The Policy Significance of Inequality Decompositions

By

Ravi Kanbur[*]
sk145@cornell.edu

August 2003

Abstract

Economists are now familiar with "between" and "within" group inequality decompositions, for race, gender, spatial units, etc. But what exactly is the normative significance of the empirical results produced by these decompositions? This paper raises some basic questions about policy interpretations of decompositions that are found in the literature.

---

1. Introduction

Consider n individuals and divide them into m mutually exclusive and exhaustive groups. Consider a variable y, which is thought to capture some aspect of wellbeing. Let µ be the mean and I an inequality index defined on y for a given population of individuals. Then, of course, µ and I can be calculated for each sub-group and for the total population. And as is well known, for the class of sub-group decomposable inequality measures, overall inequality can be written as the sum of a between group component and a within group component. Between group inequality is inequality calculated on the total population when each y in a group is replaced by the mean of y in that group. It reflects, therefore, the mean differences across the groups. Within group inequality is a weighted sum of the I's calculated for each of the groups. It reflects the inequality that exists "over and above" mean difference across groups.

The above decomposition exercise is very familiar by now. The groups can be defined by space, race, or gender, for example. The variable y is usually real per capita household expenditure of the individuals in the population, although it could be assets like land holding, years of schooling, etc. The inequality indices used are usually members of the Generalized Entropy class of measures, which includes Theil's first measure (in which case the within group component is an income share weighted sum of the Theil indices for each group) or Theil's second measure (in which case the weights are population shares of the groups). Sometimes variance or log variance are used, which

permit analysis of variance interpretations of the decomposition, but other (unattractive) properties of these two indices stand as a caveat to their use as inequality measures.

Whichever decomposition is done, it turns out that empirically, for gender (two groups) and race (usually less than five groups), the between group component is less than 15 percent. For space, it depends on how disaggregated a grouping is possible. For example, for rural Peru, it requires going below Region, below Province, and below Canton to the Parroquia level (there are 915 of these units), for the between group component still to rise only to 15 per cent[1].

2. Normative Uses of Decomposition

There is no question that inequality decompositions are an effective tool in the positive analysis of inequality. They allow useful depictions of patterns that can be a first step in identifying the proximate causes of inequality. But such inequality decompositions, and related statistics, also tend sometimes to underpin policy analysis and normative arguments, and this is where great care is needed.

For example, it could be argued, and is sometimes implicitly argued, that because the between group component of inequality for race, space and gender groupings is low, policy should be less concerned about equalizing outcomes across these categories, focusing instead on factors that determine the within group component. On race, the conclusion would be that in Malaysia, for example, policies for achieving balance in

average outcomes across Chinese, Indians and Malays should receive less weight than policies that equalize (say) educational achievements since it is the latter that account for the within race component of inequality, and it is this component which is the dominant component of total inequality.[2] A similar conclusion could be reached for gender in most developing countries.[3] On space, the conclusion might be that in rural Madagascar, where the between group component at the Fivondrona level is 15 percent, policy should focus on inequalities within Fivondronas, through community development projects and the like, and less so on equalizing mean differences between these units.[4]

There is a related tendency in some policy oriented discourses that rely on group decompositions to argue as follows. If decomposition by grouping type A leads to a larger between group component than decomposition by grouping type B, then equalizing across grouping of type A should have the higher policy priority. For example, if the between group component of education inequality is greater for a spatial decomposition, or a wealth categories decomposition, than it is for gender decomposition, then equalization across the first two groups should receive greater priority.[5]

These policy stances seem to follow naturally from the decomposition methodology. The between group component of inequality tells us, in an accounting sense, what inequality would be if there were no mean differences between groups.

---

[1] See Elbers et. al. (2002).
[2] For an early discussion on decomposing inequality in Malaysia, see Anand (1983).
[3] For a discussion of this, see Kanbur (2003).
[4] Elbers et. al. (2002).
[5] There is also a technical question of the extent to which the number of groups in any given classification affects the between groups component. Clearly, if groups are subdivided into further subgroups, the between group component will increase for this reason alone. See Kanbur (2002).

Surely this leads to high priority for equalizing across groupings that show a high between group component? But there are at least four sets of issues to be addressed—the independent role of groups in normative and policy analysis, the possibility of changing groups, the possibility of sharing within groups, and the importance of being specific on policy instruments and their costs.

3. The Significance of Groups

The individualistic roots of the economic literature on the measurement of inequality run very deep. Even the term "interpersonal inequality" shows that the key focus is on the difference between individuals, and groupings of individuals have significance only in so far as individual outcomes are aggregated across the group, and group patterns have significance only as part of the overall picture of inequality between persons. Groups qua groups have no special significance in the normative calculus. But if individual identity flows in part from group membership this may help to explain why it is the ratio of the mean incomes of two racial groups that has socio-political salience, rather than the (low) proportion of overall interpersonal inequality accounted for by the between race term in a standard inequality decomposition.

Another reason why groupings may have significance greater than their low between group contribution to inequality might suggest is that mean difference across certain types of groupings may be normatively unacceptable. For example, any income differences attributable only to race, or to gender, might be held to be abominable and

receive the highest priority, no matter what their contribution to overall interpersonal inequality. The social weight on these group differences might be far greater.

The above arguments may be easily accepted for race or gender, but what about administrative spatial units? Of course if these have racial, ethnic or linguistic differences, the earlier argument applies, but it can then be said that there is nothing special about space here, the real issue being race or ethnicity. But even without these differences, administrative units may develop their own identity through design or through practice. In terms of design, federated structures often have explicit arrangements that emphasize "equal weight" of each unit, politically and symbolically, irrespective of any other characteristics such as their population size. Thus, for example, each State in the US supplies two Senators to the US Senate; certain taxation and expenditure powers are constitutionally decentralized in Ethiopia and India, etc. Thus spatial units may develop special identities even without the basis of ethnicity, race or religion.

Such non-individualistic departures from the norm make economists uncomfortable, partly because it is a struggle to make sense of them in the framework of textbook economics. But there is increasing acceptance of the idea that social stability may well depend on these group characteristics, and the tools to think about such issues are now being developed.[6]

4. Changing Groups—Where Does Migration Fit In?

Inequality decomposition is done with fixed groups. As discussed above, part of the significance of difference across gender and race is because people do not have a choice in the group to which they belong, and it is therefore argued that differences across such groups take on special significance. But of course there is one sense in which space is fundamentally different from race. An individual's spatial location can be changed much more easily than her ethnicity or race.[7] This is perhaps why economist's intuitions lead them to "worry less" about spatial inequality than about racial inequality (even though the percentage of total inequality that these two explain is around the same). Now if spatial migration were genuinely free, then it would continue until all spatial disparities in real incomes were removed. Indeed, in the new economic geography models equilibrium is pinned down precisely by the equality of real wages of the (identical workers) in the different regions.[8] There is thus geographic concentration, but no real income inequality in these equilibria.

In reality, of course, migration is far from free. There are policy barriers to the movement of people, and adjustment to real income differentials may be slow.[9] This is all very well, but to my mind the really interesting theoretical question is--could real income (ie utility) differentials persist without restrictions on migration? On the face of it this seems impossible in our standard models. Consider two individuals in two locations,

---

[6] See Akerlof and Kranton (2000).
[7] This is not the place to get into the large literature on the evolution of ethnicity as a social construct. Suffice it to say that it is easier for a person of Chinese ethnicity in Malaysia to migrate to another town than to change ethnicity and become Malay—at least within a lifetime.

identical in every respect except that one has a higher utility level. Surely the one with the lower utility will move to take advantage of the conditions in the other place. But suppose now that each individual also has location specific human capital (social networks etc). Then we can have a situation where, comparing the two situations, neither individual wants to move, and one individual has higher utility than the other. The loss of location specific capital is what keeps each individual rationally in each place, and allows real income differences to persist. Some innovative modeling is called for here.

One exercise that might seem to come close to the above in spirit is Ravallion's (2002) identification of geographical externalities in rural China. He shows that (growth in) a household's wellbeing depends not only on that household's characteristics but also on the characteristics of the geographic region the household lives in. This does not, however, ensure that real income differences will persist, since in the logic of the model a household that migrates to another region has automatic access to ("takes on") the other region's characteristics. But what is needed for differentials to persist is that when a household moves it suffers, for some time, from a mismatch between the characteristics of the region it came from and the region it migrates to—a feature that is most strongly motivated, perhaps, for social networks. In order to test this we would have to have data on households or individuals who have moved recently.

---

[8] See Fujita, Krugman and Venables (1999).
[9] On official barriers to migration in China see Kanbur and Zhang (1999) and Au and Henderson (2002).

5. Sharing Within Groups

Continuing to stay with the individualistic tilt of the inequality measurement literature, let us stick with the assumption that what matters is individual real consumption. One reason why the 10-15 percent figure for between group inequality so goes against the basic ground level intuition that group means matter a lot to people, could be that there is more sharing within groups than is recorded in our household survey data. Thus the "within group" component as calculated from standard data is an overestimate, and between group inequality matters more than the 10-15 percent figure might suggest.

Household surveys have become progressively better at capturing sharing through interpersonal transfers, but they are some way from giving us a fully accurate picture of the benefits that individuals get from local public goods such as safety, security, sanitation, and even schools and health posts. Local institutions such as temples and mosques provide tangible and intangible public goods that are missed in our surveys. It could be this that leads to parochialism that is often commented on—a poor member of a community (suitably defined), would rather that an additional dollar went to a rich member of his or her community than that it went to a poor person of another community. There is no empirical work on this, of course, but such tensions between class and community solidarity, which we would all acknowledge, would make less sense in a world in which sharing within groups was not important.[10]

6. Policy Based Approaches

Ultimately, the inequality measurement tradition is interested in policy. Hence the jump, by some, from the observation that the contribution of spatial, or racial, inequality to total inequality is "small", to the policy conclusion that therefore policies to equalize group means across spatial or racial units are to have lower priority than other policies that would more vigorously attack the "within group" disparities (education is the favorite candidate here). But, in fact, this leap is not warranted without a closer examination of the patterns of inequality, the policy instruments to hand, their impacts, and their costs.

The key is to specify the policy instruments and work from there. Suppose we identify two classes of instruments—those that, broadly speaking, affect the group means and those that affect within group distributions (of course in practice there will not be quite such a clean division). Among the former could be differential budgetary allocation for infrastructure projects. Among the latter could be specific community development projects that reduce inequality within each group. The fact that the between group component of inequality is small might suggest that it is the latter that should be given priority. But it is easy to show that the analysis needs to be subtler than this. First of all, it can be that within group inequality is quite low for quiet a large number of the spatial units.[11] So there should in fact be targeting of the community development efforts at

---

[10] For a theoretical analysis, see Dasgupta and Kanbur (2002).
[11] This is shown by Elbers et.al. (2002).

those units where within group inequality is particularly high (if the objective is to reduce inequality), or of group transfers to communities where within group inequality is quite low (if there are concerns about elite capture within communities). Thus the real question is which of the two policy instruments will have a bigger impact on inequality per dollar of public expenditure. And the answer to this question is not clear a priori since, among other things, we do not know what the costs of each instrument are, and we have not specified the precise modalities through which they will have an effect on inequality between and within groups respectively.[12]

As mentioned earlier, another type of policy leap that is sometimes made is as follows. Grouping a population into, say, ten categories of education gives rise to a much bigger between group component than grouping them into, say, gender groups, racial groups, or spatial groups. Thus, it is implicitly argued, equalization across the ten categories of education will give a bigger "bang" than any of the other instruments. But this again falls foul of lack of specificity of policy instrument. What exactly is being proposed in each case? What are its costs?[13] What are its specific impacts on the distribution of income between and within groups? Without answers to these questions, implicitly pronouncing on policy instruments by looking at various between and within group decompositions is inappropriate.[14]

---

[12] The same applies to poverty, and this relates closely to the large literature on poverty targeting.
[13] Thus the number of categories across which the mean differences are to be removed could be a determinant of the cost side of the story, whereas, as noted earlier, continual subdivision of groups into subgroups will surely increase the between group of component by definition.

7. Conclusion: Some Interesting Research Topics

The object of these notes has been to raise a number of questions on the uses of inequality decompositions in policy analysis. The discussion suggests a number of avenues for further research:

1. Mathematical characterizations of the behavior of the between group component of inequality as the number of groups increases.

2. Philosophical characterizations of why group mean difference might matter more than just their contribution to the between group component of total interpersonal inequality.

3. Models of migration where real income differentials persist as equilibrium phenomena.

4. Models of unmeasured sharing through local public goods, and its impact on the disconnect between real and measured inequality; empirical work to better measure the benefits of group specific public goods.

5. Between and within group analysis of the impact of different instruments, and comparing the costs of alternative instruments to their benefits.

---

[14] This point is also emphasized in a different context in Kanbur, 2003.

References

Akerlof, George and Rachel Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*. 115(3): pp 715-53.

Anand, Sudhir. 1983. *Inequality and Poverty in Malaysia: measurement and Decomposition*. Oxford: Oxford University Press.

Au, Chun-Chung and Vernon Henderson. 2002. "How migration restrictions Limit Agglomeration and Productivity in China." Paper presented at the Cornell/LSE/WIDER Conference on Spatial Inequality and Development. Downloadable from www.unu.wider.edu.

Dasgputa, Indraneel and Ravi Kanbur. 2002. "Class, Community, Inequality." Cornell working paper. Downloadable from http://www.arts.cornell.edu/poverty/kanbur/Das-Kan5.pdf

Elbers, Chris, Peter Lanjouw, Johan Mistiaen, Berk Ozler and Ken Simler. 2002. "Are Neighbours Unequal? Estimating Local Inequality in Three Developing Countries." Paper presented at the Cornell/LSE/WIDER Conference on Spatial Inequality and Development. Downloadable from www.unu.wider.edu.

Kanbur, Ravi. 2003. "Education, Empowerment and Gender Inequalities." Paper presented to the Annual Bank Conference on Development Economics, April, 2002,

published in B. Pleskovic and N. Stern (eds.) *The New Reform Agenda*. Oxford

University Press for The World Bank. Downloadable from

http://www.arts.cornell.edu/poverty/kanbur/ABCDE.pdf.

Kanbur, Ravi and Xiaobo Zhang. 1999. "Which regional Inequality? The Evolution of

Rural-urban and Inland-Coastal Inequality in China." *Journal of Comparative

Economics*. 27: 686-701.

Ravallion, Martin. 2002. "Externalities in Rural development: Evidence for China."

Paper presented at the Cornell/LSE/WIDER Conference on Spatial Inequality and

Development. Downloadable from www.unu.wider.edu.

Fujita, M., Paul Krugman and A.J. Venables. 1999. *The Spatial Economy: Cities, regions

and International Trade*. MIT Press.