

The impact of income adjustments in the Casen Survey on the measurement of inequality in Chile

*El impacto de los ajustes de ingresos realizados en la Encuesta CASEN sobre la medición de la desigualdad en Chile**

DAVID BRAVO**

JOSÉ A. VALDERRAMA TORRES***

Abstract

The adjustment of the information obtained from household surveys to make the figures compatible with National Accounts is a non-standard and potentially questionable practice given that it alters the structure of income distribution. This paper analyzes the sensitivity of inequality and poverty indicators to the adjustments made by ECLAC so as to enable a consistency between what is reported by the CASEN survey and the National Accounts figures in Chile. The results reveal that this leads to important changes in the top-end of the distribution and to an overestimation in the main inequality indicators in Chile. Chile looks more unequal in international relative terms due to this adjustment.

Key words: *Inequality, Poverty, Income adjustment, Chile.*

Resumen

La práctica de ajustar los ingresos de datos provenientes de encuestas de hogares para que las cifras que provengan de éstas sean compatibles con las Cuentas Nacionales es una práctica que no sigue estándares internacionales y potencialmente criticable por alterar la estructura de la distribución de ingresos. En este artículo se analiza la sensibilidad de los indicadores de desigualdad y pobreza ante el ajuste realizado por la CEPAL para que exista consistencia

* The authors would like to thank Henry Espinoza, Osvaldo Larrañaga, Gustavo Yamada, Patricia Medrano, Esteban Puentes, and the participants at the Chilean Public Policies Society Meeting. We also thank MIDEPLAN for the information provided and Juan Carlos Feres (ECLAC) for valuable conversations. David Bravo thanks the Millennium Science Initiative (Project P07S-023-F) for financial support. The usual disclaimers apply.

** Centro de Microdatos, Departamento de Economía, Universidad de Chile. E-mail: dbravo@econ.uchile.cl

*** Dirección de Calidad del Gasto. Ministerio de Economía y Finanzas, Perú. E-mail: jvalder@ulima.edu.pe

entre lo reportado por las Encuestas CASEN y las cuentas nacionales en Chile. Los resultados revelan que dichas imputaciones provocan cambios importantes en la cola superior de la distribución, generando con ello una sobreestimación en los principales indicadores de desigualdad de Chile. Esta situación afecta las comparaciones internacionales mostrando a Chile más desigual en términos relativos, por cuanto el ajuste de ingresos no es realizado por las cifras oficiales de otros países.

Palabras clave: *Desigualdad, Pobreza, Ajuste de ingresos, Chile.*

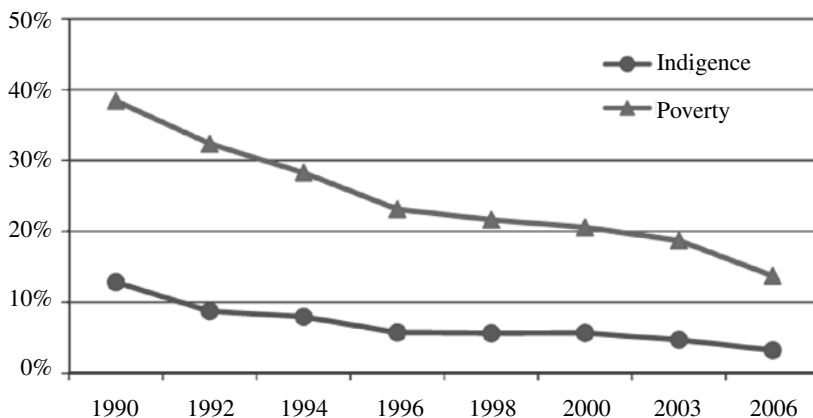
JEL Classification: *C81, D3, I32, N36, O15.*

1. INTRODUCTION

The official figures in Chile reveal an important reduction of poverty, a tendency that can be observed from the beginning of the 90's and in the 1990-2006 period. According to official sources, poverty was at its lowest level in 2006. Extreme poverty or indigence in 1990 was six times higher than poverty observed in 2006, whereas poverty in general more than tripled this figure (see Figure 1).

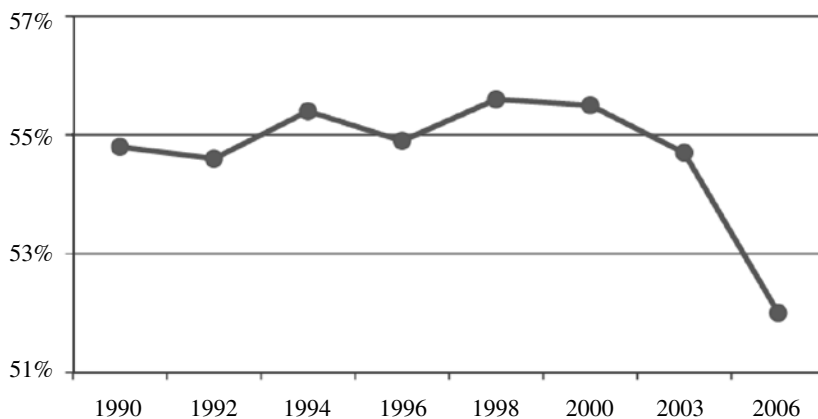
On the other hand, inequality has remained largely unchanged in the period 1990-2003. Only in 2006 there was a statistically significant drop in the indicators (see Figure 2)¹.

FIGURE 1
POVERTY RATE IN CHILE



¹ See Bravo and Contreras (2004), Bravo and Marinovic (1997) and Larrañaga (2001) for an analysis of wage inequality over a longer period of time. They show that inequality has been rather unstable with a high increase in wage inequality between 1974 and 1987.

FIGURE 2
 INCOME INEQUALITY IN CHILE
 (Gini Coefficient)



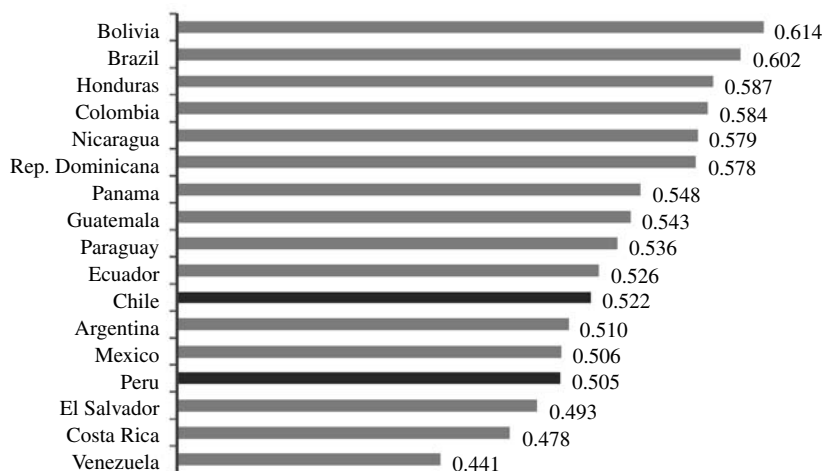
The official source for calculating inequality and poverty indicators in Chile is the National Socioeconomic Characterization Survey (CASEN) while the variable used is household income. Compared with common practices applied elsewhere, the information in Chile differs in that the data is adjusted to make the accumulated amounts consistent with those registered in the National Accounts. This adjustment, trying to correct for under-reporting, affects labor-related income of specific occupational categories, social security receipts and also has an impact on the self-reported implicit rent from own-housing. An additional amount is imputed solely to income recipients in the richest quintile to account for capital income or property income in National Accounts.

It is natural then to question the effect of these adjustments on the official inequality and poverty figures. The only study on this topic we are aware of is Pizzolito (2005). She found that the trends on inequality and poverty for the 1990-2000 period are not changed when ignoring the National Accounts adjustment (she had to use adjustment factors for aggregate items available for 1990 and 1996 instead of 2000). However, she warns that international comparisons for Chile using this data could be affected. Unfortunately, despite requests by researchers over time, the original unadjusted microdata have not been made available by Mideplan.

There are at least two reasons why analysts would prefer to use unadjusted information. First, since adjustment depends on how National Accounts behave over time and the income distribution figures have a component that is sensible to the National Accounts figures and that could distort the real behavior of indicators derived from income distribution, such as inequality and poverty levels.

Second, the practice of imputing this kind of income is precisely one of the reasons why the information is not comparable on a global scale² since this practice is neither recommended as an international standard nor used in other Latin American countries other than Chile. This also questions the veracity of the ECLAC statistics that indicate that inequality in Chile is higher, for example, than Peru, Mexico and Argentina (see Figure 3).

FIGURE 3
INEQUALITY IN LATIN AMERICA MEASURED BY THE GINI COEFFICIENT



Source: ECLAC (2007).

This paper analyzes the impact of adjustments practices based on National Accounts on the income distribution in Chile. Our study deepens Pizzolito (2005) estimating unadjusted data for 2000, 2003 and 2006 and using an array of inequality indicators as well as official poverty lines.

The study is structured in five sections. In the second part we outline the methodological aspects that justify the adjustment of income to National Accounts and discuss its disadvantages. In the third section we look into the sources of information in Chile used to calculate the indicators and compare them to universally accepted methodologies used to calculate these same indicators. In part four we analyze how sensitive Chilean inequality indicators are when they are not based on adjustment to National Accounts. We also analyze Chile with strictly comparable data from Peru. Section five is dedicated to discussion and conclusions. An appendix is also included.

² Other causes that affect international comparisons are how variables are defined (consumption or income) and the scope of the surveys (nationally representative; considering only urban areas, etc.).

2. METHODOLOGICAL ISSUES

Although ideally the microeconomic (household surveys) and macroeconomic (National Accounts) statistics should be related, in practice, the construction of National Accounts figures is not based on information from relevant units (such as households), leading to serious discrepancies in these statistics. Similarly, Ravallion (2001), referring to household consumption expenses, claims that the figures referring to those reported in national accounts are rarely based on household survey results and, in the best of cases, only some components of these are taken into account.

These discrepancies have led the National Accounts System for 1993 to point out that the results of household income surveys should be adjusted to compensate for “certain typical biases” to make them compatible with the National Accounting figures. The Canberra Group (2001) has also pointed out that it is necessary to make these adjustments:

“It undoubtedly causes considerable harm to users when two sets of statistics known as household income produce very different results which will, in turn, have a significant impact on social policies. Despite this need, national statistics offices rarely look to reconcile their results.”

In an attempt to resolve this problem and under the assumption that the National Accounting information offers more trustworthy figures than household surveys, a variety of methodologies have been developed to adjust the information to the National Accounts figures. In any case, the methodologies depend substantially on the assumptions of each author in relation to the rules of allocation that give rise to the distribution of the value of the differences between the total income of households reported by income surveys among the households or groups of households and the National Accounting. The above mentioned allocation rules have been associated with the level of household income, composition by household income sources (wage-earning, self-employment, etc.) or other combinations of the level of household income and its composition according to different sources.

Using the level of total household income as a guide, the simplest way of adjusting consists of “correcting” the figures by multiplying a constant factor corresponding to the ratio of household income for all the households in all the income levels according to National Accounts and the aggregate income of the surveyed households.

The methodology suggested by Altimir (1987) is rather more elaborated since it emphasizes the discrepancies by income source as the cornerstone of the adjustment. Altimir’s proposal, which the Economic Commission For Latin America and the Caribbean (ECLAC) has based its calculations, consists of adjusting the income of every household according to its composition, using specific adjustment factors for every income source; independently of the level of income of the household, except in what corresponds to property income, which has a zero adjustment factor for 80% of the households with less income and higher than one for 20% of richer households. The adjustment factors by source are obtained from dividing the total income reported for every category

of income from the National Accounts with those corresponding to the household survey.

Altimir attributes the differences between the household survey and National Accounts to under-reporting from the survey participant, either voluntary or involuntary, and assumes that this measurement bias is more associated with the type of income than with the level of income. It also supposes that under-reporting for each type of income can be estimated according to the discrepancy between what is reported by the survey and what is reported by the National Accounting figures, with the exception of the cases where the latter is less than the first. If what is reported in the survey is higher than what is shown in National Accounting, and there is no clear evidence of overestimation due to design, then the figures shown by the survey should be accepted as true.

Using the Altimir method, two households with the same level of income might undergo different adjustments in magnitude if the types of income are different. Consequently, this method affects the structure of the income distribution and can increase or decrease inequality and poverty, depending on the composition of household income and on the specific adjustment factors for each income source. What indeed becomes clear is that, when considered as an isolated factor, the treatment of capital income tends to increase inequality, since the adjustment in this category is carried out only in 20% of the richest households.

In addition to the selection of the percentile from which the adjustment factor for this variable (capital income) stops being neutral, other questionable characteristics of this type of adjustment is the fact that its application supposes compliance with the following assumptions:

1. That the incomes informed by the National Accounts are at least as credible as those from household income surveys.
2. That the differences between both sources are fundamentally due to problems of under-reporting (many of the survey participants say they have a lower income than the one they actually do) and not to problems of truncation (the richest households are not surveyed).
3. That there is a rule of ideal allocation that allows to distribute household income, at a macroeconomic level, to the (expanded) income of each household pertaining to the income survey sample (microeconomic level).

The first assumption implies that the figures for household income of the National Accounting system are closer to reality, since they come from a wide variety of sources and that their design makes them necessarily compatible with the rest of the components of the accounting system. On the other hand, the second assumption refers to the fact that household survey participants report less income than they really receive and that this is the only reason behind the discrepancy between the total household income reported by the surveys and the income reported by the National Accounts.

This means that it is assumed that both sources refer exactly to the same population. Nevertheless, in countries where the real inequality in income distribution is so high that an extremely small fraction of population participates with a notably high proportion of total income, it is unlikely that these few people and their incomes are represented in the survey sample. This, because in the

survey samples the probability of selection of households with extraordinarily high income is practically zero and also because the stratification of the sample is based on demographic variables estimated from population censuses and not based on income.

Consequently, if a group of population, small in number but important insofar as income is concerned is not represented in the survey sample, the entire value of their expanded income, even without sub-reporting, must be lower than the one shown in the National Accounting figures, which, because of its methodology and coverage includes in principle the income of all the recipients, without exceptions. When defined this way, the truncation in the top end of income distribution is a characteristic of the sampling that might explain one part of what would be considered as under-reporting.

This implies that if the adjustment to the National Accounts figures is made in such a way that it does not differentiate between the two components of the discrepancy, we would be statistically redistributing a higher quantity of money than the one in question between the sample households. That is to say, we would be artificially distributing the income of households that are so rich that it is highly unlikely that they would appear in the survey (the richest segment) in relation to the rest of the population, including some of the poorest households than do appear in the survey. This statistical correction of the figures, without compensating the real income received by all households, can lead to underestimations when measuring poverty.

Taking into account the above, the compliance with the second assumption is not credible because we cannot distinguish which part corresponds to sub-reporting and which part to truncation. If we were able to distinguish which part of the discrepancy corresponds solely to sub-reporting, the adjustment according to the National Accounts would consist of re-assigning only this part among the households. Nevertheless, the distinction between the amount stated in sub-reports and the amount that corresponds to the difference for truncation is not negligible and it would need at least a sample representative of those who have a higher income.

Finally, regarding the suppositions that refer to the allocation rule, by design all the available allocation rules are subjective. That is why any income distribution resulting from an adjustment to National Accounts figures is only a probable distribution whose verisimilitude depends on the validity of the assumptions that were initially chosen to make the microeconomic allocation of the macroeconomic discrepancy.

Measurement error is a serious concern when using micro-level cross-sectional or longitudinal survey data. Therefore, a National Accounts adjustment like the one used in Chile could easily introduce additional noise to the original micro-data when unadjusted records are not made available. To face mis-reporting in surveys one could develop specific studies to understand the direction and size of the biases (see Angrist and Krueger, 1999).

3. DATA

The Socio-Economic Characterization Survey (CASEN) is the principal survey used in Chile to diagnose and assess the incidence of social policies in

households and the most important programs that constitute social spending³. It is a multi-purpose survey that provides information about the socioeconomic conditions of the country's different social sectors, its most important deficiencies, the dimensions and characteristics of poverty and income distribution of the households. Additionally, the survey reports on the coverage and profile of the beneficiaries of social programs, their monetary and non-monetary contributions to household incomes which are identical to social sectors that do not have access to the above-mentioned programs, which makes it possible to calculate the related assistance shortfalls. Such information guides the design of new projects and any modifications in benefit allocation systems to improve the focus of those that have a more selective character.

The surveys have been implemented in the following years: 1985, 1987, 1990, 1992, 1994, 1996, 1998, 2000, 2003, 2006 and 2009. This is a cross-sectional survey implemented regularly in November by the Ministry of Planning (MIDEPLAN). MIDEPLAN has usually hired the University of Chile for the implementation of the fieldwork and once the data has been collected the Economic Commission for Latin America and the Caribbean (ECLAC) is responsible for making the adjustments for response errors.

4. ASSESSING THE IMPACT OF THE INCOME ADJUSTMENTS IN THE CASEN SURVEY

4.1. The process of income adjustment as carried out by ECLAC

The adjustments made by ECLAC to the survey collected data come from two sources: non-response and income misreporting. To deal with the second problem, information from National Accounts provided by the Central Bank of Chile is used and, in particular, an estimation of the principal aggregates of the income and spending accounts of households prepared specially for this task. The imputation process is shown in Figure 4.

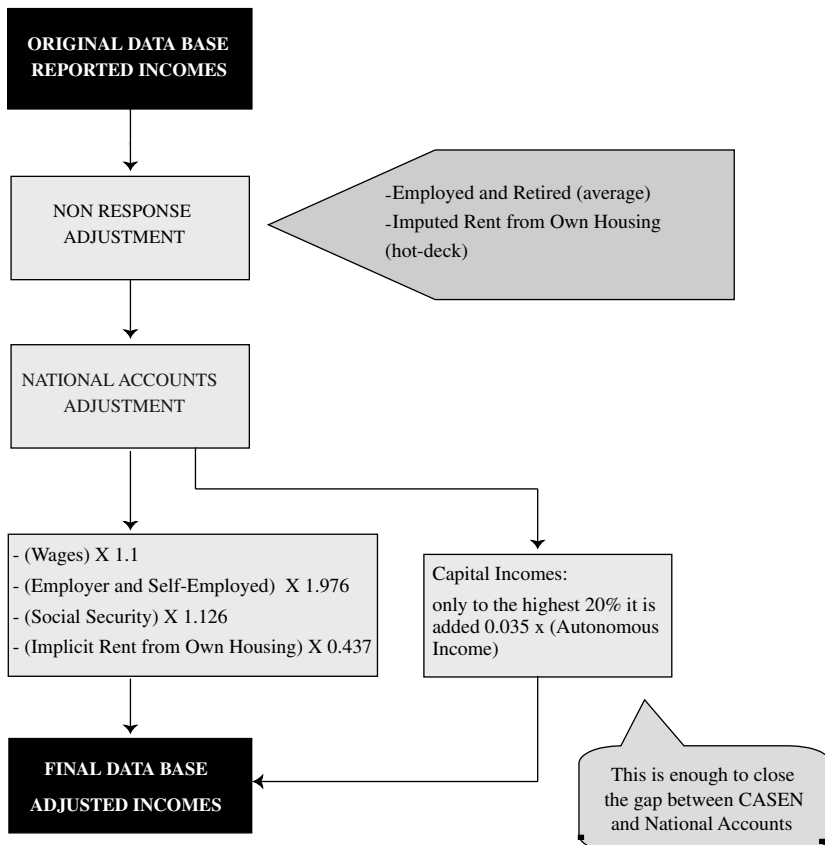
The first stage of the adjustment process is the correction of information that has been omitted, that is to say, those people who have declared that they receive some type of income, but that have not declared the amount or the corresponding total. The correction process involves assigning some type of response to the group of people who say they receive a certain income, but do not assign any values.

Three groups are considered in this process:

1. People who declare they are employed in a category different to a non-paid family member and who do not report the income received from a main occupation.
2. People who declare they are retired senior citizens or who receive pensions and who do not report this income.
3. Households that inhabit the homes that they own and do not declare the imputed rent.

³ This section is based on the methodological reports of the surveys carried out in 2003 and 2006 (Mideplan, 2005 and Mideplan 2006).

FIGURE 4
 CASEN 2006. INCOME ADJUSTMENT PROCESS CARRIED OUT BY ECLAC



The adjustment process for the employed and retired persons is carried out according to the method of averages. According to this method populations with similar characteristics to persons who have not provided any answers are selected. The average income of this group is imputed to persons who did not provide any answers. In the case of housing (Income for Imputed Rent), the Hot Deck methodology is used where households are selected according to their housing type and situation. In this case it is also necessary to correct the cases where a value for imputed rent is declared but the person is not the owner of the house.

According to this method, in every group obtained according to the characteristics of the housing type and situation, those households that provided no response have the same distribution as those that did respond.

This means that the selected variables must be related to the variable that one wants to impute –that’s why they must be variables related to the price of the house rental.

Once the base is corrected according to the non-declaration of certain incomes, it is possible to make the corresponding adjustments according to the National Accounts figures. The income is multiplied by a certain factor, so that the income figures obtained by CASEN are compatible with the information for the whole country delivered in the National Accounts System⁴.

According to the type of income, adjustment is applied to variables related to the following:

- Salaries and wages
- Income of the employer and self-employed
- Social security benefits
- Imputed rent
- Property income

For all categories but the last, the adjustment is applied directly, and the only requirement is to know which factor was used and the variables related to each type of income. In the case of property/capital income the adjustment involves calculating the total capital income of the survey (rentals, interests and dividends) and the discrepancy between what is reported by CASEN and the National Accounts figures (always in per capita terms) is attributed to all the recipients of autonomous income belonging to the last quintile, in such a way that this gap is distributed proportionally to the received autonomous income.

The autonomous income considered for this purpose is the one that is previously adjusted in all its components, including only the capital income that was declared. Finally, the additional income assigned as capital income imputed to people in the last quintile, corrects the value of the autonomous income. For example, for the 2006 figures, the recipients of income from the last quintile were imputed an additional income under the concept of capital income for the amount of 0.035 times the autonomous income registered in the survey.

Table 1 shows the different values of the adjustment factors according to the type of income. Two characteristics stand out. First, the variability over the years and second, with the exception of the imputed rent, all other variables are underestimated by the people surveyed, which is why the adjustment factor exceeds the unit. (In the case of property income, the adjustment is additive, which explains why this factor is less than one).

⁴ The comparison is carried out at an average level of per capita income, expressed in relation to the entire population of the country, because in National Accounts the number of recipients of each income type is unknown.

TABLE 1
 CASEN, 1990-2006 ADJUSTMENT FACTORS

	Wages and Salaries	Income from Self- employment	Social Security Benefits	Property/ Capital Income	Imputed Rent
1990	1.208	1.980	1.473	0.129	0.664
1992	1.071	1.992	1.633	0.067	0.548
1994	1.071	1.513	1.435	0.064	0.475
1996	0.990	2.043	1.398	0.064	0.454
1998	1.004	1.955	1.347	0.069	0.439
2000	0.957	1.826	1.471	0.054	0.449
2003	1.000	1.976	1.145	0.028	0.437
2006	1.010	1.976	1.126	0.035	0.437

Source: Mideplan.

4.2. Results for 2006

In this section we will review the 2006 CASEN survey to understand the impact of the mis-reporting imputations on income distribution⁵. We are particularly interested in investigating if the progressive character that the imputations for the implicit rent from own-housing would sufficiently compensate the regressive effects of all the other adjustments. As we have seen, in the first one there is an overestimation in what is declared and in the last there is under-reporting, which is why the net effect on the distribution is uncertain.

Additionally, by decomposing the figures from these different sources, we can assess the relevance of each type of imputation when explaining total inequality.

4.2.1. Comparison between indicators with and without adjustment

Table 2 shows the sensitivity of the per capita household income distribution to the adjustments to National Accounts figures. The results of the Gini coefficient indicate that the adjustment overstates this inequality index in 3.4 points (a 7%) for the total income. Nevertheless, as was expected, the most significant changes occur in the tail-end of the distributions, and reach an increase in inequality of up to 22% (as in the case of the decile ratio in the total income). The only indicator that reflects a descent in inequality as a consequence of the adjustment to the national accounts figures is the ratio of income deciles of the main occupation. This can be explained by the fact that self-employed workers are over-represented in the first 10%, and they are precisely the people for whom income is almost duplicated. This is why the ratio decreases after the adjustment as compared with the ratio without the adjustment.

⁵ All the estimates for 2006 were also applied to 2003, with similar results.

In the case of the poverty indicators, the extreme poverty rate changes from 3.2% with adjustment to 2.9% without adjustment; and the poverty rate would fall from 13.7% to 13.1% undoing the income adjustment.

TABLE 2
CASEN 2006. IMPACT OF THE NATIONAL ACCOUNTS INCOME ADJUSTMENT

	With adjustment	Without adjustment	Variation	% variation
Average income (\$)				
Main				
Employment	121,644	98,666	22,978	23%
Autonomous	166,556	135,733	30,823	23%
Total	176,981	157,073	19,908	13%
Gini coefficient				
Main				
Employment	0.579	0.555	0.024	4%
Autonomous	0.543	0.520	0.023	4%
Total	0.522	0.488	0.034	7%
Coefficient of variation				
Main				
Employment	1.765	1.530	0.234	15%
Autonomous	1.801	1.701	0.100	6%
Total	1.719	1.543	0.176	11%
Theil index				
Main				
Employment	0.687	0.609	0.078	13%
Autonomous	0.615	0.553	0.062	11%
Total	0.569	0.483	0.086	18%
Decile 10/decile 1				
Main				
Employment	3533	5232	-1699	-32%
Autonomous	41	38	3	9%
Total	29	24	5	22%
Poverty				
Indigence Rate	3.2%	2.9%	0.003	10%
Poverty Rate	13.7%	13.1%	0.006	4%

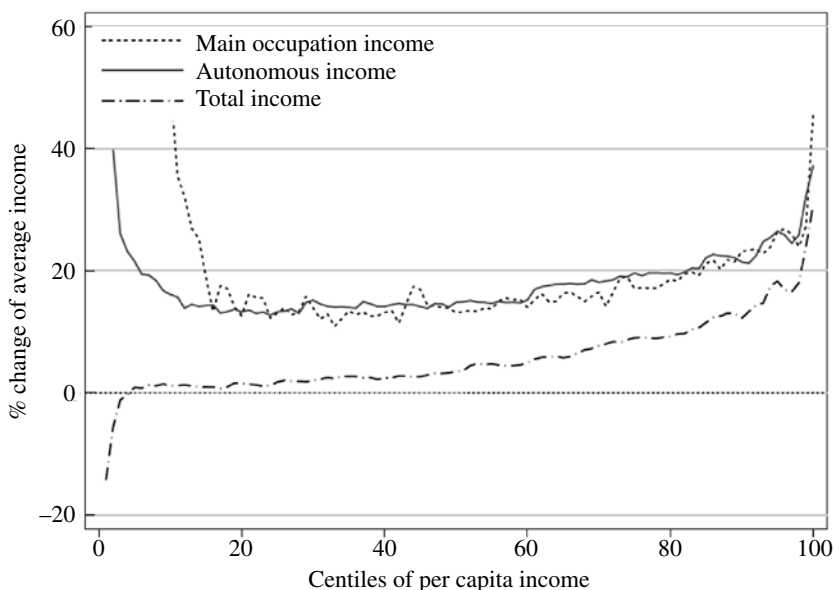
Note: All income is expressed in per capita terms. The calculations exclude income of domestic workers who live in their employers' house. The table includes the possible zeros that might be present in some households. For the case of poverty indicators, total family income per capita has been employed.

4.2.2. Changes along the distribution

As can be seen in Figure 5, there is an evident distortion at the tail-end of the distribution due to the adjustment. In the case of income from the main occupation and autonomous income (labor-related and non-labor related income), there is always an overestimation, especially in the lower end up to the 20 percentile and in the upper end from the 80 percentile onwards. In both cases the overestimation is higher than 20%.

The picture for total income (autonomous income plus monetary transfers plus imputed rent) is somewhat different: in the first centiles the National Accounts Adjustment decreases the income; up to the 40th centile there is no significant change made by this adjustment; however starting from this centile the adjustment increases the income reaching values near 30%. From that we can conclude that the net effect of the imputation on the variable is a reduction in the income of the poorest (up to the 5th percentile of distribution), no alteration to the income of households located in the lower-middle part of the distribution curve (between the 5th and 40th centiles) and an increase in the income of those who earn more (centile 40 and upwards).

FIGURE 5
PERCENTAGE OF DISCREPANCY BETWEEN DISTRIBUTIONS
WITH AND WITHOUT ADJUSTMENT



Note: The calculations exclude income of domestic workers who live in their employers' house. Included are the possible zeros that might be present in some households.

If we consider that total income incorporates the imputed rent which is decreased with the adjustment and other incomes that are increased with the adjustment, there are other conclusions to be made. In the lower part of the total income distribution predominates the effect of the downward adjustments, probably explained by over-reporting of the value of the imputed rent by such households; whereas in the high part of the distribution the same effect is exceeded by other imputations of other variables, especially of the property income category, that, as we have previously stated, is imputed only to people from the last quintile of income.

An additional detail from the graph that stands out is the “U” form shown by the autonomous income discrepancy curve and the main occupation income discrepancy curve. This can be explained by the aforementioned over-presentation of self-employed workers in the first centiles who, as we have seen, see their incomes nearly doubled by the adjustment whereas the income of wage earners scarcely changes.

4.2.3. *Impact according to the significance of each component of total income*

The progressive or regressive character and the significance that each income source has in income inequality can be formalized by the decomposition proposed by Shorrocks (1982).

According to this proposal and considering income as “Y” and its components are expressed generically by Y_f , where $Y = \sum Y_f$, an indicator of the contribution of each component to inequality is given by:

$$S_f = \rho_f * \frac{\sigma_{Y_f}}{\sigma_Y}$$

Where ρ_f is the correlation coefficient between factor Y_f , and total income “Y” and σ denotes the standard deviation. Similarly, S_f is the regression slope of Y_f on total income “Y”, where it is easy to show that $\sum S_f = 1$. Components with a positive value for S_f have a de-equalizing contribution meanwhile components with a negative value have an equalizing contribution⁶.

Considering the issue at hand, the exercise would be to establish the sign and the size of the contribution of the different types of imputations starting from the following identity:

$$Y_f = Y_o + Dif_1 + Dif_2 + Dif_3 + Dif_4 + Dif_5$$

⁶ For more details of the decomposition, see Shorrocks (1980). The appendix outlines the decomposition for two sources of income.

where Y_f and Y_o are the total income with all the adjustments for imputation applied and the income without any adjustment according to National Accounts figures, respectively. Other components are the changes in income caused by the different adjustments that were made which were grouped in the following five components:

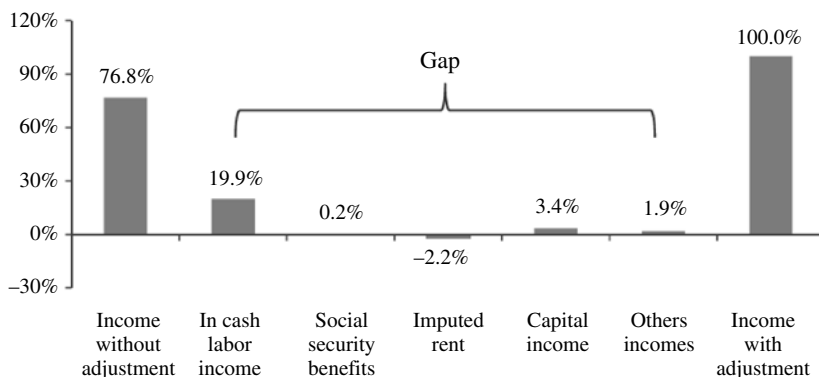
- Dif_1 : increase of labor-related income paid in cash.
- Dif_2 : increase of income due to adjustments to social security benefits.
- Dif_3 : decrease of income due to imputed rent.
- Dif_4 : increase of income due to property and capital income.
- Dif_5 : increase of income in other categories (self-consumption, previous work, etc.).

The results confirm the progressive character of the adjustment due to the rent imputed with an importance in total inequality of -2.2% .

By far the most important component is the adjustment made to labor-related income paid in cash (19.9%) followed by the adjustment for property income (3.4%), imputed rent (-2.2%) and finally other incomes with 1.9% and social security with 0.2% (see Figure 6).

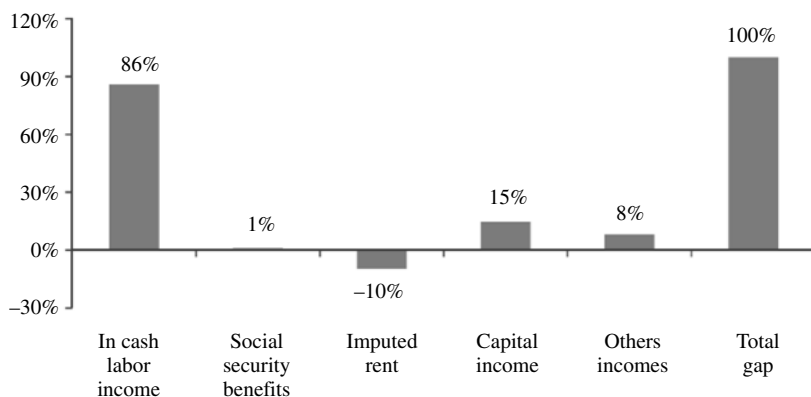
Similarly, if we only take into account the gap existing between income without adjustment and income with adjustment, we can see that imputed rent makes a negative 10% contribution whereas in cash labor-related income is the most important component in with 86% of participation (Figure 7).

FIGURE 6
IMPORTANCE OF THE COMPONENTS IN TOTAL INCOME FOR 2006
Shorrocks' Decomposition



Note: Other income includes self-consumption, withdrawal of profits, previous work and other sources of income.

FIGURE 7
IMPORTANCE OF THE IMPUTATIONS IN THE 2006 GAP.
Shorrocks' Decomposition



Note: Other income includes self-consumption, withdrawal of profits, previous work and other income.

4.3. Effects of adjustment over time

Considering that the least important components when explaining inequality are social security and “other incomes” and that we cannot recover the databases for some of the years and the impact of other variables on the inequality of total income, we have excluded these two factors when carrying out an analysis for the trend in the 1990-2006 period

Figures 8, 9 and 10 show this analysis. The comparison between the indicators shows that, with the exception of the Gini Coefficient which keeps a constant difference (7% on average), all other changes are not systematic and that the average income shows the most important changes (29% in 1990 and 6% in 2000).

In the case of poverty, there is a heterogeneous behavior: before 1996 the effect of imputations on both indicators without adjustment is an underestimation that reaches its highest value in 1990 for the case of poverty, a year in which this indicator is underestimated by about 25%. After this year the indicator is overestimated reaching its highest value in 2000 for the indigent people category with 15%. In short, not having made any imputation would have reflected a bigger reduction in the levels of poverty that those presently known.

On the other hand, Figure 10 shows that the component of income without adjustment has an increasing participation when explaining total inequality, going from 63% of total inequality in 1990 to 79% in 2006. At the same time, the importance of the imputations made in labor-related income has decreased moving from accounting for 27% in 1990 to 19% in 2006.

FIGURE 8
INEQUALITY AND AVERAGE INCOME, RATIO BETWEEN ADJUSTED
VARIABLES AND THOSE WITHOUT ADJUSTMENT

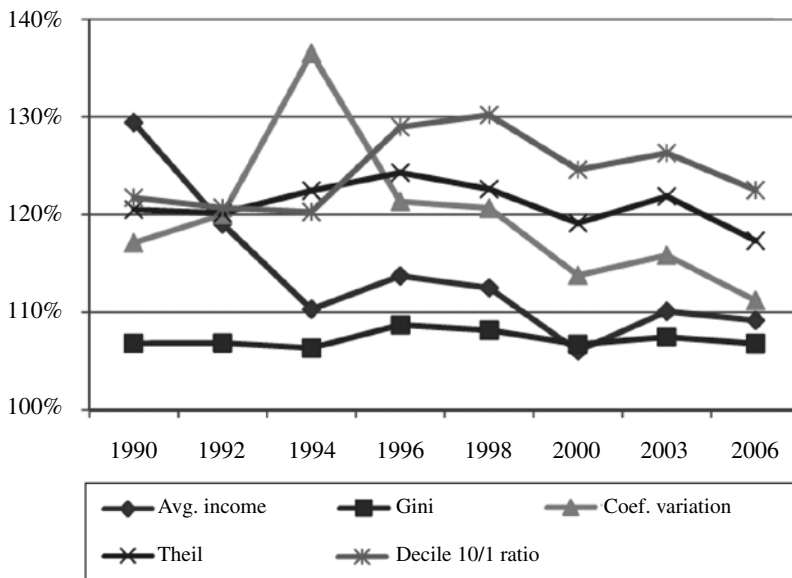


FIGURE 9
POVERTY, RATIO BETWEEN ADJUSTED AND UNADJUSTED VARIABLES

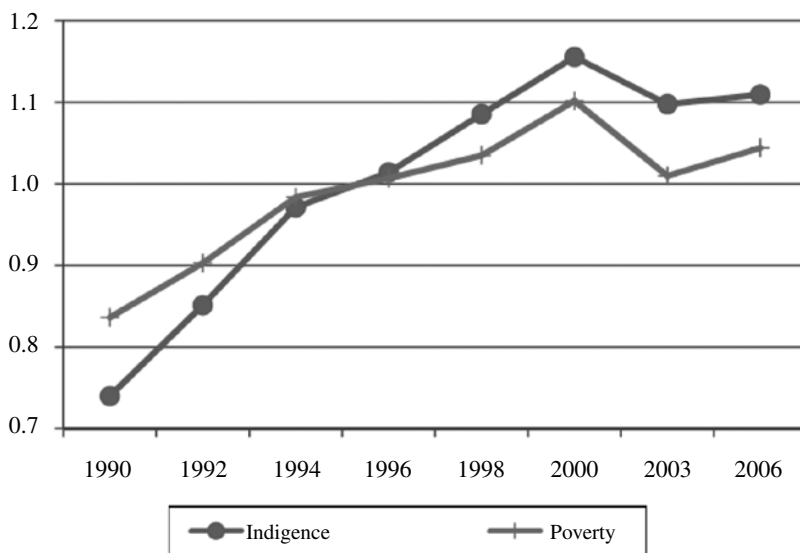
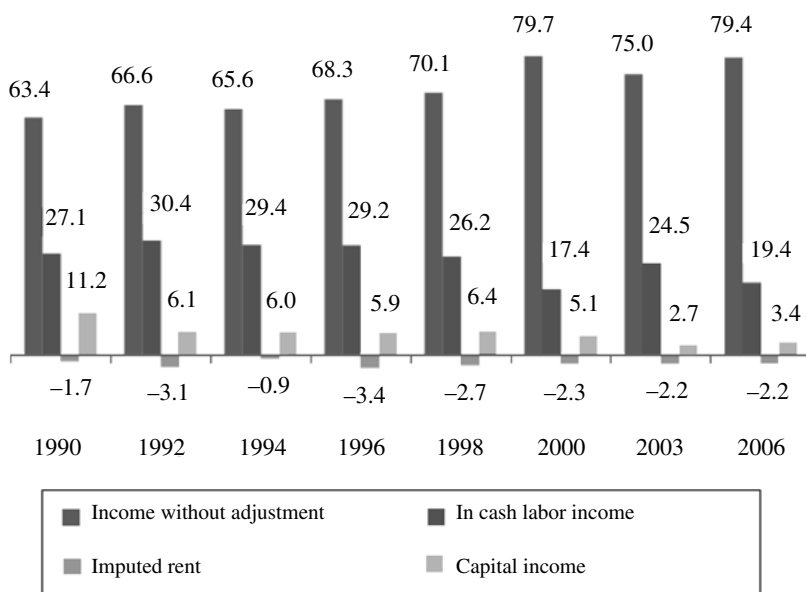


FIGURE 10
IMPORTANCE OF THE COMPONENTS FOR TOTAL INCOME IN THE 1990-2006 PERIOD
Shorrocks' Decomposition



4.4. Comparative analysis between distributions in Peru and Chile

This section compares the distributions of income in Peru and Chile using the same methodology, that is to say using per capita household income⁷ –adjusted only for non-response– as an analysis variable.

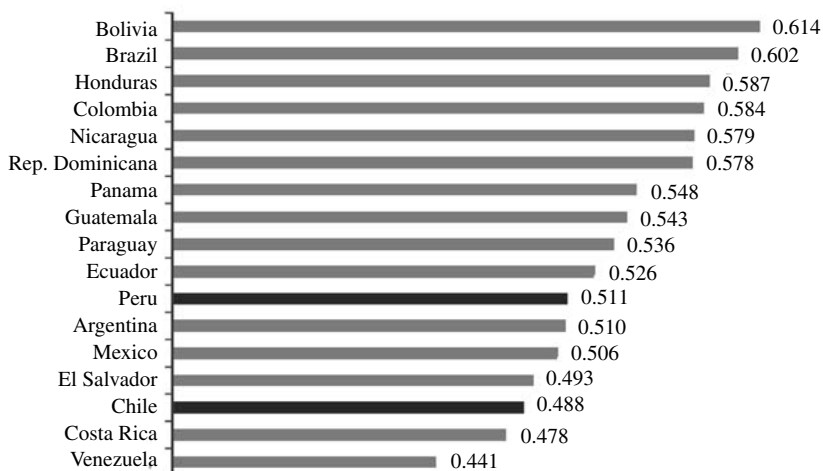
The most important results appear in the area of inequality.

Assuming that the order of the other countries in Figure 3 does not change, Chile becomes one of the countries with lower inequality in the region, overtaken only by Venezuela and Costa Rica, whereas Peru increases its level of inequality moving from a country with a low level of inequality to one that can be considered in the middle range of the regional ranking. Due to the fact that the same methodology is used for both countries, Chile and Peru exchange positions in relation to the ranking presented by ECLAC in its statistical 2007 yearbook (see Figures 3 and 11).

Table 3 shows the levels of inequality according to the sources of total income. From the table we can draw the following conclusions: it is clear that inequality of income received by self-employed workers is higher in Chile than in Peru. The situation is the opposite for wages. When we consider the net effect of both sources we see that income inequality from the main occupation is lower in Chile than in Peru.

⁷ In Peru the official information on poverty and inequality are estimated from the National Household surveys (ENAH) using the per capita consumption as the variable of analysis; nevertheless the survey also allows the income per capita to be calculated.

FIGURE 11
INEQUALITY IN LATIN AMERICA MEASURED BY THE GINI COEFFICIENT



Source: CEPAL 2007, ENAHO 2006 and CASEN 2006.

The component “others” that brings together all non-labor related income generated by the household, has a more unequal distribution in Chile than in Peru; whereas autonomous income, which is the sum of the main occupation income plus the “rest” displays a similar behavior in both countries.

Finally, considering the total income, both the Gini and the Theil coefficients indicate that inequality is higher in Peru, whereas the Variation Coefficient, more sensitive to the top-end of distribution, indicates the contrary.

TABLE 3
POVERTY AND INEQUALITY INDICATORS FOR PERU AND CHILE IN 2006

Indicator	Peru (A)	Chile (B)	Difference (B) – (A)
Extreme poverty:			
FGT0	19.8%	3.1%	-0.17
FGT1	6.8%	1.0%	-0.06
FGT2	3.2%	0.6%	-0.03
Total poverty:			
FGT0	43.4%	13.8%	-0.30
FGT1	17.2%	4.3%	-0.13
FGT2	9.1%	2.1%	-0.07

Indicator	Peru (A)	Chile (B)	Difference (B) – (A)
Gini coefficient			
–Self-employment Income	0.74	0.89	0.15
–Salaries	0.73	0.61	–0.13
Main Occupation Income	0.59	0.56	–0.03
+Rest	0.61	0.74	0.13
Autonomous Income	0.52	0.52	0.00
+Subsidies	0.70	0.82	0.12
+Imputed rent	0.72	0.61	–0.10
Total Income	0.511	0.488	–0.02
Coefficient of Variation:			
–Self-employment Income	2.57	4.01	1.44
–Salaries	2.20	1.64	–0.56
Main Occupation Income	1.65	1.54	–0.11
+Rest	2.19	4.03	1.84
Autonomous Income	1.52	1.71	0.19
+Subsidies	4.31	2.63	–1.68
+Imputed rent	2.42	1.54	–0.88
Total Income	1.480	1.550	0.07
Theil index:			
–Self-employment Income	1.17	2.03	0.86
–Salaries	1.11	0.73	–0.39
Main Occupation Income	0.67	0.61	–0.06
+Rest	0.80	1.27	0.47
Autonomous Income	0.53	0.56	0.02
+Subsidies	1.11	1.52	0.41
+Imputed rent	1.09	0.73	–0.36
Total Income	0.515	0.487	–0.03

Note: FGT is the Foster-Greer-Thorbecke metric, a generalized measure of poverty. If z : poverty line; N : the number of people in the country; H : the number of poor (those with incomes at or below z); y_i : individual incomes; α = a sensitivity parameter. Then:

$$FGT_{\alpha} = \frac{1}{N} \sum_{i=1}^H \left(\frac{z - y_i}{z} \right)^{\alpha}$$

FGT0: FGT if $\alpha = 0$ (the Headcount ratio or poverty rate used in Chile).

FGT1: FGT if $\alpha = 1$ (the average poverty gap measuring intensity of poverty).

FGT2: FGT if $\alpha = 2$ (an index that combines information on both poverty and income inequality among the poor).

See Foster, Greer and Thorbecke (1984).

5. DISCUSSION AND CONCLUSIONS

This paper has shown that inequality indicators and poverty in Chile are overestimated by virtue of the imputations for adjustment to National Accounts figures. In the case of inequality, the overestimation is nearly 22% (as in the case of the ratio between the deciles in 2006), whereas the most well-known indicator, the Gini coefficient, is overestimated in about 7%, an overestimation that is a constant in the period 1990-2006.

In the case of poverty, both extreme and non-extreme, the overestimation is about 10% and 4% respectively, which can be explained by the downward adjustment of the value of imputed rent, the only variable that has an imputation in this direction (and that according to the results would have a higher net impact on the households with a lower income). Additionally, poverty is the only characteristic that has displayed a heterogeneous behavior in the 1990-2006 period: before 1996, when the levels of poverty were higher than the ones observed in the last few years, the levels of poverty were underestimated and after 1996 they are overestimated.

This allows us to conclude that if there had been no adjustment applied to income, the reduction of the poverty in Chile in the period under analysis would have been more pronounced. The average income was overestimated in each year analyzed and this variable experienced the most important temporary changes, given that it moved from an overestimation of 29% in 1990 to 9% in 2006.

Once the structure of the imputations according to national accounts figures was known, the CASEN figures were deconstructed and the same methodology was applied both to Peru and Chile so as to obtain poverty and inequality indicators that would enable an adequate international comparison. The results indicate that income inequality is higher in Peru than in Chile.

It is important to point out that although we can recover the original information, when and if all the adjustment factors and variables involved in the adjustment are known (which does not happen in certain years), for the sake of transparency and efficiency, since the process of recovery is not negligible, the official databases should include the non-adjusted variables. This is especially important when making international comparisons, since the practice of income adjustment is not universal.

Furthermore, the advantage of congruence between the micro and the macro accounts can be overlooked for at least two reasons: First, changes in the social indicators can reflect the adjustments applied and not necessarily the real dynamics of the socioeconomic indicators. Second, the fact that the National Accounts figures and the Household Surveys refer to different universes since the sample of people with a higher income is typically underestimated in any conventional household survey (truncation in the top-end of distribution), makes adjustment between the macro and the micro accounts necessary. Since this requires specific knowledge of what part of the discrepancy of the Household Survey with the National Accounts corresponds to sub-reporting and what part corresponds to truncation and since this is not known, the allocation to the households in the survey of the total discrepancy leads to an over-correction of income in all or part of the households surveyed.

APPENDIX

SHORROCKS' DECOMPOSITION:

In its simplest version (which can be generalized more easily) when the total income (Y^T) has two sources, such as ($Y^T = Y^1 + Y^2$), we can deduce that:

$$\text{Var}(Y^T) = \text{Var}(Y^1) + \text{Var}(Y^2) + 2\text{Cov}(Y^1, Y^2)$$

$$\text{Var}(Y^T) = \text{Var}(Y^1) + \text{Cov}(Y^1, Y^2) + \text{Var}(Y^2) + \text{Cov}(Y^1, Y^2)$$

$$\text{Var}(Y^T) = \text{Cov}(Y^1, Y^T) + \text{Cov}(Y^2, Y^T)$$

$$1 = [\text{Cov}(Y^1, Y^T) / \text{Var}(Y^T)] + [\text{Cov}(Y^2, Y^T) / \text{Var}(Y^T)]$$

$$1 = S_y^1 + S_y^2$$

REFERENCES

- Altimir, O. (1987). "Income Distribution Statistics in Latin America and their Reliability". *Review of Income and Wealth*, Series 33, N° 2, June.
- Angrist, J. and A. Krueger (1999). "Empirical Strategies in Labor Economics", in O. Ashenfelter and D. Card, *Handbook of Labor Economics*, Volume 3, Chapter 23. Elsevier Science B.V.
- Bravo, D. and A. Marinovic (1997). "Wage Inequality in Chile: 40 Years of Evidence". Departamento de Economía, Universidad de Chile, August.
- Bravo, D. and D. Contreras (2004). "La distribución del ingreso en Chile 1990-1996: análisis del impacto del mercado de trabajo y las políticas sociales", in Banco Interamericano de Desarrollo, *Reformas y Equidad Social en América Latina y el Caribe*, abril.
- Canberra Group Expert on Household Income Statistics (2001). *The Canberra Group, Final Report and Recommendations* (Ottawa).
- Chez-Crespo, S. (1972). "El tratamiento de preguntas de carácter íntimo: modelo de respuesta aleatorizada". *Revista Estadística Española*, Instituto Nacional de Estadística N° 55. Período: Abr.-Jun.-72.
- Commission of the European Communities, International Monetary Fund, OECD, United Nations and World Bank (1993). *System of National Accounts 1993*, Brussels/Luxemburg, New York, Paris, Washington D.C.
- ECLAC (2007). *Anuario Estadístico de América Latina y el Caribe*. Comisión Económica para América Latina y el Caribe.
- Foster, J., J. Greer and E. Thorbecke (1984). "A class of decomposable poverty measures". *Econometrica* 52, 761-776.
- Greenberg, B., R. Kuebler Jr., J. Abernathy and D. Horvitz (1971). "Application of the Randomized Response Technique in Obtaining Quantitative Data".

- Journal of the American Statistical Association*, vol. 66, núm. 334, 243-50.
- INEI (2004). Ficha Técnica ENAHO 2003, INEI, 1-18.
- Larrañaga, O. (2001). "Distribución de Ingresos en Chile: 1958-2001". Documento de Trabajo N° 178, Departamento de Economía, Universidad de Chile.
- Mideplan (2005). "Marco metodológico CASEN 2003", Mideplan, 1-25, (2005).
- Mideplan (2006). "Marco metodológico CASEN 2006", Mideplan, 1-97, (2006).
- Ministerio de Economía y Producción (MEP) de Argentina (2007). Comunicado de prensa: Distribución Funcional del Ingreso Cuenta de Generación del Ingreso e Insumo de mano de obra. Anexo B Correcciones de ingresos a la Encuesta Permanente de Hogares, 2007.
- Pizzolito, G. (2005). "Poverty and Inequality in Chile: Methodological Issues and a Literature Review". *Centro de Estudios Distributivos, Laborales y Sociales* 20: 1-30.
- Ravallion, M. (2001). "Measuring Aggregate Welfare in Developing Countries: How well do National Accounts and Surveys Agree?". *Working Paper*, The World Bank, Washington, D.C.
- Roca, J. and M. Hernández (2004). "Evasión Tributaria e Informalidad en el Perú: una Aproximación a Partir del Enfoque de Discrepancias en el Consumo". CIES.
- Shorrocks, A.F. (1980). "The class of additively decomposable inequality measures", *Econometrica* 48 (1): 613-625.
- Shorrocks, A.F. (1982). "Inequality decomposition by factor components", *Econometrica* 50 (1): 193-211.