Luxembourg Income Study
Working Paper No. 234

What's Behind the Inequality We Measure?
An Investigation Using Latin American Data

Miguel Szekely and Marianne Hilgert

December 1999

What's Behind the Inequality We Measure?
An Investigation Using Latin American Data

by

Miguel Székely and Marianne Hilgert[*]

December 1999

Abstract: The use of income distribution indicators in the economics literature has increased considerably in recent years. This work relies on household surveys from 18 LAC countries to take a step back from the use of these indicators, and explore what's behind the numbers, and what information they convey. We find: a) that the way countries rank according to inequality measured in a conventional way is to a large extent an illusion created by differences in characteristics of the data and on the particular ways in which the data is treated; b) Our ideas about the effect of inequality on economic growth are also driven by quality and coverage differences in household surveys and by the way in which the data is treated; c) Standard household surveys in LAC are unable to capture the incomes of the richest sectors of society; so, the inequality we are able to measure is most likely a gross underestimation. Our main conclusion is that there is an important story behind each number. This story influences our judgement about how unequal countries are and about the relation between inequality and other development indicators, but it is seldom told or known. Perhaps other statistics commonly used in economics also have their own interesting story, and it might be worth trying to find out what it is.

Keywords: Inequality, income distribution, economic growth, household surveys, Latin America.

JEL Classification: D31, O12, O54.

Introduction

The use of income distribution indicators in the economics literature has increased considerably in recent years. One of the reasons is that after a long period of stability, income distribution started to deteriorate in many industrial countries, with sharper increases during the 1980s and 1990s.[1] This has attracted considerable attention and has contributed to the resurgence of income distribution measures as qualifiers of economic performance.

Another reason is the renewed interest in income distribution as a variable that can help explain other development indicators, such as the rate of economic growth.[2] This interest has been strongly influenced by the appearance of the Deininger-Squire (1996) database (DS henceforth) on inequality indicators for the world. The database mainly relies on published estimations from household surveys–which are perhaps the best source of information for the measurement of income inequality–to considerably improve on previous efforts of gathering information for comparisons across countries and over time.

As noted by Atkinson (1997), the literature on the measurement of inequality and the surge in statistical evidence on income distribution allowed by the availability of household surveys, have outpaced the development of the theory needed to interpret the facts. This situation will most probably prevail because of easy access to the DS data, the recent appearance of the World Income Distribution Data base (1999) by WIDER,[3] and the renewed efforts by countries and international agencies for holding household surveys, will result in an even more widespread and intensive use of income distribution indicators in the years to come.

The objective of this work is to take a step back from the use of household survey data for the purposes mentioned above and to explore what's behind the inequality we measure. Our approach should be viewed as a statistical or even an accounting exercise that intends to provide a clearer idea about what inequality measures taken from household surveys are really informative about. Our main goal is not to explain inequality. In this sense, the paper is similar in spirit to Atkinson and Brandolini (1999), who have subjected the DS data to careful scrutiny and pointed out the dangers of using this kind of information without knowledge about its noise-to-signal ratio.[4]

---

[1] Atkinson, Rainwater and Smeeding (1995), Atkinson (1998), Gottschalk and Smeeding (1997, 1998), and Plontchick *et al.* (1998), among others, have documented this fact.

[2] Benabou (1996), Alesina and Perotti (1996), Perotti (1994), and Panizza (1999) are among the long list of works on this issue.

[3] UNU/WIDER-UNDP World Income Inequality Database, Beta 3, 8 November, 1999, which builds on DS and offers an expanded version of this data set.

[4] For another discussion of the drawback of secondary databases, see Pyatt (1999).

We would like to stress that criticizing secondary data sources such as DS is not our objective. Rather, we try to contribute to the literature by pointing out the issues that users of indicators derived from household survey data should take into account for their analysis and for interpreting the evidence generated with these kinds of statistics.[5]

The paper takes three approaches for looking at what's behind the numbers. The first two, which are more of a statistical exploration, refer (i) to the importance of fleshing out the differences in the characteristics of the data and (ii) to the relevance of being explicit about the way in which the data is treated. The last, which is more economic-oriented, focuses on identifying what type of individuals are behind the inequality we measure.

Each of these approaches is relevant for at least four important issues. First the analysis provides a better idea of what each of the numbers is really measuring. Secondly, there are important implications for the ranking of countries. It is normally perceived that countries with relatively low inequality must be doing something right. Therefore, if our impression about the rankings changes when we take into account differences in the data or when we choose to treat the data in a different way, knowing what's behind the numbers will be crucial for making better judgments. The same argument applies to the third issue, which is the effect of inequality on other economic indicators. Finding out more about these numbers will allow for a better interpretation of the correlation between income distribution measures and other variables. Fourth, once we know better what type of information inequality indicators convey we will be in a better position to explain the causes of inequality.

Surprisingly, our analysis shows that the impression obtained about the ranking of countries in terms of inequality, and that our ideas about the effect of inequality on other development indicators, can be a mere illusion caused by differences in the characteristics of household surveys, and by the way in which the data is treated. This does *not* imply that household surveys or international data sets are not useful, but rather, that to improve the interpretation of any result, it is necessary to take these issues into account.

To look into each of the topics mentioned above, this paper uses household surveys for 18 Latin American and the Caribbean (LAC) countries. Although it is possible that our conclusions are specific to these cases, we believe that there are important lessons for other countries, and especially in the use of secondary international data sets. We focus on Latin America because we are more familiar with this data, and we restrict the analysis to household

---

[5] In their paper, Deininger and Squire (1996) themselves point out many of the caveats of their database. The authors were careful enough to warn about the comparability problems of the data they put together, but these warnings are seldom taken into account by users.

surveys–rather than, for instance, censuses or tax records–because this is the most widely used source of information for measuring inequality.

The paper is organized as follows. Section 1 presents the data. Section 2 explores the importance of accounting for differences in the characteristics of household surveys. Section 3 examines the importance of methodological consistency and making explicit the choices made in calculating inequality indexes. Section 4 shifts from a statistical to an economic-oriented approach by asking who is behind the inequality. Section 5 concludes.

1. The Data

This paper extensively uses household surveys from 18 Latin American & Caribbean (LAC) countries for the most recent year available to us.[6] The data covers 94% of the total population of the region. The earliest survey is for 1993 (Nicaragua), while we have three for 1995 (Ecuador, El Salvador and Paraguay), seven for 1996 (Argentina, Bolivia, Brazil, Chile, Dominican Republic, Honduras, and Mexico), six for 1997 (Costa Rica, Colombia, Panama, Peru, Uruguay, and Venezuela), and one for 1998 (Guatemala). Appendix Table A1 presents more information about each survey.

All but two surveys *apparently* fulfill the "good quality" requirements of the Deininger-Squire data, which are that: (i) the data has to contain information on all income sources; (ii) the unit of observation is the household or the individual; (iii) the data is representative at the national level.[7] The surveys for Argentina and Uruguay do not fulfill the last condition because they only cover urban areas, but they meet the other two. We include them because they are usually used for comparison within LAC, and since these are highly urbanized countries, the surveys cover most of their population (more than 90% in Uruguay and around 70% in Argentina). Throughout the paper we refer to Argentina and Uruguay without further clarification of their restricted urban coverage, but it is important that these differences be borne in mind.

As a benchmark for comparison in the rest of the paper, we estimate inequality in each country with what we call the "conventional" Gini. This estimate refers to the inequality of each

---

[6] Some of these surveys were obtained through MECOVI, a program sponsored by the Inter-American Development Bank, the World Bank, and ECLAC to collect and organize the existing household surveys in Latin America and to promote the implementation of new ones. The rest of the surveys were obtained directly through country statistical offices.

individual's household per capita income calculated by using sampling weights, dropping all missing and zero values, using the standard definition of the household unit, and using the widest possible definition of income. No adjustments for differences in the data are performed. We call this measure "conventional" because, among the studies in LAC that document the methods used to compute inequality (which is a minority), the Gini is estimated in this standard way.[8]

Table 1 presents the conventional Gini (multiplied by 100) and ranks LAC countries from most to least unequal.[9] There is a group of four countries (Paraguay, Brazil, Bolivia and Panama) with the greatest inequality, and which are highly unequal by any standard. For instance, if compared with all the "good quality" observations in the Deininger-Squire data beyond 1988, they appear to be among the 5 most unequal countries in the world (only after South Africa 1993). Another group of highly unequal countries follows, with Gini indexes between 55 and 57 points (Colombia, Nicaragua, Chile, Ecuador and Guatemala). Then we find Honduras, Mexico, Peru and El Salvador, all with Ginis above 50 points. Finally, the countries with a Gini of less than 50 points are Venezuela, Dominican Republic, Argentina, Costa Rica and Uruguay, all of which are still well above the world average of 40 points reported in DS.

The table also includes information for countries in other regions for reference and confirms that all the LAC countries considered are highly unequal by international standards.[10] The regional average for LAC is by far the highest and there are only two individual countries– Russia and the United States–with a Gini coefficient that is more or less comparable to the two most equal Latin American countries. The other 16 LAC countries have much higher inequality levels.

---

[7] The word "apparently" is written in italics because some of the household surveys used here also appear in the DS data presumably because they meet requirements (i) to (iii) but, as we show later, a more careful look at them reveals that some do not.

[8] Londoño and Székely (2000) review 111 studies on the measurement of poverty and inequality in the region and conclude this. By standard definition of household we mean the unit including all individuals that share the same budget, and where sub-units are counted as being part of the same household. The only transformation to the data was that incomes were deflated when necessary by using the CPI.

[9] In the cases of Colombia and Nicaragua, Honduras and Mexico, and Peru and El Salvador the differences are only marginal, but we still rank the countries differently for expositional purposes.

[10] The data for these countries was accessed through the Luxembourg Income Study (LIS). Inequality is calculated with the same methodology as for the LAC countries, but in these cases, the definition of family unit may not coincide with the standard used in LAC.

2. The Importance of Accounting for the Characteristics of the Data

For some economic indicators, such as the GDP of a country, a discussion about the quality and characteristics of the data is rather uncommon. There are international conventions that countries follow to produce this kind of information, so when the numbers are used for research or for evaluating economic performance, their origin, how they are calculated, or the characteristics of the data with which they were produced, are seldom under scrutiny. In the case of other equally important indicators such as the degree of income inequality, this is not the case. Even though there has been great improvement in the availability of inequality indicators, these numbers are still far from being produced with similar methods and are far from uniform in terms of quality and reliability. This issue is mostly overlooked in empirical research, and this section illustrates the importance of taking it into account.

We consider three sets of issues: (i) the characteristics of the sample, (ii) differences in survey quality and coverage of population groups, and (iii) differential coverage of income sources and geographic areas and differences in timing.[11] One important issue that we are not able to address due to the lack of data is that differences in policies across countries can have an important effect on our impression about how unequal countries are. For instance, if a country chooses to provide income transfers, which are captured by most surveys, the effect will be measurable and its impact will be taken into account in the computation of inequality indexes. However, if the choice is to provide price subsidies to consumption, the benefit may change the distribution of welfare, but its effect will not be captured by inequality statistics.

2.1 Differences in sampling

Two of the most basic issues commonly overlooked when using household survey data are the frame used to define sampling weights and the size of the sample. We consider them briefly, since we are unable to perform adjustments to the data to determine their importance.

Typically the most recent census is used as a sampling frame, and when the year of the survey does not coincide with a census year, the census data is projected. This would not be an issue if the most recent census were within a 5-year interval, which is the period considered for instance in the United Nations Population Statistics (1998) for their longest projections, but when the period is greater, the degree of precision of the weights declines. If the weights overstate the

6

relative importance of less or more unequal population subgroups, inequality would be under or overestimated. Table A1 in the Appendix shows that in most countries the relevant sampling frame is within 5 years of the year of the survey, but in Costa Rica and Uruguay the censuses of reference are 13 and 12 years apart, respectively. Unfortunately, we are unable to "adjust" inequality measures to account for these differences because of the difficulty in determining the sign and magnitude of the biases in the weights. So, other than documenting the facts, we can do little more than bear in mind that these two countries have been going through intensive demographic changes and have urbanized rather quickly, so it is possible that the use of outdated weights is blurring the comparison. Interestingly, these countries are ranked as most equal in Table 1.

The degree of precision of estimates from household surveys also depends on the size of the sample because the larger the size, the lower the standard error of any estimate. In LAC, samples range from 330,000 observations (individuals) in Brazil, to 11,905 in Argentina. Furthermore, the sample relative to the population also varies markedly. It ranges from almost 2%, 1.5% and 1.1% in Uruguay, Panama and Costa Rica, respectively, to .07% in Mexico. On average, the Gini index in all 18 countries has a confidence interval of .67 points, but the point estimates in Table 1 are subject to larger estimation error in Paraguay, Ecuador, Nicaragua, Argentina, Bolivia, Honduras and Peru, where the standard error is about 1 point of the Gini. Figure 1 shows the confidence interval for each conventional Gini. Some notable cases are Ecuador and Nicaragua. The point estimate for Ecuador ranks the country as the 8th most unequal in Table 1, but the conventional Gini is not statistically significantly different from Panama, which is ranked as the 4th most unequal. Nicaragua is originally ranked in 6th place, but after considering the standard error of the point estimate, it could well belong to the group of the five most unequal countries in the world. Thus, our impression about the ranking of countries in Table 1 is somewhat modified by accounting for the differences in the sample size of the household surveys from which the estimates are drawn. [12]

---

[11] Deaton (1997) and Atkinson and Micklewright (1992) also discuss some of these issues.
[12] As noted by Medina (1999), there are ways of "adjusting" inequality indexes for sampling differences. Due to the information requirements, we are not able to apply such methods here.

2.2 Quality and Coverage of Population Groups

Cross-country inequality comparisons can be blurred by differences in the quality of the information and by variations in the degree of success in defining a sample that is informative of all sectors of society. Here we explore some ways of taking these differences into account.

Measurement Error in Household Surveys

One way of obtaining an idea about the differences in quality across surveys is to assess the extent of some types of measurement error by focusing on indicators of misreporting, which are informative of the capacity of each survey to capture income with a higher degree of precision. This is actually one of the main quality concerns from the perspective of the measurement of inequality, because, for instance, if two countries have the same level of "real" inequality but in one the degree of underreporting among the rich is more severe, income distribution will appear to be better in one of them when in reality it is not.

Income misreporting is generally caused by two problems. One is that some incomes are particularly difficult to measure. This is typical of informal sector self-employment and small agricultural businesses, but also of the richest individuals, who usually have diversified portfolios with income flows that are not easy to value. Some surveys are better at minimizing these types of measurement error through the inclusion of specific questions on businesses, micro enterprises, the value of assets, and the returns to assets, but most are quite limited. The second problem is that some individuals may choose to underreport their income deliberately, even if they have a precise idea of its value. If underreporting is correlated with income or with income sources typically earned by specific sectors of society, it will introduce biases in inequality estimates.[13] Theoretically speaking the bias can be positive or negative, but generally the richest individuals are more reluctant to disclose their assets and wealth, so underreporting tends to result in an underestimation of inequality.

---

[13] Gottschalk and Smeeding (1998) discuss these issues in detail.

Gottschalk and Smeeding (1998) argue that one subgroup of the population subject to larger income measurement error due to the first problem is the self-employed. For these individuals it is usually difficult to distinguish what part of their income can be attributed to wages, what is the return to physical capital investment, and what is the value of their profit. Thus, the larger the size of this group the higher the proportion of the population potentially subject to misreporting.

The first column in Table 2 shows the proportion of self-employed in each of the 18 countries under analysis. For LAC as a whole 28.1% of the income earners are engaged in this type of activity, and in half of the cases the proportion is within one standard deviation from the LAC mean. Since surveys have limited information, if any, on the operation of micro enterprises and the finances and accounting of the self-employed, it is difficult to correct for differences in measurement error in this group by transforming the data in some way. Therefore, one of the only options is to compute the conventional Gini by excluding the group under the idea that the self-employed are suspected of being subject to larger error. Table 2 presents the result, as well as the new ranking of countries.[14] Even though the correlation between this indicator and the conventional Gini is positive and quite high (the coefficient is .86), there are some important ranking reversals. The main changes are that Guatemala and Peru appear to be relatively much more unequal than before. The position of El Salvador also deteriorates, while the position of Nicaragua improves substantially.[15]

Coverage of Socioeconomic Groups

An implicit assumption in using nationally representative income distribution indicators is that the statistics measure inequality among all individuals in society. Since household surveys use samples, it is known that not all individuals are included, but the fact that some subgroups are more likely to be underrepresented is seldom addressed. In particular, if the richest individuals are not sampled or are underrepresented, inequality will tend to be lower. Therefore, one indicator of survey quality is if the instrument is really able to capture information for the whole spectrum of

---

[14] The results are only evidence of the potential importance of differences in measurement error, because we are not able to control for the direction of the bias and for the fact that some countries are better at capturing the incomes of the self-employed than others. Furthermore, the comparison entails a significant loss of information.

[15] One way of partially evaluating the importance of misreporting due to the second problem mentioned above is by comparing the incomes in the surveys with some benchmark. In Table 2, we adjust the incomes for each labor income earner in the surveys so that the aggregate matches the labor income in the National Accounts and then add up total household income and recompute the Gini. This adjustment has only marginal implications for the ranking of countries, which suggests that underreporting of labor incomes, at least when accounted for in this way, is not an important source of variation across countries.

9

income earners. The poorest of the poor and the richest of the rich are usually the hardest to sample and if there are quality differences in the capacity to incorporate information on them, our impression about how unequal countries are can be misguided.

One way to assess the differences in population coverage across countries is to compare the lowest and the highest incomes in the surveys. This may give a good idea about the socioeconomic groups covered, and about the type of inequality each survey is measuring. Table 3 presents some information on this. The first column shows the monthly income in US dollars for the 10 poorest households in each country (we exclude all missing and zero household incomes for this comparison). The amounts are strikingly low. On average the 10 poorest households earn a total income of US$7.76 a month, which by any means is insufficient to cover the minimum necessary for survival. Although this may most likely constitute evidence of measurement error at the bottom of the distribution, it also confirms that all surveys capture at least to some extent the incomes of the poorest individuals.

The second and third columns present the monthly average income in US dollars of the 10 *richest* households and the richest person in the survey, respectively.[16] The results range from an average household income of around $38,000 in Brazil, to around $6,000 in Nicaragua, Costa Rica, Honduras and Bolivia. The income of the richest person in the survey ranges from $162,404 in Paraguay (which is a clear outlier), to a surprisingly low $7,000 in Costa Rica and Honduras. In Chile and Brazil, the income for the richest observation is also relatively high (of around $60,000), while in Mexico and Ecuador, the top observation registers an income around $40,000. In the rest, the incomes seem strikingly small, especially considering that these are regarded as the richest individuals in each country, at least in the household surveys. The case of Venezuela stands out in these comparisons. By LAC standards, this is a middle-high income country in terms of GDP per capita, and according to the surveys, the richest individual earns a monthly income of around $15,000, which seems implausible.

To obtain a better idea of how large or small these top incomes are, we include in the table the monthly disposable income in USD of a manager of a medium to large size firm obtained from an independent business survey.[17] We choose this benchmark because managers of medium to large size firms are highly educated professionals who do not typically fit the profile of the richest individuals in any of these countries. The comparison with the benchmark is quite striking (last two columns of table 3). On average, in the 16 countries for which information is

---

[16] We do not include the result for Colombia here because the Gini for this country is conventionally calculated by top coding incomes, so the highest observations are not usually used.
[17] The source is *America Economia* (1999), pg. 76-77, which reports the results of an independent survey by Price Waterhouse.

available, the total income of the 10 richest households in the survey is very similar to the average wage of a manager. In 10 countries, the average income of managers is actually higher than the income of the 10 richest households. In Brazil, Guatemala, Chile and Uruguay, the income of the richest is between 1.3 and 1.93 times greater than the income of the average manager, which is still far below expectations. Only in Ecuador and Paraguay does the income of the 10 richest households exceed the income of managers by a factor of more than 2. A similar conclusion applies to the comparison with the richest individual in the survey. In 9 out of 18 cases, the income of the richest individual is less than 1.5 times the income of an average manager. In Ecuador, the ratio is 5.76, and only in the case of Paraguay is the comparison within expectations (in this country the richest individual earns 26.56 times more than the average manager).

But even in the case of Paraguay and Ecuador it cannot be said that household surveys have full coverage of the richest individuals. To illustrate this, Figures 2a and 2b plot all incomes in the survey. Figure 2a shows how the richest individual in Paraguay is a clear outlier.[18] This income is almost 11 times greater than the second highest observation, which reports a monthly income of $16,000. In the case of Ecuador, the richest individual is also an outlier with an abyss between this observation's income and the rest. The cases of Chile and Brazil are not as extreme. These two surveys apparently capture incomes above the average of managers, and as can be seen in Figures 2c and 2d, the coverage of high incomes is more widespread than in Figures 2a and 2b. Two other interesting cases are Venezuela and Mexico. As reported in Table 3, the richest individuals in the surveys earn incomes that are much lower than the typical manager in the country. But what is more surprising is that the highest incomes in the surveys of these countries are still outliers in spite of their low value (see Figures 2e and 2f).

One important conclusion that emerges from these results is that inequality is grossly underestimated because of extremely limited information on the richest individuals, but that the degree of underestimation differs from case to case.[19] For instance, the relative position of Chile in Table 1 is rather unfavorable while the position of Venezuela is better. According to Table 3, however, this may in part simply reflect that the Chilean survey is better able to capture information on the richest segments of society. Only Brazil, Guatemala, Chile, Uruguay, Ecuador

---

[18] We confirm that the observation is not dropped in conventional computations of inequality for this country.

[19] The differences in coverage among the rich could be due to differences in survey quality, but it is also possible that they are a normal consequence of random sampling. If there are two countries with exactly the same distribution of income and exactly the same number of rich persons, the random selection of observations among the rich groups could result in higher inequality in one case if, by chance, a richer individual is included in the sample. In any case, our results show that whatever the cause, there are substantial differences in coverage of the richest sectors of society across the Latin American countries under study, and that these differences can affect income rankings.

and Paraguay seem to capture incomes above the level typically earned by managers of medium to large firms, and interestingly, with the exception of Uruguay, these are all countries with relatively high inequality according to the conventional Gini.

2.3 What do Surveys Capture, Where are they Applied, and When?

One of the most important differences across household surveys is that a common definition of income sources is not applied across countries. The omission of different sources affects inequality because some types of income are better distributed. If all surveys omitted the same source, the lack of information would not bias the comparisons and would not be an issue in their use for regression analysis. But at least for the LAC countries considered here, this is not the case.[20]

Table 4 aggregates incomes into 4 different categories: labor income, capital rents, property rents and non-monetary income. If capital rents and property incomes are grouped together in a single question that asks about all non-labor income, the shaded area appears in the second column only. We make this distinction because capital and property rents tend to be concentrated among the richest households, and when questionnaires have specific questions for these sources their accuracy is improved, and this may have an effect on inequality estimates.

As shown in Table 4, all surveys include a separate question for labor incomes, and Honduras and Nicaragua are the two cases where the surveys only capture this source. In half of the countries there is an explicit question about capital and property rents, respectively, while in 7 cases all non-labor income is grouped together in the same question. Furthermore, one of the most important differences is that only 7 countries include information on non-monetary sources, and of these, not all include estimations for imputed rents.

The differences are rather surprising since all the countries in the table (except for Argentina and Uruguay, which are urban surveys) supposedly meet the good quality requirements in the DS "high quality" data and are regarded as including all income sources. The surveys in the table, or similar ones for the same countries for previous years, are actually included in the DS database in spite of the differences in income source coverage. Some notable cases are the inclusion in the DS data of a series of Mexican surveys covering the wide range of sources

---

[20] All of the LAC countries for which we have information report after-tax incomes in household surveys so we do not need to worry about comparability in terms of net or gross income, as is common in developed countries.

specified in Table 2, along with information for Nicaragua 1993 and a series of surveys for Honduras which only consider labor incomes.

Table 5 illustrates what happens to our impression of the ranking of countries when, rather than comparing the conventional Gini, we restrict the estimation of inequality to a more comparable definition across countries. Columns 3 and 4 refer to the inequality of labor incomes, which is actually the only income definition that is strictly comparable across the 18 cases. The relative position of 8 countries changes by more than one place when we rank according to this more comparable concept. Bolivia, Panama, Nicaragua, Honduras, and especially Colombia now appear to be countries with relatively lower inequality than before, while Chile and particularly Mexico and Guatemala appear to be much more unequal.

We also estimate the Gini for labor incomes in urban areas, which is strictly speaking the minimum common denominator across all 18 countries. Column 5 and 6 in Table 5 present the ranking according to this measure. In 12 out of the 18 countries, countries are re-ranked by more than one position. These changes are somewhat expected, due to the large differences in the share of the urban population across countries, but are still illustrative about the importance of comparability across countries. Now Bolivia does not appear to be the third most unequal country, but is repositioned in 10th place. The ranking of Peru also improves significantly, since it now appears to be the second most equal country. The relative positions of Chile, Mexico, Guatemala and the Dominican Republic deteriorate substantially, and most importantly, the relative position of Argentina and Uruguay is less favorable when the countries are ranked according to the minimum common denominator.

One difference that is usually overlooked in using income distribution indicators from household surveys is that surveys are held at different times of the year in each country.[21] This becomes an issue if economic activity follows a cyclical pattern or if the composition of households changes due to seasonality.[22] For the sectors of the population that have stable employment in the formal sector and receive wages irrespectively of the time of the year this may not be an issue, but it is typically important for the self-employed and for individuals employed in

---

[21] Another potentially important issue is that recall periods for income could differ. However, Table A1 in the Appendix indicates that recall periods across the LAC countries in our sample are very similar. The only case where the period exceeds 1 month is Guatemala. In Mexico, surveys ask about income receipts in each of the 6 months prior to the survey. We only use information about income during the previous month to make it comparable with the rest of the countries in LAC.

[22] A fixed effects regression performed with data for the LAC countries in our sample for the period 1980-1998, and by using quarterly GDP (from International Financial Statistics (October, 1999)) as dependent variable and dummies per quarter as independent variables, reveals that on average GDP in the second quarter is 3.73 times larger than in the first. For the third and fourth, the difference is 0.94 and 3.39, respectively. Therefore there is high cyclicality during the year.

the agricultural sector. The larger the latter groups are, the greater the potential bias introduced in international comparisons by differences in timing.

Table 6 marks the month of the year over which each survey is performed. The first feature is that there are three cases (El Salvador, Guatemala and Uruguay) where the questionnaire is applied over the whole year. Among the remaining 15 cases, there are a few where the timing of the survey coincides. In 7 cases, there is an overlap during the month of September, but in four of them not all the observations are captured during this month. The rest of the countries are scattered around the calendar year.

In an attempt to assess, at least in a general way, the effect of these differences on country rankings, we perform a regression using the 18 surveys analyzed so far, and add observations from 37 other LAC surveys which are representative at the national level, and for which we also have direct access and full documentation (see Appendix Table A.2 for a description). The additional national surveys included in the estimation are detailed in Appendix Table A2, and refer to almost the same set of countries considered so far, but for earlier years. Argentina, Uruguay, and several earlier surveys for Bolivia that are available to us are excluded from the estimation because they refer only to urban populations. With the 55 nationally representative surveys to which we have access, we estimate the following random effects regression:

(1)

$$Gini = \underset{(6.5)}{\beta_1}(non.mon) + \underset{(9.8)}{\beta_2}(non.lab) + \underset{(-8.5)}{\beta_3}(cap.rent) + \underset{(-6.7)}{\beta_4}(Q_2) + \underset{(3.5)}{\beta_5}(Q_3) + \underset{(-3.8)}{\beta_6}(Q_4) + \underset{(10.6)}{\beta_7}(yearly) + \varepsilon_i$$

where all the independent variables are dummies. *d(non.mon)* takes a value of *1* if the survey captures non-monetary incomes and zero otherwise; *d(non.lab)* indicates that the survey reports non-labor income; *d(cap.rent)* is given a value of *1* when the survey questionnaire explicitly asks about the amount of capital incomes received, rather than grouping all non-labor income in the same question; Q_2, Q_3, and Q_4, are dummies for the quarter of the year over which the survey was held, while *d(year)* takes a value of 1 if the survey was held during the course of a whole year. ε_i is the residual.

The regression yields a gross estimate of the average effect on the conventional Gini of having or not certain household survey characteristics, while controlling for others.[23] The

---

[23] We restrict the number of independent variables in the regression because it is performed strictly for descriptive purposes, and not with the objective of explaining inequality. Anyhow, it may still be argued that there is a potential problem of omitted variables bias, and that the coefficients are also capturing the

coefficient is presented in parenthesis below each variable. The inclusion of non-monetary incomes is associated with about 6.5 Gini points more.[24] Having information on non-labor incomes is associated with 9.8 more points of the Gini, while including an explicit question for capital rents is associated with 8.5 points less. Surveys held during the second and fourth quarters of the year typically have a Gini of around 6.7 and 3.8 points less, while those held in the third quarter have 3.5 points more. Surveys carried out throughout the year have on average 10.6 points more.

We interpret the residual ε_i in each equation as the conventional Gini purged from differences in the household survey characteristics included in the regression. We call this "adjusted" Gini, and the results for each country's most recent observation (which we have been using throughout the paper) is presented in Table 5. The results should be taken with caution, since the number of observations (55) is quite reduced and the dummies for some of the household survey characteristics specified in the equation can be capturing country effects. We regard these results only as a first approximation to obtain a general idea about the importance of household survey differences. Although Argentina and Uruguay are not included in the estimation of equation (1) because they have only urban surveys, we also present the adjustment for these countries. In the case of Argentina the result should be taken with more caution because the adjustment does not take into account differences in geographic coverage. However, in the case of Uruguay, where most of the population is urban, the adjustment seems to make more sense.

The last column in Table 5 indicates the difference between the rank according to the "adjusted" Gini and the conventional Gini. In many cases the re-rankings are quite important. According to the adjusted Gini, the most equal countries are Uruguay, Costa Rica (both of which do not change rank) and El Salvador, while the most unequal are Ecuador, Paraguay, Brazil and Mexico. Surprisingly, Argentina now ranks much less favorably than before, and is not even among the 9 most equal countries. Colombia is now ranked as number 15 rather than 10, while the position of Nicaragua improves substantially. The relative positions of Panama, Guatemala

---

effect of variables that are correlated with inequality and also with some survey characteristics. Since there is no widely accepted model of inequality, we perform the regressions including parsimoniously the share of rural population (which might affect seasonality and the efforts to include specific income sources), the share of population of working age to control for demographics (which might affect the importance of transfers and self-employment), and a measure of factor endowments developed by Spilimbergo *et.al.* (1999), as well as a proxy for financial depth measured by the coefficient of M2 over GDP, to account for the importance of capital incomes. The absolute value and significance of the coefficients in both regressions change only marginally when including each of these variables at a time, or all together, so we don't present them for brevity.

[24] This result is in line with the effect found by Gottschalk and Smeeding (1998).

and Nicaragua are now much more favorable, while the position of Venezuela worsens in a significant way.

Therefore, countries that are usually regarded as relatively equal in LAC turn out to be much more unequal when the survey differences are accounted for, at least in this very general way. The rankings in Table 1 now appear more as an illusion created by differences in the household surveys from which the data is drawn, rather than being genuine disparities across countries.

2.4 Are the Differences Relevant for the Growth-Inequality Relationship?

As indicated in the introduction, one of the main reasons why the use of income inequality indicators in the economics literature has surged in recent years is because of the renewed interest in the relation between growth and inequality. During the 1950s and 1960s the conventional view was that greater inequality led to higher rates of growth. More recent evidence, however, points to the opposite conclusion: that inequality has negative effects on growth.[25] This view has been disputed by Forbes (2000), who recently showed that with the use of the DS data and the application of improved estimation techniques, the conclusion is that there is a positive relationship between the two variables, which takes us back to the idea that inequality is favorable for growth. In this subsection we add to this controversy by exploring whether incorporating more information about the household survey characteristics discussed above changes our view about the relation between these two indicators.

As a benchmark for comparison, we take the base regression in Forbes (2000), which uses the specification in equation (2). This is similar to the one used by Perotti (1996), which is one of the standard references in the literature for the negative relation between growth and inequality:

(2)

$$Growth_{it} = \partial_0 + \partial_1 Gini_{it-1} + \partial_2 LnGNPpc_{it-1} + \partial_3 Educm_{it-1} + \partial_4 Educf_{it-1} + \partial_5 PPPI_{it-1} + \boldsymbol{q}_i + \boldsymbol{h}_t + u_{i,t}$$

where *LnGNPpc* is the log of real GNP per capita in 1987 \$US for country *i* in period *t*, *Educf* and *Educm* refer to the years of secondary education of females and males over 25 years of age, and *PPPI* refers to a measure of market distortions proxied by the price level of investment. θ is a

set of country dummies and η are period dummies. Forbes' main results are obtained by estimating equation (2) for the 1966-1995 period by using growth rates over five-year intervals rather than yearly observations. The first column in Table 7 reproduces one of the basic results of the paper, obtained with a fixed effects estimation.[26] As can be seen, inequality has a strong positive effect on growth.

Ideally, we would like to re-estimate the regression by incorporating information on the specific income sources and timing of each household survey to check whether our impression about the effect of inequality on growth holds when the differences in the data are taken into account.[27] Since we do not have direct access to all the household surveys in the DS database, we restrict our analysis to the set of surveys in Appendix Table A2, which were previously used for the estimation of equation (1). [28] However, this imposes some restrictions. Since we have a more limited number of observations and countries, it is not possible to aggregate the observations in five-year intervals, so we estimate the equation by looking at the growth rate between the two periods for which a Gini index is available. The definition of the intervals will therefore depend on the availability of information on inequality.[29]

Before presenting our results with the LAC sample, we perform a regression to check whether the changes in method required for the LAC estimation have an effect on the coefficients or significance of Forbes' regression. The second regression in Table 7 is a fixed effects estimation using the same DS data for the period 1965-1995 but instead of grouping the observation in five-years intervals, it estimates the regression by defining each interval as the period between two years for which a Gini is available. Period dummies are substituted for year t dummies.[30] Even though the sample size and number of countries vary with respect to the first

---

[25] Benabou (1996) and Aghion *et al.* (1999) summarize the empirical and theoretical aspects of this literature, respectively.

[26] This corresponds to regression (1) in Table 3 in Forbes (2000). The main results of the paper are based on the fixed effects estimation as well as other techniques. We focus only on the fixed effects because the number of observations in the LAC data we use later is only suited for this method.

[27] One issue we do not examine, but which can also be important, is aggregation. As noted by Ravallion (1998) aggregation can create spurious effects of inequality on growth.

[28] 18 of these observations actually appear in the DS data set. The correlation between the 18 Gini indexes that are both in our sample and in DS, is .74.

[29] Since the timing of the Ginis in our LAC sample do not always coincide with the years for which information on education is available, we interpolate *Educf* and *Educm* to increase the number of observations in the regression. The use of all the information without grouping incorporates the effect of short-term disturbances and increases serial correlation from business cycles, but as will be shown, this has no implication for our conclusions.

[30] As in Forbes (2000), the Gini coefficient is taken from the DS good quality data set, education variables come from Barro-Lee, the log of GNP per capita is taken from the World Bank World Development Indicators 1998, while the index of price level of investment is from the World Penn Tables. The coefficients for the period dummies are not included for expositional purposes. 6.6 points are added to the

regression, the results lead to the same (although weaker) conclusion that inequality has a positive effect on growth. This assures that the change in method needed for the following estimations does not affect the overall conclusion of the positive statistically significant relation between the two variables.

The third regression in the table replicates the previous regression, but we substitute the DS data for the conventional Gini indexes computed from the LAC household surveys with national coverage, to which we have direct access (Table A2).[31] The magnitude of the coefficient on inequality, initial income and *PPPI* are smaller than in the first of the regressions in the table, while the significance level is also slightly lower. The number of observations is reduced (42 after calculating growth rates) and the number of countries is also limited to 11. However, the conclusion that inequality and growth are positively and significantly related holds.

As mentioned previously, we know that the surveys in the LAC sample used so far and many of those included in the DS data are not strictly comparable for several reasons. We are able to document the differences thoroughly in the case of income sources and timing and can therefore attempt some adjustments to increase comparability in the LAC data. One option is to reduce the definition of income to the minimum common denominator, which will guarantee that the variance in the independent variable will be "real" and not an artificial variation introduced by differential income coverage. Additionally, this will allow for a much more transparent interpretation of the results, because we know precisely what type of inequality is being measured.

The fourth equation in Table 7 replicates the third regression, but substitutes the conventional Gini for the Gini of labor income, which is comparable across countries. Surprisingly, the coefficient for inequality is reduced quite dramatically in size, and loses its statistical significance, while the sign and magnitude of the other coefficients is not modified in an important way.

We also perform the regression with the "adjusted" Gini discussed above, rather than the conventional or the labor income Gini. This has the advantage of also incorporating information on differences in timing across surveys. The fifth equation in Table 7 shows that surprisingly, our impression about the effect of inequality on growth is different than in equations (1) to (3) in this case also. The coefficient for the Gini is much smaller than before, but most importantly, it does not have a statistically significant relation with growth.

Gini coefficients that refer to the distribution of expenditures, to be consistent with the first regression in Table 7.

[31] The regression does not include period dummies due to the loss of degrees of freedom.

We would like to stress that the objective of the estimations in Table 7 is *not* to take a position regarding the effect of inequality on economic growth. In fact, the results should be interpreted with caution and cannot be regarded as solid evidence for the hypothesis that inequality is harmful or beneficial for growth. Our sample size, period coverage and number of countries are too limited for this. We also stress that we do not intend to contest the conclusions from Forbes' estimation because information on the characteristics of all the household surveys from which the DS data are derived is not available to the author or us. Our intention is simply to determine whether or not knowing more about the origin of the inequality measures is important. The results in Table 7 are quite striking because they show that our impression about the effect of inequality on other development indicators can be a mere illusion caused by differences in the characteristics of the data. Once this information is taken into account, the conclusions change.

3. Treatment of the Data

Apart from the differences in household survey coverage and characteristics, there are several important choices and assumptions behind inequality measures which are usually not accounted for in international secondary data sets and cross-country comparisons.[32] For instance, to compute the conventional Gini, we have implicitly assumed that: (i) each member of the household has the same needs; (ii) there are no economies of scale in consumption; (iii) missing and zero incomes are unreliable and are better to ignore; and (iv) the best way to summarize the information on the distribution of income is to compute the Gini coefficient. These four choices are reasonable, but they are by no means the only options. We have used them because in LAC it is standard to estimate inequality in this way, but in other regions, alternative methods, which are as valid or even more adequate, are traditionally employed. For instance, as noted by Atkinson, Rainwater and Smeeding (1995), there is a tradition in OECD countries of applying adult equivalence scales to account for some economies of scale in consumption and to impute missing incomes, but these practices are not common in LAC. This is an important issue because even in the absence of differences in the quality and characteristics of household surveys, international comparability of income distribution indicators is still not guaranteed if there are cross-country differences or variations over time in the way in which the data is treated. This section first explores how sensitive the LAC country rankings are to changing the methodology consistently across

---

[32] One of many examples is the introductory chapters to the volume by Ganuza, Morley and Taylor (1998).

countries. We then examine the implications of not applying the same method in all cases. Finally, we look into the importance of the choice of inequality measure.

Other important issues in data treatment are, first, the use of income adjustments to correct for price variations across regions. The way these adjustments are performed or their neglect can have significant effects on Gini coefficients, and, perhaps, country rankings.[33] Second, as shown by Ravallion and Chen (1999), inequality indexes are also highly sensitive to the prices used in the estimation of non-monetary incomes, and specifically the estimation of auto consumption. Third, as explained by Pyatt (1999), it is not uncommon to find that some authors reweigh the observations in household surveys to make them match, for instance, with aggregates in the National Accounts. We do not address any of these issues here due to data limitations.

3.1 Household Size, and Missing and Zero Incomes

So far in this paper household per capita incomes have been used as a welfare indicator for the measurement of inequality. Perhaps the clearest way of making explicit what this choice implies is to express each individual's income as:

(3)
$$y_{i,t} = \frac{Y_t}{[\sum_{i=1}^{n} i^{b(i)}]^a}$$

where y is the income assigned to individual i belonging to family t, Y is total household income of household t, and n is the total number of household members. Each member of the household is scaled up or down by an adult equivalence scale β specific to each individual, indicating the share of household resources allocated to i.[34] α is a parameter whose value is between 0 and 1 that accounts for economies of scale in consumption. Calculation of income per capita is equivalent to defining *b=1* and *a=1*, which assumes that each individual within the household has the same needs, and that there are no economies of scale in consumption. As restrictive as this might appear, it is standard practice in LAC and other regions. Since the household size for individuals living in households at the bottom of the per capita income distribution tends to be larger, attributing a value of 1 tends to overestimate inequality.

---

[33] Two good illustrations of the importance of this issue are Ravallion and Chen (1999) and Ferreira *et al.* (1999).

[34] β generally varies by age and gender. The extensive literature on the estimation of equivalence scales has been recently reviewed by Deaton (1997) and Cowell and Mercader-Pratts (1999).

Table 8 presents the Gini index calculated with alternative values for b and a.[35] To allow for economies of scale in consumption, we define a=.5, which is the value commonly used in studies for developed countries.[36] According to the first column in the table, the average Gini in LAC is reduced from 53.27 points to 51.11, but there are only few ranking reversals with respect to the conventional Gini. Another estimation of interest is to assign a=0, which implies the extreme assumption of "full" economies of scale. We obtain the results for this specification, but the calculation is restricted only to household heads, so it is equivalent to calculating the Gini over fractiles of households rather than individuals. In this case also, the re-rankings as well as the change in the average Gini across countries are not important, so we do not present them for brevity.

In the third column we estimate inequality using the Amsterdam equivalence scale explained by Deaton and Muellbauer (1980), which has commonly been used in the literature for developed countries.[37] The average Gini with this scale is 2.5 points lower than the conventional measure, and again, the ranking of countries is quite stable.

Thus, when a methodology is applied consistently across countries for accounting for economies of scale in consumption and equivalence scales, the country rankings do not change significantly, although the Gini is more sensitive in some cases than in others. This does *not* mean that the choice of b and a is unimportant. Rather, it illustrates that even if the methodological choice varies, it will not blur the country comparison as long as it is consistent across countries. In the following subsection we show that the pattern is considerably distorted if there is a lack of consistency across countries.

Missing and Zero Incomes

For computing the conventional Gini we follow standard practice in LAC and exclude missing and zero incomes. Usually the justification for doing so is that these observations are unreliable and introduce more noise than signal into the measurement of inequality. As noted by Juster and Smith (1998) and Pyatt (1999), missing values tend to be correlated with capital incomes, and are

---

[35] Individual incomes can be computed by using a two-parameter formula (Lanjow and Ravallion,1996), but here we have varied b and a one at a time to verify if the Gini is more sensitive to one of the two parameters.
[36] See Gottschalk and Smeeding (1998).
[37] There are few LAC-specific equivalence scales. We also performed the calculation by using the scale in Contreras (1996), which was calculated for Chile. The change in the Gini was smaller than the result with the Amsterdam scale, so we do not include it for brevity. Due to the lack of country specific scales, we

generally due to the unwillingness of the rich to disclose their wealth. Zero incomes are usually regarded as not plausible, but it is not that clear whether they are associated with specific income sources, income levels, and socioeconomic characteristics, or if they represent a precise and reliable answer to survey questions at all. It should be noted that zero incomes do not refer to individuals not earning income, but to individuals who declare being engaged in a productive activity but not receiving remuneration for it.

Excluding zero and missing values from the computation of the Gini is, however, only one among several options. In fact, Little (1988) argues that discarding missing values can entail important information losses, and the remaining observations may no longer be representative of the whole population under analysis. To reduce non-response bias it is possible to apply weighting adjustments or to use information in the household surveys to impute an estimated income for missing observations as is commonly done in developed countries (Gottschalk and Smeeding (1998). We may also choose to believe that individuals reporting zero income are providing reliable information, or alternatively, we can impute incomes for zero income respondents, rather than discarding them.

The third column in Table 8 presents the proportion of observations with missing and/or zero values in the 18 countries in our sample. In about half of the cases the number is negligible, but in others such as Nicaragua, Venezuela, and Argentina the proportion is above 16% of the sample. The last columns of the table explore the effect of either imputing missing and zero values or imputing missing values and taking zero incomes at their face value. Since there are few total non-responses in the surveys, we mostly focus on item non-responses, which are imputed by following a two step procedure. The first step consists of estimating income regressions for each source separately, where each individual's income is the dependent variable, and the independent variables are all personal and household characteristics in the surveys. The coefficients are used to predict each income-earner's income for that source, including non-respondents. The second step consists of ranking all income earners of the source in question (including non-respondents) according to their predicted income. For non-responses, the error term is calculated by averaging the residual of the observation immediately above and below, and the total imputed income corresponds to the predicted plus the estimated residual. Household incomes are added up after

---

have applied the same criteria for each of the surveys. It should be noted, however, that estimations of equivalence scales can vary substantially from country to country (see Lancaster, *et al.,* 1999).

the procedure. The same method is applied when imputations are performed on individuals reporting zero incomes.[38]

The ranking reversals in the last column of Table 8 are quite important. Nicaragua appears to be the most unequal country in LAC instead of Paraguay when, rather than discarding missing and zero values, we choose to impute the missing and "believe" the zero incomes. The relative position of Venezuela also deteriorates in an important way, while the Gini for Argentina increases by more than two points.

Table 8 treats all household surveys in the same way and shows some important reversals, but, as with the case of the choice of *b* and *a,* the effect could be magnified if a consistent methodology is not applied in each household survey.

3.2 The Importance of Methodological Consistency

Since secondary data sets based on compilations of Gini coefficients from published sources mix a range of methodological choices, at least in part, the differences in inequality in international comparisons–especially across regions with different traditions in the use of *b(i)* and *a*–may be artificial variation caused by different parameter choices rather than genuine distributive differences. Since the choices are usually not documented, it is difficult to measure the magnitude of the biases. The same issues arises with the treatment of missing and zero values. For instance, if we discard missing and zero values in Venezuela but impute missing and use zero incomes in Argentina, rather than treating the two countries in the same way, our impression about the relative inequality in these two cases would change, and they would be regarded as having practically the same Gini. Another example is that the DS data includes an observation for Nicaragua 1993. If the authors of the study from which the observation was drawn had chosen to "trust" the zero incomes reported in the survey, Nicaragua would appear to be the most unequal country in the world, and would have the highest Gini coefficient in the whole DS data. Our impression about the record of this country, and most probably other countries in terms of income distribution, is thus being blurred by the subtleties in the treatment of the data.

Since we have access to the original micro observations of each of the LAC household surveys used so far, it is possible to apply the same methods across countries, and to document what they are. Take for instance the ranking of the conventional Gini in Table 1. The three

---

[38] We tried several combinations of dropping and imputing missing and zero values one at the time. The results we present are the options under which larger variation was identified. Brick and Kalton (1996), Kalton and Kasprzyk (1986), and Rubin (1987) review other methods for dealing with missing values.

implicit choices in the calculation are *b=1, a=1,*and that zero and missing incomes are unreliable and should be thrown out of the sample. There are several ways of handling missing and zero values, but if they were restricted to two choices: 1) throw these observations out, and 2) impute missing incomes and take zero incomes at face value, then the Gini could be represented as a choice of three parameters, G(*b, a, m*), where *m* defines the treatment of missing and zero incomes, and takes a value of 1 or 2, depending on the choice. The conventional Gini is the case where G(*1,1,1*).

One way to assess the importance of the lack of methodological consistency that plagues secondary databases and compilations from published sources is through the following experiment. We take the 18 LAC countries used so far, and rather than applying a consistent methodology across countries, we assign a methodology randomly to compute the Gini. Specifically, we define four possible methods: (i) G(*1,1,1*), which is the conventional Gini; (ii) G(*.5,1,1*), (iii) G(*1,Amsterdam scale,1*),; and (iv) G(*1,1,2*). Thus, in each case we change one of the parameters and hold the others as in the computation of the conventional Gini. The process for assigning a method has two steps. First, we order countries randomly. Second, we choose the first country, and assign one of the four methods randomly to compute the Gini. Then we take the second country and replicate the process, and so on. After doing this for all 18 cases, we rank the countries according to the Gini computed with a random methodology, and compare them with the original ranking. Since secondary databases and compilations of Gini coefficients from published sources have a mixture of methodological choices, this process of random selection may in fact be closer to the variety found in those sources, than to estimates such as those presented in Table 1, with the same method applied consistently.

Table 9 presents the results from six different draws from the above experiment. The first column ranks countries according to the conventional Gini, so for every case the methodology is G(*1,1,1*). The second column introduces "noise" into the comparison by estimating the Gini of each country with a randomly assigned method (indicated next to the country code). The number of re-rankings is quite surprising, especially considering that we are only varying one parameter at a time. The relative positions of Panama, Ecuador, Guatemala, Colombia, Chile, Nicaragua, Venezuela and Mexico change significantly, and interestingly, there are only three cases where the methodology corresponds to the fourth option G(*1,1,2*), where differential treatment of missing and zero values is allowed for, and which introduces more cross-country variability. The third column presents a second draw, where there are also considerable re-rankings. The last four columns present other four draws, and in every case the relative position of many countries changes. In the last two draws, one and two cases, respectively, correspond to the fourth option,

and still we observe several re-rankings. The most notable are for the Dominican Republic, Peru, Mexico, Brazil, Nicaragua and Ecuador. The only country that is not re-ranked in these six drawsis Uruguay, which remains the most equal throughout.

Since we introduce "noise" randomly, each time we perform the experiment we obtain different results, so we tested with 50 draws and we always observed at least four re-rankings. Although the random draws will not match exactly the variability introduced in secondary data sets by the lack of methodological consistency, we think that the exercise is still illustrative about the potential "noise" introduced in these databases or compilations of Gini coefficients from published sources. The lack of documentation of the methodological preference for each observation magnifies the problem and increases the probability that country rankings are to an important extent an illusion caused by differential treatment of the data.

To verify whether this is also an important issue for the growth-inequality relationship, we perform the same experiment of random allocation of each of the four methodologies described before, but now for our complete set of 55 household surveys used in the growth regressions in Table 7. So, rather than estimating equation (3) in the Table with the conventional Gini, we estimate the same regression but the dependent variable is chosen by assigning randomly one of four different Gini coefficients (corresponding to the four methods) to every country-year observation. We experiment 50 times, and present in Table 10 the results from five different draws that illustrate the range of results obtained.

The first column corresponds to a case where the results are very similar to those presented in Table 7 with the conventional Gini as dependent variable. Thus, even after introducing "noise" in the data by using different methods, it is possible to obtain similar results as before, although the following results suggest this is only by chance. The second column presents another estimation, where the Gini has been "contaminated" by methodological inconsistency. It is interesting because, even though the conclusion about the relation between growth and inequality is the same as in the first column, the magnitude of the coefficient for the Gini index is about double the size. The third regression, corresponding to another random allocation of methodology, stands out because the coefficient for the inequality variable in the equation is statistically insignificant, and the size of the coefficient is negligible compared to the previous two. The fourth draw is the regression where we obtain the highest 't' statistic for the Gini in our 50 trials. But perhaps the most striking is the fifth draw, where inequality and growth are actually inversely related, although the coefficient is not statistically significant.[39]

---

[39] This is one of the 4 cases in our 50 draws where the coefficient for the Gini is negative.

Although our results are subject to a number of caveats, they confirm that the "noise" to signal ratio in data sets where methodological consistency is not guaranteed can be very high. Our experiment suggests that the "noise" has the potential for completely blurring the picture about how unequal countries are with respect to others, and about the relationship between inequality and other economic indicators such as the rate of growth. In secondary data sets we simply do not know how loud the "noise" is, but it definitely can be very misleading, even if the differences are restricted to only three parameters of choice.

3.3 Choice of Inequality Measure

The Gini is the most common inequality indicator because it has several advantages such as its intuitive interpretation. But it is still only one among many equally valid options for summarizing the information on the distribution of income and, as with any other inequality measure, its calculation is based on a set of assumptions. One of the main characteristics of this index is that it gives greater relative weight to the middle portion of the distribution.

One way of illustrating the range of options for measuring inequality is to use the family of entropy measures described by Cowell and Jenkins (1995). These indexes are expressed as a function of a parameter that changes in value according to the sector of the distribution to which more weight is attached. High values of the parameter yield "top sensitive" measures, while low values yield "bottom sensitive" indexes. Two countries with similar Gini coefficients can have very different top or bottom sensitive indexes if there are disparities between them at the extremes of the distribution.

Table 11 presents the ranking of countries according to two alternative indexes. The first refers to a bottom sensitive index with parameter value of approximately –1, while the second is a top sensitive index with parameter value of about 1. In the case of the first index the relative position of several countries shifts, with the greatest changes being observed in Venezuela, Ecuador, Guatemala, and Brazil. The top sensitive index also yields important ranking reversals, especially for Ecuador. Still another option is to compare differences along the whole distribution, instead of focusing on some specific section. A standard Lorenz stochastic dominance test in Figure 3 specifies which distributions unambiguously dominate the others, when comparing each centile. According to the figure, the distribution for Paraguay dominates (is more unequal than) the other 17 countries. Brazil, Bolivia and Panama dominate the rest of the countries, but their Lorenz curves intersect, so it is not possible to rank them unambiguously. A

group including Colombia, Nicaragua, Chile and Ecuador, who do not dominate each other, follows. One unambiguous conclusion is that Uruguay and Costa Rica are the countries with the lowest conventional Gini, but the most relevant aspect of the Table 11 and Figure 3 is that our impression about how unequal each country is with respect to the others might change when we consider the whole distribution, or alternative summary measures of inequality.

To explore if there are also implications for the relation between inequality and growth, we replicate regression 3 in Table 7, but by substituting the bottom and top sensitive indexes for the conventional Gini. In this case we also use our wider LAC sample of 55 surveys with national coverage. Table 10 presents the results, which are subject to the same caveats as the previous regressions in Table 7.

Surprisingly, when the estimation is performed by using a bottom sensitive entropy index the conclusion about the relation between inequality and growth changes quite drastically, and is actually reversed. Inequality now appears to have a negative effect on growth, with significance at the 10% level. However, if a top-sensitive entropy index is used instead, the conclusion is that inequality and growth are positively related.[40] So, at least in the case of LAC, equally acceptable summary measures of income distribution can lead to completely different conclusions about the effect of inequality on other development indicators. Conclusions derived with the use of international secondary data sets could be also specific to the use of the Gini or quintile shares, which are more readily available, but they might not be applicable for different ways of summarizing information on the distribution.

4. Who, and Which Incomes, are Behind LAC Inequality?

So far we have taken a statistical approach to the question of what's behind the inequality we measure. In the rest of this paper we take a more economic-oriented perspective. We explore what types of information household surveys convey and what they are really informative about. In this section we approach the issue by taking a close look at some specific sections of the distribution and at the importance of different income sources.

---

[40] In Forbes (2000) robustness tests are performed to verify if the conclusions of the paper hold when using the ratio of the top and bottom quintiles to summarize the distribution, and the author's results are confirmed. The difference between using these types of indexes and the entropy measures is that entropy indexes use information on the whole distribution, while the quintile shares ignore information on the middle sections.

4.1 Where is Income Most Concentrated?

It is well established that LAC is the most unequal region in the world, but not much has been said about the type of concentration that characterizes the distribution of these countries. One way to illustrate this is to compare income differences across deciles. Figure 4 presents these comparisons by showing the ratio of the average income of each decile and the preceding one. For instance, the second point in the horizontal axis compares the income of the second and first deciles. In Bolivia the difference is about 3.2 times and in Panama it reaches 2.7, while in Guatemala it is only 1.6 times. The figure also includes data for the United States, the United Kingdom and Canada (accessed through LIS) for reference.

Interestingly, the differences at the lower tail of the distribution are larger in the United States than in several highly unequal LAC countries, such as Brazil and Chile. The United Kingdom, where the difference is about 2 times, also registers greater disparities than about half of the LAC countries. Thus, the reason why LAC has so much more inequality is not due to income differences among the poorest deciles. The figure also plots the difference between the third and second decile, the fourth and third, the fifth and fourth, and so on. In the middle part of the distribution it is not possible to distinguish any LAC country from the USA, the UK, or Canada, which indicates that the higher LAC inequality is not due to differences among the middle deciles either.

The most important feature of the figure is that when the 10th and the 9th decile are compared, all LAC countries stand out for their large disparities. In Paraguay, the country with the highest conventional Gini, the differences at the upper tail reach 3.5 times, while in Guatemala, Chile, Ecuador, Brazil and Bolivia, the richest 10% of the population obtain incomes around 3 times greater than the next 10%. The USA, the UK, and Canada are all at the bottom of the figure with much smaller differences, not exceeding 1.6. The figure confirms the well-known fact that the USA has higher inequality than other developed countries because of the differences at the lower end of the distribution.

This evidence, added to the fact that the correlation between the conventional Gini and the $10^{th}/9^{th}$ decile income ratio is .9, suggests that high LAC inequality is mainly a reflection of the disproportionate incomes of the richest 10%. Actually, when comparing other points of the distribution, inequality is even lower than that registered in much more equal developed countries. The evidence in Section 2 suggests that the incomes at the top of the distribution in

LAC are grossly underestimated, but figure 4 shows that even so, the income of the richest individuals that the surveys are able to capture is much higher than the remaining 90%.

To obtain a more precise idea about the importance of the concentration at the top of the distribution in LAC, we estimate the Gini index by truncating the distribution at different points. For instance, in the first column of Table 12 we exclude the richest 10 households in the survey from the calculation. By excluding these observations, the average for the region drops by 1.3 points, but there are countries such as Paraguay where the Gini is much more sensitive to these few observations. Rather than being the most unequal country in LAC, Paraguay ranks as the third most unequal by excluding the ten richest households in each survey, due to a drop of almost 5 points in the Gini for this country. In contrast, the difference between the conventional and the truncated Gini is negligible in the USA (at the bottom of the table).[41]

Table 12 also includes the truncated Gini excluding the richest 1%, 2%, 5% and 10%, respectively. If the richest 1% is excluded, the average for the region drops by 5 points. Guatemala and Paraguay are the most sensitive to this truncation. The exclusion of the richest 2% results in an average decline of two additional points, while truncating the distribution in the 95th percentile implies a drop of around 11 points with respect to the conventional Gini. This reduction is significant, but still, most LAC countries continue to register higher inequality than the truncated Gini for the USA.

When the richest 10% are excluded from the calculation, the difference between the average LAC country and the USA is reduced by 3.45 points. But the most important result is that now 8 out of the 18 LAC countries present either lower inequality or a difference of less than 1.5 points with respect to the USA. This confirms that the differences between the typical LAC country and the USA is mainly a reflection of a higher concentration at the top 10% of the distribution in LAC. One of the clearest examples is Mexico, where the difference with the USA is of 8.72 points when the conventional Gini is compared, but when the richest 10% are excluded from the calculation in Mexico and the USA, Mexico appears to be slightly more equal.

According to Gottschalk and Smeeding (1997, 1998), the highest incomes in household surveys are usually measured with lower precision.[42] Interestingly, according to our results the

---

[41] As noted by Cowell and Victoria-Feser (1996), extreme observations such as the one outlier observed in Paraguay are normally regarded as data "contamination". Truncating the sample for instance, by excluding the top observations is a common procedure in developed countries to avoid the inclusion of implausibly high observations. So, the estimates presented in the table could also be interpreted as a "cleaner" ordering that does not depend on these "more unreliable" observations.

[42] In fact, in the research for developed countries it is common to top and bottom code incomes because extreme observations are regarded as less reliable, but in LAC this is usually not done. There are some exceptions such as Colombia, where the survey instrument for many years had top coding for incomes, but apart from this case, we are not able to identify any others. In the 111 studies on income distribution and

ranking of LAC countries, as well as the difference between each country and the USA, the UK, and Canada are guided mostly by the incomes at the top of the distribution. The grounds on which conclusions are made about the relative position of each country and the differences with other regions, therefore, do not seem that solid. Furthermore, the results suggest that the inclusion of observations for LAC in cross-country comparisons or in regressions with international secondary data sets will introduce cross-country variation generated from observations with higher measurement error. This might influence our impression about how LAC ranks in the world, and about the effect of inequality on other development indicators.

A Profile of the Rich

Figure 4 raises the question of who are the individuals in the top 10% "generating" most of the measurable inequality in LAC. To address this issue we obtain a profile of the richest 10% of the population in each country by estimating a probit regression. The dependent variable is a dummy indicating if an individual is located in the top decile or not, and the independent variables are dummies for all personal characteristics included in each survey (the characteristics of the household head are used as reference).

Figure 5 summarizes the results by presenting the average value of the coefficient from the country-specific regressions.[43] We present only the coefficients for the variables that are statistically significant in the majority of cases. According to the results, living in a household where the head has secondary or higher education, or reports being an employer or being occupied as a professional or technician, are associated with the highest probabilities of belonging to the richest 10%. Having primary education, being employed as director or manager, working in the financial sector and living in urban areas are also associated with a positive significant probability of belonging to the richest decile. In contrast, living in a household where the head is a rural or manufacturer worker, is an employee or lives in households with children under 15 years of age, is associated with a negative probability of belonging to the top 10%. In 5 cases, there is some information on ethnicity, and not surprisingly, belonging to an indigenous group reduces the probability of belonging to the 10th decile, by 12%.

---

poverty in LAC reviewed by Londoño and Székely (2000), very few references to this issue were found, and it seems that the vast majority of the studies there is no attempt to top code. The conventional Gini for the USA used here actually refers to a distribution truncated at the top 1%.

[43] The country-specific results are not presented for brevity, but are available from the authors. All variables were defined in the same way across countries to perform the regression. The results should only be regarded as representing associations between the dependent and independent variables, since there is potential endogeneity in the estimation.

Table 13 breaks down the income of the individuals in the top 10% into four sources: labor, self-employment, non-labor income, and non-monetary income. Income from self-employment is normally classified by respondents as labor income, but usually it contains some return to capital. It is difficult to identify the portion accruing to profits and the portion that is strictly labor payments, but this clearly underestimates the value of profits. As shown in the table, more than 80% of the incomes of the households in the richest 10% are reported to be labor income in the average LAC country, but one of the reasons for this high value may be the underestimation of profits from self-employment. To verify if this is the case, we divide the incomes of the self-employed into labor and profits by estimating an earnings regression for all non-self-employed employees reporting labor income, where the dependent variable is labor earnings, and the independent variables are dummies for personal characteristics. We use the coefficients to predict the income of the self-employed. The predicted is classified as labor income under the idea that these are the earnings that the individual would obtain in the labor market, and the residual is classified as profits. But even when performing this imputation, around 60% of the incomes of the richest 10% are on average classified as labor. In Honduras, Colombia, Venezuela and Costa Rica the proportion reaches around 70%, while there are only three cases (Paraguay, Guatemala and the Dominican Republic) where labor income represents less than 50% of the total.

These results, added to Figure 5, confirm that the profile of the average individual in the top 10% of the distribution, where incomes are most concentrated, is closer to the prototype of highly educated professionals obtaining labor incomes, rather than capital owners living on profits. This does not imply that in LAC inequality is not generated by a small number of individuals at the top of the distribution typically earning profits from capital investments. What it reveals is that the inequality that we are able to measure with household surveys in LAC is informative about a spectrum of society that does not include the richest households.

4.2 Which Types of Incomes are Behind the Concentration?

It is also of interest to determine which incomes distribution indicators are informative about. So far we have provided evidence that the incomes of the rich may be grossly underestimated, and one of the reasons might be the failure of the surveys to capture profits and capital incomes

accurately. To assess the proportion of inequality accounted for by the different types of incomes, we apply the decomposition by factor components suggested by Shorrocks (1982).[44]

Table 14 presents the results. If self-employment incomes are classified as reported in the surveys (that is, mostly as labor), on average, 80% of the overall inequality would be associated with labor incomes. This high share is not surprising, and it is not exclusive for the LAC region. For example, Gottschalk and Smeeding (1997) have argued that household surveys also capture mainly labor incomes in developed countries.

Due to the possible overestimation of the role of labor in some activities, in this case we also break down incomes from self-employment into labor and profits through imputation. According to the table, on average 57.4% of the total inequality is accounted for by labor earnings by employees, while 6.6% is from labor incomes from the self-employed. So, even when we assign most of the income from self-employment to profits, almost two thirds of total inequality is accounted for by labor incomes. The most extreme cases are, on the one hand, Paraguay, Panama, Nicaragua, Ecuador, Venezuela, and Costa Rica, where labor incomes account for more than 73% of total inequality. On the other hand, Peru is the only country where labor incomes account for less than 50%, but this is a clear exception to the rule.

Thus, most of the inequality that we are able to measure in LAC is from labor incomes. This does not mean that in reality profits or capital incomes do not play a predominant role in the generation of inequality, but rather that the household surveys have a limited capacity to incorporate that type of information.

5. Conclusions

This paper takes a step back from the use of income distribution indicators estimated from household surveys to explore what these numbers really are. Inequality indexes are being increasingly used to evaluate economic performance and as explanatory variables for other development indicators, but the quality of the data, the assumptions behind the calculations, and the capacity to cover all population groups in society and all income sources, are far from homogeneous. As a consequence, an important part of the cross-country differences observed are

---

[44] See Székely (1998) for a more detailed explanation, and Lerman (1999) for a review of alternative methods. The advantage of following Shorrocks' procedure is that the results are independent from the choice of inequality index.

simply due to disparities in the way in which the data is originated and treated, rather than genuine differences in distribution.

We divide our exploration into two broad parts. First we take a statistical approach to investigate the differences in the characteristics and the treatment of the data. Second, we focus on a more economic-oriented perspective by examining which sectors of the population and which types of income distribution measures are more informative about.

The analysis is based on recent household surveys for 18 LAC countries, most of which are typically included in secondary data sets. In terms of the characteristics and comparability of the data, the analysis covers survey frames, sample size, quality and coverage of population subgroups at the extremes of the distribution, as well as differences in income sources and timing of the survey. The results suggest that LAC household surveys severely underestimate inequality because they do not include information on the richest sectors of society, and that a considerable part of the cross-country disparities are driven by differences in the way the data is collected, rather than genuine distributive differences. Furthermore, once data differences are accounted for, at least in a very general way, our understanding of the relation between inequality and growth changes from a strong positive connection to no statistical relation at all.

The tradition in LAC is to focus on the distribution of household per capita income by discarding zero and missing values and using the Gini coefficient as a summary measure of inequality. But by adopting other equally valid assumptions about economies of scale in consumption, adult equivalence scales, and the treatment of missing and zero incomes consistently in the same way in all countries, we observe that some distributions are more sensitive to these choices. But in reality, secondary data and other compilations of Gini coefficients from published sources are far from having methodological consistency. Rather, they mix methodological choices by different authors that are usually not even documented. We perform an experiment to assess at least in a general way the effect of lack of methodological consistency. We find that inconsistency has the potential for completely blurring our impression about how unequal countries are, and about the relation between growth and inequality. The choice of inequality measure can also change country rankings and the view about the effect of inequality on economic growth.

A close examination of each distribution reveals that although surveys fail to cover the richest individuals, the Latin American inequality we observe is still mainly reflecting a disproportionate concentration of income among the top 10% of the population covered by the surveys. The profile of individuals in the richest 10% in the surveys corresponds to highly educated professionals whose main income source is labor income earned as employees. A

decomposition of inequality by sources reveals that, similarly to other countries, the LAC inequality is mainly informative about the distribution of labor incomes.

In sum, our main conclusion is that although household surveys are the best source of data for the measurement of inequality, it is crucial to remember when using these numbers that there is an important story behind each number. This story influences our judgement about how unequal countries are and about the relation between inequality and other development indicators, but it is seldom told or known. Perhaps other statistics commonly used in economics also have their own interesting story, and it might be worth trying to find out what it is.[45]

---

[45] A recent example in this line is the paper by Ward and Dikhanov (1999), who present some interesting insights into the use of inflation measures.

References

Aghion, P., Caroli, E. and García Penaloza, C., 1999. "Inequality and Economic Growth: The Perspective of New Growth Theories". *Journal of Economic Literature*. 37 (4): 1615-1660.

Alesina, A., and Perotti, R. 1996. "Income Distribution, Political Instability and Investment". *European Economic Review*. 40 (6): 1203-1228.

América Economía. 1999. "Informe especial: salarios de los ejecutivos en América Latina." *América Economía*. 160: 65-78.

Atkinson, A. 1997. "Bringing Income Distribution in from the Cold". *The Economic Journal*, 107 (441): 297-321.

-----. 1998. "Equity Issues in a Globalizing World: the Experience of OECD Countries". Working Paper prepared for Conference on Economic Policy and Equity. Washington, DC, United States: International Monetary Fund.

Atkinson, A. and Brandolini, A. 1999. "Promise and Pitfalls in the Use of 'Secondary' Data-Sets: Income Inequality in OECD Countries". Mimeographed document.

Atkinson, A. and Micklewright, J. 1992. *Economic Transformation in Eastern Europe and the Distribution of Income*. New York, United States: Cambridge University Press.

Atkinson, A., Rainwater L, and Smeeding, T. 1999. "Income Distribution in European Countries". DAE Working Paper 9506 (Amalgamated Series No. 9535). Cambridge, United Kingdom: University of Cambridge, Department of Applied Economics.

----. 1995. "Income Distribution in OECD Countries". Organization for Economic Co-Operation and Development. Social Policy Studies No. 18. Paris, France: Organization for Economic Co-Operation and Development.

Bénabou, R. 1996. "Inequality and Growth". In: B.S. Bernanke and J.J. Rotemberg, editors. *NBER Macroeconimics Annual*. Volume 9. Cambridge, United States: MIT Press.

Brick, J. M., and Kalton, G. 1996. "Handling missing data in survey research". *Statistical Methods in Medical Research*. 5 (2): 215-238.

Contreras, D. 1996. "Household and Individual Welfare: Evidence from less Developed Economies". Working Paper. Los Angeles, United States: University of California. Mimeographed document.

Cowell, F. and Jenkins, S. 1995. "How Much Inequality Can Be Explained: A Methodology and An Application to the USA". *Economic Journal*. 102 (2): 421-430.

Cowell, F., and Victoria-Feser, M.P. 1996. "Robustness Property of Inequality Measures". *Econometrica*. 64 (1): 77-101.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, United States: World Bank/Johns Hopkins University Press.

Deaton, A., and Muellbauer, J. 1980. *Economics and Consumer Behavior.* New York, United States: Cambridge University Press.

Deininger, K. and Squire, L. 1996. "Measuring Income Inequality: A New Data Base." *World Bank Economic Review.* 10 (3): 565-91.

Ferreira, F., Lanjow, P., and Neri, M. 1999, "The Urban Poor in Brazil in 1996: A New Poverty Profile Using PPV, PNAD, and Census Data". Rio de Janeiro, Brasil: PUCI. Mimeographed document.

Forbes, K. 1999. "A Reassessment of the Relationship between Inequality and Growth". *American Economic Review.* Forthcoming 2000.

Ganuza, E., Taylor, L., and S. Morley, S. 1998. *Política Macroeconómica y Pobreza en América Latina y el Caribe.* Madrid, Spain: Mundi Prensa Libros.

Gottschalk, P. and Smeeding, T. 1997. "Cross-National Comparisons of Earnings and Income Inequality". *Journal of Economic Literature.* 35 (2): 633-687.

----. 1998. "Empirical Evidence on Income Inequality in Industrialized Countries". Working Paper. Mimeographed document.

Juster, T. and Smith, J. 1998. "Enhancing the Quality of Data on Income and Wealth: Recent Development in Survey Methodology". Paper presented at the 25th General Conference of The International Association for Research in Income and Wealth, Cambridge, United Kingdom.

Kalton G. and Kasprzyk, D. 1986. "The Treatment of Missing Survey Data". *Survey Methodology.* 12 (1): 1-16.

Lancaster, G., Ray, R., and Valenzuela, M.R. 1999 "A Cross-Country Study of Equivalence Scales and Expenditure Inequality on Unit Record Household Budget Data". *Review of Income and Wealth.* 45 (4): 435-454.

Lanjouw, P. and Ravallion, M. 1995. "Poverty and Household Size". *Economic Journal.* 105 (433): 1415-1434.

Lerman, R. 1999. "How do Income Sources Affect Income Inequality?" In: J. Silber, editor. *Handbook of Income Inequality Measurement.* Kluwer Academic Publishers.

Little, R. 1988. "Missing-Data Adjustments in Large Surveys". *Journal of Business & Economic Statistics.* 6 (3).

Londoño, J.L., and Székely, M. "Persistent Poverty and Excess Inequality: Latin America 1970-1995". *Journal of Applied Economics.* Forthcoming 2000.

Medina, F. 1999. "Los errores de muestreo en las encuestas complejas: usos y abusos de la información". Paper presented at the meeting of International Association of Statistics and Surveys and the International Association of Official Statistics.

Perotti, R. 1996. "Growth, Income Distribution and Democracy: What the Data Say". *Journal of Economic Growth*. 2 (1): 149-187.

Plotnick, R., Smolensky, E., Evenhouse, E., *et al.* 1998. "The Twentieth Century Record of Inequality and Poverty in the United States". Paper presented at the 25th General Conference of the International Association for Research on Income and Wealth, Cambridge, United Kingdom.

Pyatt, G. 1999. "The Distribution of Living Standards Within Countries: An Historical Perspective on a New International Data Base". The Hague, Netherlands: Institute of Social Studies. Mimeographed document.

Ravallion, M. 1998. "Does Aggregation Hide the Harmful Effects of Inequality on Growth?" *Economics Letters*. 61 (1): 73-77.

Ravallion, M., and Chen, S. 1997. "What Can New Survey Data Tell Us about Recent Changes in Distribution and Poverty?" *World Bank Economic Review*. 11 (2): 357-82.

----. 1999. "When Economic Reform is Faster than Statistical Reform: Measuring and Explaining Income Inequality in Rural China". *Oxford Bulletin of Economics and Statistics*. 61 (1): 33-56.

Rubin, D. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York, United States: John Wiley and Son.

Shorrocks, A. 1982. "Inequality Decomposition by Factor Components". *Econometrica*. 50 (1): 193-211.

Spilimbergo, A., Londoño, J.L. and M. Székely, M. 1999. "Income Distribution, Factor Endowments and Trade Openness". *Journal of Development Economics*. 59 (1): 77-101.

Smeeding, T. and Coder. J. 1995, "Income Inequality in Rich Countries During the 1980s". *Journal of Income Distribution*. 5 (1): 13-29.

Székely, M. 1998. *The Economics of Poverty, Inequality and Wealth Accumulation in Mexico.* London, United Kingdom: MacMillan.

UNU/WIDER-UNDP. 1999. "World Income Inequality Database". Beta 3, November 8, 1999.

Ward, M., and Dikhanov, Y. 1999. "What is Inflation?" Washington, DC, United States: World Bank. Mimeographed document.

Figure 1

Standard errors for the Gini index
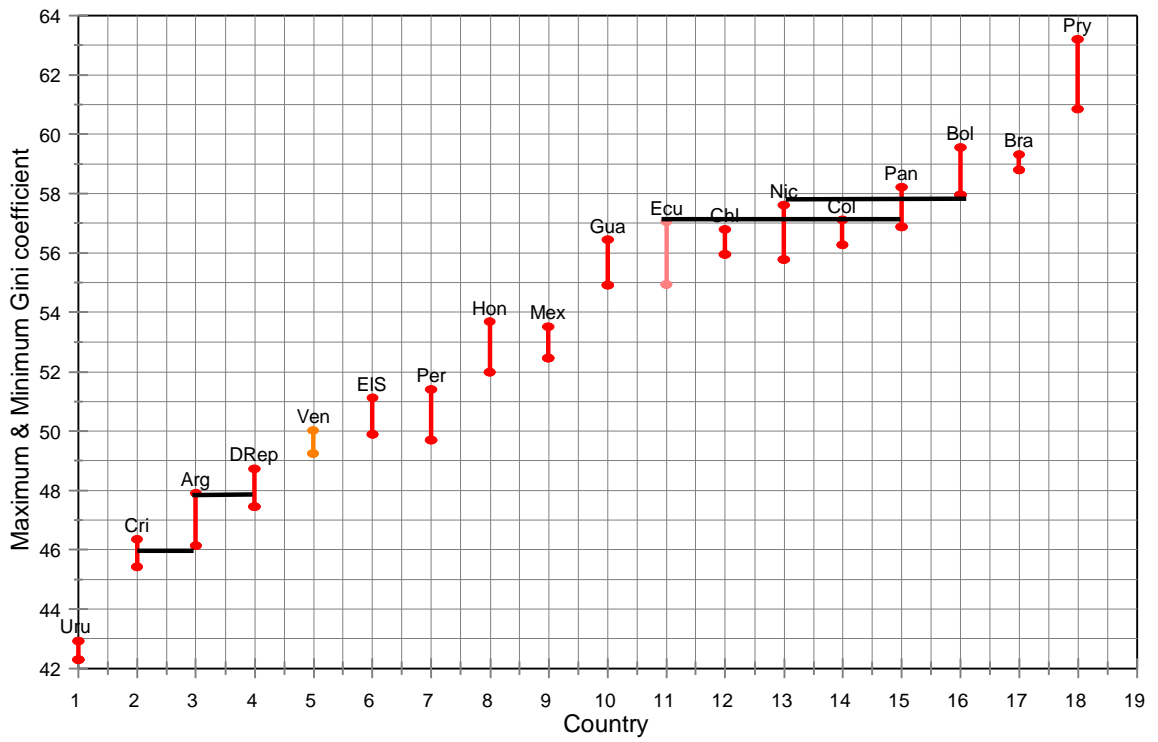
Figure 2
Figure 2a: Paraguay
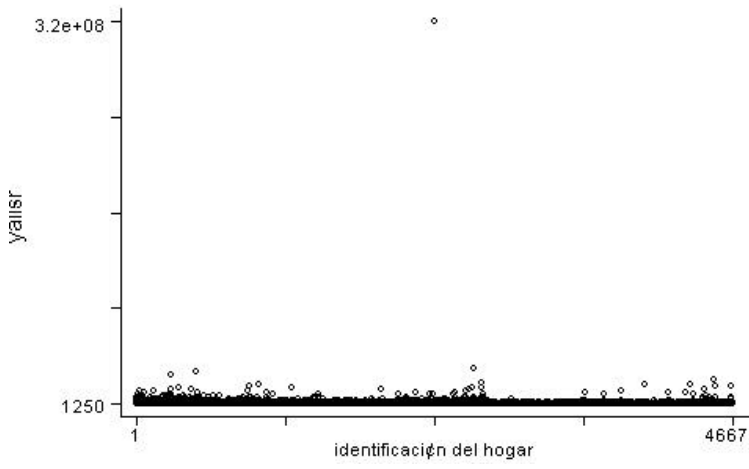
F igure 2b: Ecuador
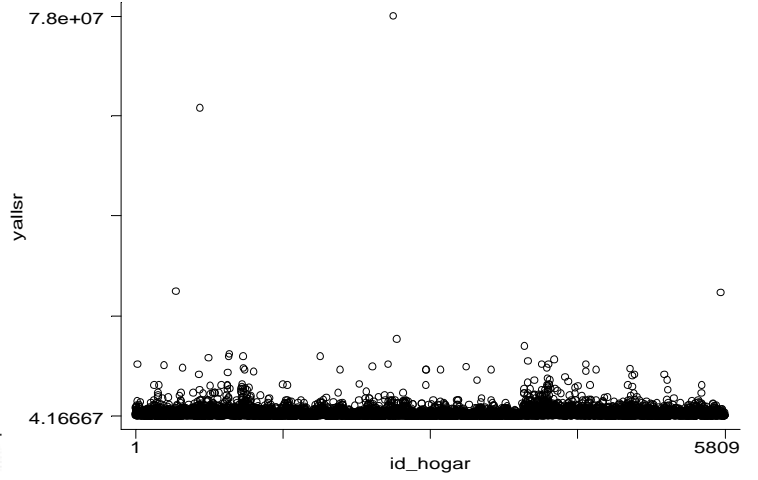
Figure 2c: Chile
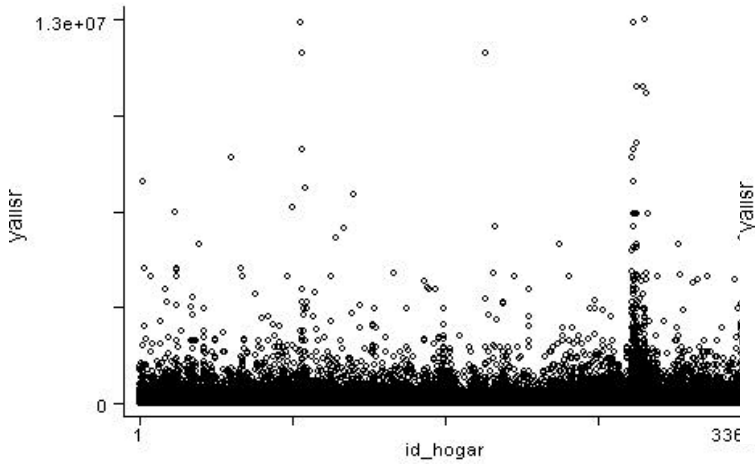
Figure 2d: Brazil
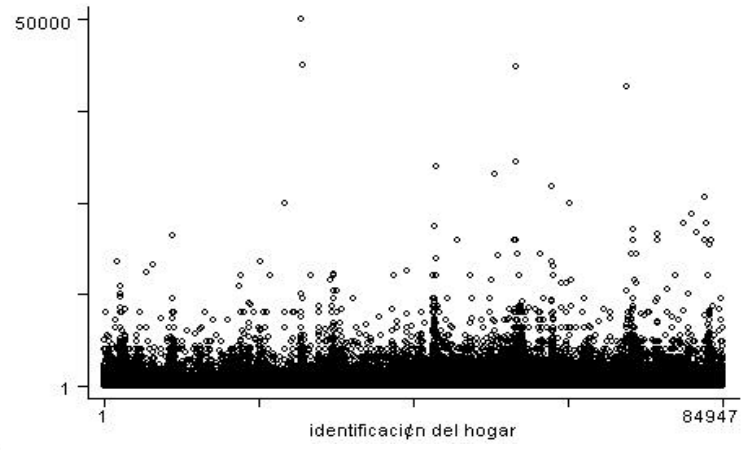
Figure 2e: Venezuela

Figure 2e: Mexico

39
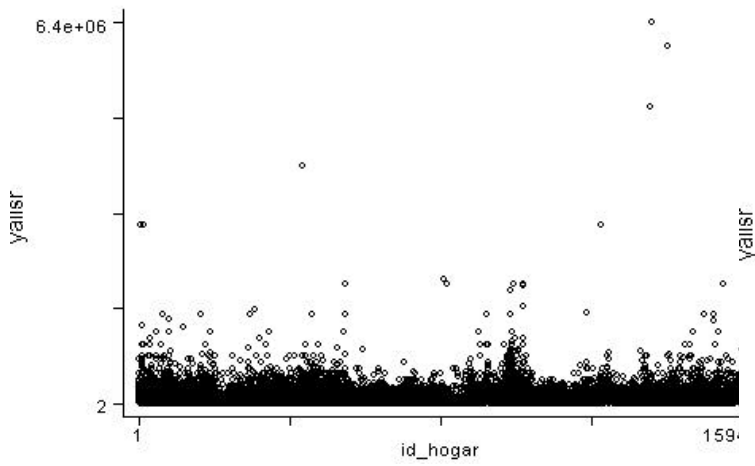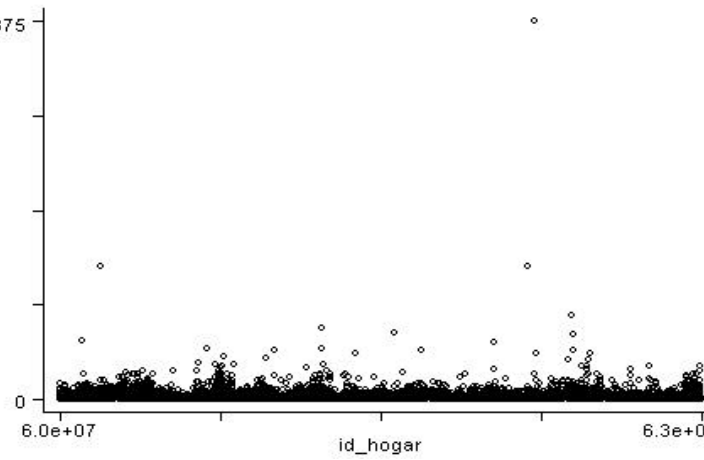
Figure 3

Ranking for Stochastic Dominance Tests

Figure 4

Income Ratio of Successive Deciles
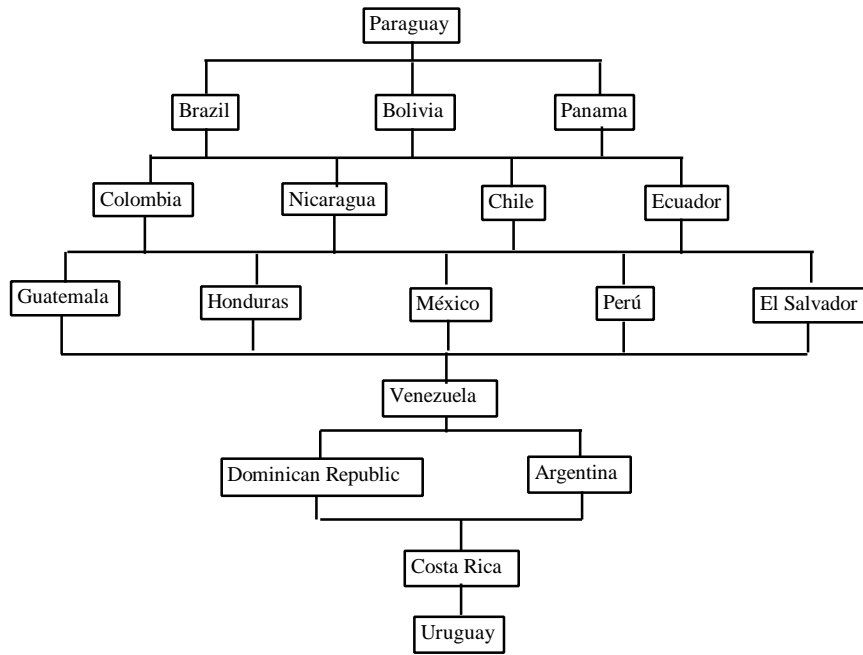
Figure 5

Head's Characteristics and Probability
of Being at the Top 10%

Characteristic	Value
Higher education	42.62
Secondary complete	26.35
Employer	18.13
Secondary incomplete	17.55
Proffessional/technitian (occup)	17.38
Primary complete	12.66
Primary incomplete	8.72
Director, high level manager (occup)	8.42
Financial services sector	7.92
Urban areas	6.38
Rural workers (occup)	-7.88
Manufacturing worker (occup)	-9.27
Working as employee in services (occup)	-9.39
Children under 15 in hh	-10.07
Ethnicity (indigenous)	-11.97

Table 1

Income Distribution in Latin America and Other Regions
Conventional Gini Coefficient

Rank	Country	Year	Gini
Latin America			
1	Paraguay	1995	62.03
2	Brazil	1996	59.06
3	Bolivia	1996	58.77
4	Panama	1997	57.55
5	Colombia	1997	56.70
6	Nicaragua	1993	56.69
7	Chile	1996	56.38
8	Ecuador	1995	56.00
9	Guatemala	1998	55.69
10	Honduras	1996	52.84
11	Mexico	1996	52.76
12	Peru	1997	50.55
13	El Salvador	1995	50.50
14	Venezuela	1997	49.63
15	Dominican Republic	1996	48.10
16	Argentina*	1996	47.02
17	Costa Rica	1997	45.89
18	Uruguay*	1997	42.62
	Average		53.27
OECD Countries			
	United States	1994	44.04
	United Kingdom	1995	39.65
	Australia	1994	39.15
	Italy	1995	36.21
	Canada	1994	35.44
	Germany	1994	34.64
	Norway	1995	29.35
	Finland	1995	29.17
	Sweden	1995	28.84
	Luxembourg	1994	26.94
	Average		34.34
Eastern Europe			
	Russia	1995	45.25
	Polland	1995	37.49
	Hungary	1994	33.20
	Average		38.65
South East Asia			
	Taiwan	1995	30.27

Source: Authors'calculations.
*Urban coverage only.

Table 2
Quality Indicators of Household Surveys

	Country	Self-employment			Comparison with National Accounts		
		Self-employed in informal sector as (%) of all income earners	Gini excluding the self-employed		Labor income in HH survey as (%) of labor income in NA	Gini after adjusting labor income to labor income in NA	
			Rank	Gini		Rank	Gini
1	Paraguay	24.8	1	65.44	0.86	1	62.08
2	Brazil	25.1	4	60.28	1.15	3	59.17
3	Bolivia	23.1	3	61.42	1.47	2	59.52
4	Panama	20.9	7	58.17	0.92	4	57.57
5	Colombia	34.1	8	56.48	0.96	6	56.68
6	Nicaragua	31.6	13	52.09		5	56.69
7	Chile	19.1	5	60.12	1.28	7	56.21
8	Ecuador	30.4	11	56.25	0.79	9	55.57
9	Guatemala	32.6	2	62.36		8	55.69
10	Honduras	41.1	12	54.56	1.05	10	52.84
11	Mexico	24.0	10	56.34	0.83	11	51.80
12	Peru	46.6	6	58.24	1.57	13	49.47
13	El Salvador	25.7	9	56.35	1.08	12	50.35
14	Venezuela	31.6	15	51.58	0.68	14	49.02
15	Dominican Republic	30.7	14	51.78	0.66	16	48.25
16	Argentina	22.0	16	48.58	0.57	15	48.50
17	Costa Rica	20.2	17	47.06	0.75	17	45.95
18	Uruguay	21.6	18	45.57	1.24	18	42.82
	LAC Average	28.1		55.7	1.0		53.2
	Correlation with conventional Gini			0.86			0.98

Source: Authors'calculations.

Notes: GDP figures taken from IDB (1999). Figures for labor income as a share of GDP
are from CEPAL (1998). Exchange rates are from IDB (1998).

Table 3
Incomes at the Top and Bottom of the Distribution

Country	Monthly Incomes from household survey (USD)			Price Waterhouse+		
	10 poorest households in survey	10 richest households in survey	Richest Individual in survey	Monthly income for typical Manager	Richest 10 households in survey vs. typical manager	Richest individual in survey vs. typical manager
Paraguay*	3.0	28,335	162,404	6,114	4.63	26.56
Ecuador*	1.5	14,957	40,500	7,034	2.13	5.76
Uruguay	49.9	17,241	27,930	8,950	1.93	3.12
Chile	1.5	32,359	65,510	16,757	1.93	3.91
Guatemala	17.7	18,436	25,683	9,706	1.90	2.65
Brazil	2.2	37,974	62,871	29,050	1.31	2.16
Panama**	2.3	12,559	17,548	14,455	0.87	1.21
El Salvador	3.1	7,096	11,516	8,086	0.88	1.42
Argentina	12.1	22,232	37,050	32,212	0.69	1.15
Peru	14.1	6,693	10,302	9,718	0.69	1.06
Bolivia*	0.6	5,782	8,189	9,248	0.63	0.89
Mexico	11.8	11,821	39,210	22,570	0.52	1.74
Honduras	3.1	5,170	7,607	8,426	0.61	0.90
Costa Rica	7.7	5,594	7,039	11,619	0.48	0.61
Nicaragua	2.3	6,030	16,852	12,106	0.50	1.39
Venezuela	0.6	8,737	14,878	29,167	0.30	0.51
Colombia	1.2			16,711		
Dominican Republic	5.0	6,717	9,357			
LAC Average	7.76	14,573	33,203	14,819	1.25	3.44

Source: Authors'calculations. Price Waterhouse incomes are taken from America Economia (1999).
*Since incomes for managers from the largest firms were unavailable, managerial incomes from the largest possible firms were used
instead. An adjustment was made by taking the average difference by firm size in the other countries.
**Salaries for general managers were not reported. The salary for financial managers was taken instead, and an adjustment was
based on the income difference between general and finance managers of countries with similar income per capita.

Table 4
Income Sources Covered by Each Survey

Country	Labor Income	Non-Labor Income	Capital Rents	Property Rents	Non-Monetary
Guatemala	X	X	X	X	X
Ecuador	X	X	X	X	X
Mexico	X	X	X	X	X
Uruguay	X	X	X	X	X
Bolivia	X	X	X	X	
Brazil	X	X	X	X	
Dominican Republic	X	X	X	X	
Panama	X	X	X	X	
Paraguay	X	X	X	X	
Chile	X	X	X		X
Peru	X	X	X		X
El Salvador	X	X	X		X
Costa Rica	X	X	X		
Venezuela	X	X	X		
Colombia	X	X	X		
Argentina	X	X	X		
Honduras	X	X			
Nicaragua	X				

Table 5
Inequality with Comparable Sources, Geographic Coverage and Timing

Rank for Conventional Gini	Country	Labor Income		Labor Income Urban Areas		Adjusted Gini		
		Rank	Gini	Rank	Gini	Rank	Gini	Change in Rank
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	Paraguay	1	62.67	1	57.44	2	57.21	-1
2	Brazil	2	59.64	3	57.26	3	56.26	-1
3	Bolivia	5	58.05	10	51.42	7	51.51	-4
4	Panama	7	56.78	8	52.88	9	49.51	-5
5	Colombia	10	56.60	6	53.97	15	46.46	-10
6	Nicaragua	9	56.69	7	53.05	11	48.24	-5
7	Chile	4	58.45	2	57.42	5	54.08	2
8	Ecuador	8	56.72	9	52.56	1	57.69	7
9	Guatemala	3	59.59	5	54.23	14	47.12	-5
10	Honduras	12	53.01	13	49.36	13	47.52	-3
11	Mexico	6	58.00	4	55.79	4	55.00	7
12	Peru	11	53.01	17	46.37	8	50.16	4
13	El Salvador	13	51.16	16	47.02	16	41.93	-3
14	Venezuela	14	49.51	12	49.51	10	48.51	4
15	Dominican Republic	16	48.41	11	50.06	12	48.24	3
16	Argentina	15	48.54	14	48.54	6	51.81	10
17	Costa Rica	18	45.82	18	44.34	17	41.15	0
18	Uruguay	17	47.92	15	47.92	18	34.05	0
	LAC Average		54.48		51.62		49.25	

Source: Authors'calculations.

Table 6

Month of the Year When Survey is Held

Country	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
El Salvador	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
Guatemala	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
Uruguay	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
Venezuela								▓	▓	▓	▓	▓
Nicaragua		▓	▓	▓	▓	▓						
Ecuador								▓	▓	▓	▓	
Mexico								▓	▓	▓	▓	
Paraguay								▓	▓	▓	▓	
Peru									▓	▓	▓	
Argentina				▓	▓	▓						
Bolivia						▓	▓					
Brazil									▓			
Chile											▓	
Colombia									▓			
Costa Rica							▓					
Dominican Republic			▓									
Honduras									▓	▓		
Panama								▓				

Table 7

Estimation of Growth Regressions
(Dependent Variable: Annual Per Capita Growth)

	Regression with world-wide DS Data		Regression with LAC Data**		
Independent Variable	Forbes Regression (1) (5-year periods)	Forbes Regression no grouping by period**	Conventional Gini	Gini Labor Income	Adjusted Gini
	(1)	(2)	(3)	(4)	(5)
Gini	0.0036	0.0018	0.00165	0.00092	0.00117
	(2.4)	(1.94)	(1.9)	(1.07)	(0.30)
Ln GNP per capita	-0.076	-0.0289	-0.02353	-0.02368	-0.01954
	(3.8)	(2.9)	(2.5)	(-2.5)	(2.3)
Male Education	-0.014	-0.0516	-0.08121	-0.07754	-0.08237
	(0.45)	(1.55)	(0.84)	(0.82)	(0.91)
Female Education	0.07	0.0684	0.09829	0.09482	0.10476
	(2.18)	(2.2)	(1.76)	(1.83)	(2.2)
PPPI	-0.0008	-0.0009	-0.00093	-0.00013	-0.00038
	(2.6)	(3.75)	(1.15)	(2.26)	(2.1)
R-squared	0.67	0.59	0.27	0.27	0.24
Countries	45	56	11	11	11
Observations	180	462	42	42	42
Period	1965-95	1965-95	1979-1998	1979-1998	1979-1998

*Source: Forbes (2000), Table 3.
t' Statistics in parenthesis
**Source: Authors'calculations

Table 8

Gini Coefficient with Different Ways of Treating the Data

Rank for Conventional Gini	Country	Gini with economies of scale in consumption	Gini with adult equivalence scales	% obs. with missing and/or zero incomes	Impute missing & zero's	Impute missing & use zero incomes Rank	Impute missing & use zero incomes Gini
1	Paraguay	59.80	59.36	1.9	62.06	2	62.40
2	Brazil	57.24	57.12	4.3	59.44	3	60.05
3	Bolivia	57.32	56.70	8.0	58.87	4	59.87
4	Panama	55.70	55.43	7.1	57.72	5	58.32
5	Colombia	54.87	55.02	3.3	56.69	6	57.92
6	Nicaragua	54.67	54.64	22.6	57.72	1	65.37
7	Chile	54.82	53.58	0.5	56.37	7	56.45
8	Ecuador	54.08	54.01	5.9	55.95	8	56.00
9	Guatemala	52.61	51.71	0.0	55.69	9	55.69
10	Honduras	50.70	50.56	9.1	53.33	10	54.82
11	Mexico	48.87	49.16	0.0	51.70	13	51.71
12	Peru	48.70	47.91	0.6	50.54	14	50.57
13	El Salvador	48.28	48.00	0.7	50.54	12	52.19
14	Venezuela	47.25	47.22	23.3	49.82	11	53.12
15	Dominican Republic	46.58	46.40	0.5	48.11	16	48.34
16	Argentina	44.45	44.20	11.3	47.99	15	49.10
17	Costa Rica	44.03	43.52	16.8	45.73	17	46.17
18	Uruguay	40.02	40.10	0.0	42.62	18	42.64
LAC Average		51.11	50.81	6.45	53.38		54.48

Source: Authors'calculations.

Table 9

Gini Coefficient with Methodology Chosen Randomly

Conv.	Gini	1st. Draw		Gini	2nd. Draw		Gini	3rd. Draw		Gini	4th. Draw		Gini	5th. Draw		Gini	6th. Draw		Gini
Par	62.03	Par	3	59.36	Par	1	62.03	Bol	4	59.87	Par	4	62.40	Bra	4	60.05	Par	4	62.40
Bra	59.06	Pan	1	57.55	Bra	4	60.05	Par	3	59.36	Bra	4	60.05	Par	2	59.80	Bol	1	58.77
Bol	58.77	Bra	3	57.12	Pan	4	58.32	Pan	4	58.32	Pan	4	58.32	Bol	1	58.77	Col	4	57.92
Pan	57.55	Bol	3	56.70	Col	4	57.92	Bra	2	57.24	Bol	2	57.32	Pan	1	57.55	Bra	2	57.24
Col	56.70	Ecu	4	56.00	Bol	2	57.32	Nic	1	56.69	Nic	1	56.69	Col	1	56.70	Chi	1	56.38
Nic	56.69	Gua	4	55.69	Nic	1	56.69	Chi	1	56.38	Ecu	4	56.00	Nic	3	54.64	Pan	2	55.70
Chi	56.38	Col	3	55.02	Chi	1	56.38	Ecu	4	56.00	Col	3	55.02	Ecu	2	54.08	Gua	1	55.69
Ecu	56.00	Nic	3	54.64	Ecu	4	56.00	Col	2	54.87	Chi	3	53.58	Chi	3	53.58	Nic	3	54.64
Gua	55.69	Chi	3	53.58	Hon	4	54.82	Hon	1	52.84	Gua	3	51.71	Gua	2	52.61	Ecu	2	54.08
Hon	52.84	Hon	2	50.70	Mex	1	52.76	ElS	4	52.19	Hon	2	50.70	Mex	4	51.71	Hon	1	52.84
Mex	52.76	Per	1	50.55	Gua	2	52.61	Gua	3	51.71	Ven	1	49.63	Hon	2	50.70	Per	1	50.55
Per	50.55	Ven	1	49.63	ElS	4	52.19	Per	1	50.55	Mex	3	49.16	DRep	1	48.10	ElS	1	50.50
ElS	50.50	Mex	3	49.16	Per	4	50.57	Mex	2	48.87	DRep	4	48.34	ElS	3	48.00	Mex	2	48.87
Ven	49.63	ElS	2	48.28	Ven	3	47.22	Ven	3	47.22	ElS	3	48.00	Per	3	47.91	Ven	3	47.22
DRep	48.10	DRep	1	48.10	DRep	2	46.58	DRep	2	46.58	Per	3	47.91	Ven	3	47.22	Arg	1	47.02
Arg	47.02	Cri	1	45.89	Arg	3	44.20	Cri	1	45.89	Arg	2	44.45	Arg	3	44.20	DRep	2	46.58
Cri	45.89	Arg	2	44.45	Cri	2	44.03	Arg	2	44.45	Cri	3	43.52	Cri	2	44.03	Cri	3	43.52
Uru	42.62	Uru	4	42.64	Uru	4	42.64	Uru	4	42.64	Uru	3	40.10	Uru	1	42.62	Uru	3	40.10
Correlation of ranks																			

Source: Authors' calculations.

Table 10
Gini with Random Methodology, and Entropy Indexes

Independent Variable	Dependent Variable: Gini Coefficient					Dependent Variable: Entropy	
	First draw	Second draw	Third draw	Fourth draw	Fifth draw	Top-sensitive Entropy (a=1)	Bottom-sensitive Entropy (a=-1)
Inequality index	0.0017	0.0031	0.0001	0.0018	-0.0010	0.0055	-0.0038
	(1.9)	(2.04)	(0.09)	(2.42)	(0.73)	(2.4)	(1.7)
Ln GNP per capita	-0.0255	-0.0287	-0.0198	-0.0261	-0.0166	-0.0273	-0.0215
	(2.8)	(3.0)	(2.2)	-3.1000	(1.93)	(3.2)	(2.6)
Male Education	-0.0771	-0.1377	-0.0232	-0.0704	-0.0121	-0.0363	-0.0231
	(0.95)	(1.4)	(1.26)	(0.92)	(1.3)	(0.45)	(0.26)
Female Education	0.0966	0.1508	0.0466	0.0869	0.0136	0.0663	0.0501
	(1.87)	(1.77)	(1.89)	(1.92)	(1.6)	(1.93)	(1.82)
PPPI	-0.0006	-0.0007	-0.0002	-0.0005	-0.0006	-0.0002	-0.0003
	(1.04)	(0.4)	(1.65)	(1.78)	(1.3)	(1.58)	(1.48)
R-squared	0.32	0.34	0.24	0.35	0.26	0.38	0.32
Countries	11	11	11	11	11	11	11
Observations	42	42	42	42	42	42	42
Period	1979-1998	1979-1998 9-1	1979-1998	1979-1998	1979-1998	1979-1998	1979-1998

t' Statistics in parenthesis
Source: Authors'calculations

Table 11

Inequality measured with alternative indices

Rank for Conventional Gini	Country	Entropy with a=-1			Entropy with a=1		
		Rank	E(-1)	Change in Rank	Rank	E(1)	Change in Rank
1	Paraguay	4	1.63	3	1	1.02	0
2	Brazil	8	1.34	6	4	0.69	2
3	Bolivia	7	1.34	4	3	0.69	0
4	Panama	3	1.82	-1	7	0.64	3
5	Colombia	5	1.56	0	2	0.83	-3
6	Nicaragua	2	2.29	-4	6	0.64	0
7	Chile	9	1.05	2	5	0.66	-2
8	Ecuador	1	2.36	-7	18	0.21	10
9	Guatemala	14	0.74	5	8	0.62	-1
10	Honduras	11	0.83	1	10	0.57	0
11	Mexico	13	0.75	2	9	0.59	-2
12	Peru	12	0.80	0	11	0.50	-1
13	El Salvador	10	0.96	-3	16	0.31	3
14	Venezuela	6	1.47	-8	12	0.47	-2
15	Dominican Republic	15	0.74	0	13	0.43	-2
16	Argentina	16	0.66	0	14	0.42	-2
17	Costa Rica	17	0.63	0	15	0.38	-2
18	Uruguay	18	0.45	0	17	0.28	-1

Source: Authors'calculations.

Table 12

Inequality with Truncated Distributions

	Country	Excluding					Difference with USA	
		Richest 10 hhs	Richest 1%	Richest 2%	Richest 5%	Richest 10%	Conventional	Truncated at 10%
1	Paraguay	57.36	53.61	51.19	47.02	42.88	17.99	7.62
2	Brazil	58.91	54.60	52.40	48.14	43.75	15.02	8.49
3	Bolivia	57.37	53.84	51.93	48.37	44.66	14.72	9.39
4	Panama	56.98	53.15	51.12	47.44	43.89	13.50	8.63
5	Colombia	54.19	49.71	47.67	43.75	39.68	12.66	4.42
6	Nicaragua	55.21	51.47	49.38	45.51	42.29	12.65	7.03
7	Chile	55.75	51.10	48.39	43.47	38.71	12.33	3.45
8	Ecuador	53.79	50.15	48.03	43.87	39.83	11.95	4.56
9	Guatemala	52.50	47.87	45.32	40.99	36.69	11.64	1.42
10	Honduras	51.75	47.60	45.57	42.18	38.68	8.80	3.42
11	Mexico	52.08	46.21	42.08	39.01	34.00	8.72	-1.27
12	Peru	49.38	45.89	43.75	40.03	36.84	6.50	1.57
13	El Salvador	49.38	45.82	43.64	39.98	36.55	6.45	1.29
14	Venezuela	48.97	45.60	43.86	40.51	37.06	5.58	1.79
15	Dominican Republic	47.44	44.52	42.68	39.05	35.34	4.06	0.08
16	Argentina	47.48	44.19	42.58	39.34	35.94	2.98	0.67
17	Costa Rica	45.53	42.73	41.19	38.22	35.10	1.85	-0.17
18	Uruguay	42.46	39.89	38.46	35.61	32.52	-1.42	-2.75
	LAC Average	52.03	48.22	46.07	42.36	38.58	9.22	3.31
	USA	43.96	42.03	40.52	37.82	35.27		

Source: Authors'calculations.

Table 13

Income Sources for Households
In the Top 10% of the Per Capita Distribution

			Shares ot Total Income				
	Country	Labor income employees	Income from self-employment*			non-labor income	non-monetary income
			Total	labor	capital		
1	Paraguay	28.2	59.5	13.1	46.4	12.3	
2	Brazil	40.4	38.7	15.8	23.0	20.9	
3	Bolivia	35.9	44.2	15.0	29.2	19.9	
4	Panama	60.7	12.5	5.7	6.8	26.8	
5	Colombia	52.8	32.2	14.3	17.8	15.0	
6	Nicaragua	51.4	48.6	13.8	34.9		
7	Chile	34.0	47.4	17.0	30.4	14.6	4.0
8	Ecuador	37.8	37.6	13.6	24.0	16.1	8.5
9	Guatemala	35.1	39.3	13.3	25.9	13.3	12.3
10	Honduras	53.7	46.3	16.4	29.9		
11	Mexico	45.2	21.1	7.2	13.9	10.3	23.4
12	Peru	43.0	24.4	11.4	13.0	18.9	13.7
13	El Salvador	48.0	31.2	11.8	19.4	17.5	3.3
14	Venezuela	57.9	36.4	16.8	19.6	5.7	
15	Dominican Republic	34.1	45.6	13.8	31.8	20.3	
16	Argentina	53.3	24.9	9.5	15.4	21.8	
17	Costa Rica	62.4	23.6	11.5	12.1	14.0	
18	Uruguay	36.9	22.5	16.4	6.1	25.4	15.1
	LAC Average	45.0	35.3	13.1	22.2	17.0	11.5

*Incomes from self-employment are split into labor and capita through imputation methods
Source: Authors' calculations

Table 14

Decomposition of Inequality by Income Source
(%) of Total Inequality

	Country	Labor income employees	Income from self-employment*			non-labor income	non-monetary income
			Total	labor	capital		
1	Paraguay	80.55	13.96	2.15	11.81	5.49	
2	Brazil	60.13	21.92	8.33	13.59	17.95	
3	Bolivia	46.54	27.66	7.77	19.89	25.79	
4	Panama	70.83	8.06	3.48	4.58	21.11	
5	Colombia	59.63	21.97	9.47	12.50	18.40	
6	Nicaragua	83.57	16.43	5.11	11.32		
7	Chile	62.50	12.89	3.59	9.30	20.80	3.81
8	Ecuador	78.03	8.69	2.42	6.27	8.32	4.96
9	Guatemala	46.95	25.89	6.84	19.05	13.91	13.25
10	Honduras	61.08	38.92	10.77	28.15		
11	Mexico	64.06	12.97	4.55	8.42	5.56	17.41
12	Peru	37.29	15.59	5.67	9.92	37.86	9.26
13	El Salvador	55.59	20.67	8.09	12.58	19.94	3.80
14	Venezuela	67.76	26.01	10.54	15.47	6.23	
15	Dominican Republic	48.14	31.90	10.93	20.97	19.96	
16	Argentina	57.16	18.74	10.20	8.55	24.09	
17	Costa Rica	64.04	18.59	9.07	9.52	17.38	
18	Uruguay	46.60	14.92	8.64	6.28	24.55	13.93
	LAC Average	60.6	19.8	7.1	12.7	18.0	9.5

*Incomes from self-employment are split into labor and capita through imputation methods
Source: Authors' calculations

Appendix Table A1

		Household Survey General Description				Des
Country	Year	Survey Name (Origninal Name in Spanish (Portugese))	Sample Frame (Census Year)	Recall Periods (Months)	Geographic Coverage	
Argentina	1996	Encuesta Permanente de Hogares	1991	1	Urban	
Bolivia	1996	Encuesta Nacional de Empleo	1992	0	National	
Brazil	1996	Pesquisa Nacional por Amostra de Domicilios	1991	0	National	
Chile	1996	Encuesta de Caracterización Socioeconómica Nacional	1992	1	National	
Costa Rica	1997	Encuesta de Hogares de Propósitos Múltiples	1984	0	National	
Colombia	1997	Encuesta Nacional de Hogares - Fuerza de Trabajo	1993	0	National	
Dominican Republic	1996	Encuesta Nacional de Fuerza de Trabajo	1993	0	National	
Ecuador	1995	Encuesta de Condiciones de Vida	1990	1	National	
El Salvador	1995	Encuesta de Hogares de Propósitos Múltiples	1992	0	National	
Guatemala	1998	Encuesta Nacional de Ingresos y Gastos Familiares	1994	3	National	
Honduras	1996	Encuesta Permanente de Hogares de Propósitos Múltiples	1988	1	National	
Mexico	1996	Encuesta Nacional de Ingreso Gasto de los Hogares	1995	1	National	
Nicaragua	1993	Encuesta Nacional de Hogares Sobre Medicion de Niveles de Vida	1991	0	National	
Paraguay	1995	Encuesta de Hogares - Mano de Obra	1992	0	National	
Panama	1997	Encuesta de Hogares	1990	1	National	
Peru	1997	Encuesta Nacional de Hogares sobre Medición de Niveles de Vida	1993	1	National	
Uruguay	1997	Encuesta Continua de Hogares	1985	1	Urban	
Venezuela	1997	Encuesta de Hogares por Muestra	1990	1	National	

Table A2

Household Surveys Used in Growth Regressions

Country	Year	Name of the survey	Reference Month for Incomes	Households	Individuals	Labor	Property Rent	Capital Rent	Transfers	Non-Monetary	Imputed R...
Bolivia	96	Encuesta Nacional de Empleo	June	8,311	35,648	X	X	X	X	n.a.	n.a.
	97	Encuesta Nacional de Empleo	November	8,461	36,752	X	X	X	X	n.a.	n.a.
Brazil	81	Pesquisa Nacional por Amostra de Domicilios	September	103,193	481,480	X	X	X	X	n.a.	n.a.
	83	Pesquisa Nacional por Amostra de Domicilios	September	113,599	511,147	X	X	X	X	n.a.	n.a.
	86	Pesquisa Nacional por Amostra de Domicilios	September	65,277	289,533	X	X	X	X	n.a.	n.a.
	88	Pesquisa Nacional por Amostra de Domicilios	September	68,833	298,031	X	X	X	X	n.a.	n.a.
	92	Pesquisa Nacional por Amostra de Domicilios	September	78,188	317,145	X	X	X	X	n.a.	n.a.
	93	Pesquisa Nacional por Amostra de Domicilios	September	80,054	322,011	X	X	X	X	n.a.	n.a.
	95	Pesquisa Nacional por Amostra de Domicilios	September	85,167	334,106	X	X	X	X	n.a.	n.a.
	96	Pesquisa Nacional por Amostra de Domicilios	September	84,862	331,142	X	X	X	X	n.a.	n.a.
Chile	87	Encuesta de Caracterización Socioeconómica Nacional	November	22,719	97,044	X	X	X	X	X	X
	90	Encuesta de Caracterización Socioeconómica Nacional	October	25,793	105,189	X	X	X	X	X	X
	92	Encuesta de Caracterización Socioeconómica Nacional	October	27,666	110,555	X	X	X	X	X	X
	94	Encuesta de Caracterización Socioeconómica Nacional	October	45,379	178,057	X	X	X	X	X	X
	96	Encuesta de Caracterización Socioeconómica Nacional	October	33,636	134,262	X	X	X	X	X	X
Colombia	95	Encuesta Nacional de Hogares - Fuerza de Trabajo	September	18,255	79,012	X	Xb	Xb	Xb	n.a.	n.a.
	97	Encuesta Nacional de Hogares - Fuerza de Trabajo	September	32,442	143,398	X	Xb	Xb	Xb	n.a.	n.a.
Costa Rica	81	Encuesta Nacional de Hogares - Empleo y Desempleo	July	6,604	22,170	X	n.a.	n.a.	n.a.	n.a.	n.a.
	83	Encuesta Nacional de Hogares - Empleo y Desempleo	July	7,132	23,449	X	n.a.	n.a.	n.a.	n.a.	n.a.
	85	Encuesta Nacional de Hogares - Empleo y Desempleo	July	7,351	23,960	X	n.a.	n.a.	n.a.	n.a.	n.a.
	87	Encuesta de Hogares de Propósitos Múltiples	July	7,510	34,591	X	Xb	Xb	Xb	n.a.	n.a.
	89	Encuesta de Hogares de Propósitos Múltiples	July	7,637	34,368	X	Xb	Xb	Xb	n.a.	n.a.
	91	Encuesta de Hogares de Propósitos Múltiples	July	8,002	35,565	X	Xb	Xb	Xb	n.a.	n.a.
	93	Encuesta de Hogares de Propósitos Múltiples	July	8,696	37,703	X	Xb	Xb	Xb	n.a.	n.a.
	95	Encuesta de Hogares de Propósitos Múltiples	July	9,631	40,613	X	Xb	Xb	Xb	n.a.	n.a.
	97	Encuesta de Hogares de Propósitos Múltiples	July	9,923	41,277	X	Xb	Xb	Xb	n.a.	n.a.
Dominican Republic	96	Encuesta Nacional Fuerza de Trabajo	February	5,548	24,041	X	X	X	X	n.a.	n.a.
Ecuador	95	Encuesta de Condiciones de Vida	July-Sept.	5,810	26,941	X	X	X	X	X	X
El Salvador	95	Encuesta de Hogares de Propósitos Múltiples	1995	8,482	40,004	X	Xb	Xb	Xb	n.a.	X
Honduras	89	Encuesta Permanente de Hogares de Propósitos Múltiples	August	8,727	46,672	X	n.a.	n.a.	n.a.	n.a.	n.a.
	92	Encuesta Permanente de Hogares de Propósitos Múltiples	August	4,757	24,704	X	n.a.	n.a.	n.a.	n.a.	n.a.
	96	Encuesta Permanente de Hogares de Propósitos Múltiples	August	6,428	33,172	X	n.a.	n.a.	n.a.	n.a.	n.a.
	98	Encuesta Permanente de Hogares de Propósitos Múltiples	February	6,493	32,696	X	Xb	Xb	Xb	n.a.	n.a.
Mexico	84	Encuesta Nacional de Ingreso Gasto de los Hogares	Third quarter	4,735	23,985	X	X	X	X	X	X
	89	Encuesta Nacional de Ingreso Gasto de los Hogares	Third quarter	11,531	57,289	X	X	X	X	X	X
	92	Encuesta Nacional de Ingreso Gasto de los Hogares	Third quarter	10,530	50,862	X	X	X	X	X	X
	94	Encuesta Nacional de Ingreso Gasto de los Hogares	Third quarter	12,815	60,365	X	X	X	X	X	X
	96	Encuesta Nacional de Ingreso Gasto de los Hogares	Third quarter	14,042	64,916	X	X	X	X	X	X
Nicaragua	93	Encuesta Nacional de Hogares Sobre Medicion de Niveles de Vida	February to June	4,455	24,542	X	X	X	X	X	X
Panama	79	Encuesta Continua de Hogares - Mano de Obra	July	8,593	24,284	X	n.a.	n.a.	n.a.	n.a.	n.a.
	91	Encuesta Continua de Hogares - Mano de Obra	July	8,867	38,000	X	Xa	Xa	X	n.a.	n.a.
	95	Encuesta Continua de Hogares	July	9,875	40,320	X	Xa	Xa	X	n.a.	n.a.
	97	Encuesta de Hogares	July	9,897	39,706	X	Xa	Xa	X	n.a.	n.a.
Paraguay	95	Encuesta de Hogares - Mano de Obra	August to November	4,667	21,910	X	X	X	X	n.a.	n.a.
Peru	85-86	Encuesta Nacional de Hogares sobre Medición de Niveles de Vida	July 1985 to July 1986	5,108	26,323	X	X	X	X	X	X
	91	Encuesta Nacional de Hogares sobre Medición de Niveles de Vida	August-October	2,308	11,507	X	Xa	Xa	X	X	X
	94	Encuesta Nacional de Hogares sobre Medición de Niveles de Vida	April-June	3,623	18,662	X	Xa	Xa	X	X	X
	96	Encuesta Nacional de Hogares sobre Niveles de Vida y Pobreza	April-June	16,744	88,863	X	Xa	Xa	X	X	X
	97	Encuesta Nacional de Hogares sobre Niveles de Vida y Pobreza	August-October	3,843	19,575	X	Xa	Xa	X	X	X
Venezuela	81	Encuesta de Hogares por Muestra	Second semester	45,421	239,649	X	n.a.	n.a.	n.a.	n.a.	n.a.
	86	Encuesta de Hogares por Muestra	Second semester	129,713	682,636	X	n.a.	n.a.	n.a.	n.a.	n.a.
	89	Encuesta de Hogares por Muestra	Second semester	61,385	315,650	X	n.a.	n.a.	n.a.	n.a.	n.a.
	93	Encuesta de Hogares por Muestra	Second semester	61,477	306,629	X	n.a.	n.a.	n.a.	n.a.	n.a.
	95	Encuesta de Hogares por Muestra	Second semester	18,702	92,450	X	Xb	Xb	Xb	n.a.	n.a.
	97	Encuesta de Hogares por Muestra	Second semester	15,948	76,965	X	Xb	Xb	Xb	n.a.	n.a.

an not separate between property and capital rent.

an not separate between property rent, capital rent, and transfers.

51