share of group i in the population. The index RQ reaches maximum when two equally sized groups face each other and declines as the configuration of groups differs more and more from this half and half split. The authors also show that this index is highly correlated with ethno-linguistic fractionalization (ELF) at low levels of ELF, uncorrelated at intermediate levels, and negatively correlated at high levels. In a cross country regression analysis, they find that ethnic polarization has a positive impact on the likelihood that a civil war occurs, and a negative effect on a country's growth rate. They do not find an independent effect of ethnic fractionalization. Using a different data set, Alesina et al. (2003) compare the results of the polarization index RQ and the fractionalization index ELF, and find that fractionalization works slightly better as a determinant of policies and economic outcomes. While the apparent inconsistency between the two sets of results may be due partly to different parameterization and partly to different data sources, it is difficult to gauge the statistical significance of the difference due to the high correlation between the two measures at low levels of fragmentation.

In the context of studying segregation in cities, Douglas Massey and Nancy Denton (1988) provide an excellent summary of a variety of indices that capture aspects of segregation and distribution of groups within a certain geographical area. While these indices are certainly useful, the data requirements may be insurmountable for large cross country studies; at the very least they would require a large investment in data construction.

Another important issue is whether all groups should be treated symmetrically, as they are in the fragmentation index and to an extent in the polarization index.[43]

[42] James Madison (Federalist Papers n.11) used this argument to convince skeptics that a multi-ethnic Unites States was viable, precisely because a complex web of cross group cleavages would make it more stable. Whether or not the history of the US with the Civil War confirmed Madison's views is a much debated question. The importance of "non-linearities" in the effects of ethnic diversity is also stressed by Robert Bates (2000) in the African context.

[43] In the formula for RQ the deviation of each group from the maximum polarization share of 0.5

Alternatively, one may want to assign weights to the *distance* between groups. Two theoretical contributions that axiomatize measures of diversity based on the "dissimilarity" of the categories under consideration are those by Weitzman (1992) and by Bossert, Pattanaik and Xu (2003). To our knowledge, however, these measures have not yet been implemented in empirical analysis, and the only applications containing proxies for the distance among groups are of a more heuristic nature. For example, Laitin (2000) and Fearon (2003) use measures of distance between languages to measure how different linguistic groups are in several countries. Caselli and Coleman (2002) stress the importance of ethnic distance in theory and in work in progress are trying to measure it, using surveys of anthropologists. Another dimension of distance is income: for example, two ethnic groups are especially far apart if their average income is also very distant. This is an approach followed by work in progress by Philippe Aghion, Alberto Alesina and Francesco Trebbi (2004b) for US cities.

6 Conclusions and policy implications

We proposed a model in which public good provisions was lower in fragmented societies while productivity may be positively related to variety. Is the evidence consistent with it? We certainly found overwhelming evidence supporting the first part of the preposition. As for the productivity effects of diversity the picture is complex. It is somehow easy to point to economic failures of fractionalized societies, but this is not a general phenomenon. Rich democratic societies work well with diversity, in the case of the US very well in terms of growth and productivity. Even within the developing world, similar levels of ethnic diversity are associated with very different degrees of conflict and inter-ethnic cooperation. Useful theoretical progress would incorporate in a model like this more realistic institutional features that would distinguish cases in which the economy manages to actually take advantage more or less well of the potential benefits of variety in production.

What are the policy implications of all of the above? The issue is quite difficult and politically charged and it is relevant in at least two areas: immigration policies and local policies that may increase or decrease racial integration. The implication of promoting racial homogeneity is unappealing and probably incorrect both in the short and in the long run. David Laitin (1994) provides an interesting example concerning language in Ghana. After independence this country faced the question of which language to adopt as the official one. Using English had the advantage of being understood by most and of not favoring one ethnic group over another. On the other hand it was the language of a colonizer. Laitin argues that a solution with multiple languages may dominate that of

is weighted by the group's own share. However, underlying that formula is the assumption that the "distance" between each group (continuous, as originally conceived by Esteban and Ray (1994)) is discrete and it is the same.

a single homogenous language. The benefit of homogeneity had to be traded off against other considerations (national pride, ethnic balance; etc.).[44] On the other hand peaceful separation and country breakdown may be perfectly reasonable solutions to racial or cultural diversity.

Globalization also has important implications for ethnic politics. To the extent that small countries can prosper in a world of free trade, then peaceful separatism of certain minorities should not be viewed as threatening, at least from an economic point of view. As far as domestic social policy is concerned, the question is to what extent favoring racial mixing (say with affirmative action) promotes harmony, an issue that would require an entirely separate paper. The starting point would be Arend Lijphart's (1977) seminal contribution that provides a notion of power sharing denoted as "consociational democracy". The key features of this type of democracy should be a coalition government in which "all significant segments of the plural society"[45] are represented, with a proportionality system, a mutual veto, and a federalist structure. He highlights the conditions under which power sharing is likely to succeed, namely, a relative balance of power and economic equality among the different groups. Most importantly, he argues that different groups are most likely to find an agreement when they have to face *external* threats. This makes power sharing schemes difficult to implement and ultimately unstable in some developing countries (e.g., Africa) where most threats to the State come from within. Among recent examples of power sharing agreements that have failed due to internal conflicts are those of Angola and Rwanda. On the other hand, South Africa and Somaliland have managed to successfully implement consociationalist schemes. Spears (2002) reports that, in addition to the presence of an "external" threat (Mogadishu), in the case of Somaliland a deeply rooted tradition of power sharing among the elders of local clans may have contributed to the viability of such schemes. However, this calls into question the effectiveness of power sharing as a means of *generating* inter-ethnic cooperation: indeed power sharing may well be the *result* of pre-existing attitudes towards inter-ethnic cooperation. Aghion, Alesina and Trebbi (2004a) in fact report that racial and ethnic fractionalization are empirically inversely related to forms of consociativism and widespread proportional representation.[46]

The issue of multi ethnicity is especially relevant for current Europe. In fact while the United States have been a melting pot throughout most of its history, Western European countries have been much more ethnically homogeneous. However, with the opening of borders within the European Union and its expansion to the East, in addition

[44]For a recent application to language diversity in the European Union and a measure of the "disenfranchisement" that would arise from a reduction in the number of EU working languages, see Victor Ginsburgh and Shlomo Weber (2004).

[45]Lijphart (1977), p.25.

[46]Note that while proportional representation and consociationalist schemes may diffuse racial tension, their presence is also empirically associated with difficulties in pursuing adeguate fiscal policies, larger budget deficits and macroeconomic policy instability. For extensive empirical evidence see Torsten Persson and Guido Tabellini (2003).

to increasing migration from Africa and other neighboring areas, members countries of the European Union will become less and less homogeneous; in fact the issue of multi ethnicity will be one of the major challenges for Europe in the near future.

With this survey we have tried to asses costs and benefits of ethnic fragmentation and the policy issues arising in diverse societies. In a more and more integrated world, the question of how different people can peacefully interact is the critical problem for the next many decades.

References

[1] Ades, Alberto F. and Edward L. Glaeser (1995), "Trade and circuses: explaining urban giants", *Quarterly Journal of Economics,* 110(1), *195-227.*

[2] Ades, Alberto F. and Edward L. Glaeser (1999), "Evidence on growth, increasing returns and the extent of the market", *Quarterly Journal of Economics*, 114(3), 1025-1046.

[3] Aghion, Phillipe, Alberto Alesina and Francesco Trebbi (2004a), "Endogenous Political Institutions", *Quarterly Journal of Economics, 119(2), 565-612.*

[4] Aghion Phillipe, Alberto Alesina and Francesco Trebbi (2004b), "Choosing Electoral Rules: Evidence from US Cities," unpublished.

[5] Alcala, Francisco and Antonio Ciccone (2004), "Trade and productivity" *Quarterly Journal of Economics, 119(2), 613-646.*

[6] Alesina, Alberto, Raza Baqir and William Easterly (1999), "Public Goods and Ethnic Divisions", *Quarterly Journal of Economics*, 114(4), 1243-1284.

[7] Alesina, Alberto, Raza Baqir and William Easterly (2000) "Redistributive Public Employment" *Journal of Urban Economics,* 48(2), 219-241.

[8] Alesina, Alberto, Raza Baqir and Caroline Hoxby (2004), "Political Jurisdictions in Heterogeneous Communities", *Journal of Political Economy, 112(2), 348-396.*

[9] Alesina, Alberto, Robert J. Barro and Silvana Tenreyro (2002) " Optimal Currency Areas" in Mark Gertler and Kenneth Rogoff (eds.)*NBER Macroeconomics Annual* MIT Press Cambridge Mass

[10] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg (2003), "Fractionalization", *Journal of Economic Growth,* 8(2), 155-194.

[11] Alesina, Alberto and Edward Glaeser (2004), *Fighting poverty in the US and Europe: a world of difference*, Oxford University Press, Oxford, UK.

[12] Alesina, Alberto, Edward Glaeser and Bruce Sacerdote (2001), "Why Doesn't the US Have a European Style Welfare State?" *Brookings Papers on Economic Activity,* Fall.

[13] Alesina, Alberto and Eliana La Ferrara (2002), "Who Trust Others?" *Journal of Public Economics,* 85(2), 207-234.

[14] Alesina, Alberto and Eliana La Ferrara (2000), "Participation in Heterogeneous Communities", *Quarterly Journal of Economics,* 115(3), 847-904.

[15] Alesina, Alberto and Enrico Spolaore (1997), "On the Number and Size of Nations," *Quarterly Journal of Economics,* 112(4), 1027-1056.

[16] Alesina, Alberto, Enrico Spolaore and Romain Wacziarg (2000), "Economic Integration and Political Disintegration", *American Economic Review*, 90(4), 1276-1296.

[17] Alesina, Alberto and Romain Wacziarg (1998), "Openness, Country Size and Government", *Journal of Public Economics,* 69(3), 305-321.

[18] Anderson, Benedict (1983), *Imagined Communities*, London: Verso.

[19] Arcand, Jean-Louis, Patrick Guillaumont and Sylviane Guillaumont Jeanneney (2000), "How to Make a Tragedy: On the Alleged Effect of Ethnicity on Growth", *Journal of International Development*, 12(7), 925-938.

[20] Bannon, Alicia, Edward Miguel and Daniel Posner (2004), "Sources of Ethnic Identification in Africa", mimeo, University of California at Los Angeles.

[21] Barr, Abigail (2003), "Trust and expected trustworthiness: experimental evidence from Zimbabwean villages", *Economic Journal*, 113(489), 614-630.

[22] Bates, Robert H. (2000), "Ethnicity and Development in Africa: A Reappraisal," *American Economic Review,* 90(2), 131-134.

[23] Berman, Eli (2000), "Sect, Subsidy and Sacrifice: An Economist's View of Ultra-Orthodox Jews", *The Quarterly Journal of Economics*, 115(3), 905-954.

[24] Bernard, Tanguy, Alain de Janvry and Elisabeth Sadoulet (2004), "Social Resistance to Institutional Change: Explaining the Emergence of Differentiating Organizations in Rural Senegal", mimeo, University of California at Berkeley.

[25] Biggs, Tyler, Mayank Raturi, and Pradeep Srivastava (2002), "Ethnic Networks and Access to Credit: Evidence from the Manufacturing Sector in Kenya", *Journal of Economic Behavior and Organization*, 49(4), 473-486.

[26] Bigsten, Arne, Peter Kimuyu, and Karl Lundvall (2000), "Informality, Ethnicity, and Productivity. Evidence from Small Manufacturers in Kenya", Department of Economics Working Paper 27, Goteborg University.

[27] Blanchard, Olivier J. and Lawrence F. Katz (1992), "Regional Evolutions" *Brookings Papers on Economic Activity, 1,* 1-75.

[28] Bloch, Francis and Vijavendra Rao (2001), "Statistical Discrimination and Social Assimilation," *Economics Bulletin*, 10(2), 1-5.

[29] Bolton, Patrick and Gerard Roland (1997), "The break up of Nations: A Political economic analysis," *Quarterly Journal of Economics, 112(4),* 1057-1090.

[30] Bossert, Walter, Prasanta K. Pattanaik and Yongsheng Xu (2003), "Similarity of Options and the Measurement of Diversity", *Journal of Theoretical Politics*, 15(4), 405-421.

[31] Brender Adi (2004) "Ethnic Segregation and the Quality of Local Governments in the Minority's Localities: Local Tax Collection in the Israeli- Arab Municipalities" unpublished

[32] Buchanan, James M. and Roger L. Faith (1987), "Secession and the Limits of Taxation: Toward a Theory of Internal Exit", *American Economic Review,* 77(5), 1023-1031.

[33] Burns, Nancy (1994) *The Formation of American Local Governments* Oxford University Press, New York

[34] Calabrese, Stephen, Glenn Cassidy and Dennis N. Epple (2002), "Local governments, fiscal structure and metropolitan consolidation", in W.G. Gale and J. Rothenberg Pack (eds.), *Brooking Wharton Papers on Urban Affairs 2002.*

[35] Canning, David and Marianne Fay (1993), "The Role of Infrastructures in Economic Growth", unpublished.

[36] Casella, Allesandra and James E. Rauch (Eds.), (2001), *Networks and Markets*, Russell Sage Foundation.

[37] Casella, Allesandra and James E. Rauch (2003), "Overcoming Informational Barriers to International Resource Allocation: Prices and Ties", *Economic Journal*, 113(484), 21-42.

[38] Caselli, Francesco and Wilbur J. Coleman (2002), "On the Theory of Ethnic Conflict", unpublished, Harvard University.

[39] Collier, Paul (2000), "Ethnicity, Politics and Economic Performance", *Economics and Politics, 12(3), 225-246.*

[40] Collier, Paul (2001), "Implications of Ethnic Diversity", *Economic Policy, 32(16), 127-166.*

[41] Collier, Paul and Jan W. Gunning (1999) "Explaining Africa's Economic Performance" *Journal of Economic Literature, 37(1), 64-111.*

[42] Costa, Dora L. and Matthew E. Kahn (2003a), "Civic Engagement and Community Heterogeneity: An Economist's Perspective" *Perspectives on Politics, 1(1), 103-111.*

[43] Costa, Dora L. and Matthew E. Kahn (2003b), "Understanding the decline in American Social Capital, 1952-1998", *Kyklos, 56(1), 17-46.*

[44] Dayton-Johnson, Jeff (2000), "The determinants of collective action on the local commons: A model with evidence from Mexico", *Journal of Development Economics*, 62(1), 181-208.

[45] Demange, Gabrielle (2005), "Group Formation: The Interaction of Increasing Returns and Preferences Diversity", in Demange and Wooders (eds) *Group Formation in Economics. Networks, Clubs, and Coalitions*. Cambridge University Press, forthcoming.

[46] Demange, Gabrielle and Myrna Wooders (2005), *Group Formation in Economics. Networks, Clubs, and Coalitions*. Cambridge University Press, forthcoming.

[47] Dudley, Geoff (1991), "Scale, Aggregation, and the Modifiable Area Unit Problem", *The Operational Geographer*, 9(3), 28-33.

[48] Easterly, William (2001), "Can Institutions Resolve Ethnic Conflict?", *Economic Development and Cultural Change*, 49(4), 687-706.

[49] Easterly, William and Ross Levine (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions", *Quarterly Journal of Economics,* 111(4), *1203-1250.*

[50] Ellickson, Bryan, Birgit Grodal, Suzanne Schotchmer and William R. Zame (1999), "Clubs and the Market", *Econometrica*, 67(5), 1185-1218.

[51] Epple, Dennis and Thomas Romer (1991) "Mobility and Redistribution" *Journal of Political Economy, 99(4), 828-858.*

[52] Esteban, Joan-Maria and Debraj Ray (1994), "On the Measurement of Polarization", *Econometrica*, 62(4), 819–851.

[53] Fafchamps, Marcel (2000), "Ethnicity and Credit in African Manufacturing", *Journal of Development Economics*, 61(1), 205-235.

[54] Fafchamps, Marcel (2004) *Market Institutions in Sub-Saharan Africa* MIT Press Cambridge Mass

[55] Fearon, James D. (1998) " Commitment Problem and the Spread of Ethnic Conflict" in David Lake and Donald Rothchild (eds) *The International Spread of Ethnic Conflict: Diffusion and Escalation* Princeton University Press, Princeton NJ

[56] Fearon, James D. (1999), "Why Ethnic Politics and "Pork" Tend to Go Together", unpublished, Stanford University.

[57] Fearon, James D. (2002), "Fractionalization and civil wars", unpublished, Stanford University.

[58] Fearon, James D. (2003), "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth,* 8(2), 195-222.

[59] Fearon, James D. and David D. Laitin (1996), "Explaining inter Ethnic Cooperation", *American Political Science Review, 90(4), 715-735.*

[60] Fearon, James D. and David D. Laitin (2003) "Ethnicity Insurgency and Civil War" *American Political Science Review,* 97(1), 75-90.

[61] Fisman, Raymond (1999), "Trade Credit and Productive Efficiency in Developing Economies", mimeo, Columbia University.

[62] Fisman, Raymond (2003), "Ethnic Ties and the Provision of Credit: Relationship-Level Evidence from African Firms", *Advances in Economic Analysis and Policy*, 3(1), Article 4.

[63] Florida, Richard (2002a), "Bohemia and Economic Geography", *Journal of Economic Geography*, 2(1), 55-71.

[64] Florida, Richard (2002b) "The Economic Geography of Talent", *Annals of the Association of American Geographers, 92(4), 743-55.*

[65] Friedman, David (1977), "A Theory of the Shape and Size of Nations", *Journal of Political Economy, 85(1), 59-77.*

[66] Garfinkel, Michelle (2004), "On The Stability of Group Formation: Managing The Conflict Within," *Conflict Management and Peace Science*, 21(1), 43-68.

[67] Ginsburgh, Victor and Shlomo Weber (2004), "Languages Disenfranchisement in the European Union," FEEM Working Paper 04-2004.

[68] Glaeser, Edward L. (2002), "The Political Economy of Hatred", unpublished, Harvard University.

[69] Glaeser, Edward L., David Laibson, Jose A. Scheinkman and Christine L. Soutter (2000), "Measuring trust",

[70] *The Quarterly Journal of Economics*, 811-846.

[71] Glaeser, Edward L, Jose A. Scheinkman, and Andrei Shleifer (1995), "Economic Growth in a Cross Section of Cities," *Journal of Monetary Economics*, 36(1), 117-143.

[72] Goldin, Claudia and Lawrence F. Katz (1999), "Human Capital and Social Capital: The Rise of Secondary School in America, 1910 to 1940", *Journal of Interdisciplinary History*, 29, 683-723.

[73] Greif, Avner (1993), "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition", *American Economic Review*, 83(3), 525-548.

[74] Gurr, Ted (1996) *Minorities at risk dataset,* University of Maryland.

[75] Habyarimana, James, Macartan Humphreys, Daniel Posner and Jeremy Weinstein (2004), "Ethnic Identifiability: An Experimental Approach", mimeo, UCLA.

[76] Hong, Lu and Scott E. Page (1998), "Diversity and Optimality", Santa Fe Institute Working Paper 98-08-077.

[77] Horowitz, Donald L. (1985), *Ethnic Groups in Conflict*, Berkeley, CA: University of California Press.

[78] Horowitz, Donald .L. (2001), *The Deadly Ethnic Riot*, Berkeley, CA: University of California Press.

[79] Humphreys, Macartan and Habaye ag Mohamed (2002), "Senegal and Mali", paper presented at the World Bank/Center for United Nations Studies Conference, Yale University, 12-15 April.

[80] Humphreys, Macartan, Daniel N. Posner and Jeremy M. Weinstein (2002), "Ethnic Identity, Collective Action, and Conflict: An Experimental Approach", unpublished, Harvard University and UCLA.

[81] Huntington, Samuel P. (1998) *The Clash of Civilizations and the Remaking of the World Order,* New York, NY: Simon & Schuster, Inc.

[82] Jackson, Susan E. and Marian N. Ruderman (1996) (eds), *Diversity in Work Teams*, Washington. D.C.: Americal Psychological Association.

[83] Jehiel, Phillipe and Suzanne Scotchmer (2001), "Constitutional Rules of Exclusion in Jurisdiction Formation," *Review of Economic Studies*, 68(2), 393-413.

[84] Karlan, Dean S. (2003), "Social Capital and Group Banking", mimeo, Princeton University.

[85] Katzenstein, Peter J. (1985), (ed.) *Small States in World markets: Industrial Policy in Europe,* Cornell University Press, Ithaca NY

[86] Kaufmann, Chaim D.(1996), "Possible and Impossible Solutions to Ethnic Civil Wars", *International Security, 20(4), 136-175.*

[87] Kaufmann, Chaim D.(1998), "When All Else Fails: Ethnic Population Transfers and Partitions in the Twentieth Century", *International Security, 23(2), 120-156.*

[88] Keefer, Philip and Stephen Knack (2000), "Polarization, Politics, and Property Rights: Links between Inequality and Growth", Policy Research Working Paper 2418, The World Bank.

[89] Khwaja, Asim I. (2000), "Can good projects succeed in bad communities? Collective action in the Himalayas", mimeo, Harvard University.

[90] Kochan, Thomas, Katerina Bezrukova, Robin Ely, Susan Jackson, Aparna Joshi, Karen Jehn, Johnatan Leonard, David Levine and David Thomas (2002), "The Effects of Diversity on Business Performance: Report of the Diversity Research Network", mimeo, MIT Sloan School of Management.

[91] Kyriacou, Andreas P. (2004), "Rationality, Ethnicity and Institutions: A Survey of Issues and Results", *Journal of Economic Surveys*, forthcoming.

[92] La Ferrara, Eliana (2002a), "Inequality and Participation: Theory and Evidence from Rural Tanzania", *Journal of Public Economics*, 85(2), 235-273.

[93] La Ferrara, Eliana (2002b), "Self-help groups and income generation in the informal settlements of Nairobi", *Journal of African Economies*, 11(1), 61-89.

[94] La Ferrara, Eliana (2003a), "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana", *American Economic Review*, 93(5), 1730-1751.

[95] La Ferrara, Eliana (2003b), "Solidarity in Heterogeneous Communities," in Phillipe van Parjis (ed.), *Cultural Diversity versus Economic Solidarity*, Francqui Scientific Library, Brussels: Deboeck Universite'.

[96] La Ferrara, Eliana and Angelo Mele (2003), "Racial Segregation and Public School Expenditure", mimeo, Bocconi University.

[97] La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer and Robert Vishny (1999), "The Quality of Government", *Journal of Law, Economics and Organization*, 15(1), 222-279.

[98] Laitin, David D. (1994), "The tower of babel as a coordination game: political linguistic in Ghana" *American Political Science Review, 88(3), 622-634.*

[99] Laitin, David (1995), "Marginality: A Microperspective", *Rationality and Society, 7(1), 31-57.*

[100] Laitin, David (1998), *Identity in Formation: The Russian Speaking Populations in the Near Abroad* Cornell University Press, Ithaca NY

[101] Laitin, David (2000), "What is a Language Community?" *American Journal of Political Science, 44(1), 142-54.*

[102] Lazear, Edward P. (1999a), "Globalization and the market for team-mates", *Economic Journal,* 109(454), 15-40.

[103] Lazear, Edward P. (1999b), "Culture and Language", *Journal of Political Economy,* Supplement, *107(S6), 95-126.*

[104] Levinson, David (1998), *Ethnic Groups Worldwide,* Phoenix, AZ: Oryx Press.

[105] Lijphart, Arend (1977), *Democracy in Plural Societies: A Comparative Exploration*, New Haven, CT: Yale University Press.

[106] Lijphart, Arend (1999), *Patterns of Democracy,* New Haven, CT: Yale University Press.

[107] Luttmer, Erzo F. P. (2001), "Group Loyalty and the Taste for Redistribution," *Journal of Political Economy*, 109(3), 500-528.

[108] Macours, Karen (2003), "Ethnic Divisions, Interlinkages and Search Costs in the Guatemalan Land Rental Market," unpublished, John Hopkins University.

[109] Madison, James, "Federalist Papers n.11".

[110] Massey, Douglas S. and Nancy A. Denton (1988) : "The Dimensions of Racial Segregation", *Social Forces, 67(2), 281-315.*

[111] McMillan, Margaret (2003) *Paris 1919: Six Months that Changed the World.* Random House, New York NY.

[112] Miguel, Edward, and Mary Kay Gugerty (2004), "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya", *Journal of Public Economics*, forthcoming.

[113] Milchtaich, Igal and Eyal Winter (2002), "Stability and Segregation in Group Formation", *Games and Economic Behavior*, 38(2), 318-346.

[114] Montalvo, Jose Garcia and Martha Reynal-Querol (2002), "Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth", unpublished, Universitat Pompeu Fabra.

[115] Nopo, Hugo, Jaime Saavedra and Maximo Torero (2002), "Ethnicity and earnings in urban Peru", mimeo, GRADE.

[116] Okten, Cagla and Una Okonkwo Osili (2004), "Contributions in Heterogeneous Communities: Evidence from Indonesia", *Journal of Population Economics*, forthcoming.

[117] Olson, Mancur (1965), *The logic of collective action,* Harvard University Press, Cambridge, MA.

[118] Openshaw, Stan and Peter J. Taylor (1979), "A Million Or So Correlation Coefficients: Three Experiments on the Modifiable Area Unit Problem", in N. Wrigley (ed), *Statistical Applications in the Spatial Sciences*, London: Pion, 127-144.

[119] O'Reilly, Charles L., Katherine Y. Williams and Sigal G. Barsade (1997) "Demography and group performance", unpublished.

[120] Ottaviano, Gianmarco and Giovanni Peri (2003), "The Economic Value of Cultural Diversity: Evidence from US Cities", mimeo, UC Davis.

[121] Ottaviano, Gianmarco and Giovanni Peri (2004), "Cities and Cultures", unpublished.

[122] Persson Torsten and Guido Tabellini (2003) *The Economic Effects of Constitutions* MIT Press Cambridge Mass

[123] Posner, Daniel N. (2004a), "Measuring Ethnic Fractionalization in Africa," *American Journal of Political Science*, 48(4), 849-863.

[124] Posner, Daniel N. (2004b), "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi", *American Political Science Review*, forthcoming.

[125] Poterba, James M. (1997), "Demographic Structure and the Political Economy of Public Education", *Journal of Policy Analysis and Management*, 16(1), 48-66.

[126] Prat, Andrea (2002), "Should a team be homogeneous?", *European Economic Review,* 46(7), 1187-1207.

[127] Rappaport, Jordan M. (1999), "Local Growth Empirics", Center for International Development at Harvard University Working Paper No. 23.

[128] Rauch, James E. (2001), "Business and Social Networks in International Trade", *Journal of Economic Literature*, 39, 1177-1203.

[129] Ray, Debraj and Rajiv Vohra (1999), "A Theory of Endogenous Coalition Structure," *Games and Economic Behavior,* 26, 286–336.

[130] Ray, Debraj and Rajiv Vohra (2002), "Coalitional Power and Public Goods," *Journal of Political Economy* 109, 1355–1384.

[131] Richard, Orlando, Thomas Kochan and Amy Mcmillan-Capehart (2002), "The Impact of Visible Diversity on Organizational Effectiveness: Disclosing the Contents in Pandora's Black Box", *Journal of Business and Management*, 8(3), 265-292.

[132] Rogowski, Ronald (1987), "Trade and the Variety of Democratic Institutions" *International Organization, 41(2), 203-223.*

[133] Sambanis, Nicholas (2000), "Partition as a Solution to Ethnic War: An Empirical Critique of the Theoretical Literature", *World Politics*, 52(4), 437-483.

[134] Sethi, Rajiv and Rohini Somanathan (2004), "Inequality and Segregation", *Journal of Political Economy*, 112(6), 1296-1322.

[135] Spears, Ian S. (2002), "Africa: The Limits of Power Sharing", *Journal of Democracy*, 13(3), 123-136.

[136] Spolaore, Enrico and Romain T. Wacziarg (2002), "Borders and Growth", *NBER Working Paper No. W9223.*

[137] Srinivas, Mysore Narasimhachar (1966), *Social Change in Modern India,* Berkley, CA: University of California Press.

[138] Tajfel, Henri, Michael Billig, Robert P. Bundy and Claude Flament (1971), "Social Categorization and Intergroup Behavior", *European Journal of Social Psychology*, 1, 149-178.

[139] Tavares, Jose and Romain T. Wacziarg (2001), "How Democracy Affects Growth," *European Economic Review*, 45(8), 1341-1379.

[140] Tiebout, Charles M. (1956), "A Pure Theory of Local Public Expenditure", *Journal of Political Economy*, 64(5), 416-424.

[141] Vigdor, Jacob L. (2004) "Community Composition and Collective Action:Analyzing Initial mail responses to 2000 Census" *Review of Economics and Statistics,* forthcoming.

[142] Weitzman, Martin L. (1992), "On Diversity", *The Quarterly Journal of Economics*, 107(2), 363-405.

[143] Wilkinson, Steven (2002), "Memo on Developing Better Indicators of Ethnic and Non Ethnic Identities", presented at LICEP 5th Meeting, Stanford University.

[144] Williams, Katherine Y. and Charles L. O'Reilly (1998), "Demography and Diversity: A Review of 40 Years of Research", in B. Staw and R. Sutton (eds.), *Research in Organizational Behavior*, vol. 20, Greenwich, CT: JAI Press.

7 Appendix

The data we use in this paper come from Alesina, Devleeshauer, Easterly, Kurlat and Wacziarg (2003). The authors use the Encyclopedia Britannica (2001). The variable "language" that underlies the fractionalization index based on it, refers to "the shares of languages spoken as mother language based upon national census data." Other sources for language data are the CIA World Factbook (which however is available for only a smaller set of countries) and the Ethnologue project that lists approximately 6,800 languages. Alesina et al. (2003) report that fractionalization for languages based upon these alternative sources are highly correlated with Encyclopedia Britannica.

The variable ethnic fractionalization combines the language variable above with other information about racial characteristics (normally skin color). Groups were classified as different if they spoke a different language and/or had different physical characteristics. Sources for physical differences were the Encyclopedia Britannica, CIA Factbook (2000) for 25 countries, Levinson (1998) for 23 countries and Minority Rights Group International (1997) for 13 cases. The rule used for data collection was: "if two or more sources for the index of ethnic fractionalization were identical to the third decimal point, we used these sources (the recorded source in this case was normally Encyclopedia Britannica). If sources diverged up to the second decimal point, we used the source were reported ethnic groups covered the largest share of the population"..The resulting ethnicity data covers 650 different ethnic groups in 190 countries, and is available on the web.

Table 1: Fractionalization and Long-run growth
(dependent variable is growth of per capita GDP)

	ETHNIC		LANGUAGE	
Variable	**1**	**2**	**3**	**4**
Dummy for the 1960s	0.059	0.153	0.065	0.156
	(3.357)	(5.144)	(3.563)	(5.248)
Dummy for the 1970s	0.057	0.158	0.062	0.161
	(3.093)	(5.222)	(3.280)	(5.333)
Dummy for the 1980s	0.036	0.141	0.042	0.145
	(1.940)	(4.601)	(2.213)	(4.725)
Dummy for Sub-Saharian Africa	-0.008	-0.016	-0.009	-0.014
	(-1.630)	(2.853)	(-2.026)	(-2.595)
Dummy for Latin America and the Caribbean	-0.016	-0.011	-0.019	-0.018
	(-4.458)	(-2.923)	(-5.252)	(-4.201)
Log of initial income	-0.004	-0.018	-0.004	-0.018
	(-1.499)	(-3.767)	(-1.660)	(-3.724)
Log of schooling	0.012	0.005	0.011	0.008
	(2.767)	(1.092)	(2.627)	(1.669)
Assassinations		-21.342		-13.988
		(2.212)		(-1.010)
Financial Depth		0.012		0.010
		(1.798)		(1.652)
Black Market premium		-0.021		-0.022
		(4.738)		(-4.953)
Fiscal Surplus/GDP		(0.128)		0.132
		3.369		(3.474)
Log of telephones per worker				
		(0.006)		0.004
		2.078		(1.488)
Fractionalization	**-0.020**	**-0.014**	**-0.019**	**-0.021**
	(-3.005)	**(-1.795)**	**(-2.979)**	**(-2.881)**
No of Observations	82; 88; 94	40; 69; 66	82; 86; 92	39; 68; 65
R-squared	.23; .17; .35	.32; .43; 54	.21; .21; .30	.36; .47; .52

(t-statistics in parentheses)
Estimated using Seemingly Unrelated Regressions: a separate regression for each 10 year period.

Table 2: Fractionalization and Long-run growth
(dependent variable is growth of per capita GDP)

	ETHNIC		LANGUAGE	
Variable	**1**	**2**	**3**	**4**
Dummy for the 1960s	0.064	0.220	0.098	0.253
	(2.522)	(5.116)	(3.910)	(6.827)
Dummy for the 1970s	0.061	0.226	0.096	0.260
	(2.369)	(5.179)	(3.735)	(6.897)
Dummy for the 1980s	0.041	0.209	0.077	0.245
	(1.542)	(4.757)	(2.951)	(6.411)
Dummy for Sub-Saharian Africa	-0.007	-0.014	-0.007	-0.011
	(-1.574)	(-2.479)	(-1.478)	(-2.138)
Dummy for Latin America and the Caribbean	-0.016	-0.013	-0.021	-0.019
	(-4.386)	(-3.233)	(-5.517)	(-4.787)
Log of initial income	-0.005	-0.027	-0.008	-0.031
	(-1.297)	(-4.253)	(-2.420)	(-5.523)
Log of schooling	0.012	0.006	0.011	0.009
	(2.775)	(1.112)	(2.599)	(1.966)
Assassinations		-21.880		-16.919
		(-2.311)		(-1.303)
Financial Depth		0.011		0.008
		(1.649)		(1.385)
Black Market premium		-0.021		-0.020
		(-4.736)		(-4.729)
Fiscal Surplus/GDP		0.136		0.146
		(3.618)		(4.048)
Log of telephones per worker		0.007		0.005
		(2.532)		(1.969)
Fractionalization	**-0.031**	**-0.129**	**-0.083**	**-0.214**
	(-0.655)	**(-2.319)**	**(-1.851)**	**(-4.382)**
Fractionalization * log of initial income	**0.001**	**0.015**	**0.008**	**0.025**
	(0.227)	**(2.084)**	**(1.279)**	**(3.977)**
No of Observations	82; 88; 94	40; 69; 66	80; 86; 92	39; 68; 65
R-squared	.23; .18; .35	.27; .48; .55	.22; .25; .28	.36; .55; .56

(t-statistics in parentheses)
Estimated using Seemingly Unrelated Regressions: a separate regression for each 10 year period.

Table 3: Fractionalization, Democracy and Long-run growth (dependent variable is growth of per capita GDP)

	ETHNIC		LANGUAGE	
Variable	**1**	**2**	**3**	**4**
Dummy for the 1960s	0.059	0.153	0.073	0.159
	(3.290)	(5.090)	(3.897)	(5.331)
Dummy for the 1970s	0.056	0.155	0.069	0.162
	(2.869)	(4.983)	(3.418)	(5.220)
Dummy for the 1980s	0.035	0.137	0.050	0.146
	(1.790)	(4.358)	(2.420)	(4.632)
Dummy for Sub-Saharian Africa	-0.008	-0.014	-0.006	-0.010
	(-1.628)	(-2.493)	(-1.371)	(-1.805)
Dummy for Latin America and the Caribbean	-0.016	-0.012	-0.020	-0.017
	(-4.521)	(-3.017)	(-5.324)	(-4.087)
Log of initial income	-0.004	-0.019	-0.006	-0.019
	(-1.619)	(-3.933)	(-2.274)	(-4.029)
Log of schooling	0.012	0.007	0.013	0.010
	(2.842)	(1.351)	(3.108)	(1.959)
Assassinations		-23.495		-14.057
		(-2.423)		(-1.045)
Financial Depth		0.012		0.012
		(1.951)		(1.897)
Black Market premium		-0.021		-0.023
		(-4.828)		(-5.169)
Fiscal Surplus/GDP		0.117		0.131
		(3.060)		(3.520)
Log of telephones per worker		0.006		0.004
		(2.185)		(1.610)
Fractionalization	**-0.014**	**-0.002**	**-0.017**	**-0.008**
	(-1.856)	**(-0.233)**	**(-2.187)**	**(-0.877)**
Democracy	**0.001**	**0.003**	**0.002**	**0.002**
	(0.867)	**(1.833)**	**(1.390)**	**(2.064)**
Fractionalization * Democracy	**-0.002**	**-0.005**	**-0.003**	**-0.005**
	(-1.230)	**(-1.871)**	**(-1.885)**	**(-2.489)**
No of Observations	82; 87; 93	40; 69; 66	80; 85; 90	39; 68; 65
R-squared	.23; .19; .34	.33; .46; .53	.21; .26; .27	.35; .52; .52

(t-statistics in parentheses)

Estimated using Seemingly Unrelated Regressions: a separate regression for each 10 year period.

Table 4: Fractionalization, Democracy and Long-run growth (dependent variable is growth of per capita GDP)

Variable	ETHNIC 1	LANGUAGE 3
Dummy for the 1960s	0.118	0.138
	(4.689)	(5.593)
Dummy for the 1970s	0.115	0.135
	(4.356)	(5.197)
Dummy for the 1980s	0.096	0.117
	(3.562)	(4.426)
Dummy for Sub-Saharian Africa	-0.005	-0.003
	(-1.053)	(-0.668)
Dummy for Latin America and the Caribbean	-0.017	-0.020
	(-4.793)	(-5.267)
Log of initial income	-0.012	-0.014
	(-3.398)	(-4.247)
Log of schooling	0.012	0.012
	(2.878)	(2.979)
Fractionalization	**-0.149**	**-0.170**
	(-3.510)	**(-4.135)**
Fractionalization * log of initial income	**0.017**	**0.020**
	(3.233)	**(3.769)**
Democracy	**0.001**	**0.001**
	(0.665)	**(1.228)**
Fractionalization * Democracy	**-0.002**	**-0.003**
	(-1.067)	**(-1.944)**
No of Observations	82; 87; 93	80; 85; 90
R-squared	.21; .33; .30	.20; .39; .25

(t-statistics in parentheses)
Estimated using Seemingly Unrelated Regressions: a separate regression for each 10 year period.

Table 5: Fractionalization and Population Growth in US Counties (dependent variable is growth in log of population 1970-2000)

Variable	1	2	3	4
Intercept	-0.088	0.902	-0.088	0.906
	(-1.600)	-18.66	(-1.600)	(18.690)
Log of population 1970	0.034	-0.034	0.033	-0.036
	(5.610)	(-6.860)	(5.170)	(-6.910)
Income per capita 1970[(a)]	.095	-0.071	0.100	-0.068
	(6.300)	(-5.630)	(6.360)	(-5.290)
Growth in log of population 1960-1970		1.619		1.620
		(44.730)		(44.740)
Northeast	-0.396	-0.273	-0.396	-0.271
	(-12.480)	(-10.960)	(-12.410)	(-10.800)
Central	-0.413	-0.318	-0.413	-0.316
	(-19.740)	(-19.330)	(-19.610)	(-19.080)
South	-0.115	-0.137	-0.116	-0.143
	(-5.220)	(-7.930)	(-5.010)	(-7.890)
Fractionalization 1960			**0.019**	**0.042**
			(0.370)	**(1.080)**
No of Observations	3133	3120	3120	3120
Adj. R-squared	.17	.50	.18	.50

(t-statistics in parentheses)
(a) Coefficient multiplied by 10^3.

Table 6: Fractionalization, Income and Population Growth in Counties (dependent variable is growth in log of population 1970-2000)

Variable	growth 1960-2000 1	2	growth 1970-2000 3	4	growth 1970-2000 5	6	growth 1980-2000 7	8
Intercept	-0.221	0.679	0.043	1.026	-0.096	0.974	-0.397	-0.229
	(-2.500)	(7.170)	(0.660)	(18.360)	(-1.450)	(16.930)	(-9.550)	(-7.160)
Log of population 1960	0.019	-0.038						
	(2.250)	(-4.510)						
Log of population 1970			0.029	-0.039	0.038	-0.036		
			(4.560)	(-7.470)	(6.020)	(-6.710)		
Log of population 1980							0.051	0.019
							(13.630)	(6.900)
Income per capita 1970[(a)]	0.216	0.049	0.059	-0.105	0.086	-0.094		
	(8.810)	(1.940)	(3.060)	(-6.790)	(4.440)	(-6.000)		
Income per capita 1980[(a)]							0.028	0.003
							(5.100)	(0.740)
Growth in log of population 1950-1960		0.965						
		(20.880)						
Growth in log of population 1960-1970				1.618		1.622		
				(44.830)		(44.720)		
Growth in log of population 1970-1980								1.060
								(52.880)
Northeast	-0.378	-0.236	-0.385	-0.261	-0.405	-0.269	-0.253	-0.017
	(-9.110)	(-5.990)	(-12.030)	(-10.390)	(-12.63)	(-10.660)	(-12.460)	(-1.080)
Central	-0.446	-0.369	-0.410	-0.314	-0.420	-0.317	-0.252	-0.040
	(-16.360)	(-14.260)	(-19.530)	(-18.990)	(-19.87)	(-19.010)	(-18.680)	(-3.800)
South	-0.085	-0.048	-0.130	-0.155	-0.101	-0.138	-0.067	-0.028
	(-2.500)	(-1.670)	(-5.530)	(-8.490)	(-4.410)	(-7.720)	(-4.990)	(-2.890)
Fractionalization 1960	**-0.727**	**-0.906**	**-0.581**	**-0.505**				
	(-3.270)	**(-4.280)**	**(-3.380)**	**(-3.770)**				
Fractionalization 1960 * Income per capita 1970[(a)]	**0.415**	**0.471**	**0.297**	**0.271**				
	(3.950)	**(4.690)**	**(3.650)**	**(4.270)**				
Fractionalization 1970					**-0.130**	**-0.335**		
					(-0.770)	**(-2.520)**		
Fractionalization 1970 * Income per capita 1970[(a)]					**0.001**	**0.151**		
					(0.020)	**(2.410)**		
Fractionalization 1980							**0.068**	**0.015**
							(0.670)	**(0.190)**
Fractionalization 1980 * Income per capita 1980[(a)]							**-0.033**	**0.025**
							(-1.970)	**(1.740)**
No of Observations	3120	3102	3120	3120	3133	3120	3137	3133
Adj. R-squared	.19	.29	.18	.50	.18	.50	.20	.58

(t-statistics in parentheses)
(a) Coefficient multiplied by 10^3.