

# Improving the Measurement of the Relationship between Opportunities and Income: Evidence from Longitudinal Data from Chile

Dante Contreras, Osvaldo Larrañaga, Esteban Puentes and Tomás Rau\*

*The relationship between exogenous circumstances faced in early childhood (opportunities) and earnings has been increasingly studied in the past decade. In this article we assess the sensitivity of this relationship to different income measures. Typically the empirical evidence is based on short-term income measures that suffer from variability and measurement error. Using longitudinal data from Chile, we find that when four- and seven-year earnings are used, the relationship between inequality of opportunity and income inequality is significantly higher than that obtained with yearly measures. Monte Carlo simulations with several data-generating processes confirm this result. This supports policies targeted to reduce long-term income inequality via providing equal opportunity to individuals at early stages in life.*

**Key words:** Earning inequality, inequality of opportunity

## 1 Introduction

The ideal of equality of opportunity appeals to many people. A society that achieves equality of opportunity is seen as one that has successfully combined fairness and economic efficiency. Nevertheless, monitoring and evaluating the distribution of opportunity has been hindered in the past by lack of suitable indicators. However, this situation has started to change. Following Roemer's (1998) approach that classifies how socio-economic outcomes are related to exogenous circumstances and choices, a variety of methodologies have been developed to measure the distribution of opportunity (see, for example, Bourguignon et al., 2007; Ferreira and Gignoux, 2011; Checchi and Peragine, 2005). Essentially, they estimate the impact of family-background and other circumstance variables on outcome or advantage variables.

---

\*Respectively, Departamento de Economía y Centro de Microdatos, Universidad de Chile, (contreras.dante@gmail.com); PNUD (Programa de las Naciones Unidas para el Desarrollo); Departamento de Economía y Centro de Microdatos, Universidad de Chile and Instituto de Economía, Pontificia Universidad Católica de Chile. They would like to thank the participants of the Chilean Economic Society meetings and workshops at Universidad de Chile and Universidad Diego Portales, and especially Claudia Martínez, Patricia Medrano, José Molinas, Jaime Ruiz-Tagle and two anonymous referees. Dante Contreras, Esteban Puentes, and Tomás Rau are grateful for financing provided by the Millennium Science Initiative from the Ministry of Economy, Development and Tourism to Microdata Center, Project NS100041. Dante Contreras also acknowledges funding from the Centre for Social Conflict and Cohesion Studies (CONICYT/FONDAP/15130009).

A recent study by the World Bank uses these methodologies with a sample of seven Latin American countries, deriving a set of estimates that allows the ranking of these countries according to how opportunity is distributed equally or unequally (Paes de Barros et al., 2009).<sup>1</sup> This can be done by means of a so-called top-down approach, which relates current socio-economic outcomes to past circumstances, or a bottom-up approach that assesses how opportunity is distributed among children and young people.

In this article we investigate the sensitivity of inequality of opportunity measures to long-term earnings data and its relationship with income inequality. We estimate the share of inequality of opportunity in income inequality in the Chilean economy using a top-down approach focused on the distribution of individual labour earnings. A companion paper addresses the bottom-up approach (Contreras et al., 2012).

Recent evidence from Latin America has shown that inequality of opportunity explains a significant fraction of total income inequality, and its reduction would be a significant factor in improving welfare. However, previous research has been implemented using monthly measures of earnings that are subject to measurement error and seasonality. This is particularly problematic, given the high percentage of the population working in the informal labour market, who are subject to significant fluctuations in income and employment opportunities, segregation in the labour market, and other issues (Perry et al., 2007). At the same time, yearly income reduces seasonality, but is still affected by measurement error and the persistence of income shocks. We consider that welfare indicators should be computed using more precise estimates of permanent income. In a country like Chile with high levels of income inequality, it is particularly important to measure the level of opportunity inequality properly, especially when considering how to improve living conditions for the majority of the population.

This article contributes to the previous literature by examining several income measures for the Chilean economy. We take advantage of a longitudinal survey available for Chile and construct four-year and seven-year earnings averages, which are a more precise estimate of permanent income than traditional monthly income and yearly data.<sup>2</sup> In addition, given the low level of female labour-force participation, we focus our analysis only on men.<sup>3</sup> We find that income inequality decreases when using longer-term measures. At the same time, the level of inequality of opportunities rises slightly, which implies that the share of inequality of opportunities increases as well, from 16% when using monthly and yearly data to 20% when using seven-year data. These differences show that short-term income measures can underestimate by 25% the level of inequality of opportunities. This is the first estimation in the literature that proves the importance of having longer-term measures to assess accurately the relationship between income inequality and inequality of opportunities.

- 
1. The seven countries analysed are Brazil, Colombia, Ecuador, Guatemala, Mexico, Panama and Peru.
  2. Ferreira and Gignoux (2011) mostly use monthly data to measure wage earnings, apart from self-employment in Ecuador and Guatemala where income is measured yearly, and they use yearly data when analysing consumption. Despite the fact that consumption is a better measure of welfare than income, there are no data on consumption for Chile.
  3. We observe wages only for women who work, and potential earnings are therefore unobserved for many women who do not have earnings. In addition, female labour-force participation has been around 42% between 2003 and 2009 (Ministry of Planning, 2009). Selection in the labour market could substantially affect the results for women. For this reason we restrict our analysis to men.

The data available for Chile consist of a longitudinal study with three waves, which collects information about labour histories. For each history, individuals report the length of the labour trajectory and an average wage, which allows us to construct proper measures of yearly, four-year and seven-year income measures. Also, monthly wages can be imputed from labour histories; however, by construction, monthly income is less noisy in our data than in traditional survey data. In order to analyse the potential biases that data collection could have in our calculations, we proceed to estimate Monte Carlo simulations to compare our results with different earning-generation processes. We find that our results are consistent with the Monte Carlo simulations.

The results presented in this article provide important inputs for public policy perspective and the related effort to compensate for inequality, which should be increased in at least two dimensions. First, public policy should be targeted more intensively, using more innovative actions and instruments. Second, a greater funding is needed to correctly compensate for the influence of household background on health, schooling, training, and other variables that affect income and welfare. For instance, Ferreira and Gignoux (2011) show that there is a negative relationship of inequality of opportunities in educational outcomes for 15-year-old students and the share of public spending on primary education. Even though the relationship is not causal, it suggests the importance of public policy and inequality of opportunities. In addition, these results provide ground for a conceptual discussion on compensation. Indeed, providing a foundation for redistributing current income is one way to address unfortunate childhood experiences. Given that an important fraction of income inequality can be traced back to the distribution of circumstances in early life, it can be argued that justice requires compensating for negative childhood circumstances. This argument is, of course, at the conceptual level. Its implementation in the public policy arena would require taking into account other considerations such as the roles of information, incentives, and targeting.

The article is organised in six sections. In Section 2, we present the methodology for estimating the impact of opportunity inequality on socio-economic outcomes. Section 3 presents the data used in the article. Section 4 presents and discusses the results of the inequality of opportunity measures. Section 5 presents the Monte Carlo simulations and Section 6 presents the final remarks.

## 2 Methodology

In this section we briefly explain the methodology implemented to measure inequality of opportunity. We closely follow Bourguignon et al. (2007) and Ferreira and Gignoux (2011).

Roemer (1998) provides the foundation for understanding the role of inequality of opportunity in overall inequality. He postulates that advantage variables, for example labour earnings, are the result of two types of variables: circumstances and effort. Circumstances are variables over which the individual has no control; for instance, the educational level of parents. On the other hand, effort is based solely on the individual. Implicitly or explicitly, these approaches are based on a model of advantage such as:

$$y = g(C, E, u) \quad (1)$$

where  $y$  denotes the outcome of interest or advantage;  $C$  denotes a vector of circumstance variables that are assumed to be exogenous;  $E$  denotes a vector of effort variables; and  $u$  denotes pure luck or random factors.

Following Bourguignon et al. (2007), we can approximate an advantage model by following a linear system:

$$\ln y = C\alpha + E\beta + u \quad (2)$$

$$E = C\lambda + v \quad (3)$$

where the structural equation relates income with circumstance and effort variables in a log-linear way. There is also a linear equation for the effort variables that respond to circumstances and an unobservable term.

The reduced form of this linear system is:

$$\ln y = C(\alpha + \beta\lambda) + v\beta + u \quad (4)$$

which can be estimated by OLS, given the exogeneity of circumstances as:

$$\ln y = C\theta + \varepsilon \quad (5)$$

Under these functional form assumptions, a parametrically standardised distribution is estimated by:

$$\bar{y}_i = \exp(\bar{C}\theta_{ols} + \varepsilon_{i,ols}) \quad (6)$$

where  $\bar{C}$  corresponds to the population average of circumstances,  $\theta_{ols}$  are the parameters obtained by OLS and  $\varepsilon_{i,ols}$  is the individual residual.

This parametrically standardised distribution is a counterfactual distribution of income, where all differences in circumstances are eliminated and only luck and effort remain. A measure of inequality of opportunity follows naturally from these estimates:

$$S = 1 - D(\bar{y}_i) / D(y_i) \quad (7)$$

where  $D(\ )$  is any inequality index calculated over the income distribution.  $S$  measures the share of inequality of opportunities in total inequality. As mentioned in Ferreira and Gignoux (2011),  $S$  is a lower-bound measure of inequality of opportunities, because the part of the total distribution of income explained by all circumstances does not decrease when adding more circumstances. Then, if there are omitted circumstances, the result is a lower  $S$ . We shall also focus on the level of inequality of opportunity, which will be measured as:

$$IO = D(y_i) - D(\bar{y}_i) \quad (8)$$

Ferreira and Gignoux (2011) discuss the use of non-parametric methods to measure inequality of opportunities; these methods do not rely on functional form assumptions. We prefer a parametric method because, as the vector of observed circumstances becomes larger and/or the number of categories within each variable increases, sampling variances become very large, making non-parametric estimates of inequality of opportunity too imprecise. This places an upper bound in the number of circumstances that can be included and in the number of categories within each circumstance that we can define. On the other hand, a multivariate regression framework would use data more efficiently, allowing a more detailed specification of circumstances and categories, but at the expense of imposing a given functional form on the relation between variables.<sup>4</sup>

Ferreira and Gignoux (2011) show how the parametric approach permits the estimation of the partial effect of one (or a subset) of the circumstance variables while controlling for others by constructing alternative counterfactual distributions, such as:

$$\bar{y}_i^K = \exp(\bar{C}_i^K \theta_{ols}^K + C_i^{k \neq K} \theta_{ols}^{k \neq K} + \varepsilon_{i,ols}) \quad (9)$$

This allows one to compute circumstance J-specific inequality shares:

$$S^K = 1 - [D(\bar{y}_i^K) / D(y_i)] \quad (10)$$

As mentioned in Ferreira and Gignoux (2011), the calculation of partial shares assumes that the parameters are unbiased, which will not hold if the omitted opportunities in the main equation are correlated with the observed circumstances. Thus the interpretation of the shares must be done with care.

Turning to the choice of inequality indicators, the literature on inequality measurement shows that only some of the usual inequality indexes can decompose additively between-group inequality and within-group inequality (defined as the level of inequality of outcome measured within a given sub-group distribution). The best-known family of additively decomposable measures is the generalised entropy measures, among which the mean log deviation (GE(0)) and the Theil entropy (GE(1)) indexes are used in this article.<sup>5</sup>

### 3 Data

The data used in this article come from *Encuesta de Protección Social (EPS, Social Protection Survey)* from the 2004, 2006 and 2009 waves. The *EPS* is a longitudinal survey and follows approximately 20,000 individuals. It contains information about household composition, including individual information such as age, gender, education, and a wide range of labour-market variables. In addition, it incorporates parental data of the given individual, such as education, labour participation and characteristics of childhood household, which are key to estimating the level of inequality of opportunity.<sup>6</sup> One of the

4. See Ferreira and Gignoux (2011). For further discussion on nonparametric estimation of systems of equations see Rau (2013).

5. GE(1) results are presented in the article and GE(0) results in the appendix.

6. For a detailed description see Bravo et al. (2008).

main advantages of using the *EPS* instead of other data is that we can construct four-year and seven-year labour incomes, which are more precise estimates of permanent income than monthly and yearly income, the usual variables in the literature.

In order to study how measures of inequality of opportunities change with short-term or long-term income, we use the same sample of individuals to eliminate composition effects from our calculations. This restriction implies that we consider only individuals with positive earnings for the whole period. Certainly this will bias the sample to the most skilled workers. Since the sample is more homogeneous in background characteristics, this could diminish the variance of the circumstances, but, at the same time, income variance could also decrease. The final effect on inequality of opportunities is unclear. However, if the probability of having always-positive earnings is positively related to circumstances, then we are calculating a lower bound in the calculation of inequality of opportunities. Nevertheless, our purpose is not to derive a comparable measure of inequality of opportunity for different samples, but to examine the effects of short-term vs. long-term income on a given inequality measure.<sup>7</sup> In addition, it is expected that skilled workers have less income variation on time, which makes it more difficult to find income-inequality differences using different length-of-income measures.

The issue of controlling for composition effects has been considered in several methodologies that seek to decompose changes of inequality in several components, such as in Juhn et al. (1993), di Nardo et al. (1996), and Gosling et al. (2000) among others. We approach the composition issue by considering the same sample for all calculations. In addition, it is not possible to include individuals with zero earnings at any point in time, since it would not be possible to calculate inequality measures considering those earnings.<sup>8</sup>

An important point to consider is the nature of the *EPS*, which does not allow us to utilise a pure monthly income measure. Individuals are asked about periods of labour activities and not for a strict monthly labour status. For instance, in the 2004 *EPS*, participants were asked about their employment status since January 2002. Thus, if one person had the same job from January 2002 until the time of the interview, he reports only one wage for the entire period. Individuals report a change in wages only when they change jobs. Using occupational information, we compute a monthly wage by using the wage reported for each labour history; then if an individual had the same job from January 2002 to December 2003, the monthly wage will be the same for those 24 months. In the case that an individual changes jobs, a different wage will be reported and a new monthly wage will be imputed. In each new wave individuals are asked again about their labour histories.<sup>9</sup>

Using the 2004, 2006 and 2009 waves of the *EPS*, we are able to build labour histories for the period January 2002 to December 2008 and compute monthly wages from the same period. Then we calculate yearly, four-yearly, and seven-yearly measures of earnings. The monthly wage used in our calculations corresponds to June 2008, the yearly wage corresponds to wages in 2008, the four-year wage covers from January 2005 to December

---

7. The percentage of men between 30 and 49 years of age who have always-positive earnings over the number of men between 30 and 49 who have at least one monthly earning greater than zero is 61%.

8. In the case of the Theil index, it requires the evaluation of the logarithm of earnings, which is not possible when earnings are equal to zero. Similar calculations are made by Ferreira and Gignoux (2011).

9. For instance, in the 2006 *EPS* round, individuals are asked about their labour status from January 2004 until the time of the interview; this implies a most likely change in the reported wage, even if individuals have the same job, which creates variance in wages per individual.

2008, and the seven-year wage covers from January 2002 to December 2008.<sup>10</sup> The characteristic of the survey tends to decrease the variance of wages per individual over time, making it more difficult to find differences in measures of inequality when using long-term income measures. Thus our findings can be considered as a lower bound, since more differences should be expected when using traditional measures of monthly income.

To better understand the potential biases of our estimates, we proceed to simulate the wages using Monte Carlo simulations of monthly income for a period of seven years and then calculate the indicators for monthly, yearly, four-year and seven-year income measures. The results are presented in Section 5 and show that our results with the *EPS* are consistent with the main trends of the Monte Carlo simulations.

Considering the nature of the data, our measures of monthly and yearly income are very similar by construction, and so the more relevant comparisons in the article are between yearly and four-year and seven-year income measures.

One issue that is often raised when measuring income inequality is that it is difficult to measure high-income groups. This is a potential problem not only for the *EPS*, but also for other Chilean surveys, such as the *Socio-Economic Characterisation Survey* or *Encuesta de Caracterización Socio-Económica (CASEN)*.<sup>11</sup> However, Feres (1998) and Bravo and Valderrama (2011) show that the main concern for *CASEN* is not labour earnings, but capital earnings, and, at the same time, the report of labour earnings tends to be consistent with the National Accounts. Thus, we expect this not to be a major problem for our estimates.

**Table 1: Inequality for different income measures, men aged 30-49**

	Monthly income	Yearly income	Four-Year income	Seven-Year income
Gini coefficient	0,397	0,396	0,380	0,366
Mean log deviation	0,265	0,264	0,239	0,220
Theil index	0,301	0,301	0,275	0,253
Coefficient of variation	0,970	0,970	0,921	0,872

Notes: Seven-year income measure is the average wage from 2002 to 2008; four-year income the average wage from 2005 to 2008; yearly income measured is the wage average for 2008, and monthly wage is the wage during June 2008. Number of observations is 1,611. Own elaboration.

Table 1 shows how different income inequality measures change when we move from monthly income to yearly, four-year, and then seven-year income data.<sup>12</sup> It is observed that the index change is very small when comparing monthly and yearly income, due to the data characteristics described above. However, the indexes calculated using the four-year and seven-year earnings are lower than the indexes calculated using yearly and monthly measures,

10. Yearly, four-year and seven-year income corresponds to income averages in nominal terms. A correction in real terms is not straightforward since individuals report income for time periods and not month by month. Also, when an individual reports more than one labour history, the wage reported is updated in real terms. All of these imply that the average of nominal earnings is close to the average of real earnings.

11. The *CASEN* is the official survey to measure poverty in Chile.

12. We present the result for the Gini coefficient and Generalised Entropy measures because it allows us to compare with the relevant literature.

which supports the view that long-term measurements of income are less affected by transitory components, even in a highly homogenous sample such as the one we are considering.

The level of the inequality measures shown in Table 1 is low compared with Latin American standards. Paes de Barros et al. (2007) reports that the lowest Theil index is for Panama (0.485), and the differences are related to different data collection methods. In Panama the earnings data are for the previous month, while, in our case, the *EPS* data are labour histories.

As circumstances, we include the following variables: mother's and father's education, father's occupation, region where the individual attended primary school, and number of siblings raised in the household. Father's occupation is a dummy variable equal to one if the father was a white-collar worker, zero otherwise. The region of the primary school serves as a proxy for location of birth. Region is south if it is to the south of Santiago and north if it is north of Santiago. We do not have access to ethnicity or income data for the household where the individual was born. This implies that the level and share of inequality of opportunities include the information of all the variables used in the estimation, plus the potential correlation with the variables not included, such as ethnicity or parent's income when the individual was a child.

Table 2 shows the main descriptive statistics of the variables used in the decomposition. As usual we observe that mean income is higher than median income, fathers and mothers have similar schooling levels, and mean number of siblings is higher than the median, implying that household size is the log normally distributed.

**Table 2: Descriptive statistics, men aged 30-49**

Variable	Mean	Median	Standard Deviation
Monthly income	357,351	250,000	351,615
Yearly income	357,063	250,000	350,988
Four-year income	340,943	247,500	317,712
Seven-year income	311,968	235,714	274,883
Father was white-collar	0.11	0.00	0.31
Father's years of schooling	5.31	5.00	4.66
Mother's years of schooling	5.19	5.00	4.35
Lived in the South	0.22	0.00	0.42
Lived in the North	0.47	0.00	0.50
No. of siblings	3.97	3.00	2.77

Notes: Seven-year income measure is the average wage from 2002 to 2008, four-year from 2005 to 2008. Yearly income measured is the wage average for 2008, and monthly wage the wage during June 2008. Income measures are in Chilean Pesos, in June 2008 the average exchange rate was 493 pesos = 1 US\$. Number of observations is 1,611. Own elaboration.



## 4 Results

In this section we present the results for men between 30 and 49 years old, in the appendix we show results for men between 25 and 54 years old. The age restriction allows us to partially control for the life-cycle profile of wages.<sup>13</sup>

Table 3 shows the main results, including estimates of overall labour income inequality and inequality of opportunity levels for the GE(1) measure, i.e. Theil index, and for four different income measures: seven-year and four-year average earnings, yearly average earnings and monthly earnings.<sup>14</sup>

**Table 3: Inequality and inequality of opportunity for different income measures, men aged 30-49**

Income measure	Wage inequality: Theil Index	Inequality of opportunity: IO	Share
Seven-year	0,252 (0,014)	0,052 (0,011)	0,206 (0,041)
Four-year	0,274 (0,016)	0,050 (0,011)	0,184 (0,041)
Yearly	0,300 (0,017)	0,050 (0,012)	0,166 (0,042)
Monthly	0,301 (0,017)	0,049 (0,012)	0,164 (0,041)

Notes: Seven-year income measure is the average wage from 2002 to 2008, four-year from 2005 to 2008. Yearly income measured is the wage average for 2008, and monthly wage the wage during June 2008. Standard error in parenthesis, calculated using 200 bootstrap repetitions. Own elaboration.

From Table 3 we can observe that income inequality is marginally lower (but not significantly so) when using yearly instead of monthly income. The differences in income inequality are statistically significant for longer-term incomes. The level of inequality of opportunity marginally increases with long-term income measures, but the differences are not statistically significant. Finally, the share of income inequality is statistically the same for the monthly and yearly measures, but the differences are statistically different from zero for the rest of the comparisons. The fact that indexes using monthly and yearly income do not have significant differences is related to the data-collection process, which makes it difficult to separate monthly and yearly income. Since the available data better permit measures of four-year and seven-year income measures, we observe statistically significant differences with the expected sign, decreasing overall inequality and increasing share of inequality of opportunities.

The fact that income inequality decreases with long-term income measures can be explained by the fact that income measured in a broader period of time contains less transitory

13. We use bootstrap to construct standard errors, using 200 repetitions. All calculations consider the weighting factor of the 2009 survey.

14. In the appendix we report results for GE(0) as well.

components. The marginal increment of inequality of opportunity suggests that circumstances are correlated to measurement error in wages and to transitory shocks, and that when these factors are taken into account, the level of inequality of opportunities increases. However, since the differences are not significant, further evidence is needed to support that hypothesis.

The share of inequality of opportunity increases about four percentage points or 25%, when considering long-term income measures. This result shows the importance of using long-term income measures to study inequality of opportunity, since the changes in the indicators can be of relatively large importance. This is a significant underestimation, which may affect policy recommendations in countries with high levels of inequality. In addition, the increment in the share of inequality of opportunity indicates that measures of permanent income tend to be more correlated with circumstances of the individuals. A higher correlation with permanent income is consistent with theories of human capital formation as mentioned in Cunha and Heckman (2007), where early investment is shown to have cumulative effects on ability.

The result that inequality of opportunity increases with longer-term incomes is novel in the recently developed literature of measuring opportunities in terms of early circumstances. However, it relates to a similar result found in the literature of estimating intergenerational income elasticity. The first computations for the US economy were based on short-term income data for parents and sons, rendering estimates of the elasticity of about 0.20, which depicted an optimistic picture of social mobility in that country (Behrman and Taubman, 1985). Later, the availability of longitudinal data allowed the computation of intergenerational elasticity based on the long-term or permanent income of parents and sons. Now the estimates were close to 0.50, sobering the views about mobility in the US (Solon, 1989).

In a Latin American context, the share of inequality of opportunities in Chile for the monthly and yearly data (16%) is higher than for Peru (10%) and Panama (14.8%), but lower than for Colombia (19.7%), Guatemala (22.7%), Ecuador (23.7%) and Brazil (29.5%).<sup>15</sup>

In Table 4 we present the contribution of each circumstance variable to the inequality of opportunity. Corresponding OLS estimates are shown in the appendix.

**Table 4: Decomposition by circumstance; Theil Index**

Income measure	Father's occupation	Mother's education	Father's education	Region	Siblings
Seven-year	0,036 (0,010)	0,101 (0,018)	0,101 (0,020)	0,030 (0,014)	0,037 (0,005)
Four-year	0,029 (0,008)	0,089 (0,018)	0,098 (0,021)	0,025 (0,013)	0,028 (0,005)
Yearly	0,025 (0,007)	0,085 (0,018)	0,089 (0,021)	0,022 (0,013)	0,024 (0,005)
Monthly	0,024 (0,007)	0,085 (0,018)	0,088 (0,021)	0,021 (0,013)	0,024 (0,005)

Notes: As for Table 3.

15. Ferreira and Gignoux (2011). However, since the income measures are not directly comparable, the differences in the figures include measurement and actual opportunities differentials.

The two most important circumstances contributing to inequality of opportunity are paternal and maternal education levels. Both variables increase their importance when using long-term income measures. Interestingly, we find that mother's education is not the key variable as is shown in the literature (Paes de Barros et al., 2009). Parent's education is likely to be highly correlated with parental household income, which is a key determinant in the formation of early human capital through its impact on nutrition, health and education.<sup>16</sup> Another important feature of Table 4 is the fact that the share of each circumstance individually increases when using four- and seven-year income, showing that all these factors are correlated with permanent income.

Finally, from the results reported in Tables 3 and 4, it can be argued that circumstances affect inequality of opportunity in a significant manner and inequality of opportunity translates into income inequality. We can note that circumstances such as father's and mother's education account for more than 10% of inequality of opportunity and that the share related to income inequality is about 20%.

The results in this article identify that calculation bases in short-term income measures can be misleading about the level of inequality of opportunities, and a more thorough empirical investigation must be made. Due to data limitations we cannot include ethnicity, which has been indicated also as an important circumstance in the Latin American context (Paes de Barros et al., 2009). At the same time, we must resolve how to obtain credible results for women, which might involve another important circumstance related to earnings and schooling.

## 5 Monte Carlo simulations

In this section we present the results of Monte Carlo simulations to assess the robustness of our results to different data-generating processes and the relationship of longer-term income measures and inequality of opportunity. We find that our results are consistent with several income processes, and that the data limitations of the *EPS* are only one factor in obtaining monthly measures. However, we are able to capture the dynamics of yearly, four-year, and seven-year earnings and inequality of opportunities measures.

To develop the Monte Carlo simulations, we use the equations developed in Section 2, and assume different data-generation processes for the error term, which for convenience is separated into two components:

$$\varepsilon = f + \eta \quad (11)$$

where  $f$  is a permanent component, an unobserved individual fixed effect, and  $\eta$  is a transitory, monthly shock.

The simulation uses three circumstance variables, which are constructed from a trivariate normal distribution, with the covariance matrix similar to that calculated for

---

16. Nuñez and Tartakowsky (2011) use the *CASEN* and follow Bourguignon et al. They calculate that 15% of total inequality is due to inequality of circumstances. They use monthly income, a different age range, and a set of different circumstance variables, but also find that parent's education is the largest contributor to inequality of opportunity.

father's occupation, father's education, and mother's education for the *EPS*. The coefficients of the log-wage equation correspond to the ones obtained also from the actual *EPS* data.

We have four data-generation processes. First, we assume two scenarios for the correlation between  $f$  and the circumstances: zero correlation and different from zero.<sup>17</sup> Second, we assume that  $\eta$  is *iid* or *AR(1)*. These assumptions give four different scenarios. Monthly wages for 7-year periods are generated for a sample of 2,000 individuals, and then inequality and inequality of opportunities indexes are calculated. We replicate the results 200 times for each data-generation process.

Table 5 shows the results of the Monte Carlo experiment when correlation is allowed between the error term and the circumstances and  $\eta$  follows an *AR(1)* process.

**Table 5: Monte Carlo simulation with *AR(1)* error term with no correlation between  $f$  and circumstances**

Income measure	Wage inequality	Inequality of opportunity:	Share
	Theil Index	IO	
Seven-year	0,221 (0,009)	0,028 (0,004)	0,125 (0,018)
Four-year	0,223 (0,009)	0,027 (0,004)	0,121 (0,017)
Yearly	0,232 (0,010)	0,027 (0,004)	0,118 (0,017)
Monthly	0,244 (0,009)	0,026 (0,004)	0,108 (0,017)

Note: Standard error in parenthesis, calculated using 200 bootstrap repetitions. Own elaboration.

Table 5 shows that inequality decreases when long-term income measures are used; also the level of inequality of opportunities decreases and the share of inequality of opportunities increases, which is consistent with our findings. We observe that inequality decreases when we move from monthly to yearly income data, which, as expected, is not observed in our results. Inequality continues to decrease when we use four-year and seven-year income measures, which is consistent with our data. Similar results are found for the rest of the scenarios (see Appendix).

## 6 Final remarks

This article contributes to the existing literature by examining the sensitivity of inequality of opportunity measures to long-term earnings and its relationship with income inequality. We use four alternative measures of income for the analysis. The first two are monthly and yearly income, which are the usual measure of labour income used in the literature and are

17. We assume that  $f$  has correlation 0.1 with one of the circumstances and close to zero with the other two. The correlation between number of siblings and father's occupation is 0.1.

heavily exposed to transitory shocks and measurement error. The third measure is four-year income and the fourth is seven-year income, both of which are more precise estimates of permanent income, attenuating measurement error. Hence, the last two measures will allow us to more accurately understand the relationship between an individual's early life circumstances and his economic outcomes. In a country like Chile with high levels of income inequality, it is of particular interest to properly measure the level of inequality of opportunities, especially when considering how to improve the living conditions of most of the population.

Our results show that using monthly or yearly measures of income can produce an underestimation of 25% in indexes of inequality of opportunities, indicating that measures of permanent income are particularly important when measuring distribution of opportunity. This is the first estimation in the literature of the importance of having longer-term measures to assess accurately the relationship of income inequality and inequality of opportunities.

In addition to measurement issues, our results show that circumstances in individuals' early lives have an impact on their adult socio-economic outcomes. Our results find that at least 20% of labour income is correlated to circumstances in the parental household and that paternal education is the most highly correlated circumstance.

What about policy implications? Can we draw policy recommendations from an issue that has been determined in the past, as has the case of the effects of early life circumstances of the current adult population? Most policy implications about opportunity policies are related to current and future children, an issue that is addressed in a companion paper (Contreras et al., 2012). Our results show that current measures of inequality of opportunities could be underestimated, indicating that the efforts made to reduce inequality of opportunities have to scale-up. This should be done with innovative policies targeted on children with poor family backgrounds, especially in health and schooling. However, the budget dedicated to it must increase considerably. For example, developing countries could increase the length and scope of Conditional Cash Transfers, especially to the conditions of children's health and education, but the provision of those services has to improve to decrease effectively the levels of inequality of opportunity for future generations. Ferreira and Gignoux (2011) show negative relationships between public expenditure in education and inequality of opportunities of educational outcomes, suggesting that inequality indexes can be useful to measure the impact of public policy. However, in order to do so, a correct measure of inequality of opportunity is needed.

For adults, a public policy focusing on a late compensation for some previous injustices could be proposed. This is mainly the 'redressing principle' discussed by Rawls (1971) in his famous *Theory of Justice*. In this work, the author claims that a just society is one where institutions maximise the life prospects of the most disadvantaged. To do so, society has to level the socio-economic playing field to ensure that individuals can develop their talents fully. In addition, individuals who were not favoured with talents by nature should be allocated income from redistribution financed out of taxation. On the other hand, Nozick (1974) in his response to Rawls explicitly advocates a redressing of past injustices, with the main exception in his postulates that the state should confine itself to the protection of individual rights and not intervene in the allocation of income by market forces.

Note that not every low-income individual qualifies for redressing, but only those who experienced unfavourable circumstances early in life. According to our results, individuals whose parents had less education could receive some compensation for the disadvantages they faced during their childhood.

The previous argument is made at the conceptual level. Its implementation in the public-policy arena requires taking into account other considerations such as the roles of information, incentives and targeting. For example, it can be discussed whether a high-income individual should also be compensated if he/she experienced disadvantages early in life. From a more pragmatic point of view, some initiatives that have been advanced to improve the income distribution in the country, like those that originated in the *Consejo de Equidad y Trabajo* (2008), can be related to the issues of compensation and redressing, in addition to more traditional arguments.

*first submitted October 2012*

*final revision accepted December 2012*

## References

- Behrman, J. and Taubman, P. (1985) 'Intergenerational Earnings Mobility in the US. Some estimates and a test of Becker's intergenerational endowment model', *Review of Economics and Statistics* 67(1): 144-51.
- Bourguignon, F., Ferreira, F.H.G. and Menéndez, M. (2007) 'Inequality of Opportunity in Brazil.', *Review of Income Wealth* 53(4): 585-618.
- Bravo, D. and Valderrama, J. (2011) 'The Impact of Income Adjustment in the Casen Survey on the Measurement of Inequality in Chile', *Estudios de Economía* 38(1): 43-65.
- Bravo, D.; Vásquez, J.; Berhman, J.; Mitchell, O.S. and Todd, P. (2008) *Encuesta de Protección Social 2006: Presentación General y Principales Resultados*. Working Paper 273. Department of Economics, Universidad de Chile.
- Checchi, D. and Peragine, V. (2005) *Regional Disparities and Inequality of Opportunity: The case of Italy*. IZA Discussion Paper No. 1874/2005. Bonn: Institute for the Study of Labour (IZA).
- Consejo Asesor Presidencial de Equidad y Trabajo* (2008) *Final Report*. <http://www.iadb.org/intal/intalcdi/PE/2009/02516.pdf> (accessed 12 December 2013).
- Contreras, D.; Larrañaga, O.; Puentes, E. and Rau, T. (2012) 'Inequality of Opportunity for Children in Chile', *Cepal Review* 106 (August): 107-24.
- Cunha, F. and Heckman, J. (2007) 'The Technology of Skill Formation', *American Economic Review* 97(2): 31-47.
- Di Nardo, J., Fortin, N. and Lemieux, T. (1996) 'Labor Market Institutions and the Distribution of Wages, 1973-1992: A semiparametric approach', *Econometrica* 64(5): 1001-44.
- Feres, J. (1998) 'La Medición de Ingresos en la Encuesta Casen-1996', in *Medición de Ingresos en las Encuestas de Hogares*. Capítulo 29. Taller 2. Programa para el mejoramiento de las Encuestas y la Medición de las Condiciones de Vida en América Latina y El Caribe. <http://www.eclac.cl/deype/mecovi/docs/TALLER2/29.pdf> (accessed 12 December 2013).
- Ferreira, F.H.G. and Gignoux, J. (2011) 'The Measurement of Inequality of Opportunity: Theory and an application to Latin America', *Review of Income and Wealth* 57(4): 622-57.

- Gosling, A., Machin, S. and Meghir, C. (2000) 'The Changing Distribution of Male Ages in the U.K.', *Review of Economic Studies* 67(4): 635-66.
- Juhn, C., Murphy, K. and Pierce, B. (1993) 'Wage Inequality and the Rise in Returns to Skill', *Journal of Political Economy* 101(3): 410-42.
- Ministry of Planning (2009) *Encuesta de Caracterización Socioeconómica Nacional: Mujeres*. Santiago: Ministry of Planning.
- Nozick, R. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- Nuñez, J. and Tartakowsky, A. (2011) 'The Relationship between Inequality of Outcomes and Inequality of Opportunities in a High-Inequality Country: The case of Chile', *Applied Economic Letters* 18(4): 359-69.
- Paes de Barros, R., De Carvalho, M. and Franco, S. (2007) 'Preliminary Notes on the Measurement of Socially-Determined Inequality of Opportunity when the Outcome is Discrete.' Unpublished manuscript. Rio de Janeiro: IPEA.
- Paes de Barros, R.; Ferreira, F.; Molinas, J. and Saavedra, J. (2009) *Measuring Inequality of Opportunities in Latin America and the Caribbean*. Washington, DC: World Bank.
- Perry, G.; Maloney, W.; Arias, O.; Fajnzylber, P.; Mason, A. and Saavedra, J. (2007) *Informality: Exit and inclusion*. Washington, DC: World Bank.
- Rau, T. (2013) 'Modeling Structural Equations with Endogenous Regressors and Heterogeneity through Derivative Constraints', *Quantitative Economics* 4(1): 125-48.
- Rawls, J. (1971) *A Theory of Justice*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Roemer, J. (1998). *Equality of Opportunity*. Cambridge, MA: Harvard University Press.
- Solon, G (1989) 'Biases in the Estimation of Intergenerational Earnings Correlations', *Review of Economics and Statistics* 71(1): 172-4.

## Appendix

**Table A1: Inequality results, men 30-49**

Income	Wage inequality		Inequality of opportunities		Share	
	GE(0)	GE(1)	GE(0)	GE(1)	GE(0)	GE(1)
Seven-year	0,220	0,252	0,046	0,052	0,208	0,206
	(0,011)	(0,014)	(0,007)	(0,011)	(0,027)	(0,041)
Four-year	0,239	0,274	0,046	0,050	0,191	0,184
	(0,012)	(0,016)	(0,007)	(0,011)	(0,028)	(0,041)
Yearly	0,263	0,300	0,046	0,050	0,176	0,166
	(0,013)	(0,017)	(0,007)	(0,012)	(0,027)	(0,042)
Monthly	0,264	0,301	0,046	0,049	0,175	0,164
	(0,013)	(0,017)	(0,007)	(0,012)	(0,027)	(0,041)

Notes: Seven-year income measure is the average wage from 2002 to 2008. Four-year income is the average wage from 2005 to 2008. Yearly income measured is the wage average for 2008, and monthly wage is the wage during June 2008. Standard error in parenthesis, calculated using 200 bootstrap repetitions. Own elaboration.

**Table A2: Inequality results, men 25-54**

	Wage inequality		Inequality of opportunities		Share	
	GE(0)	GE(1)	GE(0)	GE(1)	GE(0)	GE(1)
Seven-year	0,218 (0,010)	0,252 (0,014)	0,038 (0,005)	0,044 (0,008)	0,176 (0,020)	0,176 (0,027)
Four-year	0,239 (0,011)	0,277 (0,015)	0,038 (0,005)	0,043 (0,008)	0,160 (0,020)	0,154 (0,029)
Yearly	0,263 (0,012)	0,302 (0,016)	0,038 (0,006)	0,041 (0,009)	0,145 (0,020)	0,136 (0,029)
Monthly	0,264 (0,012)	0,303 (0,016)	0,038 (0,006)	0,041 (0,009)	0,144 (0,020)	0,134 (0,029)

Notes: As for Table A1.

**Table A3: Analysis by circumstance, men 25-54**

Income measure	Father's occupation	Mother's education	Father's education	Region	No. of siblings
Seven-year	GE(1) 0,025 (0,006)	GE(1) 0,077 (0,012)	GE(1) 0,097 (0,015)	GE(1) 0,012 (0,008)	GE(1) 0,027 (0,005)
Four-year	0,020 (0,005)	0,066 (0,012)	0,090 (0,016)	0,009 (0,008)	0,023 (0,005)
Yearly	0,016 (0,005)	0,063 (0,012)	0,082 (0,016)	0,006 (0,007)	0,020 (0,005)
Monthly	0,016 (0,005)	0,063 (0,012)	0,081 (0,016)	0,006 (0,008)	0,020 (0,005)

Notes: Ibid.

**Table A4: Regression results, men 30-49**

	Seven-year	Four-year	Yearly	Monthly
Father's occupation	0.127 (0.045)**	0.127 (0.048)**	0.122 (0.051)*	0.122 (0.051)*
Father's education	0.028 (0.004)**	0.028 (0.004)**	0.028 (0.004)**	0.028 (0.004)**
Mother's education	0.024 (0.004)**	0.026 (0.004)**	0.027 (0.004)**	0.027 (0.004)**
South	-0.158 (0.037)**	-0.144 (0.039)**	-0.129 (0.042)**	-0.132 (0.042)**
North	-0.211 (0.032)**	-0.207 (0.034)**	-0.212 (0.036)**	-0.214 (0.036)**



**Table A4 cont.**

	Seven-year	Four-year	Yearly	Monthly
No. of siblings	-0.020 (0.005)**	-0.018 (0.005)**	-0.017 (0.006)**	-0.017 (0.006)**
Constant	12.343 (0.038)**	12.392 (0.040)**	12.403 (0.043)**	12.404 (0.043)**
Observations	1611	1611	1611	1611
R <sup>2</sup>	0.20	0.19	0.17	0.17

Notes: Standard errors in parentheses; \* significant at 5%; \*\* significant at 1%.

**Table A5: Regression results, men 25-54**

	Seven-year	Four-year	Yearly	Monthly
Father's occupation	0.111 (0.041)**	0.106 (0.043)*	0.098 (0.046)*	0.098 (0.046)*
Father's education	0.024 (0.003)**	0.024 (0.003)**	0.024 (0.004)**	0.024 (0.004)**
Mother's education	0.025 (0.004)**	0.026 (0.004)**	0.027 (0.004)**	0.027 (0.004)**
South	-0.127 (0.033)**	-0.125 (0.034)**	-0.113 (0.036)**	-0.115 (0.037)**
North	-0.157 (0.028)**	-0.154 (0.030)**	-0.155 (0.032)**	-0.156 (0.032)**
No. of siblings	-0.018 (0.004)**	-0.018 (0.005)**	-0.017 (0.005)**	-0.017 (0.005)**
Constant	12.307 (0.033)**	12.368 (0.035)**	12.378 (0.038)**	12.379 (0.038)**
Observations	2128	2128	2128	2128
R <sup>2</sup>	0.17	0.16	0.14	0.14

Notes: As for Table A4.

**Table A6: Monte Carlo exercise, iid error term and no correlation between *f* and circumstances**

Income measure	Wage inequality		Inequality of opportunity		Share	
	GE0	GE1	GE0	GE1	GE0	GE1
Seven-year	0,227 (0,008)	0,221 (0,008)	0,027 (0,003)	0,026 (0,004)	0,121 (0,013)	0,119 (0,014)
Four-year	0,226 (0,007)	0,220 (0,007)	0,028 (0,003)	0,026 (0,004)	0,122 (0,013)	0,120 (0,015)

**Table A6 cont.**

Income measure	Wage inequality		Inequality of opportunity		Share	
Yearly	0,228	0,222	0,027	0,026	0,118	0,116
	(0,007)	(0,007)	(0,003)	(0,004)	(0,012)	(0,015)
Monthly	0,250	0,245	0,028	0,026	0,111	0,106
	(0,008)	(0,009)	(0,004)	(0,005)	(0,014)	(0,017)

Note: Standard error in parenthesis, calculated using 200 bootstrap repetitions. Own elaboration.

**Table A7: Monte Carlo exercise, iid error term with correlation between  $f$  and circumstances**

Income measure	Wage inequality		Inequality of opportunity		Share	
	GE0	GE1	GE0	GE1	GE0	GE1
Seven-year	0,222	0,219	0,028	0,029	0,126	0,131
	(0,008)	(0,009)	(0,003)	(0,004)	(0,015)	(0,018)
Four-year	0,222	0,220	0,028	0,028	0,125	0,129
	(0,008)	(0,010)	(0,003)	(0,004)	(0,012)	(0,016)
Yearly	0,223	0,222	0,028	0,029	0,124	0,129
	(0,008)	(0,010)	(0,003)	(0,004)	(0,014)	(0,018)
Monthly	0,241	0,237	0,028	0,027	0,114	0,115
	(0,008)	(0,010)	(0,004)	(0,005)	(0,015)	(0,020)

Note: As for Table A6.

**Table A8: Monte Carlo exercise, AR(1) error term with no correlation between  $f$  and circumstances**

Income measure	Wage inequality		Inequality of opportunity		Share	
	GE0	GE1	GE0	GE1	GE0	GE1
Seven-year	0,229	0,223	0,027	0,026	0,119	0,118
	(0,008)	(0,008)	(0,003)	(0,004)	(0,013)	(0,015)
Four-year	0,230	0,224	0,028	0,026	0,120	0,118
	(0,007)	(0,007)	(0,003)	(0,004)	(0,013)	(0,015)
Yearly	0,238	0,231	0,025	0,023	0,105	0,100
	(0,007)	(0,008)	(0,003)	(0,004)	(0,013)	(0,016)
Monthly	0,252	0,246	0,026	0,023	0,102	0,093
	(0,008)	(0,008)	(0,004)	(0,004)	(0,014)	(0,016)

Note: Ibid.

**Table A9: Monte Carlo exercise, AR(1) error term with correlation between *f* and circumstances**

Income measure	Wage inequality		Inequality of opportunity		Share	
	GE0	GE1	GE0	GE1	GE0	GE1
Seven-year	0,224 (0,008)	0,221 (0,009)	0,027 (0,003)	0,028 (0,004)	0,121 (0,014)	0,125 (0,018)
Four-year	0,226 (0,008)	0,223 (0,009)	0,027 (0,003)	0,027 (0,004)	0,118 (0,013)	0,121 (0,017)
Yearly	0,234 (0,008)	0,232 (0,010)	0,026 (0,003)	0,027 (0,004)	0,113 (0,013)	0,118 (0,017)
Monthly	0,249 (0,008)	0,244 (0,009)	0,026 (0,004)	0,026 (0,004)	0,105 (0,014)	0,108 (0,017)

Note: Ibid.