Shapley Decomposition by Components of a Welfare Aggregate

September 2012

Joao Pedro Azevedo

Viviane Sanfelice

Minh Cong Nguyen

Abstract

We present an exercise to account for changes in an indicator into components of a welfare aggregate. The decomposition presented in this paper focus on the distribution of the welfare variable rather than a decomposable indicator. Thus, it is a statistical exercise in which we assume that we can in fact modify one factor at a time and keep everything else constant. This paper contributes to the existing literature by applying the concept of a Shapley decomposition to deal with the path dependence that arises from changing one element at a time. Although, the interpretation changes, the proposed decomposition is applicable to both the case of panel data as well as repeated cross section and allows accounting the contribution of each component to change in any given indicator.

Keywords: Microdecompositions, Shapley value, Poverty, Inequality

The authors would like to thank Gabriela Inchauste, Samuel Freije, Sergio Oliveri, Samuel Franco, Andres Castaneda, Gabriel Facchini and Jordan Solomon for their valuable suggestions.

1 Introduction

In the last two decades, the economic literature related to poverty and inequality measures has dedicated efforts in techniques to decompose those measures into certain factors. Such decompositions have been useful by assigning contribution to a specific factor. The straightest are those in which the measure can be decomposable by itself, such as, the FGT poverty indices that can be partitioned by subgroup of the population. Similar property is observed for some inequality measures, such Theil index and allows the decomposition by between and within group effect. However, since these decompositions rely on desirable proprieties of the indicator, they may faces problem such as those listed by Shorrocks (2012). First, related to the interpretation of a factor, it is not always true that a factor will have a meaningful and intuitive interpretation. Secondly, some indicators cannot be decomposable in a way that the effect of the factors sums up the total indicator value; the Gini coefficient subgroup decomposition is an example. Another point is the amount and mix of factors that decompositions allow, in the most part of them, the number of factors has to be constrained.

Focusing the distribution of income rather than decomposable propriety of an indicator, Datt and Ravallion (1992) proposed to decompose changes in poverty into the two main mechanisms to reduce it: growth and redistribution. Later on, Kolenikov and Shorrocks (2003) observed the same decomposition, but added a third component to capture the effect of prices on change in poverty and proposed a new methodology based on Shapley[1] value.

The concept of the Shapley decomposition can be describe as the “marginal effect on *[the indicator]* of eliminating each of the contributory factors in sequence, and then assigns to each factor the average of its marginal contributions in all possible elimination sequences. This procedure yields an exact additive decomposition of *[the indicator]* into *[the number of]* contributions” (Shorrocks, 2012 p.3). This paper utilized Shapley technique to decompose the change in a given indicator over two periods of time by components of the welfare aggregate.

The knowledge of which source of income has been important to decrease poverty and/or inequality seen to be relevant. For instance, a government may be interested in assessing how much of the observed reduction in poverty was due to the implementation of a cash conditional

[1] See Shorrocks (1999, 2012), Sastre and Trannoy (2001) for more details about the Shapley technique in decomposition of poverty and inequality indicators.

transfer (CCT program) or how much labor income contributes to decrease inequality. The decomposition follows on the exercise proposed by Paes de Barros et alii (2006) and takes advantage of the additive property of the welfare aggregate to construct counterfactual unconditional distributions of the welfare aggregate by changing each component at a time to calculate their contributions to the observed changes in poverty or/and inequality index. The Shapley concept is used in order to deal with the path dependence[2] that arises from the stepped decomposition. Additionally, it relies, in the absence of panel data, on the rank correlation of the welfare aggregate observed in each period in order to transpose the distributions from one period to other.

In the literature, different approaches are utilized in order to decompose an indicator into components of income. However, the most part of decompositions rely on decomposable propriety of the indicator and then, become dependable of hypothesis. In the context of inequality measure[3], for instance, Shorrocks (1982) showed that the amount of constraints may lead to a restrict decomposition framework when using more sophisticated index as Gini.

A positive point of the decomposition is that it can be applied using any welfare measure as well as any indicator which is calculated based on the welfare aggregate. Besides that, it clearly provides the contribution of each component to the observed change in the indicator.

Next section presents the Shapley decomposition. Then, a more detailed description of the approach used to transpose the distribution is provided and possible problems about the data are discussed in the same section. Section 4 presents applications of the decomposition and external validation when repeated cross section data is used in the decomposition. Finally, section 5 concludes.

2 Shapley Decomposition

Let Y be a welfare aggregate given by a function $f: C^n \rightarrow R$ that maps components $c_1, \ldots, c_n$ to Y

$$Y = f(c_1, c_2, \ldots, c_n) \tag{1}$$

[2] See Essama-Nssah (2012), Fortin et al (2011) and Ferreira (2010) for recent reviews of the literature.
[3] See Shorrocks (1982), Fournier (2001), Paul (2004) and Mussini (2013).

Let I be an indicator that can be calculated using Y

$$I = I(Y) = I(f(c_1, c_2, \dots, c_n), .) \tag{2}$$

We are interested in decomposing the change in I over two periods of time[4], $t=0,1$, into n contributions, σ_i, attributed to each of the n components, such that

$$\sigma_i = I\big(f(c_i{}^{t=1}, .)\big) - I\big(f(c_i{}^{t=0}, .)\big) \quad \forall\, i = 1, \dots, N \tag{3}$$

However, if we merely take the first marginal change in I when switching the distribution of the component c_i from $t=0$ to $t=1$ and then computing the difference on the indicator, there is no guarantee that, at the end, the sum of all n contributions is going to return the total change in I from period 0 to period 1.

The Shapley decomposition deals with the problem by calculating all $n!$ possible ways of decomposing I by eliminating each component at once and then taking the average of the contributions of the component. Shapley (1953) presented this concept in the context of Game Theory. He establishes that the relationship between the sequential elimination of players and their final contribution is given by a weighted average.

Figure 1 shows the structure of a decomposition of n components. Note that the components are changed sequentially and there will be $n!$ different paths to go from I^0 to I^1.

In our case, the goal is to separate the change in an indicator into n factors attributed to each components individually, thus the final contribution of the component c_i , when the Shapley decomposition is used, is going to be determined by the following weighted average.

$$\sigma_i = \sum_{s=0}^{n-1} \sum_C \frac{s!(n-s-1)!}{n!} \left[I\left(f\left(c_i{}^{t=1}, C_{n-1,s+1} - \{c_i\} \right) \right) - I\left(f\left(c_i{}^{t=0}, C_{n-1,s} - \{c_i\} \right) \right) \right] \tag{4}$$

in which s indicates how many components have been already changed from period 0 to period 1. C denote all the combinations of the others n-1 components that have already change from t=0 to t=1. The Equation (4) establishes the relationship between the structure presented in Figure 1 and the final contribution of the component c_i after calculating all the paths from I^0 to I^1.

[4] We will denote it here as period in order to make the understating easier, but it is also possible to make the decomposition comparing between two states. For instance, to compare the difference in the poverty headcount between municipality A and municipality B.

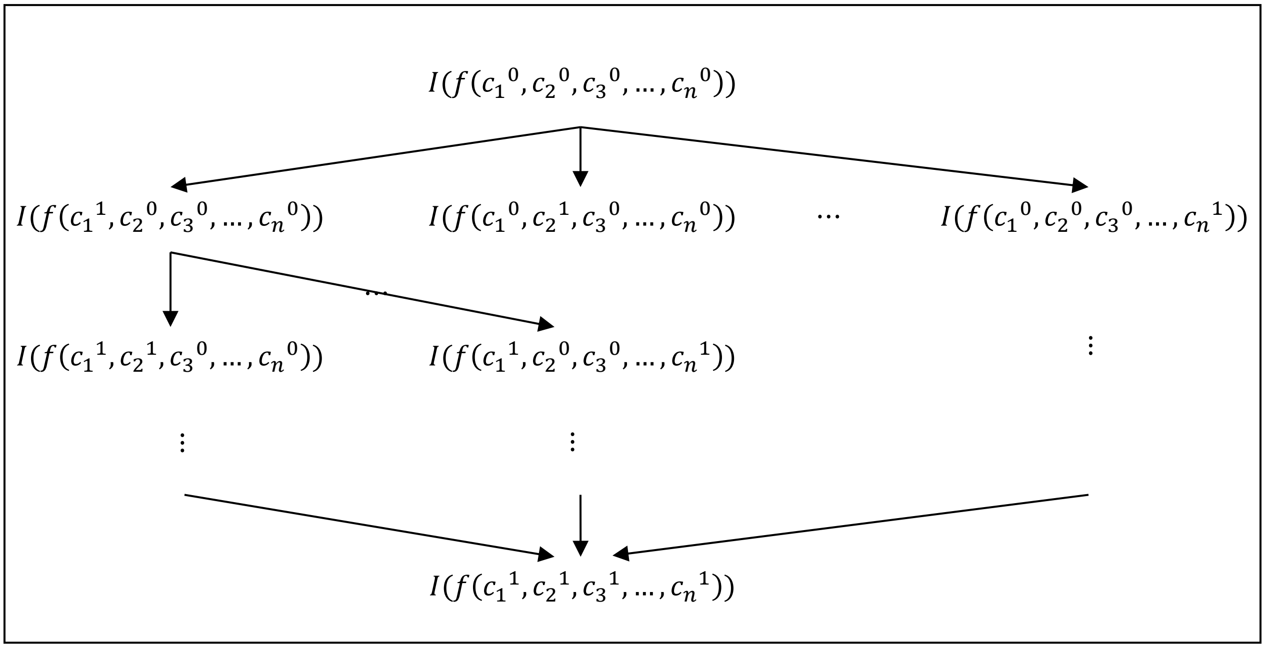


Figure 1 – Shapley Decomposition of n components.

Suppose a decomposition of three components: c_1, c_2 and c_3. The Figure 2 shows the 3! ways to decomposition the change in a indicator based on the welfare measure, $Y=f(c_1,c_2,c_3)$. Using the notation presented before, the contribution of the component c_1, for instance, is given by the following weighted average

$$\sigma_1 = \sum_{s=0}^{2} \sum_{C} \frac{s!(3-s-1)!}{3!} \left[I\left(f\left(c_1{}^{t=1}, C_{2,s+1} - \{c_1\} \right) \right) - I\left(f\left(c_1{}^{t=0}, C_{2,s} - \{c_1\} \right) \right) \right] \tag{5}$$

$$\begin{aligned} \sigma_1 = {} & \frac{2}{6}\left[I\left(f(c_1^1, c_2^0, c_3^0) \right) - I\left(f(c_1^0, c_2^0, c_3^0) \right) \right] \\ & + \frac{1}{6}\left[I\left(f(c_1^1, c_2^1, c_3^0) \right) - I\left(f(c_1^0, c_2^1, c_3^0) \right) \right] + \frac{1}{6}\left[I\left(f(c_1^1, c_2^0, c_3^1) \right) - I\left(f(c_1^0, c_2^0, c_3^1) \right) \right] \\ & + \frac{2}{6}\left[I\left(f(c_1^1, c_2^1, c_3^1) \right) - I\left(f(c_1^0, c_2^1, c_3^1) \right) \right] \end{aligned} \tag{6}$$

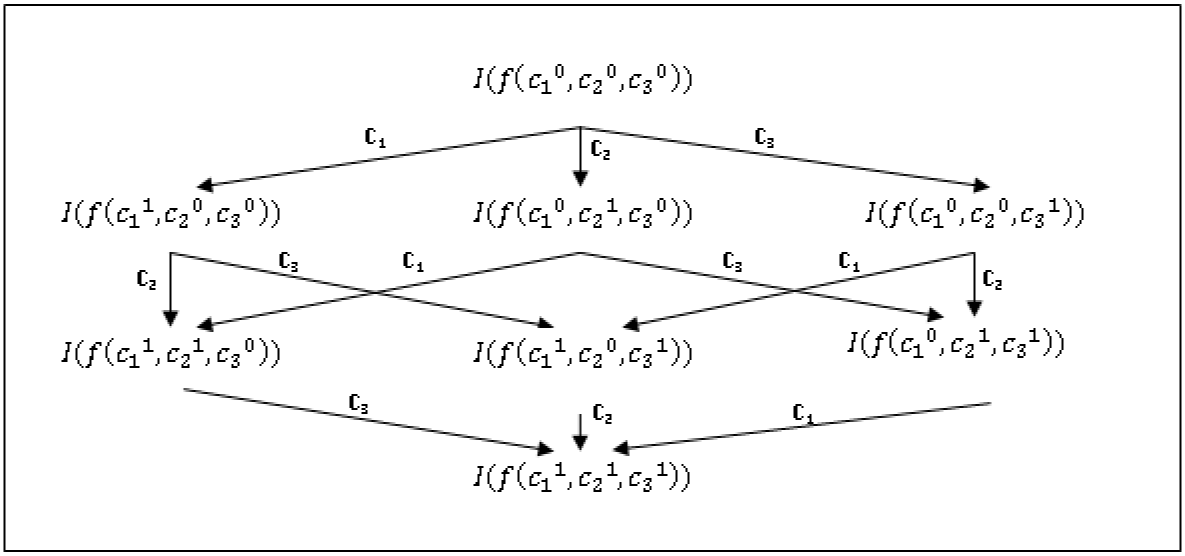


Figure 2 – Shapley Decomposition of 3 components.

A disadvantage of the Shapley decomposition is that the principle of independence of aggregation may not be satisfied. It means that the results may be different depending on the level of disaggregation of a component. For example, the contribution of transfer income probably is not going to be equal to the sum of the contribution in a decomposition that has transfer split between public and private as separately components.

3 The Decomposition

As a description of the method suggested here, consider that the distribution of an observable welfare aggregate (i.e. income or consumption) for period 0 and period 1 is known and that we can calculate this welfare variable using an equation based on other variables called components. As illustrated in Figure 1, the decomposition consists in constructing counterfactual distributions by substituting the observed distribution of components of the welfare aggregate in period 1, one at a time, until a completed change from period 0 to period 1. For each counterfactual distribution, we can compute the poverty or inequality measures, and interpret those counterfactuals as the poverty or inequality level that would have prevailed in the absence of a change in that component. Nevertheless, since the counterfactuals generated are not result of an economic equilibrium, we need to keep in mind that the method suggested here is more an accounting exercise rather than a behavioral economic model.

For instance, substitute the distribution of labor earnings in period 0 by the earnings from period 1 and keep all the other sources of income constant in period 0. Thus, we can compute how poverty would be if the labor income in period 0 was equal to labor income in period 1. The difference in poverty, when use earnings in period 0 and transfer in period 1, would be a contribution of the income from labor to reduce/increase poverty.

Figure 3 illustrates the example describe above for labor income using data from Paraguay in 1999 and 2009. As we can see, a counterfactual distribution of the per capita income in 1999 is computed using the labor income observed in 2009. The effect on poverty attributed to labor would be the difference between the area below green curve and limited by the poverty line in Fig 3 A.2 and the area below blue curve and limited by the gray line in Fig 3 A.1.

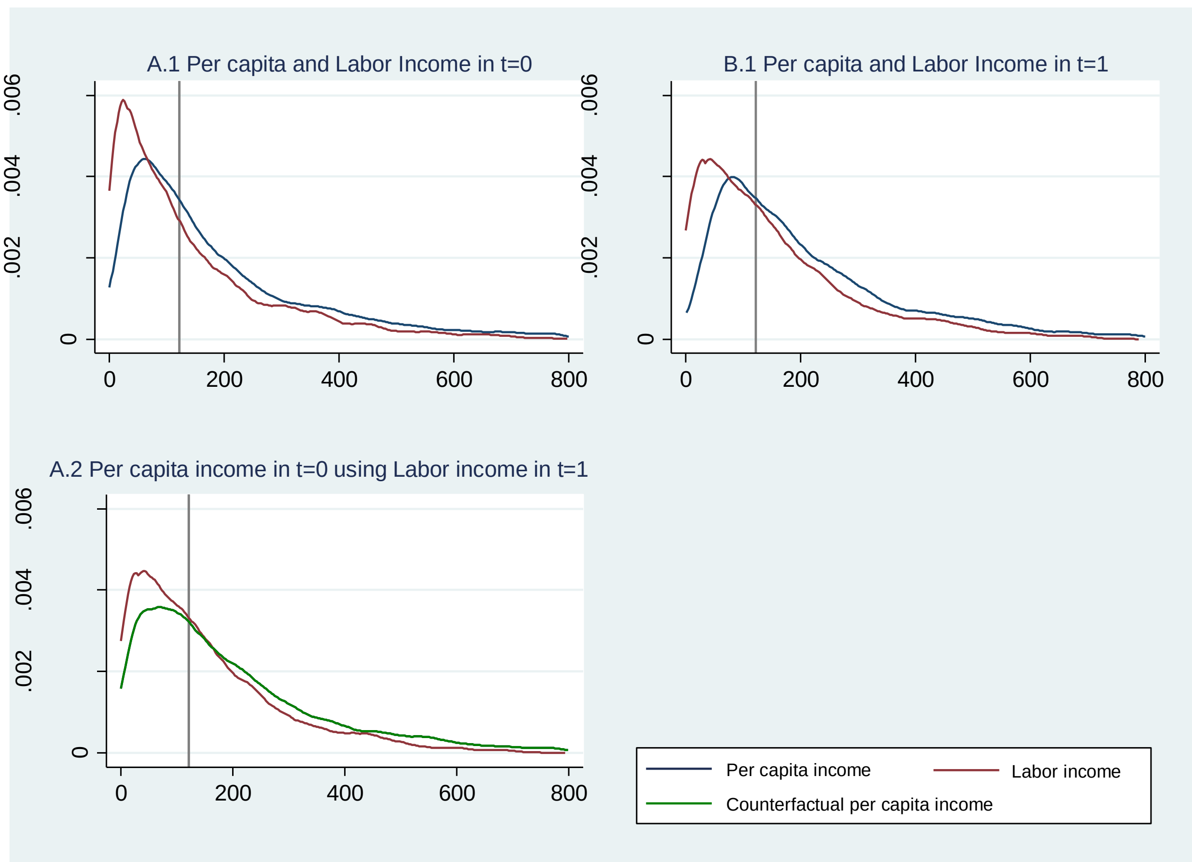


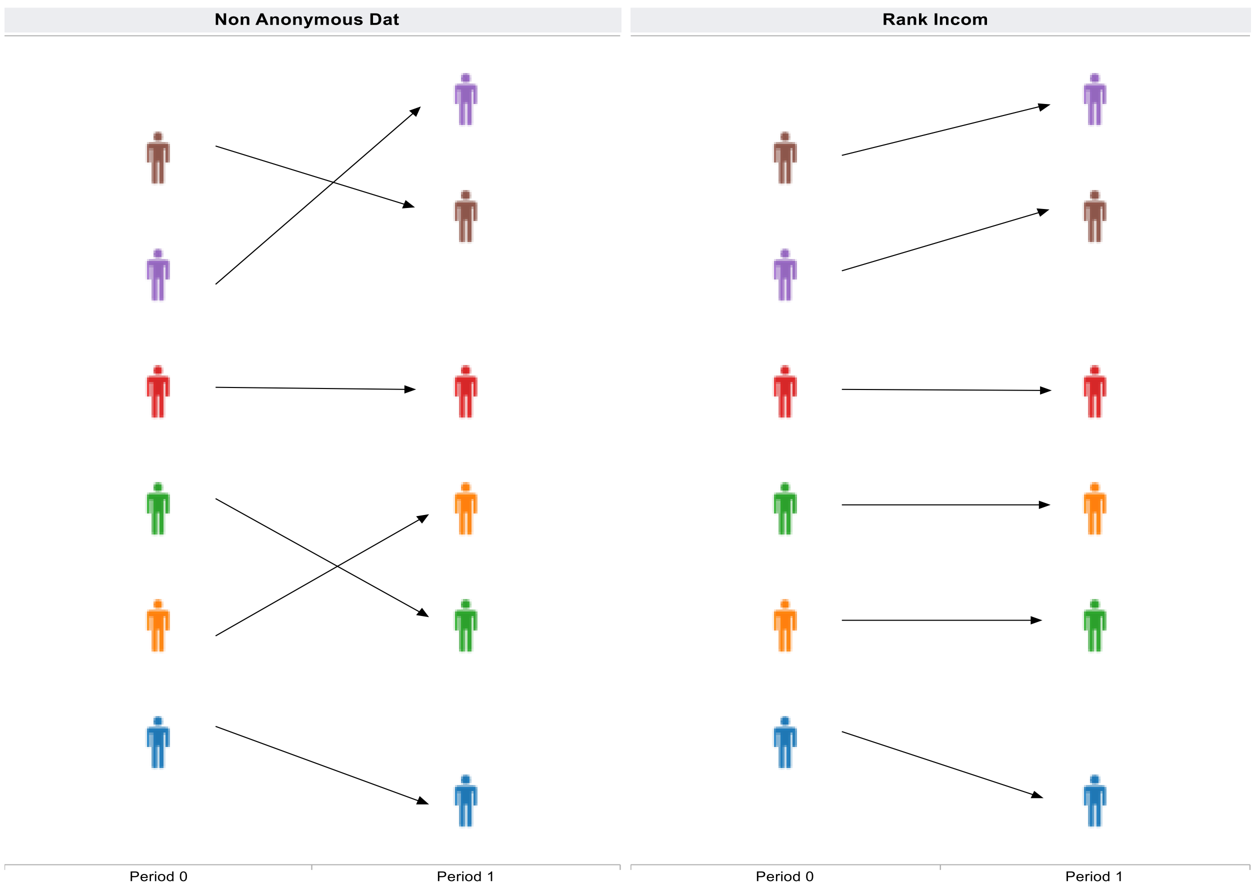
Figure 3 – Counterfactual distribution of income after changing labor income.

The change of the distribution of one element of the welfare aggregate from period to the other is a marginal change in the distribution of the welfare variable. Thus, in terms of application, the transposition of distribution of the components between the two periods holds an important aspect. The main concern is how the values are assigned to the unit of observations. In the case of Datt-Ravallion decomposition, for instance, there is not such issue, since the entire distribution of the welfare aggregate is switched. Here, we substitute only one component of the welfare aggregate and then, the distribution of the welfare aggregate need to be reconstructed using the new distribution of that component, therefore the value assigned to the unit of observations matters.

The use the *anonymous* and *non-anonymous* data has been discussed and analyzed due to possible disparities on the results when one type of data is used in contrast to the other. This discussion has taking place especially in the context of Growth Incidence Curve (GIC) (See Bourguignon, 2011 and Jenkins and van Kern, 2006) in which the percentile of the per capita income is used in anonymous context in order to compare this variable in two periods. Bourguignon (2011) shows that the way the units of analysis are tracking can drive to different results. According to him, the use of the rank of the income to compare the individuals ignores income mobility, once that the comparison occurs between individuals that are in the same position in each period. Below, we discuss the use of panel data and repeated cross section in the context of the decomposition proposed here.

In the case of panel data, in which the value of the welfare and components variables is observed in both periods for each unit of observation, there is no concern and the matching should use the observed value that the individual has in each period. Note that even when balanced panel data is available, the problem of path dependence remains and the decomposition using the Shapley procedure is still very useful to solve this issue.

However, in most countries panel data are not available or are still relatively short. Therefore, we need to make an assumption about the welfare aggregate through the time. Here, we propose to use the concept of rank correlation, in which the rank of variables is observed rather than the statistic correlation among them. This concept was utilized by Fournier (2001) who suggests decomposing inequality into sources of income and their rank correlation. This approach is convenient because it does not rely on any parametrical assumptions. In our case, we propose to match the unit of analysis based on their observed rank of the welfare aggregate in each period. In this case, the relationship between the component and welfare aggregate observed in each period is kept. In fact, it is the situation of distributional dynamics without exchange mobility in which each individual keeps his rank across the periods, from poorest to richest (Yalonetzky, 2012). The figure below gives a better idea about the matching based on the rank of the individuals according a welfare variable. The individuals are sorted by income in each period. When panel data is available we can track the individuals through the time. Otherwise, the matching is performed by the observed distribution of income in each period.



Note: This figure is based on the illustration of Yalonetzky, 2012 Fig 1.

Figure 4 – Individual trajectories in case of panel data and in case of matching using welfare rank.

It is important to highlight that using the rank of the welfare measure to transpose the distribution of the components, we are not tracking individuals, but the entire observed distribution of the welfare aggregate and its components. Thus, we are interested in knowing to what extension changes in the mean and the distribution of each component of the welfare aggregate between the period 0 and period 1, using the observed structure of correlation of the components in relation to the welfare aggregate in each period, affect the variation observed in indicators of poverty and/or inequality.

In terms of practical data issue, two main problems arise when we need to transpose the components distribution between the periods, especially because a value has to be assigned for each unit of observation, i.e., individuals or household. The first concerning refers to the size of the database; very often there is not the same number of observations in the dataset of each period. Secondly, in most part of surveys are sample design, in which a expansion factor is attributed to each individual.

To deal with the first problem, we suggest a rescale technique. It literally rescales the number of observation in one year in order to have the same range as the other year. Suppose that the rank of the welfare aggregate Y is equal $1,\ldots,N_0$ in the initial period and equal to $1,\ldots,N_1$ for $t=1$. Usually $N_0 \neq N_1$, then a matching based on the rescale rank of Y is made in order to switch the distribution between the periods. Equation (6) provides the rescale formula for period 0.

$$rescaled\ rank\ Y^{t=0} = round\left(\frac{N_1}{N_0} * rank\ Y^{t=0}\right) \tag{6}$$

The rescale method is going to be a matching with repetition for period 0 (period 1) and some observations is not going to be used for period 1 (period 0) when $N_0>N_1$ ($N_1>N_0$). Therefore, after the matching we suggest to recover the original mean by multiplying the whole distribution by the original mean and diving by the mean generated after the matching. Thus, we get a very close distribution and preserve the mean.

Another possibility frequently utilized is to divide the distribution of Y in m percentiles in both periods and then using the mean of the components variables in each bin. The problem with this technique is that indicators sensible to the tail of the distribution, such as Gini, would not be measured correctly, when the number of percentile is not close to the total number of observation. Indeed, the percentile option is similar to rescale, when m goes to $min(N_0,N_1)$.

As an external validation exercise, the Kolmogorov-Smirnov test about equality of distributions was implemented using database from several countries in Latin America to circa 2000. The rescale and the 200 percentiles methods were compared. With exception of labor income in Argentina, for all the others sources of income (pensions, transfers and others) and countries, the rescale matching methodology performs very well in a way that the test rejects difference in the distribution at the significance level of five percent. The same is not true for the percentile method using 200 percentiles, for practically all sources of income the test points out difference in the distributions.

The second data issue cited at the beginning of this section regards the sample weight is probably softened when the decomposition is made in both direction, using *t=0* as initial period and then *t=1* as initial point. In this case, counterfactuals using weights from either period are accounted.

Both limitations cited before are smoothed when the survey has similar design on the two periods analyzed.

4 Application and Validation

Suppose we are interested in knowing the main contributors to reduce poverty in Latin American and Caribbean (LAC) countries in the last decade. Or by how much have males and females contributed to the observed reduction in inequality. Or even, why poverty rates between states X and Z are different. The decomposition proposed here allows answering these questions by accounting change in a given indicator into components of a welfare variable. Below, some simple applications[5] are demonstrated as well as some external validation of the decomposition when repeated cross section data is used.

Poverty and Inequality Indicators

In order to decompose the contribution of each factor to poverty reduction, we need some structure that would allow us to measure the contribution of each factor to the total change in poverty. We begin modeling household per capita income as:

$$Y_{pc} = \frac{Y_h}{n} = \frac{n_A}{n}\left(\frac{1}{n_A}\sum_{i\in A}^{n} y_i^L + \frac{1}{n_A}\sum_{i\in A}^{n} y_i^{NL}\right) \tag{7}$$

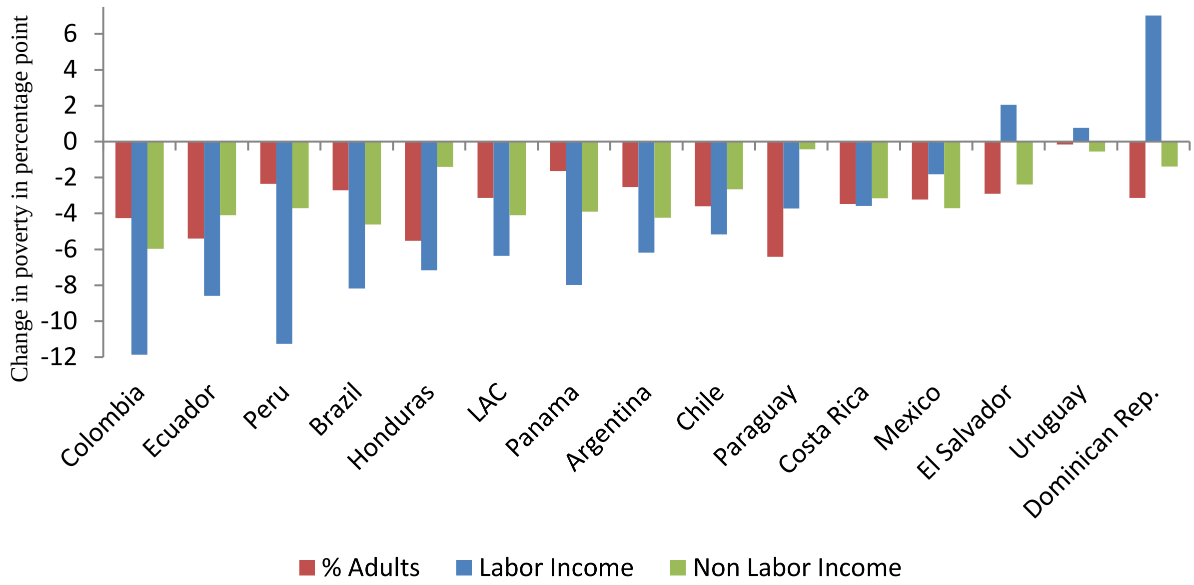
in which, *Y* denotes income, and *n* number of members in the household. *A* represent adult, *L* labor and *NL* non labor.

[5] To others applications of the decomposition, see Azevedo, Inchauste and Sanfelice (2012) and World Bank (2012) "The effect of women's economic power in Latin America and the Caribbean" Washington, D.C.: LCSPP Poverty and Labor Brief, No. 4.

Thus, income per capita is the sum of each individual's income and will depend on the number of household members. If in addition we recognize that only individuals adults contribute to family income, income per capita will in fact depend on the number of adults in the family, n_A. Income per adult, in turn, depends on labor income, y_i^L, and non-labor income, y_i^L, where non labor income includes public social transfers, pensions, remittances and other private transfers.

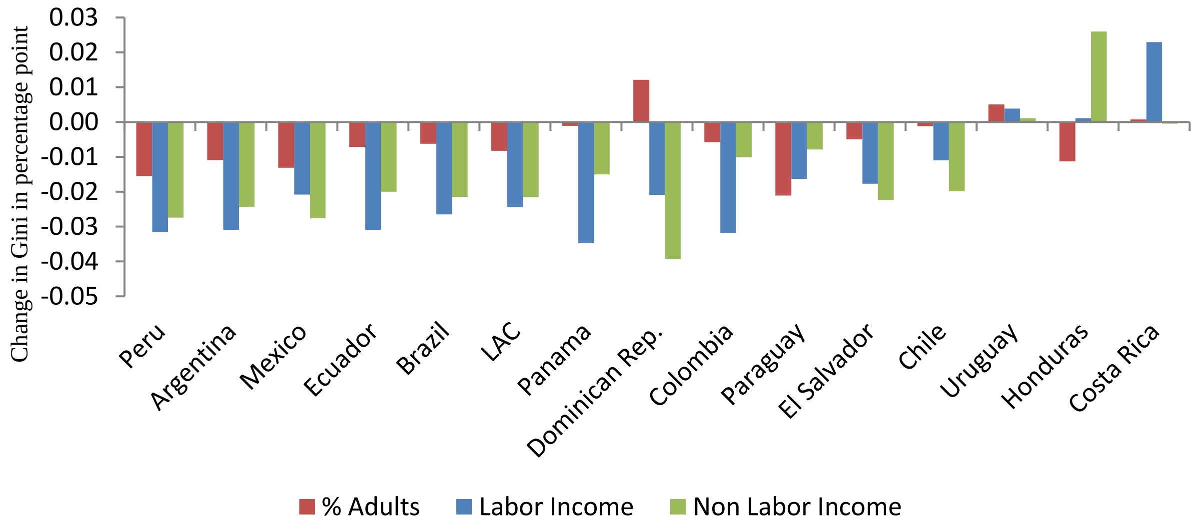
Figure 5 shows the results for poverty headcount in LAC countries and region. In general, labor income accounts for the most part of the reduction in poverty. For LAC region, it is around 47 percent. In fact, the countries that show greater reduction in poverty are those whose labor income had high participation. The demographic factor is reflected in the contribution of the percentage of adults in the household. The growth in the number of adults helped with 23 percent of the decrease of poverty for LAC region. Non labor income that includes transfers, pensions (contributory and non contributory) and rent income contributed with 30 percent.

Considering inequality measures through the Gini coefficient, the percentage of adults is not as important as it is to poverty reduction, the number of adult in the household accounts for only 15 percent of the reduction in inequality in LAC region. Labor income continues to be the main contributor with participation of 45 percent. Interestingly, non labor income was a fundamental factor to decrease inequality in LAC from 2000 to 2010, contributing with 40 percent.



Source: Estimates based on data from SEDLAC (CEDLAS and the World Bank).

Figure 5 – Contribution of percentage of adults, labor and non-labor income in percentage points to reduce moderate poverty ($4USD a day) in Latin American countries, circa 2000 to 2010.

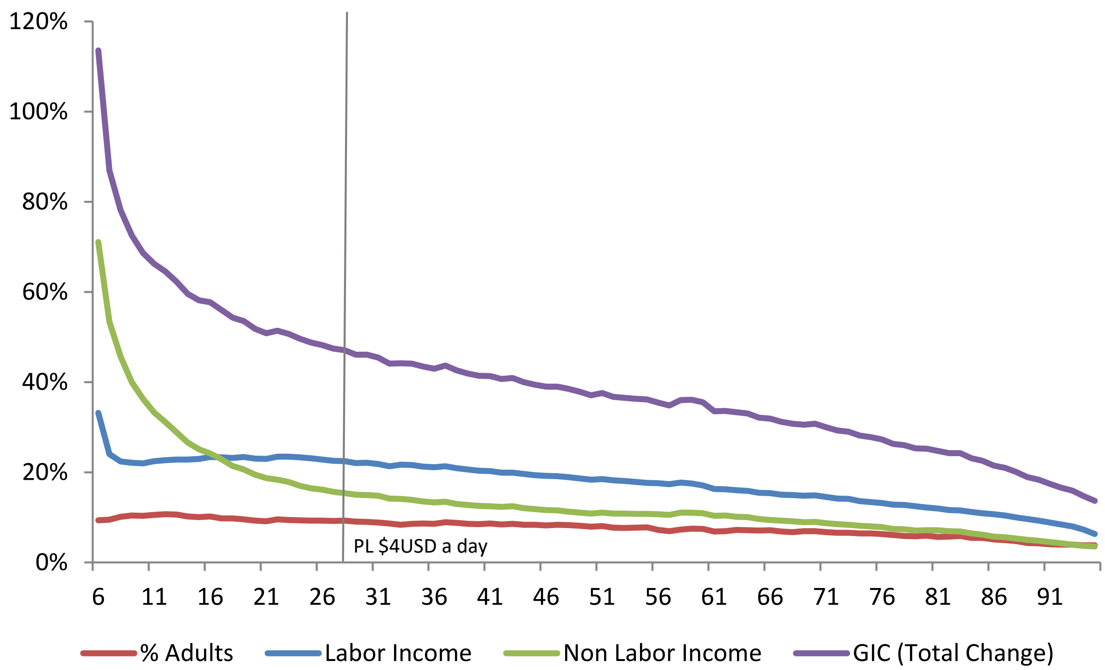


Source: Estimates based on data from SEDLAC (CEDLAS and the World Bank).

Figure 6 – Contribution of percentage of adults, labor and non-labor income in percentage points to reduce Gini in Latin American countries, circa 2000 to 2010.

Growth Incidence Curve

As poverty and inequality index, we can also use the change in mean income by percentile of the income per capita as an indicator in order to analysis change in the entire distribution of the welfare aggregate. It means to decompose the GIC into components of the income per capita.



Source: Estimates based on data from SEDLAC (CEDLAS and the World Bank).

Figure 7 – Contribution of percentage of adults, labor and non-labor income in the GIC Latin American region, circa 2000 to 2010.

The Figure 7 shows the results using the disaggregation in the equation (7) for LAC region. We can see that the percentage of adults presents a slightly decrease through the percentiles, while labor income and non-labor income are extreme regressive. These results are consistent with contribution of the components for the poverty and inequality index reported before. Labor income is the greater contributor to the increase in the income per capita on the bins around the poverty line of $4USD a day, while the non-labor income is the greater contributor to raise income on the poorest percentile.

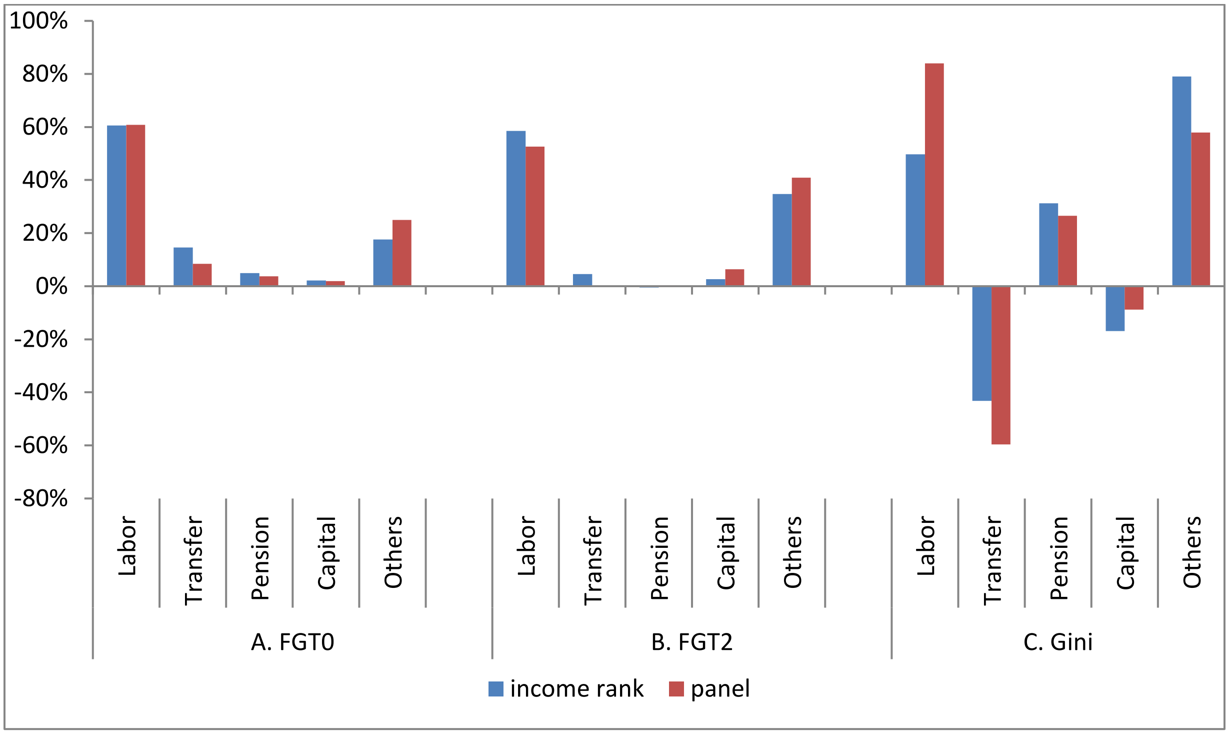
External Validation

Below we report decompositions of poverty headcount, poverty severity and inequality by sources of income using the panel data from Encuesta Nacional de Hogares (ENAHO) from Peru in 2009 and 2010, Encuesta Nacional sobre Niveles de Vida de los Hogares (ENNVIH) from Mexico in 2002 and 2005 and data from Chilean survey, Encuesta de Caracterización Socioeconómica Nacional (CASEN) in 1996 and 2001. The goal is to analysis the result when panel data is used and when the dataset is treated as repeated cross section, i.e. the decomposition is made using non-anonymous and anonymous data.

First, we assess the approach based on the rank of the welfare aggregate to match the individuals in case of repeated cross section data, i.e., knowing that the individuals in the sample are the same, how would be the findings in case of matching the individuals using the rank of the welfare variable. This analysis is made using data from ENAHO and ENNVIH. As the Figure 8 and 9 show, the results are qualitative consistent through the procedure adopted to transpose the distribution between the years, especially for poverty headcount. For more sensible measure as FGT(2) and Gini index, the two approaches present shares with different magnitude, but it does not modify substantially the interpretations.

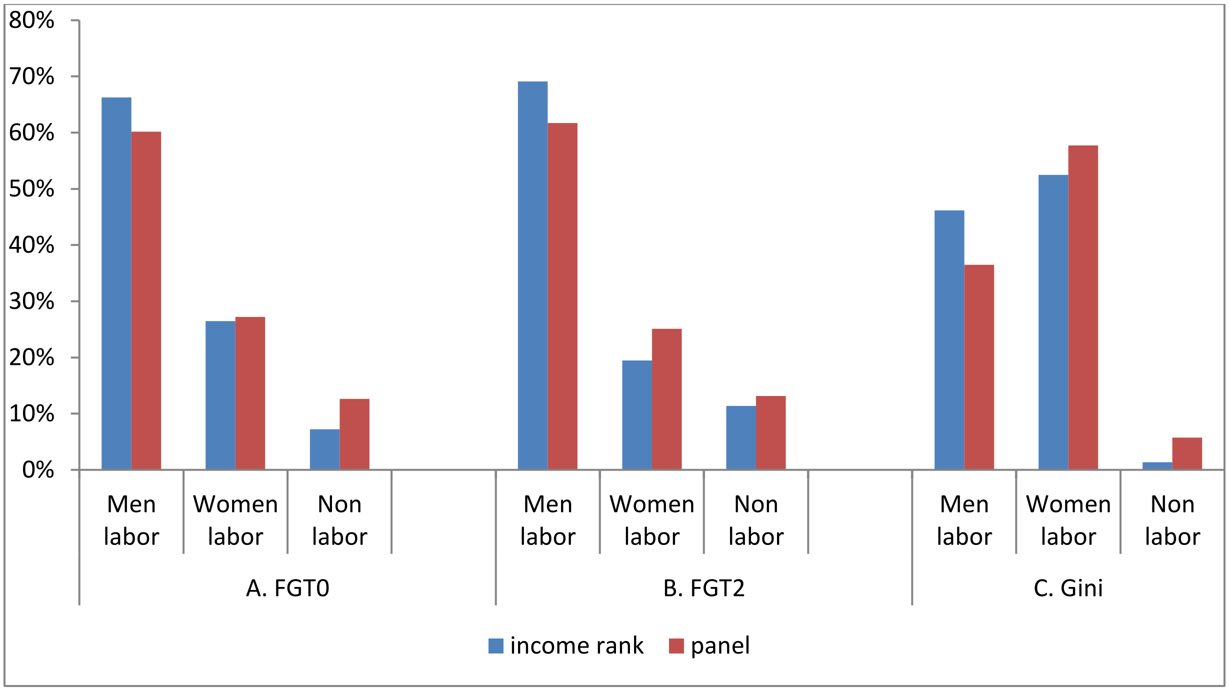
In the decomposition presented in Figure 10 using panel data from Chile and a range of five years, we are interested in knowing whether repeated cross section data provides similar picture as panel data. Thus, we randomly divided the individuals in two groups, A and B. The non-anonymous decomposition was made using individuals in group A, while for the anonymous decomposition we used individuals in group A in period 0 and individuals in group B in period 1. Thereby, when the panel data is used as repeated cross section, we ensure that the dataset has different individuals in each year. Indeed, we would expect that this data structure would lead for more unlike findings, since the individuals compared are not the same. However, as we can observe below, the results are quite robust within the method utilized, even for more sensible measure as FGT(2) and Gini.

In his empirical application of mobility decomposition using data from Peru, Yalonetzky (2012) actually found that exchange mobility, which is the one that we exclude when using the rank of income to transpose the distribution, is actually responsible for about 13 percent or less of the total mobility. His findings are in agreement with ours, if the exchange mobility does not account too much in the total mobility, we should not expect serious differences in the results using anonymous or non-anonymous approach.



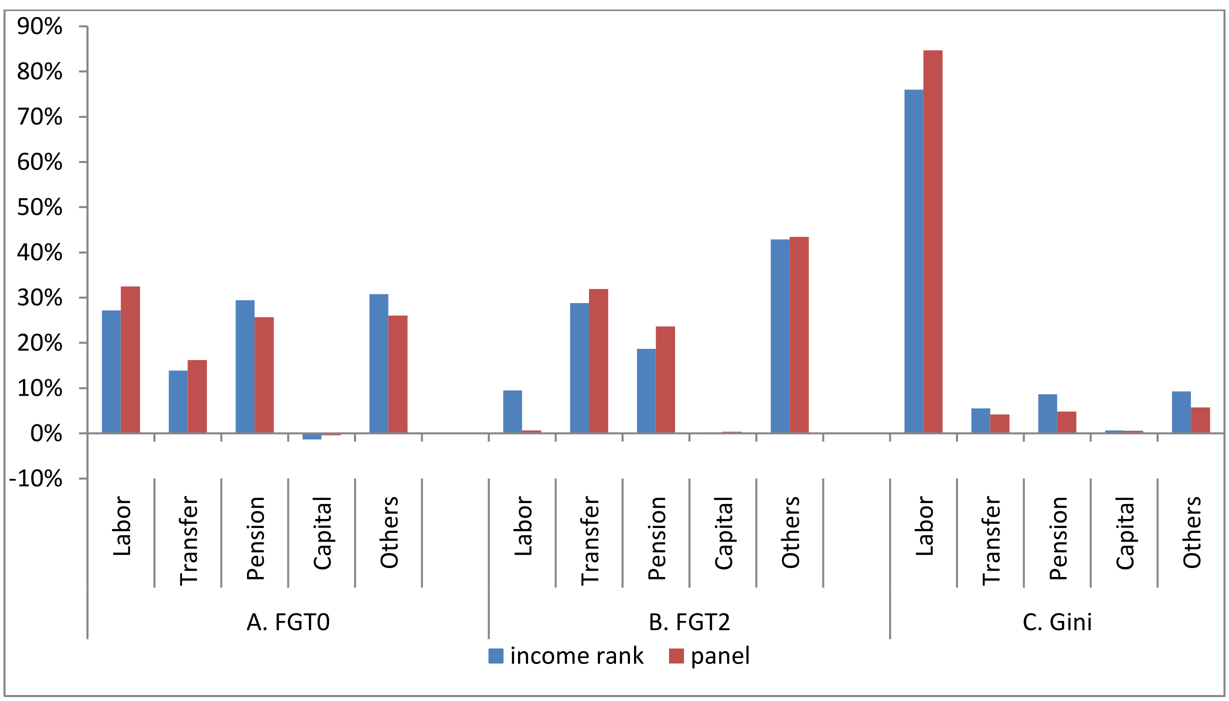
Source: Estimates based on data from SEDLAC (CEDLAS and the World Bank). Note: Poverty line of $4USD a day.

Figure 8 – Decomposition of per capita income into sources of income, using ENAHO data from Peru, 2009 and 2010.



Source: Estimates based on data from ENNVIH. Note: Poverty line of $4USD a day.

Figure 9 – Decomposition of per capita income into sources of income, using data from Mexico, 2002 and 2005.



Source: Estimates based on data from SEDLAC (CEDLAS and the World Bank). Note: Poverty line of $4USD a day.

Figure 10 – Decomposition of per capita income into sources of income, using CASEN data from Chile, 1996 and 2001.

5 Conclusion

We present a decomposition by welfare's components to capture the contribution of each component on the change of an indicator. The Shapley concept is used to deal with the existence of path dependence that arises when each components is changed at once and sequentially. The decomposition proposed has a statistical profile once that, the counterfactual distributions obtained when changing a component and kept the others constant suffer from equilibrium-inconsistency.

A non-parametric method to transpose the distribution, in absence of panel data, is presented. It relies on the observed rank correlation of the welfare aggregate on the two periods analyzed and utilized the observed correlation between the component and welfare aggregate when transposing the distribution from one period to the other.

Possible data limitations were discussed as well as some applications. The results for poverty and inequality indexes in LAC were presented. Labor income seems to be the main contributor to reduce poverty in most part of the countries. For inequality, non labor income becomes as important as labor income. Additionally, the relevance of each component to the distribution of per capita income was analyzed through the decomposition of the GIC.

In the last section, comparison using anonymous and non anonymous data was made in order to validate the matching based on the rank of the welfare aggregate in case of repeated cross section data. The results were consistent. However, it is important to highlight that when anonymous data is used the interpretation of the results should be different, since we are not tracking individuals but the entire observed distribution of the welfare aggregate and its components.

The attractive aspect of the decomposition is that it provides the contribution of each components variable for the change of an indicator. Therefore, it should be utilized as an accounting exercise in complement to others analysis and economic models.

References

Azevedo, J.P.; Inchauste, G.; Sanfelice, V. *Decomposing the Recent Decline in Income Inequality in Latin America* in VIII Annual Meetings of the LACEA / IADB / WB/ UNDP Research Network on Inequality and Poverty, 2012.

Bourguignon, F. *Non-anonymous growth incidence curves, income mobility and social welfare dominance.* Journal of Economic Inequality 9:605–627, 2011.

Bourguignon, F.; Ferreira, F. H.G.; Leite, P. G. *Beyond Oaxaca-Blinder : Accounting for Differences in Households Income Distributions*. Journal of Economic Inequality, Vol. 6, No.2, 117-148, 2008.

Bourguignon, F.; Ferreira, F. H.G.; Lustig, N. *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, pp. 17-46 Washington, D.C.: The World Bank.

Bruno, M.; Ravallion, M.; Squire, L. *Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues*. In Vito Tanzi and Ke-young Chu, eds., Income Distribution and High-quality Growth. Cambridge, Mass.: MIT Press 1998.

Chen, S.; Ravallion, M. *The Developing World is Poorer Than We Thought, But No Less Successful in the Fight Against Poverty* World Bank Policy Research Working Paper 4703 2008.

Datt, G.; Ravallion, M. *Growth and Redistribution Components of Changes in Poverty Measures: A Decomposition with Applications to Brazil and India in the 1980s*. Journal of Development Economics, 38: 275-296, 1992.

Dollar, D.; Kraay, A. *Growth is good for the poor* Journal of Economic Growth 7: 195-225, 2002.

Essama-Nssah, B. *Identification of Sources of Variation in Poverty Outcomes*, World Bank Policy Research Working Papers, No. 5954, 2012.

Ferreira, F. H.G. *Distributions in Motion: Economic Growth, Inequality and Poverty Dynamics*. World Bank Policy Research Working Paper No. 5424. The World Bank, Washington, D.C, 2010.

Fortin, N.; Lemieux, T.; Firpo, S. *Decomposition Methods in Economics*. In: Ashenfelter Orley and Card David (eds) Handbook of Labor Economics, Vol. 4A , pp. 1-102. Northolland, Amsterdam, 2011.

Fournier, M. *Inequality decomposition by factor component : a "rank-correlation" approach illustrated on the Taiwanese case*, Recherches économiques de Louvain, 2001/4 Vol. 67, p. 381-403, 2001.

Jenkins, S.; van Kerm, P. *Trends in income inequality, pro-poor income growth and income mobility*. Oxf. Econ. Paper 58, 531–548, 2006.

Juhn, C.; Murphy, K. M.; Pierce, B. *Wage Inequality and the Rise of Returns to Skill*. Journal of Political Economy, Vol. 101, No.3: 410-442, 1993.

Mussini, M. *A matrix approach to the Gini index decomposition by subgroup and by income source.* Applied Economics, 45:17, 2457-2468, 2013.

Paes de Barros, R.; Carvalho, M; de Franco, S.; Mendoça, R. *Uma Análise das Principais Causas da Queda Recente na Desigualdade de Renda Brasileira* In: Revista Econômica. Volume 8, número 1, p.117-147. Universidade Federal Fluminense. Rio de Janeiro, 2006.

Paul, S. *Income sources effects on inequality*. Journal of Development Economics, 73: 435-451, 2004.

Ravallion, M. *Growth, Inequality and Poverty: Looking Beyond Averages*. World Development, Vol. 29: No. 11: 1803-1815, 2001.

Ravallion, M.; Chen, S. *China's (Uneven) Progress Against Poverty*. Journal of Development Economics, 82 (1): 1-42, 2007.

Ravallion, M.; Chen, S. *Measuring Pro-Poor Growth*. Economic Letters 78, 93-99, 2003.

Ravallion, M.; Chen, S. What *can new survey data tell us about recent changes in distribution and poverty?* World Bank Economic Review 11 (2): 357-382, 1997.

Sastre, M. and Trannoy, A. *Shapley inequality decomposition by factor components: some methodological issues*. Journal of Economics (in press), 2001.

Shapley, L. *A value for n-person games* in: H. W. Kuhn and A. W. Tucker (eds.), Contributions to the Theory of Games, Vol. 2 (Princeton, N.J.: Princeton University Press), 1953.

Shorrocks, A. F. *Decomposition procedures for distributional analysis: a unified framework based on the Shapley value*. Journal of Economic Inequality, 2012.

Shorrocks, A. F. *Inequality Decomposition by Factor Components.* Econometrica, Vol. 50, No. 1, January, pp. 193-211, 1982.

Shorrocks, A.F.; Kolenikov, S. A *Decomposition Analysis of Regional Poverty in Russia*, Discussion Paper No. 2003/74 United Nations University, 2003.

Yalonetzky, G. *Mobility Decomposition over Several Periods: Theory and Evidence on Consumption Mobility from Peru*. Present in 32nd General Conference of the International Association for Research in Income and Wealth, Boston, 2012.