Long-Run Labour Income Distribution Dynamics: The Case of Chile 1990-2017*

La dinámica de largo plazo de la distribución de ingresos laborales: el caso de Chile 1990-2017

Richard Blundell**
Vicente Corral***
Andrés Gómez-Lobo****

Abstract

We analyse the long-run evolution of the labour income distribution for Chile. To this end, we use thirteen waves of the CASEN household socioeconomic survey from 1990 to 2017. During this period hourly earnings inequality measured by the Gini coefficient fell from 0.47 to 0.40. We use a RIF regression approach similar to Ferreira et al. (2021) for Brazil to decompose changes in average earnings and earnings inequality. We do not find observable variables that explain –either through an endowment effect or through a structural price change– a significant part of the decrease in hourly earnings inequality.

Key words: *Income Distribution, Inequality Dynamics, RIF Decomposition, Chile.*

JEL Classification: *D30, D31, D39, J31.*

---

* We thank comments by Claudio Sapelli and conference participants at Faculty of Economics and Business of the University of Chile. The usual disclaimers apply.

** Professor of Economics, University College London and Institute for Fiscal Studies. E-mail: r.blundell@ucl.ac.uk

*** Research Assistant, Department of Economics, University of Chile. E-mail: vcorral@fen.uchile.cl

**** Associate Professor of Economics, Department of Economics, University of Chile. E-mail: agomezlo@fen.uchile.cl

Received: August, 2023 Accepted: August, 2024

Resumen

Este trabajo analiza la evolución de largo plazo de la distribución de ingresos laborales para Chile. Con este fin, se usan trece olas de la encuesta de hogares CASEN desde 1990 a 2017. Durante este período, la desigualdad en los ingresos por hora, medida por el coeficiente Gini, cayó desde 0.47 a 0.40. Usamos regresiones RIF, similar a Ferreira et al. (2021) para Brasil, para descomponer cambios en los ingresos promedio y la desigualdad de ingresos. No encontramos variables observables que expliquen –ya sea a través de un efecto dotación o cambio estructural de precios– una parte significativa de la disminución en la desigualdad de los ingresos por hora.

Palabras clave: *Distribución de Ingresos, Dinámica de Desigualdad, Descomposición RIF, Chile.*

Clasificación JEL: *D30, D31, D39, J31.*

1. Introduction

In this paper we analyse the long-run evolution of the distribution of labour incomes for Chile. To this end, we use thirteen waves of the CASEN household socioeconomic survey from 1990 to 2017, a 28-year time span.

Chile is an interesting case to analyse since it is often touted as a regional economic success story. Real per capita GDP grew from 9,702 (PPP, constant 2017 international $) in 1990 to 24,547 in 2017.[1] Absolute poverty rates decreased from 38.6% in 1990 to 8.5% in 2017 and extreme poverty rates from 13% to 1.5% during the same period (MSDF and UNDP (2020)).[2] According to the same source, multidimensional poverty (household with unsatisfied needs in at least three dimensions or more) fell from 34.7% to 7.4% between 1990 and 2017.

Despite these impressive figures, income inequality has remained stubbornly high. According to OECD data, Chile's Gini coefficient of household disposable income was 0.46 in 2017, the highest among all 37 OECD coun-

---

1 World Development Indicators database, World Bank, https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?locations=CL. Last accessed on August 7, 2023.

2 The absolute poverty rate is the percentage of households whose disposable income is below the poverty line, while extreme poverty is the percentage of households whose income is below the price of a basic basket of foodstuff. Different official poverty lines have been defined through time resulting in slightly different poverty dynamics during this period. However, they all point in the same direction: a significant fall in poverty rates measured by income. See MSDF and UNDP (2020) for more details.

tries, and only slightly lower than the maximum value of 0.48 (2009) recorded for Chile in this dataset.[3]

That Chile has an unequal income distribution is well known and researched.[4] In part, this is due to a tax system with low progressivity with almost half of revenues coming from VAT taxation, together with government transfers that have been unable to reduce the post tax-transfers Gini coefficient by more than a few basis points.[5]

The distribution of labour earnings in Chile has been studied by Behrman (2011) who simulates the potential impact of different human capital policies on earning inequality. Sapelli (2011) analyses income distribution by age cohort between 1957 and 2004 and shows that there is an inverted U shape of income inequality by cohort through time. He decomposes these changes and suggests that the fall in income inequality for younger cohorts seems to be related to a flattening of the age-income profile and thus a reduction in the returns to experience. Sapelli (2016) argues that owing to the educational attainment of recent generations, income inequality by cohort should continue to decrease in the future. UNDP (2017) (Chapter 7) also analyses the distribution of earnings inequality. It notes that inequality was quite stable until it began to fall between 2003 and 2015. It attributes half of this fall to the higher supply of more educated workers; that is, a fall in the educational wage premium.[6] Contreras and Gallegos (2011) decompose earnings inequality for a set of Latin American countries (including Chile) and find that changes in educational attainment is the most important factor explaining the evolution of wage inequality between 1990 and 2000.

In this paper, we contribute to the literature by revisiting the long-term dynamics of the distribution of hourly labour incomes. Our data covers the 1990 to 2017 period, a more recent time frame than previous studies (e.g. Sapelli (2011)) and longer than most other studies.[7] We show that common time effects by educational group are important factors in the evolution of earnings by cohort. Once these common time effects are controlled for, younger cohorts with low levels of education have a similar wage-age profile as their older peers. However, younger cohorts of both female and male workers with medi-

---

3 OECD (2020), Income inequality (indicator). https://doi.org/10.1787/459aa7f1-en. Last accessed on December 28, 2020.

4 An encompassing and up-to-date analysis is the book UNDP (2017). For a historical analysis going back to the mid nineteenth century see Rodríguez (2017) and Eyzaguirre (2019). For the top income dynamics from 1964-2017 see Flores et al. (2019).

5 According to the same OECD data cited above, Chile's market Gini coefficient (market incomes before taxes and transfers) in 2017 was 0.495, so the tax and benefit system reduced inequality (measured by the Gini coefficient) by only 3.5 points that year.

6 UNDP (2017) also analyses wage inequality according to employer and finds that a significant proportion of wage inequality is explained by firm attributes.

7 UNDP (2017) is the exception. In some of their analysis they use the same data from 1990 to 2015.

um levels of education seem to earn less at the same age as older cohorts once these common time effects are controlled for. This is even more marked for workers with a high level of education. The educational wage premium has declined once educational group time effects are considered. That common time effects are important and modify the interpretation of wage-age profiles and income distribution dynamics has been found for the United States by Heathcote et al. (2005) and more recently by Blundell et al. (2023). In this paper we discuss the possible impacts that common educational group time effects may have on the labour income distribution dynamics in Chile.

In addition, we decompose the changes in earnings between 1990 and 2017 using a recentered influence function (RIF) approach (Firpo et al. (2009)) to delve deeper into the potential structural factors that may have influenced the evolution of labour income inequality during this period. This is similar in spirit to Sapelli (2011, 2016) who decomposes the variance in (log) income to explain the evolution of earnings inequality in Chile. Although complementary to our approach, there are several limitations to the variance decomposition that justify exploring a different methodology (see Section 4.1 of Fortin et al. (2011)). First, only a limited set of covariates can be used since the quadratic form in the compositional effect variables will generate interaction terms that are difficult to interpret. In Sapelli (2011, 2016) only education (and the evolution of the returns to education) are analysed. In contrast, in this paper we use a large set of potential explanatory variables. Second, under the reasonable assumption of heteroskedasticity of the error term in the Mincer equation, the conditional moment for the variance of unobservables must also be specified and estimated. Finally, our approach can be used to decompose other inequality measures, such as the Gini, Theil index and interquantile ratios.

During the period of our data there was a marked increase in the educational attainment of the workforce. The percentage of workers with at least 12 years of schooling (approximately a high-school degree) increased from 41.2% in 1990 to 70.2% in 2017. Also, the percentage of workers with 17 years of schooling or more (approximately a higher education degree) increased from 6.6% to 13.5%. On the other hand, the minimum wage during this period increased 159% in real terms, more than the 113% increase in mean labour income from the main occupation of workers. Another structural change was the female participation rate that grew from 33.6% in 1990 to 52.9% in 2017. Our aim is to discern what these and other exogenous or policy changes may have had on the distribution of (hourly) labour incomes.

In spirit and methodology, our approach follows closely the analysis undertaken for Brazil by Ferreira et al. (2021). They find that between 1995 and 2012 labour income inequality declined in Brazil. Rising educational levels does not seem to explain this evolution. The composition effect of a higher skilled labour force moving workers into the more convex part of the skill-wage premium curve more than compensated for the equalizing effect of higher educational endowments. Minimum wage policies may have contributed to the decline in inequality in Brazil, but only in the second half of the period (2003-2012). In the first half, it may have increased inequality by increasing informal activities and thus increasing the wage difference between informal workers earning below the minimum wage and formal workers earning a higher minimum wage. The main explanations for the decline in labour income inequality put forward by Ferreira et al. (2021) are a reduction in the returns to potential experience and the closing of the wage gap by gender, race, location and formal work status.

We undertake a similar approach, using a simple decomposition to explore possible explanations for the changes in average earnings in Chile between 1990 and 2017. We find that while the increase in educational attainment explains part of the increase in average earnings it was countered by a fall in the returns to education with a small overall effect. Other observable characteristics, such as gender, economic sector, potential experience, rural or urban workers and several institutional variables of the labour market (formal contract, firm size and minimum wage) explain but a small proportion in the increase in average labour incomes during the period. Most of the change is explained by returns to unobservable skills.

As for earnings inequality measured by the Gini coefficient, most of the observed characteristics are estimated to have been inequality increasing or neutral, with the possible exception of minimum wage policies, worker formalization and some regional convergence. We do not find observable variables that help to explain –either through an endowment effect or through a structural price change– a significant part of the decrease in the Gini coefficient of hourly earnings.

However, we do find that certain observable variables explain the reduction in the p90/p10 and p95/p5 interquantile ratios of the wage distribution. Changes in the return to potential experience, equalizing regional earnings differences and possibly the increase in the minimum wage, reduced this indicator and more than compensated the impact of rising educational attainment. However, for the wage at the 10th and 20th quantile in the earnings distribution, the results are similar to those for the Gini coefficient.

These results, together with our initial finding that educational group time effects are important explanatory variables for income dynamics, may suggest that the unexplained fall in labour income inequality may be related to these time effects. However, we present evidence to indicate that this does not seem to be the case. Therefore, it is still an open question as to what factors explain most of the decrease in earnings inequality.

The paper is organized as follows. Section 2 presents the data used in this study. In Section 3 we present a variety of summary statistics of the evolution of labour income distributions and other socioeconomic data. We also describe life cycle average income by cohort and show that stripping out common time effects drastically changes these wage profiles compared to the raw data.

Once the main patterns of the data have been described, Section 4 presents the RIF methodology, a decomposition technique in the spirit of the Oaxaca-Blinder approach but generalized to assess the impact of changes in variables over different distributional statistics. Then Section 5 presents the results of the decompositions. Lastly, the paper concludes in Section 6 summarizing the results and discussing the policy implications.

2. Data

The data used in this paper come from the *Encuesta de Caracterización Socioeconómica Nacional* (CASEN).[8] The CASEN is a nationally and regional representative household survey, covering both urban and rural areas of Chile. The survey was fielded every two years from 1990 to 2000, every three years from 2000 to 2009, and every two years again from 2009 to 2017, which amount to a total of thirteen waves of the CASEN household socioeconomic survey for the 1990–2017 period.[9]

Before 2013 –that is, in the period 1990-2011– the income information in the survey was corrected by the Economic Commission for Latin America and the Caribbean (ECLAC) for non- response and adjustments to national account aggregates. Starting with the 2013 CASEN, this approach (also called the 'historical methodology') was no longer applied. Therefore, to have comparable data for the whole period, our database was constructed using the supplementary CASEN databases available from the MSDF for the period 1990-2011 that record income variables without the adjustments applied by ECLAC.

Our working sample comprises all workers between the ages of 18 and 65 who reported strictly positive hourly labour earnings from their main occupation during the reference year of the survey. Information on total monthly earnings from all jobs is also provided in the surveys. All labour income measures are expressed in real terms using the *Unidad de Fomento* (UF) deflator with base-year 2017. Our primary focus is on the hourly labour earnings measure (henceforth, HLE), which is constructed by dividing the main occupation labour earnings by the weekly hours worked. Further, HLEs are trimmed at the 1st and 99th percentiles by year. Altogether, the full dataset is a pseudo-panel that contains information on 883,399 workers (about 68,000 workers per CASEN wave, on average).

Our analysis includes continuous variables such as the earnings measures

---

8 National Socioeconomic Characterization Survey. This survey is administered by the Ministerio de Desarrollo Social y Familia (MDSF), that is, the Ministry of Social Development and Family.

9 The specific years in which each survey was conducted are 1990, 1992, 1994, 1996, 1998, 2000, 2003, 2006, 2009, 2011, 2013, 2015, and 2017.

just described, schooling, and potential experience in the labour market.[10] Schooling and potential experience are measured in years. The remaining variables are categorical. Demographic characteristics of workers are recorded by a gender dummy and a dummy that differentiates between urban and rural workers. Workers are also classified into 9 different economic sectors. A variable indicating the worker's firm size is also included.[11] We also explore the influence of labour unions by including an indicator variable if the worker is a member of a labour union or not.

We differentiate between three types of workers. First, workers are classified whether they are self-employed or employees according to their answer to a specific question regarding their labour market status. Workers are further divided into formal employees if they reported having a labour contract, if they contributed to the pension system or if they emitted bills or invoices. Those without a contract, who did not contribute to the pension system or did not emit bills or invoices are considered informal workers.[12] Finally, we have also included an indicator variable for workers whose HLE in their main occupation was below the national minimum wage in a given year.

3. Descriptive Statistics, Inequality, And Education

Table 1 presents descriptive statistics for four years of our data (1990, 2000, 2011, and 2017). Average household size of workers decreased from 4.7 in 1990 to 3.7 in 2017. Also, the population has been aging with the average age of workers rising from 36.2 years to 41.1 years in the sample period. The proportion of women in the sample increased from 32.6 % to 43.9 %. Hours worked, on the other hand, has fallen from close to 49.9 hours per week in 1990 to 43.0 hours per week in 2017.[13] The proportion of rural workers fell from 15.1 % to 10.8 % during the period. The number of years of schooling and potential experience increased during the 28-year period. The proportion of formal workers rose from 61.5 % in 1990 to 72.0 % in 2017 while informality fell from 38.5 % to 28.0 % (self-employed workers are added to informal workers). There was a significant increase in real labour incomes during the sample period as well as in the (gross) minimum wage.[14]

10 Potential experience is computed from age and schooling variables as: *exp = age-educ-6 if educ* > 9 and *exp = age-15 if educ* < 9.

11 Firms are classified into 5 groups according to the number of workers: 1, 2-9, 10-49, 50-199 or 200 and more.

12 Many self-employed workers may also be informal, and they are often grouped along with informal workers to estimate the level of informality in the labour market.

13 The legal working week was decreased from 48 to 45 hours per week in 2001.

14 Using the minimum wage net of social security contributions does not change the empirical results of this paper. But, for reasons that will become apparent below, we used the gross minimum wage in what follows.

TABLE 1

DESCRIPTIVE STATISTICS FOR 1990, 2000, 2011 AND 2017

	1990		2000		2011		2017	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Persons per Household	4.96	2.01	4.42	1.89	4.1	1.8	3.73	1.7
Age	36.2	11.7	38.3	11.3	40.1	12.3	41.1	12.6
Women (%)	32.6	46.9	37	48.3	40.3	49.1	43.9	49.6
Hours worked	49.9	15.6	47	15.7	43.3	13.4	43	13.7
Rural (%)	15.1	35.8	10.2	30.3	11	31.3	10.8	31.1
Education (years)	9.86	4.3	11	4	11.5	3.7	12.3	3.67
Potential experience (years)	19	12.3	20.4	11.9	22	13.2	22.3	13.6
Self employed (%)	0	0	21.8	41.3	8.58	28	8.79	28.3
Formal (%)	61.5	48.7	65.1	47.7	71..9	45	72	44.9
Informal (%)	38.5	48.7	13.1	33.8	19.6	39.7	19.2	39.4
Labour income (CLP$/month)	208,167	230,414	366,505	405,139	395,259	418,508	489,170	460,169
Hourly income (CLP$/hour)	1,133	1,247	1,094	2,267	2,423	2,523	2,998	2,789
Total income (CLP$/month)	221,769	264,077	402,309	500,753	436,126	478,226	546,432	573,043
Household income (CLP$/month)	602,487	697,429	979,788	1,115,177	1,107,300	1,084,351	1,320,237	1,498,773
Household per capita income (CLP$/month/person)	147,761	200,717	261,224	339,385	313,549	357,978	415,155	509,579
Minimum wage (CLP$/month)	102,018	0	170,223	0	206,943	0	264,000	0
Below hourly minimum wage (%)	34.4	47.5	30.3	45.9	22.7	41.9	19.9	39.9
N	32,325		71,011		71,113		82,385	

Note: Hourly Labour Earnings (HLEs) trimmed at the 1st and 99th percentiles by year. Labour income and hours worked refer to the main occupation while total income is the sum of all labour earnings. All monetary figures expressed in 2017 Chilean Pesos (CLP$). Sample weights used to calculate descriptive statistics.

Panel A in Figure 1 presents the evolution of monthly average and median income from the main occupation and household per capita income (from all sources of income of household members). There was a steady rise in labour incomes during the 90's but then stagnated during most of the following decade. This was probably a consequence of the international financial crisis that affected Latin America starting in 1998 and earnings only started to rise again after 2010. However, average household per capital disposable income rose continually during this period. Changing demographics, higher government monetary transfers or higher labour market participation rates for women in the household could explain the difference between average household incomes and the evolution of individual earnings.

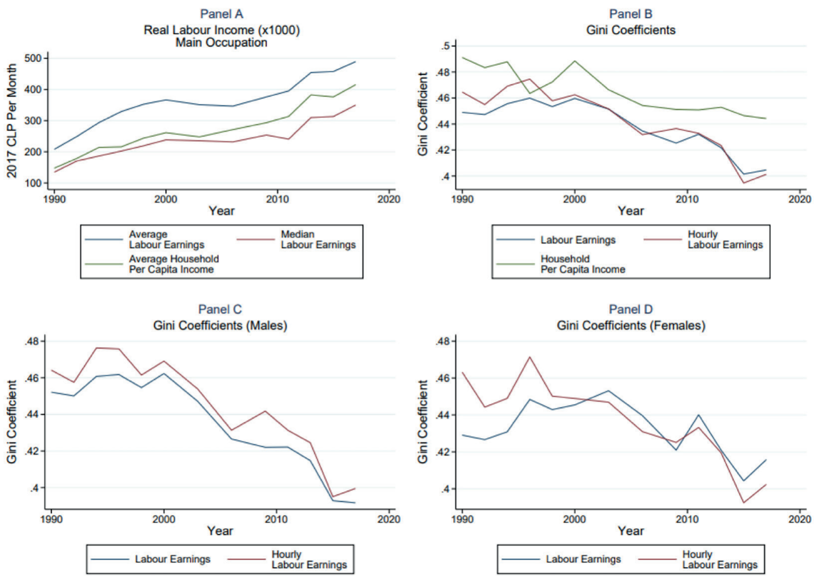
Panel B in Figure 1 presents the evolution of the Gini coefficient using total individual labour income from the main occupation, HLE from the main occupation and total household per capita income. This figure suggests an optimistic view of income inequality dynamics. All indices fell during the period indicating a steady decrease in labour income inequality.

The fall in inequality measured using HLE is significant, from 0.47 to 0.40, close to 15%. Also, after 1998, HLE and total labour earnings from the main occupation have a similar level and evolution, suggesting that in the aggregate the distribution of hours worked did not change significantly.

However, if we analyse the earnings distribution separately by gender, we find there are differences between the distribution of total labour earnings and HLE for females. It can be seen from Panel C of Figure 1 that for males, both Gini coefficients have the same level and evolution, and similar to the aggregate measures shown earlier. Nevertheless, Panel D of Figure 1 for females shows a somewhat different picture. Although the fall in the inequality of hourly earnings is comparable to that for males, the fall in inequality of total labour earnings is lower, indicating that the distribution of hours worked among women changed during the period.

FIGURE 1

LABOUR AND HOUSEHOLD REAL MONTHLY INCOMES AND GINI COEFFICIENTS



Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Labour earnings refer to monthly earnings reported in the main occupation. Hourly earnings is this last figure divide by monthly hours worked. Households per capita income includes all incomes perceived by household members

Figure 2 shows the evolution of other inequality measures. Panel A show the same tendency as noted above. All indices declined substantially after a peak in the mid 90's. Panel B and C show the same information separately for males and females, respectively.

Panels D and E show the evolution of ratios taken at other points of the income distribution. The p95/p90 and p99/p95 ratios exhibit a slight downward tendency during the period. How- ever, inequality seems to have stagnated at the bottom end of the distribution with the p10/p5 and p5/p1 ratios decreasing and then increasing towards the end of the period. Despite this last tendency, these ratios were lower in 2017 than in 1990.

FIGURE 2

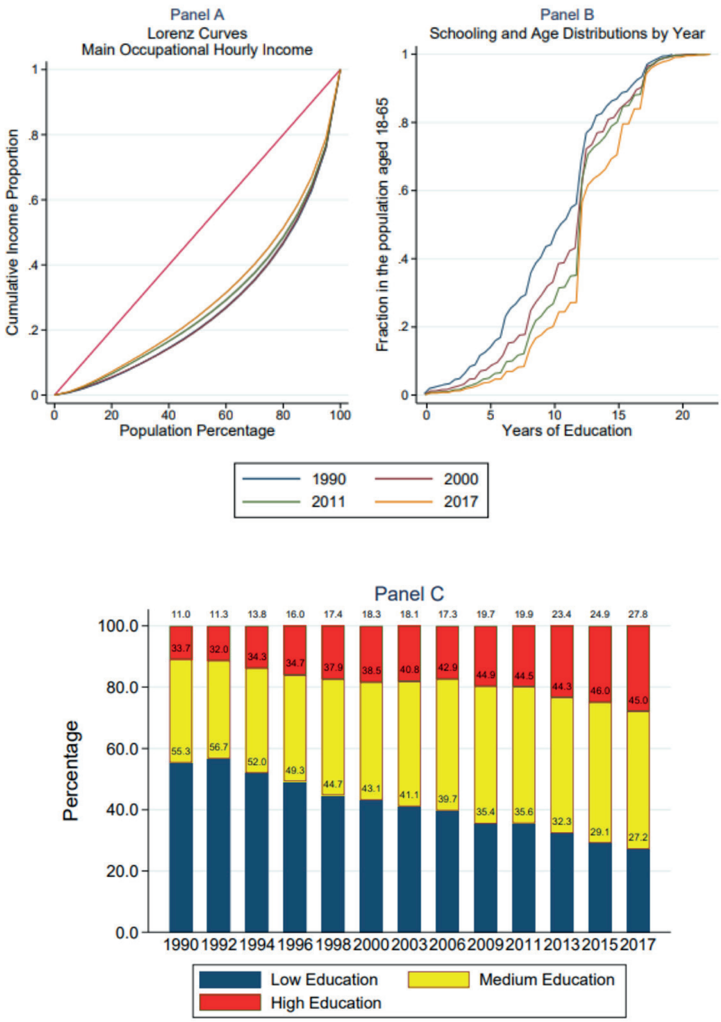
OTHER INEQUALITIES MEASURES (HOURLY EARNINGS)

Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Labour earnings refer to monthly earnings reported in the main occupation.

In sum, the earnings distribution seems to have improved in Chile over the 1990-2017 period. However, this improvement seems to be coming more from the centre of the distribution rather than the tails, as we have just discussed. Graphing the Lorenz curves for selected years (Figure 3, Panel A) shows that the curves for 1990 and 2000 are indistinguishable, however, by 2017 there is a noticeable improvement, particularly in the middle of the curve.

FIGURE 3

LORENZ CURVE (MAIN OCCUPATIONAL HOURLY INCOME), SCHOOLING AND AGE DISTRIBUTIONS BY YEAR AND EDUCATIONAL ATTAINMENT BY YEAR



Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Labour earnings refer to monthly earnings reported in the main occupation. The distributions of years of education are calculated for all the individuals over the estimating sample (formal, informal and self-employed of ages 18-65). The empirical cumulative distribution function of schooling is approximated by an adaptative kernel density estimate to produce a kernel smoothed cumulative distribution function. Low education are workers without a high school degree, medium education are workers with a high school degree and high education are workers with more than high school degree. For every year, observations with negative Hourly Labour Earnings (HLE) and the 99th and 1st percentiles of HLE are trimmed.

The results discussed so far raise several questions. First, some commentators have expressed concerns that income distribution measures in Chile may be biased due to the increasing difficulty of measuring incomes of top earners using the CASEN survey. In Appendix A we explore this hypothesis and conclude that there is no evidence that this is the case.[15]

Another interesting question is how inequality has been affected by the significant increase in the educational attainment of the labour force. Panel B of Figure 3 shows the distribution of schooling for four years of our sample. There was a significant rise in the years of education in the population. While in 1990 nearly 60% of workers had 12 years or less of schooling, by 2017 this proportion had fallen to approximately 30%.

Using the OECD classification of educational attainment, we see that the proportion of workers without a high-school degree (low education) fell from 55.4% in 1990 to 27.2% in 2017 (Figure 3, Panel C). On the other hand, the group with medium educational attainment in- creased from 33.7% to 45% while those with high educational attainment increased from 10.9% to 27.8% between 1990 and 2017.

Whether this sharp increase in the supply of skilled labour affected the educational wage premium is a hypothesis we explore empirically further below.[16] In this section we present the HLE of each educational group by birth cohort in panels A, B and C of Figure 4 for women and panels D, E and F for men.[17]

There is a clear cohort effect for low and medium education female and male workers. Hourly earnings increase with age and for younger cohorts. The age-wage profile is also steeper for younger cohorts. For high education female workers there is also a cohort effect up to the 1980-1985 cohort, but less clear for later cohorts. In the case of high education males, the age-wage profiles are quite similar across cohorts.

These age-wage profiles may be affected by common time effects. This period was characterized by rapid economic growth in earnings. Therefore, in the spirit of Blundell et al. (2023) we explore how the cohort age-wage profiles shown in Figure 4 change if we eliminate educational group specific

15 UNDP (2017) also argues that there is no evidence that under-reporting of top incomes may be generating a dynamic bias in inequality measures. Given the number of means-tested welfare benefits in Chile, there is also the issue of whether low incomes are over reported as low-income individuals are not willing to report their true (higher) earnings in a survey if they think it may affect their eligibility for future benefits. We do not explore this possibility here.

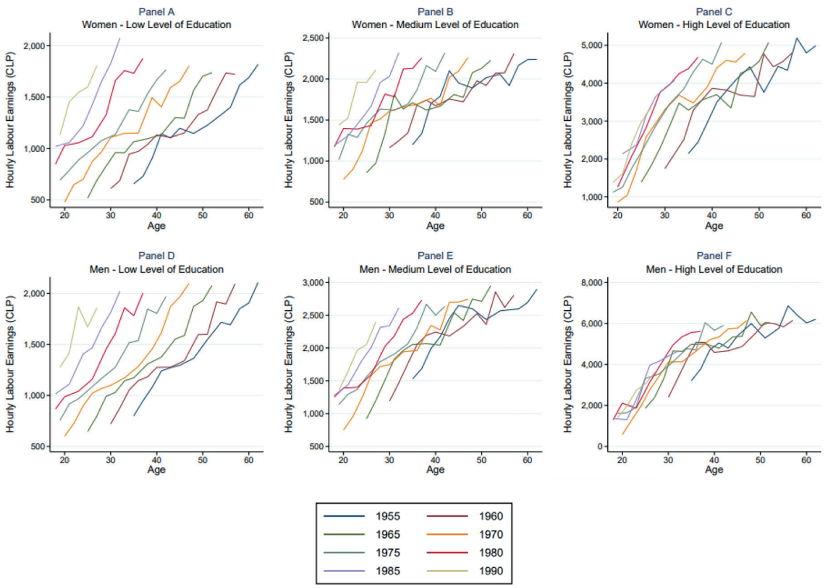
16 This was found to be the case in Brazil by Ferreira et al. (2021) where the increase in educational attainment was neutralized by a decrease in the wage premium.

17 Cohorts are defined by 5-year birth periods starting with the year of the name of the cohort. Cohort 1955, for example, includes all individuals born between 1955 and 1959.

time effects.[18] To this end, for each educational group we ran a regression of individual hourly earnings on cohort specific age effects and common yearly time dummies.[19]

FIGURE 4

AVERAGE HOURLY EARNINGS AGE PROFILES BY COHORT AND WORKERS' EDUCATIONAL LEVEL



Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Cohorts are defined by year groups starting with the year of the name of each cohort.

Figure 5 graphs the cohort specific wage-age profiles for each gender and educational group without the educational group common time effects.[20] These profiles now have the expected concave shape. Comparing with Figure 4 it can see that educational group common time effects have an important impact on these profiles. For low education workers the raw data shows younger cohorts have higher hourly earnings at the same age as older cohorts. However, this seems to be due to the favourable economic conditions faced by these workers early in their working careers. Striping out the common time effects for low educated workers shows that the earning profiles do not change much across cohorts for women. For men, younger cohorts of low education workers earn roughly the same as older cohorts at the same age. For medium and high educated workers, younger cohorts earn less at the same age than older cohorts.

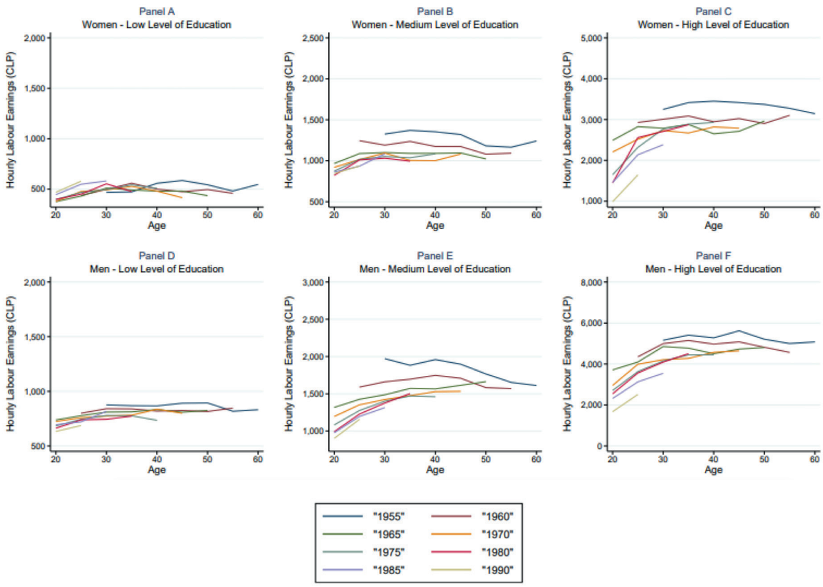
[18] However, unlike Blundell et al. (2023) we do not control for sample selectivity in the observed distribution of wages.

[19] Regression results are available upon request.

[20] The scale in these figures is the same as their counterpart in Figure 4.

FIGURE 5

AVERAGE HOURLY EARNINGS AGE PROFILES WITHOUT EDUCATIONAL GROUP TIME EFFECTS



Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Cohorts are defined by 5 year groups starting with the year of each cohort.

The above comparison suggests that the educational wage premium has fallen through time but owing to favourable economic conditions (common time effects) younger cohorts earn more than their older peers at the same age and educational attainment.

In sum, earnings distribution has improved in Chile between 1990 and 2017. This seems to have come about through less inequality in the middle of the income distribution rather than the tails. It is also associated with an increase in workers with medium and high educational attainment. Hourly earnings for low and medium educational workers are on average higher for younger cohorts. This last effect seems to be related to common time effects within each educational group. For workers with high education, younger cohorts earn about the same as older cohorts at the same age. However, if common time effects are excluded, younger cohorts would earn less than what their older peers earned at the same age.

We now turn to the empirical analysis to ascertain to what extent the reduction in earnings inequality in Chile is related to observable characteristics, either through endowment or price effects.

4. Methodology

Standard regression models focus on measuring the marginal effect of covariates on the conditional expectation of the variable of interest. Therefore, these methods are not useful when the interest lies in the effects of covariates over other aspects of the distribution of the dependent variable, such as its quantiles or higher-order moments.

Recentered influence function (RIF) regressions offer an alternative to estimate marginal effects of covariates over more complex statistics of the distribution of the variable of inter- est. In what follows, we briefly explain the RIF regression and decomposition methods used in this paper.

4.1 RIF Regressions

Let $F_Y(y)$ be the cumulative distribution function (CDF) of a real-valued random variable Y (labour income in our case) which means that $F_Y(y) = \Pr(Y \leq y)$. Henceforth, we omit the argument y in $F_Y(y)$ for the sake of notational simplicity. Furthermore, let $v(F_Y)$ be a functional (or distributional statistic) of the labour income distribution (e.g. mean income, Gini coefficient or Theil index).[21]

Now, if we want to assess the effect on the value of the functional $v(F_Y)$ of an infinitesimal perturbation at the point y, one way to do this is to use the influence function (IF) which is defined as:[22]

(1) $$IF\left\{y, v\left(F_Y\right)\right\} = \lim_{t \to 0} \frac{v\left(\left(1-t\right)F_Y + t\delta_y - v\left(F_Y\right)\right)}{t}$$

where $0 \leq t \leq 1$ and δ_y is a distribution that puts mass only at the value y.

Nevertheless, for an attractive statistical reason stated below in equation (4), Firpo et al. (2009) propose instead the so-called ‘recentered influence function’ (RIF):

(2) $$RIF\left\{y, v\left(F_Y\right)\right\} = v\left(F_Y\right) + IF\left\{y, v\left(F_Y\right)\right\}$$

Following closely Rios-Avila (2020), the foremost statistical properties of the IF and RIF functions are:

---

21 Roughly speaking, functionals are functions where the inputs are themselves functions. The terms functional and distributional statistic are interchangeable throughout this work.

22 More precisely, the influence function corresponds to the one-sided Gâteaux derivative of $v(.)$ at F_Y, in the direction of δ_y. Meanwhile, the Gâteaux derivative is a generalization of the concept of directional derivative, but for functionals instead of functions.

(3) $$\int IF\{y, v(F_Y)\}\, dF_Y = 0$$

(4) $$\int RIF\{y, v(F_Y)\}\, dF_Y = v(F_Y)$$

Thus, equation (3) states that the expected value of IF is 0. Consequently, equation (4) implies that the expected value of RIF is equal to the functional itself, and is the main reason why Firpo et al. (2009) propose using RIF regressions instead of IF regressions.[23]

However, when our approach is a conditional analysis of $v(F_Y)$, and we are interested in exploring the effect of a vector of covariates X on Y, we must be capable of generalizing equation (4).

More formally, let X be a random vector with CDF $F_X(x)$, then when applying the law of iterated expectations to equation (4) to incorporate the effect of covariates, we have that:

(5) $$v(F_Y) = \int RIF\{y, v(F_Y)\}\, dF_Y = \int E[RIF\{y, v(F_Y)\} | X = x]\, dF_X$$

In practice, given an observation yi of the sample data ($y_1, y_2, \ldots, y_N$), the empirical and linear counterpart of the previous equation is a linear regression of the form:

(6) $$RIF\{y_i, v(F_Y)\} = X_i^{'}\beta + \varepsilon_i$$

where ϵ_i is an error term with $E(\epsilon_i) = 0$.

The final step is to calculate the correct marginal effects of the model. Firpo et al. (2009) show that the unconditional partial effect on $v(F_Y)$ induced by a small translation of the distribution of X is given by:

(7) $$\alpha(v) = \int \frac{dE[RIF\{y, v(F_Y)\} | X = x]}{dx}\, dF_X$$

where $\alpha(v)$ is the vector of partial effects of 'small location shifts' in the distribution of X, that is, the effect on $v(F_Y)$ of moving each coordinate of X separately as a location shift assuming that the conditional distribution of Y given X remains constant.

The preceding equation has significant implications as it allows for the determination of the partial effects resulting from a slight shift in the CDF of X in three steps: (1) Regress the RIF of the distributional statistic of interest, on the

[23] Note that estimating the variance of the IF is equivalent to estimating the variance of the RIF, since adding the mean to the IF is just a constant shift in values.

vector of covariates X (RIF regression); (2) compute the marginal effects; and (3) integrate over the values of X. In essence, the partial effect of a covariate on an unconditional quantile of Y can be expressed as a weighted average (over the distribution of X) of the conditional partial effects.[24]

4.2 RIF Decomposition

As RIF regressions are a generalization of the standard regression methods, RIF decompositions are a generalized framework of the standard Oaxaca-Blinder approach for decomposing changes in average earnings (Blinder (1973); Oaxaca (1973)). In fact, a RIF decomposition applied to the mean (i.e. average earnings) yields exactly the Oaxaca-Blinder decomposition. Nevertheless, RIF regressions allow extending this decomposition to any statistic of the earnings distribution v(FY), not just the mean.

Henceforward, we suppress the argument F_Y in $v(F_Y)$. Given two groups (or time periods) indexed by $t = 1, 2$; the overall change in the distributional statistic v (i.e. $\Delta v = v_2 - v_1$) for a counterfactual $v_c = v\left(F_Y^c\right) = v\left(\int F_{Y|X,t=1} dF_{X,t=2}\right)$ can be disaggregated into two components as:

$$\Delta v = v_2 - v_1 = \underbrace{v_2 - v_c}_{\Delta v_S} + \underbrace{v_c - v_1}_{\Delta v_X} \tag{8}$$

where the first term v_S is the structural or price effect that measures how changes over $v(F_Y)$ can be explained by changes in returns or premiums, while the second term v_X is the composition or endowment effect that explains changes in $v(F_Y)$ attributable to changes in covariate composition. Notice that when the counterfactual chosen is $v_c = \overline{X_2'}\hat{\beta}_1$, (8) leads to the standard Oaxaca-Blinder decomposition when the baseline wage structure is given by $\hat{\beta}_1$.

However, the previous RIF decomposition strategy may yield an incorrectly identified counterfactual statistic v_c because the distribution of outcomes and covariates of the counterfactual scenario are not directly observed. Considering this limitation, we use a semiparametric reweighting approximation proposed by Rios-Avila (2020) to identify the counterfactual distribution based on the observed data.

After reweighting factors are estimated using a probit or logit model;[25] $v_c = \overline{X'_c}\hat{\beta}_c$ is estimated by weighted least squares and the new reweighted RIF decomposition components are given by:

---

[24] When the RIF regression is linear with respect to the X variables, the estimated β coefficients can be interpreted in a manner similar to traditional linear regression models. The only difference is that these coefficients must be interpreted as the marginal effect of a slight change in the mean value of X ($E[X]$) on the distributive statistic.

[25] We direct readers to Rios-Avila (2020) for a comprehensive overview of this method.

(9) $$\Delta v = \underbrace{\overline{X}_2'\left(\hat{\beta}_2 - \hat{\beta}_c\right)}_{\Delta v_S^p} + \underbrace{\left(\overline{X}_2' - \overline{X}_c'\right)\hat{\beta}_c}_{\Delta v_S^e} + \underbrace{\left(\overline{X}_c' - \overline{X}_1'\right)\hat{\beta}_1}_{\Delta v_X^p} + \underbrace{\overline{X}_c'\left(\hat{\beta}_c - \hat{\beta}_1\right)}_{\Delta v_X^e}$$

The first term is a pure price or structural effect. The second term, Δv_S^e, is a reweighting error that should go to zero in large samples and can be used as a specification test for the reweighting strategy. If large and significant, this term will be indicating that the counterfactual is not well identified. The third term is a pure endowment or compositional effect while the last term, Δv_X^e, is a specification test for the RIF model. A large and significant value for this error may be indicating that the model is not well specified, and the RIF regression is not providing a good approximation to the distributional statistic v. Furthermore, when the estimated counterfactual v_c coincides exactly with $\overline{X}'_2\hat{\beta}_c$, equation (9) collapses to equation (8) because the counterfactual was correctly specified, and the error terms vanish (i.e. $\Delta v_S^e = \Delta v_X^e = 0$). Notice that if an intercept is included, so that $X_{t,1} = 1$ for $t = 1, 2$; the price effect presents a component given by $\left(\hat{\beta}_{2,1} - \hat{\beta}_{1,1}\right)$ that reflects changes in average returns to unobservable covariates.

5. Results

We begin by following Ferreira et al. (2021) and present some simple Mincer wage equations by year. Table 2 presents the results of OLS regressions on the logarithm of hourly earnings for four different years in our sample. Specifically, we estimate the following model for each year t = 1990, 2000, 2011 and 2017:

(10) $$\ln y_{it} = \alpha + \beta_{1,S} S_{it} + \sum_{n=2}^{4}\left(\beta_{n,S}\frac{S_{it}^n}{10^n}\right) + \beta_{1,E} E_{it} + \sum_{n=2}^{4}\left(\beta_{n,E}\frac{E_{it}^n}{10^n}\right) + D_{it}'\delta + \varepsilon_{it}$$

where y_{it} is hourly earnings, S_{it} is the number of years of education, E_{it} is the years of potential experience, and D_{it}' is a vector of dummy variables for the following categorical variables: gender, minimum wage status, formal employment, rural area, economic sector, and region.[26] For these last variables the

[26] To conserve space, the results for these last two categorical variables are not shown in the table.

omitted categories in each equation are: male, above minimum wage, self-employed (or informal), urban, construction, and Santiago Metropolitan region.

Although difficult to see directly given the fourth order polynomial in years of education, Panel A of Figure 6 graphs the simulated returns to education from the parameters estimated in the equations of Table 2. Interestingly, the educational wage premium did not change much between 1990 and 2011, however the following decade shows an important decrease in this premium.

The opposite occurred with the experience premium as can be seen from Panel B of Figure 6. Although it also decreased during the sample period, this occurred mostly between 1990 and 2011. Other results from Table 2 indicate that the gender wage gap has remained relatively constant over time near 10%. The earnings penalty from having an hourly earnings rate below the minimum wage has decreased by around 12 percentage points.[27] Figure 7 shows the distribution of wages and the minimum wage for this period. It suggests that this minimum may have become more binding through time, reducing the dispersion in the lower end of the distribution, and generating a peak close to this limit.

The formal employment premium was reduced and became negative by 2017. We do not have an explanation for this except that perhaps self-employed workers who do not have to pay social security contributions (pensions and health) may have a net income higher than formal sector workers. The urban rural wage gap also seems to have fallen from 4.8% in 1990 to 2.9% in 2017.

Finally, the last row of Table 2 presents the root MSE of each regression. This statistic decreases through time, indicating that wage dispersion has decreased even after controlling for explanatory variables.

[27] Unlike Dube (2019) for the US, we cannot include the minimum wage directly in the specification since there is no cross-section variation in the value of this variable given that it is applied simultaneously across the whole country.

TABLE 2

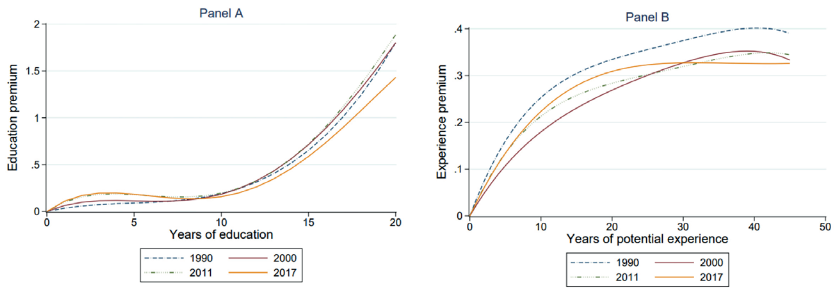
MINCER EQUATIONS LOG-HOURLY EARNINGS (1990-2017)

	(1)	**(2)**	**(3)**	**(4)**
	1990	2000	2011	2017
Education	0.042***	0.080***	0.124***	0.138***
	(0.012)	(0.013)	(0.017)	(0.012)
Education^2/100	-0.733**	-1.814***	-2.658***	-3.049***
	(0.295)	(0.296)	(0.357)	(0.215)
Education^3/1000	0.521**	1.464***	1.972***	2.301***
	(0.255)	(0.247)	(0.277)	(0.154)
Education^4/10000	-0.017	-0.266***	-0.359**	-0.471***
	(0.071)	(0.067)	(0.071)	(0.037)
Potential Experience	0.041***	0.026***	0.034***	0.032***
	(0.005)	(0.005)	(0.005)	(0.003)
Potential Experience^2/100	-0.198***	-0.104***	-0.162***	-0.115***
	(0.045)	(0.040)	(0.041)	(0.028)
Potential Experience^3/1000	0.048***	0.026**	0.038***	0.017**
	(0.014)	(0.012)	(0.013)	(0.008)
Potential Experience^4/10000	-0.004***	-0.003**	-0.003**	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Female	-0.098***	-0.087***	-0.104***	-0.097***
	(0.009)	(0.009)	(0.009)	(0.006)
Below Hourly Minimum Wage	-0.938***	-0.891***	-0.819***	-0.819***
	(0.007)	(0.006)	(0.009)	(0.005)
Formal Employee	0.021***	0.017**	-0.053***	-0.038***
	(0.008)	(0.008)	(0.010)	(0.006)
Rural	-0.046***	-0.023***	-0.026***	-0.020***
	(0.010)	(0.007)	(0.008)	(0.007)
Constant	6.441***	6.983***	7.067***	7.330***
	(0.031)	(0.031)	(0.033)	(0.027)
N	32,203	70,803	71,071	81,539
Adjusted R^2	0.622	0.646	0.592	0.555
Root MSE	0.496	0.468	0.465	0.455

Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. For the categorical variables in each equation the omitted categories are: male, above minimum wage, self-employed (or informal), urban, construction, and Santiago M.A. region. Survey frequency weights used in estimation. * *p<0.10*, ** *p<0.05*, *** *p<0.01*.

FIGURE 6

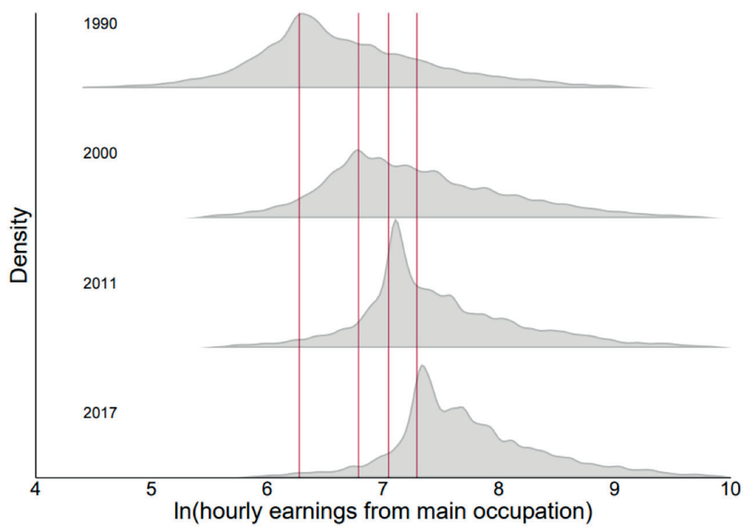
EDUCATION AND EXPERIENCE WAGE PREMIUMS BY YEAR



Note: Simulations based on the results of Table 2.

FIGURE 7

HOURLY WAGE DENSITY AND MINIMUM WAGE BY YEAR



Note: Vertical lines correspond to the logarithm of gross minimum hourly wage for the years 1990, 2000, 2011 and 2017. Gaussian kernel densities estimated. Hourly labour earnings from main occupation trimmed at the 1st and 99th percentiles by year. Survey frequency weights used.

We next present decomposition results for log HLE. The specification of the Mincer equations serves as a guide to the potential explanatory variables to use in the RIF decompositions. In what follows then, we use these same variables to determine possible factors driving the changes in the various inequality statistics.

Correct standard error estimation requires bootstrapping.[28] Since the CASEN surveys have a stratified and clustered sampling design, knowledge of this sampling structure was required to implement the bootstrapping method. However, for the 1990 CASEN survey the publicly available database does not include a strata variable nor a cluster variable. Therefore, in what follows, we estimate the decomposition using data from 1992 to 2017.

Table 3 presents the results of a classical Oaxaca-Blinder decomposition of the average log-hourly earnings. The first column presents the results for the whole period (1992-2017) while the three other columns present the results for sub-periods (1992-2000, 2000-2011 and 2011-2017, respectively).[29]

The decomposition for the whole period shows that log HLE increased by close to 0.91 between 1992 and 2017 (or 148% increase in real hourly earnings). Close to 41% of this difference was due to a composition effect (explained) while the other 59% points was due to changes in premiums (unexplained). This same pattern appears in the sub-periods with the exception of the 2000-2011 period where the compositional effect was larger than the change in premiums. Note that the specification error is small or statistically insignificant, while the reweighting error is significant but much smaller than the unexplained effect.

As for the compositional effect, years of education account for an important positive impact on earnings between 1992 and 2017. Rising years of education with the returns to education of the counterfactual scenario would have increased mean log hourly earnings by 0.21 log points. However, returns to education fell between 2011 and 2017 (pure unexplained), reducing the impact of rising educational attainment on wages during this last period. Thus, increasing years of education in the population did have an impact on average log earnings but due to the fall in the educational premium the net effect was reduced almost by half.

Compositional effects of potential experience also had a positive effect on mean earnings. However, this effect is much smaller than changes in educational levels and seems to be present mostly in the 1992-2000 period, when there were still high returns to potential experience. The fall in the returns to experience was much larger than the compositional effect, implying that the net effect of experience on average earnings was negative.

The increase in the female labour participation would have decreased mean log earning a bit. The gender wage gap during the period does not seem to have changed much as already noted above with the Mincer equation results.

---

28 Estimations were undertaken using the oaxaca_rif module in Stata. See Rios-Avila (2020). The Probit reweighing option was used for the counterfactual. Categorical variables (economic sector and regional variables) were normalized as in Yun (2005).

29 It should be noted that these results are robust to changes in the excluded category.

As less workers earned an hourly wage below the minimum wage during the period, the compositional impact on hourly wages is positive. In addition, average earnings conditional on being below the minimum wage also rose during the period. This suggests a potential effect of minimum wages on hourly incomes for this group of workers.

There was a slight decrease in wages due to compositional changes in formality (workers with a formal contract). However, there was an increase in the formality wage premium during the whole period that more than compensated for the compositional effect.

The rural-urban earnings gap did not contribute much to the change in mean log earnings during the period, while there was a very small but positive compositional effect. The regional and economic sector did contribute something to the increase in log hourly wages through the compositional effect, although in the case of region it was more than compensated by the pricing effect.

By far the largest impact on mean log HLE is the change in the non-observable skills premium (the constant in the unexplained results).

In summary then, the decomposition of mean HLE during the 1992-2017 period is explained by a compositional effect due to an increase in years of education, an increase in years of experience, and a fall of workers earning below the minimum wage, together with a net effect due to changes in structural returns, being the returns to unobserved skills by far the most important, while the observed structural changes had a negative effect on earnings (mainly a fall in the returns to education and potential experience).

TABLE 3
RIF DECOMPOSITION – MEAN OF LOG HOURLY EARNINGS (1992-2017)

	(1) 1990-2017	(2) 1992-2000	(3) 100-2011	(4) 2011-2017
Overall				
Post	7.747***	7.295***	7.489***	7.747***
	(0.003)	(0.003)	(0.003)	(0.003)
Counterfactual	7.204***	6.941***	7.420***	7.610***
	(0.004)	(0.004)	(0.004)	(0.003)
Pre	6.838***	6.838***	7.295***	7.489***
	(0.004)	(0.004)	(0.003)	(0.003)
Difference	0.910***	0.457***	0.193***	0.259***
	(0.005)	(0.005)	(0.004)	(0.004)
Explained	0.366***	0.103***	0.125***	0.121***
	(0.006)	(0.006)	(0.005)	(0.004)
Unexplained	0.544***	0.354***	0.068***	0.138***
	(0.005)	(0.005)	(0.005)	(0.004)
Explained				
Total	0.366***	0.103***	0.125***	0.121***
	(0.006)	(0.006)	(0.005)	(0.004)
Pure Explained	0.351***	0.105***	0.123***	0.121***
	(0.005)	(0.004)	(0.004)	(0.003)
Specification Error	0.015***	-0.002	0.002	0.001
	(0.004)	(0.004)	(0.003)	(0.003)
Pure Explained				
Education	0.211***	0.086***	0.036***	0.089***
	(0.003)	(0.002)	(0.002)	(0.002)
Experience	0.026***	0.016***	0.004***	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Region	0.002***	0.002***	0.002***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Sector	0.013***	0.008***	0.005***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Size	0.019***	0.013***	0.011***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Female	-0.008***	-0.003***	-0.004***	-0.005***
	(0.001)	(0.000)	(0.000)	(0.000)
Below Hourly Minimum Wage	0.090***	-0.017***	0.072***	0.038***

	(0.003)	(0.003)	(0.002)	(0.002)
Formal Employee	-0.004***	-0.001***	-0.003***	-0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Rural	0.002***	0.001***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained				
Total	0.544***	0.354***	0.068***	0.138***
	(0.005)	(0.005)	(0.005)	(0.004)
Reweighting Error	-0.057***	-0.009**	-0.009**	-0.039***
	(0.004)	(0.004)	(0.004)	(0.003)
Pure Unexplained	0.601***	0.364***	0.078***	0.177***
	(0.003)	(0.003)	(0.003)	(0.003)
Pure Unexplained				
Education	-0.102***	0.002	-0.009	-0.108***
	(0.034)	(0.024)	(0.029)	(0.028)
Experience	-0.188***	-0.100***	-0.036***	-0.023*
	(0.016)	(0.017)	(0.014)	(0.013)
Region	-0.040***	0.002	-0.026***	-0.003
	(0.005)	(0.004)	(0.004)	(0.004)
Sector	0.041***	0.025***	-0.003	0.025***
	(0.005)	(0.004)	(0.004)	(0.004)
Size	-0.004***	0.000	-0.014***	-0.003***
	(0.001)	(0.001)	(0.002)	(0.001)
Female	-0.002	0.003	-0.004	0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
Below Hourly Minimum Wage	0.036***	0.020***	0.013***	0.002*
	(0.002)	(0.002)	(0.002)	(0.001)
Formal Employee	0.029***	0.019***	-0.036***	0.012**
	(0.007)	(0.005)	(0.006)	(0.006)
Rural	0.000	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.830***	0.391***	0.194***	0.268***
	(0.039)	(0.031)	(0.034)	(0.032)
Post *N*	71,431	67,213	61,188	71,431
Counterfactual *N*	43,374	43,374	67,213	61,188
Pre *N*	43,374	43,374	67,213	61,188

Note: Education and experience are the sum of the coefficients from fourth order polynomials in years. Regional and economic sector categorical variables are normalized in the estimation and coefficients results are summed. A logit reweighting model is used to define the counterfactual. * *p<0.10*, ** *p<0.05*, *** *p<0.01*.

We next turn to the RIF decomposition of the Gini coefficient of hourly earnings. The results are shown in Table 4.[30] The specification error is significant in all the columns, but small in the three sub-periods. The reweighting error is small in all models indicating a good reweighting strategy defining the counterfactual for the decomposition.

The first thing to note is that there was a fall in the Gini coefficient of HLE during the whole period from 0.456 to 0.405, explained by decrease from 2000 to 2017 (it increased from 1992 to 2000). The second thing to note is that the effect of compositional changes would have increased inequality during the whole period, driven mainly by the years of education but countered to some extent by the fall in workers earning below the minimum wage.

Overall, earnings inequality only fell because there was a large effect from structural changes (price or premium changes) that more than compensated for the compositional effects.

Rising educational attainments would have significantly increased earnings inequality given the counterfactual returns to education. This is not wholly unexpected as more workers may shift to the convex part of the educational wage premium curve, increasing earnings inequality. This is what Ferreira et al. (2021) call the "The Paradox of Progress" (Bourguignon et al. (2005)). However, what is striking is that the change in the educational earnings premium would also have increased inequality. This implies that the fall in the educational premium was not equality enhancing.

Something similar can be said regarding potential experience. Both the compositional and the structural effects imply a more unequal earnings distribution although its magnitude is smaller than in the case of education.

---

[30] In addition, we further explore the effects of labour unionization in Appendix B, Table 8. The worker's labour union variable can only be constructed in the 1994 and 2017 CASEN waves. Our estimates suggest that unionization of workers does not have a significant effect on reducing labour income inequality. We also estimated the RIF models including the interaction of the economic sector dummies with the copper price for the respective year. The results were unchanged to those presented here.

TABLE 4

RIF DECOMPOSITION – GINI COEFFICIENT OF HOURLY EARNINGS (1992-2017)

	(1) **1990-2017**	**(2)** **1992-2000**	**(3)** **100-2011**	**(4)** **2011-2017**
Overall				
Post	0.405***	0.464***	0.438***	0.405***
	(0.001)	(0.001)	(0.001)	(0.001)
Counterfactual	0.483***	0.494***	0.449***	0.450***
	(0.001)	(0.002)	(0.001)	(0.001)
Pre	0.456***	0.456***	0.464***	0.438***
	(0.002)	(0.002)	(0.001)	(0.001)
Difference	-0.050***	0.009***	-0.026***	-0.033***
	(0.002)	(0.002)	(0.002)	(0.002)
Explained	0.027***	0.038***	-0.015***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Unexplained	-0.077***	-0.030***	-0.011***	-0.045***
	(0.002)	(0.002)	(0.002)	(0.002)
Explained				
Total	0.027***	0.038***	-0.015***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)
Pure Explained	0.086***	0.050***	-0.013***	0.019***
	(0.002)	(0.002)	(0.001)	(0.001)
Specification Error	-0.059***	-0.012***	-0.002	-0.008***
	(0.003)	(0.002)	(0.002)	(0.002)
Pure Explained				
Education	0.095***	0.036***	0.009***	0.034***
	(0.002)	(0.001)	(0.001)	(0.001)
Experience	0.009***	0.005***	0.001***	0.000
	(0.001)	(0.000)	(0.000)	(0.000)
Region	0.001***	0.001***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Sector	0.005***	0.002***	0.003***	-0.001***
	(0.001)	(0.001)	(0.000)	(0.000)
Size	0.005***	0.004***	-0.004***	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Female	-0.002***	-0.001***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Below Hourly Minimum Wage	-0.023***	0.004***	-0.017***	-0.009***

	(0.001)	(0.001)	(0.001)	(0.000)
Formal Employee	-0.004***	-0.001***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Rural	-0.000	-0.000	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Unexplained				
Total	-0.077***	-0.030***	-0.011***	-0.045***
	(0.002)	(0.002)	(0.002)	(0.002)
Reweighting Error	0.004***	0.000	0.003***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Pure Unexplained	-0.082***	-0.030***	-0.014***	-0.049***
	(0.002)	(0.002)	(0.002)	(0.002)
Pure Unexplained				
Education	0.119***	0.026*	0.020	0.048***
	(0.017)	(0.014)	(0.018)	(0.017)
Experience	0.066***	-0.002	0.034***	0.011
	(0.008)	(0.010)	(0.009)	(0.008)
Region	-0.016***	-0.009***	-0.000	-0.009***
	(0.002)	(0.003)	(0.003)	(0.002)
Sector	-0.002	0.004	-0.010***	0.007***
	(0.002)	(0.003)	(0.002)	(0.002)
Size	0.002***	0.000	0.002	-0.000
	(0.000)	(0.000)	(0.001)	(0.000)
Female	0.0002	-0.003*	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Below Hourly Minimum Wage	-0.016***	-0,007***	-0.012***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Formal Employee	-0,024***	-0.004	-0.016***	-0.027***
	(0.003)	(0.003)	(0.004)	(0.004)
Rural	0.001	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.214***	-0.034*	-0.034	-0-076***
	(0.020)	(0.019)	(0.021)	(0.020)
Post *N*	71,431	67,213	61,188	71,431
Counterfactual *N*	43,374	43,374	67,213	61,188
Pre *N*	43,374	43,374	67,213	61,188

Note: Education and experience are the sum of the coefficients from fourth order polynomials in years. Regional and economic sector categorical variables are normalized in the estimation and coefficients results are summed. A logit reweighting model is used to define the counterfactual. * *p<0.10*, ** *p<0.05*, *** *p<0.01*.

As already noted, rising minimum wages seems to have compressed the bottom of the earnings distribution. This may explain the fall in the number of workers earning below the minimum wage and decreasing inequality. In addition, the reduction in the wage penalty for these workers adds to this last effect, leading to an overall decrease in the Gini coefficient.

The rise in formal employment also reduced the Gini coefficient but only by a small amount. However, there was a significant effect from the change in the formality wage premium, which became negative at the end of our sample period.

By far the most important impact on the Gini coefficient was the negative and large effect of non-observable skills premium. Therefore, apart from the minimum wage and formality, the other observable variables that we are controlling for in the decomposition do not explain the fall in HLE inequality in Chile.

Table 5 presents the decomposition results for other distributional statistics of the earnings distribution for the 1992-2017 period. The first column presents the results for the interquantile ratio p90/p10, while the second column those for the interquantile ratio p95/p5. The last two columns present results for quintile p20 and p10, respectively.

In the case of the interquantile ratios, they both fell during the period. The p90/p10 ratio fell from 7.02 to 5.45, a drop of 22.4% while the p95/p5 ratio fell from 12.71 to 10.17 a drop of 20%. However, the composition effect was positive in both cases, indicating that that changes in the composition of workers would have increased these ratios, and they only fell due to structural (price) changes. These are mainly associated with changes in the experience premium, the fall in the wage penalty of earning below the minimum wage and some regional convergence in earnings during the period.

Notably the fall in the p90/p10 ratio is more strongly related to the minimum wage effect than in the other results, although changes in unobservable skills was also important. The same patterns hold for the p95/p5 ratio.

TABLE 5
RIF DECOMPOSITION – OTHER DISTRIBUTIONAL STATISTICS
OF HOURLY EARNINGS (1992-2017)

	(1) p90/p10	(2) p95/p5	(3) p20	(4) p10
Overall				
Post	5.453***	10.171***	1467.396***	1127.574***
	(0.045)	(0.119)	(2.770)	(4.985)
Counterfactual	9.878***	19.267***	627.839***	472.747***
	(0.081)	(0.179)	(3.131)	(2.818)
Pre	7.018***	12.712***	520.106***	392.625***
	(0.071)	(0.200)	(2.239)	(2.181)
Difference	-1.565***	-2.541***	947.290***	734.950***
	(0.084)	(0.233)	(3.561)	(5.441)
Explained	2.860***	6.555***	107.733***	80.122***
	(0.108)	(0.268)	(3.849)	(3.564)
Unexplained	-4.425***	-9.096***	839.557***	654.828***
	(0.093)	(0.215)	(4.180)	(5.726)
Explained				
Total	2.860***	6.555***	107.733***	80.122***
	(0.108)	(0.268)	(3.849)	(3.564)
Pure Explained	4.047***	10.072***	84.403***	55.920***
	(0.087)	(0.243)	(2.784)	(2.566)
Specification Error	-1.186***	-3.517***	23.329***	24.202***
	(0.122)	(0.335)	(2.634)	(3.468)
Pure Explained				
Education	3.962***	9.512***	2.765**	4.399***
	(0.075)	(0.207)	(1.165)	(1.578)
Experience	0.442***	0.973***	-0.524	0.173
	(0.030)	(0.083)	(0.569)	(0.770)
Region	0.013	0.050*	0.223	1.244***
	(0.009)	(0.027)	(0.202)	(0.311)
Sector	0.183***	0.622***	-1.655***	-0.851
	(0.031)	(0.086)	(0.599)	(0.808)
Size	0.367***	0.679***	-1.262**	-2.489***
	(0.026)	(0.075)	(0.553)	(0.748)
Female	-0.040***	-0.210***	-1.328***	-3.510***
	(0.015)	(0.043)	(0.311)	(0.434)
Below Hourly Minimum Wage	-0.711***	-1.069***	83.733***	51.596***

	(0.025)	(0.056)	(2.295)	(1.474)
Formal Employee	-0.165***	-0.422***	2.238***	4.015***
	(0.013)	(0.036)	(0.227)	(0.348)
Rural	-0.005	-0.062*	0.213	1.344***
	(0.012)	(0.034)	(0.245)	(0.337)
Unexplained				
Total	-4.425***	-9.096***	839.557***	654.828***
	(0.093)	(0.215)	(4.180)	(5.726)
Reweighting Error	0.148***	0.326***	-44.149***	-27.687***
	(0.057)	(0.098)	(3.908)	(2.700)
Pure Unexplained	-4.573***	-9.422***	883.706***	682.514***
	(0.083)	(0.203)	(0.896)	(4.480)
Pure Unexplained				
Education	2.106**	7.481***	-23.699***	-31.154
	(0.852)	(2.119)	(8.864)	(44.510)
Experience	-1.318***	-2.782***	-35.457***	62.789***
	(0.405)	(1.006)	(4.254)	(20.749)
Region	-0.730***	-1.543***	-1.665	-3.682
	(0.122)	(0.303)	(1.279)	(6.268)
Sector	0.347***	0.970***	1.762	8.169
	(0.115)	(0.287)	(1.177)	(6.219)
Size	-0.055***	0.050	-0.373**	-0.759
	(0.018)	(0.043)	(0.187)	(0.884)
Female	0.308***	-0.068	-0.240	-19.005***
	(0.081)	(0.201)	(0.848)	(4.167)
Below Hourly Minimum Wage	-2.632***	-3.659***	-42.011***	-197.932***
	(0.050)	(0.117)	(0.580)	(2.724)
Formal Employee	-0.402**	-0.715*	8.067***	206.044***
	(0.170)	(0.424)	(1.771)	(8.959)
Rural	-0.008	-0.273***	-0.887**	-0.885
	(0.034)	(0.084)	(0.364)	(1.666)
Constant	-2.191**	-8.883***	978.207***	658.930***
	(0.996)	(2.477)	(10.385)	(51.779)
Post *N*	71,431	71,431	71,431	71,431
Counterfactual *N*	43,374	43,374	43,374	43,374
Pre *N*	43,374	43,374	43,374	43,374

Note: Education and experience are the sum of the coefficients from fourth order polynomials in years. Regional and economic sector categorical variables are normalized in the estimation and coefficients results are summed. A logit reweighting model is used to define the counterfactual. * *p<0.10*, ** *p<0.05*, *** *p<0.01*.

As for the earning of the first quintile and decile, these both increased significantly during the period. However, in both cases endowment changes explain only a small fraction of this increase. Rather almost all of the increase can be attributed to structural changes of which the returns to non-observable skills is by far the most important, together with formalization of workers in the case of the 20th quintile.

Given the educational group common time effects discussed in Section 3, a question arises as to what extent these time effects may explain the results found in this section related to increases in unobservable skill premiums. To this end, separate mincer equations were estimated for each educational group including a common time effect. Then, the wage rate for each observation was stripped of its corresponding educational group year effect. With these last observations, the Gini coefficients were recalculated and compared to those of the original data. The results are presented in Figure 8.

FIGURE 8

GINI COEFFICIENTS OF HOURLY WAGES WITH AND WITHOUT EDUCATIONAL GROUP COMMON TIME EFFECTS

Note: All measures are calculated over the estimating sample (formal, informal, and self-employed of ages 18-65). Hourly labour earnings trimmed at the 1st and 99th percentiles by year. Labour earnings refer to monthly earnings reported in the main occupation. Hourly earnings is the last figure divide by monthly hours worked.

It can be seen from this last figure that the Gini coefficient without the common educational group time effects has a similar evolution compared to those form the raw data. If at all, inequality would have been lower without the common time effects during the last period of the data. Therefore, the unexplained fall in the Gini coefficient of labour income inequality in Chile does not seem to be related to these time effects, although more research is warranted on this issue.

6. Conclusions

In this paper we have described the changes in workers' hourly earnings distribution for Chile from 1990 to 2017. Consistent with other Latin American experiences, there was a significant fall in earnings inequality after the year 2000. The Gini coefficient on hourly labour earnings decreased by around 15% between 1990 and 2017. This period is also characterized by a significant increase in educational attainments and average earnings. While the increase in educational attainment explains part of the increase in average earnings, we find this was countered by a fall in the returns to education with a small overall effect.

A decomposition of the changes in earnings inequality reveals that most of the fall in in- equality cannot be attributed to changes in observed endowments of workers (compositional changes) or changes in the returns to observable skills (structural changes). The significant increase in educational attainment and the fall in the educational earnings premium, would both have increased earnings inequality. Something similar occurs for experience. Overall, these variables do not explain the fall in earnings inequality. Neither does the observed increase in female labour participation or the decrease in the gender wage gap.

There does seem to be an effect related to the minimum wage, formalization, and some regional convergence in reducing earnings inequality. However, most of the fall (as well as the increase in average earnings) is explained by unobservable effects. This may due to changes in policy or premiums in skills that we are not controlling for. Common educational group time effects do not seem to explain this fall either.

One possible hypothesis for our results are changes in the quality of education. There is some evidence that standardized test scores have narrowed between children of families in the upper- and lower-income quintiles (Wales et al. (2014)). Thus, it may well be that wages among workers of the same educational level change among cohorts. This is also suggested by Figure 4. Low education workers' wages have increased for younger cohorts while they have remained the same for high educational workers. However, due to space limitations we leave this issue for future research.

Another possibility is that a rise in public employment or in the public sector wage premium occurred during this period and this may help explain the fall in inequality.[31] It is common in Latin America for the public sector to pay higher salaries to less educated persons than those paid by the private sector, and lower salaries for qualified professionals. This generates two very different wage distributions (private versus public) with the overall distribution being the sum of both.[32]

We tested this idea by first identifying public sector workers in our data (either in the central or local governments). This was only available from 2000 to 2017. During this last period, we see a monotonic fall in the share of public sector workers, from 8.6% in 2000 to a minimum of 5,9% in 2009, rising monotonically reaching 8.7% in 2017. Therefore, public sector employment had a U-shaped dynamic with no increase in the share of public employment between 2000 and 2017. To test whether a wage premium effect might be present, we estimated the RIF regressions for the mean of log incomes using a public sector dummy variable and although both the compositional effect and the wage premium effect were statistically significant, the coefficient estimates were very small and did not change our overall conclusions.

Although the idea that public sector employment generates two different distributions does not seem to be consistent with the data, the general idea of a mixture of distributions of subpopulations is intriguing. Chumacero and Paredes (2005) find that the heterogeneity in the population in Chile can be characterized by, at least, two populations, with different returns to schooling and different volatilities. This could be explained by differences in the quality of education for each group. Depending on how changes in the composition of the population evolved, this could also account for the decrease in inequality. For example, if more of the population transitioned towards the first distribution, one would expect decreases in inequality because of the reduced variance of the unobservable component of that distribution.

In sum, it is still an open question as to what caused the unexplained fall of labour earnings inequality in Chile. More research is required to ascertain what these unobservable effects are. For example, they could be related to improvements in the quality of elementary and secondary education or technological change that favoured the unobservable skills of younger less educated workers, or changes in the composition of sub-populations characterized by different

31 We thank Claudio Sapelli for suggesting this explanation. He commented that his research for Uruguay seems to point in this direction, with a bi-modal distribution owing to the public sector wage structure.

32 For the modelling of the bi-modal income distribution in Chile using a mixed distribution see Chumacero and Paredes (2005).

distributions of unobservables. Another possibility is that the local approximation of the RIF decomposition approach is not flexible enough to capture the long-term dynamics of labour earnings inequality in Chile.

References

Behrman, J. R. (2011). How much might human capital policies affect earning inequality and poverty? *Estudios de Economía*, Vol. 38(1), 9–41.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, Vol. 8(4), 436–455.

Blundell, R., Lopez, H. and Ziliak, J. P. (2023). Labour market inequality and the changing life cycle profile of male and female wages. IFS Working Paper 23/16, Institute for Fiscal Studies, London.

Bourguignon, F., Ferreira, F. H. and Lustig, N. (2005). The microeconomics of income distribution dynamics in East Asia and Latin America.

Chumacero, R. and Paredes, R. (2005). Characterizing income distribution for poverty and inequality analysis. *Estudios de Economía*, Vol. 32(1), 97–117.

Contreras, D. and Gallegos, S. (2011). Desigualdad salarial en américa latina: Una década de cambios. *Revista Cepal*, Vol. 103(1), 27–45.

Dube, A. (2019). Minimum wages and the distribution of family incomes. *American Economic Journal: Applied Economics*, Vol. 11(4), 268–304.

Eyzaguirre, N. (2019). Desigualdad: Raíces históricas y perspectivas de una crisis. Penguin Random House Grupo Editorial, Santiago.

Ferreira, F. H., Firpo, S. P. and Messina, J. (2021). Labor market experience and falling earnings, inequality in Brazil: 1995–2012, *The World Bank Economic Review*, Vol. 36(1), 37–67.

Firpo, S. P., Fortin, N. M. and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, Vol. 77(3), 953–973.

Flores, I., Sanhueza, C., Atria, J. and Mayer, R. (2019). Top incomes in Chile: A historical perspective on income inequality, 1964-2017. *The Review of Income and Wealth*, Vol. 66(4), 1–25.

Fortin, N. M., Lemieux, T. and Firpo, S. P. (2011). Decomposition methods in economics, in O. Ashenfelter and D. Card (Eds.), *Handbook of labor economics*. Vol. 4a, 1–102. Great Britain: North Holland.

Heathcote, J., Storesletten, K. and Violante, G. (2005). Two views of inequality over the life cycle. *Journal of the European Economic Association*, Vol. 3(2/3), 765–775.

MSDF and UNDP (2020). Evolución de la pobreza: 1990-2017: ¿Cómo ha cambiado chile? Ministerio de Desarrollo Social y de la Familia, y Programa de las Naciones Unidas para el Desarrollo.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, Vol. 14(3), 693–709.

Rios-Avila, F. (2020). Recentered influence functions (rifs) in stata: Rif regression and rif decomposition. *The Stata Journal*, Vol. 20(1), 51–94.

Rodríguez, J. (2017). Desarrollo y desigualdad en chile (1850-2009): Historia de su economía política. Santiago: Dirección de Bibliotecas, Archivos y Museos, Centro de Investigaciones Diego Barros Arana, Colección Sociedad y Cultura.

Sapelli, C. (2011). A cohort analysis of the income distribution in Chile. *Estudios de Economía*, Vol. 38(1), 223–242.

Sapelli, C. (2016). Chile: ¿más equitativo? una mirada a la dinámica social del chile de ayer, hoy y mañana. Ediciones UC, Santiago.

UNDP (2017). Desiguales: Orígenes, cambios y desafíos de la brecha social en chile. Informe, Programa de las Naciones Unidas para el Desarrollo.

Wales, J., Ali, A and Nicolai, S. (2014). Improvements in the quality of basic education: Chile's experience [with F. Morales and D. Contreras]. Case Study Report, Overseas Development Institute.

Yun, M. S. (2005). A simple solution to the identification problem in detailed wage decompositions. *Economic Inquiry*, Vol. 43(4), 766–772.

Appendix

A. Top household incomes

In this Appendix, we explore whether the CASEN survey data may contain a bias due to an increasing difficulty in measuring top incomes. The point we want to explore here is a dynamic one. It is probable that the richest 1% of households are underrepresented in the CASEN survey or, when surveyed, they do not reveal their true income. This would bias the level of any income inequality measure but not its evolution –which is our interest in this paper– unless this under-reporting has increased through time.

Table 6 presents the number of observations in each year of the survey of households earning more than 5 million CLP per month. This number represents roughly 1,5% of the highest household income level in 2017.

TABLE 6
PERCENTAGE OF HOUSEHOLDS REPORTING OVER 5 MILLION CLP

	(1) %	(2) %	(3) %	(4) %
1990	0.30	0.35	0.00	0.00
1992	0.59	0.75	0.00	0.00
1994	0.47	0.78	0.00	0.00
1996	0.46	0.98	0.00	0.00
1998	0.54	0.87	0.00	0.00
2000	0.46	1.51	0.18	0.52
2003	0.64	1.30	0.13	0.30
2006	0.59	1.17	0.08	0.18
2009	0.43	1.50	0.15	0.50
2011	0.91	1.55	0.32	0.56
2012	1.18	2.02	0.55	1.02
2015	1.52	1.88	0.72	0.88
2017	1.72	2.00	0.92	1.00
Trimming	No	No	Yes	Yes
Exp. Factors	No	Yes	No	Yes

Note: The percentages in the third and fourth columns are calculated trimming household income at the 1st and 99th percentiles each year.

The first two columns of the table present the percentage of households earning more than 5 million CLP before trimming the 1st and 99th percentiles of each years' data. The first column presents the raw percentages while the second uses the survey expansion factors. There is no evidence that the number of high-income households in the survey has decreased during the period. Naturally, as incomes have risen the number of high-income households has increased in tandem. The third and fourth columns of the table show the same information but after trimming the data each year. Again, there is no evidence that the top earning households have diminished in the sample.

The data of Table 6 may not be very informative as incomes have risen through time. However, Table 7 shows the average income of households earning more than 5 million CLP per month. There is no discernible pattern that points to a potential under-reporting of incomes for this group. Average monthly incomes have remained fairly constant throughout this period for this group of households.

TABLE 7

PERCENTAGE OF HOUSEHOLDS REPORTING OVER 5 MILLION CLP

	(1) **CLP**	**(2)** **CLP**	**(3)** **CLP**	**(4)** **CLP**
1990	7,704,447	7,468,650		
1992	7,628,787	7,571,488		
1994	10,196,431	9,796,880		
1996	7,380,248	7,285,126		
1998	8,968,972	8,356,178		
2000	8,374,909	7,753,978	5,487,549	5,540,792
2003	9,385,429	8,792,003	5,267,166	5,266,357
2006	8,114,511	7,548,902	5,153,000	5,183,402
2009	7,188,909	7,575,726	5,418,999	5,345,266
2011	7,439,235	7,217,802	5,441,190	5,379,669
2012	7,846,929	7,679,414	5,652,764	5,700,459
2015	7,740,785	7,735,459	5,669,175	5,689,454
2017	8,237,981	8,454,179	5,781,648	5,816,401
Trimming	No	No	Yes	Yes
Exp. Factors	No	Yes	No	Yes

Note: The percentages in the third and fourth columns are calculated trimming household income at the 1st and 99th percentiles each year.

B. Additional specifications

TABLE 8

RIF DECOMPOSITION – LABOUR UNION EFFECTS (1994-2017)

	Mean log-Hourly Earnings			Gini Coefficient of Hourly Earnings		
	(1)	(2)	(3)	(4)	(5)	(6)
	1992-2017	1994-2017	1994-2017	1992-2017	1994-2017	1994-2017
Overall						
Post	7.747***	7.747***	7.747***	0.406***	0.406***	0.406***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Counterfactual	7.204***	7.346***	7.354***	0.472***	0.486***	0.487***
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Pre	6.838***	7.023***	7.023***	0.449***	0.464***	0.464***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Difference	0.910***	0.724***	0.724***	-0.043***	-0.058***	-0.058***
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Explained	0.366***	0.322***	0.322***	0.023***	0.021***	0.023***
	(0.006)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Unexplained	0.544***	0.402***	0.402***	-0.066***	-0.080***	-0.081***
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Explained						
Total	0.336***	0.322***	0.322***	0.023***	0.021***	0.023***
	(0.006)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Pure Explained	0.351***	0.320***	0.319***	0.115***	0.108***	0.109***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Specification Error	0.015***	0.003	0.003	-0.092***	-0.087***	-0.086***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Pure Explained						
Education	0.211***	0.202***	0.203***	0.129***	0.129***	0.129***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Experience	0.026***	0.0025***	0.024***	0.012***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Region	0.002***	0.005***	0.005***	-0.003***	-0.001	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Sector	0.013***	0.014***	0.014***	0.006***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Size	0.019***	0.020***	0.019***	0.008***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	-0.008***	-0.010***	-0.010***	-0.003***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Below Hourly Minimum Wage	0.090***	0.069***	0.069***	-0.029***	-0.033***	-0.033***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Formal Employee	-0.004***	-0.005***	-0.005***	-0.006***	-0.007***	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rural	0.002***	0.001***	0.001***	0.000	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Labour Union			-0.000***			0.000
			(0.000)			(0.000)
Unexplained						
Total	0.554***	0.402***	0.402***	-0.066***	-0.080***	-0.081***
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Reweighting Error	-0.057***	-0.062***	-0.063***	0.007***	0.014***	0.014***
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)
Pure Unexplained	0.601***	0.464***	0.465***	-0.074***	-0.094***	-0.095***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Pure Unexplained						
Education	-0.102***	-0.121***	-0.125***	0.105***	0.126***	0.126***
	(0.034)	(0.031)	(0.031)	(0.017)	(0.017)	(0.017)
Experience	-0.188***	-0.058***	-0.057***	0.062***	0.072***	0.072***
	(0.016)	(0.015)	(0.015)	(0.009)	(0.009)	(0.008)
Region	-0.040***	-0.062***	-0.059***	-0.008***	-0.004***	-0.004***
	(0.005)	(0.005)	(0.005)	(0.001)	(0.001)	(0.001)
Sector	0.041***	0.027***	0.026***	0.007***	-0.002	-0.001
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Size	-0.004***	0.000	0.000	0.001*	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Female	-0.002	0.029***	0.030***	0.001	0.009***	0.009***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Below Hourly Minimum Wage	0.036***	0.047***	0.047***	-0.020***	-0.022***	-0.022***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Formal Employee	0.029***	0.019***	0.022***	-0.024***	-0.027***	-0.026***
	(0.007)	(0.007)	(0.007)	(0.003)	(0.003)	(0.003)
Rural	0.000	-0.001	-0.002	0.001	0.001	0.001

	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Labour Union			0.012			-0.011***
			(0.007)			(0.004)
Constant	0.830***	0.584***	0.572***	-0.198***	-0.247***	-0.240***
	(0.039)	(0.036)	(0.037)	(0.020)	(0.020)	(0.020)
Post *N*	71,431	71,431	71,431	71,431	71,431	71,431
Counterfactual *N*	43,374	52,856	52,856	43,374	52,856	52,856
Pre *N*	43,374	52,856	52,856	43,374	52,856	52,856

Note: Education and experience are the sum of the coefficients from fourth order polynomials in years. Regional and economic sector categorical variables are normalized in the estimation and coefficients results are summed. A logit reweighting model is used to define the counterfactual. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.