# Unions and Income Distribution: a Study of the Chilean Case

*Raimundo MacAuliffe*[*], *Santiago Rosselot*[†]
Profesor Guía: *Dante Contreras*
Facultad de Economía y Negocios
Universidad de Chile

November 13, 2015

**Abstract**

This paper attempts to measure the effect of trade unions on wage inequality in Chile using panel data. Income distribution analysis is performed through a variance decomposition approach, while income equations are estimated via instrumental variables. Union history of the individual is used as an instrument. We find that the union premium is 21%, but it is due to non-observable individual characteristics rather than union structure. Once we control for these characteristics, the union premium loses significance. Therefore, the effect on income distribution is null. Comparative evidence shows that Chile lags behind OECD countries in terms of labour legislation, even if the labour reform in process is implemented.

**Keywords:** Trade unions, income distribution, labour economics, endogeneity **JEL:** J31, J51, J80

[*]rmacauli@fen.uchile.cl
[†]srosselo@fen.uchile.cl

# 1 Introduction

This paper attempts to measure the effect of trade unions on income distribution in Chile. Chile is a high income country where collective negotiation is restricted, as is worker participation in the mentioned process. Union trade density reaches 15.3% and only 8% of workers participate in the collective bargaining process. It's also one of the most unequal countries in the world. With a 50.8 Gini index, Chile leads the inequality ranking within the OECD. Another striking fact is that the richest 1% of the population concentrates 30.5% of the country's wealth (López et al., 2013). Our belief is that most of these inequalities are generated in the labour market. Here is where trade unions become relevant; they can shift income distribution by negotiating higher wages. Also, they can push for more and better jobs. Both the former and the latter might have a positive effect on diminishing Chilean inequality.

To measure the effect of unions on inequality an income equation is estimated first. We then use the variance of log-income as an inequality measure, and decompose it following Fields (2002) to find how it is affected by unions in particular.

It is widely recognized in union premium literature that the union sorting process is endogenous. Workers are not assigned randomly into union and non-union sectors. Participation in unions is modeled as an endogenous variable for the first time by Ashenfelter and Johnson (1972). Solutions to this problem have been given by Heckman (1979) or Lee (1978). On the other hand, an instrumental variable procedure is developed in Duncan and Leigh (1985).

The Chilean case has been studied before. Landerretche et al. (2013) estimated the union premium via a two-stage procedure using panel data, finding it to be between 18 and 24%. We find this result surprising considering how flexible and deregulated the Chilean labour market is and the harsh reality that unions have to face. Section 2 develops this idea in detail.

We face the endogeneity problem through an instrumental variables approach. Union history of the individual is used as excluded instruments and union status is estimated via maximum likelihood. Two procedures proposed by Jeffrey Woolridge are implemented to correct for *initial conditions* and *forbidden regression* problems, both detailed in section 4.

Panel data from the Chilean Social Protection Survey for years 2002, 2004, 2006 and 2009 is used. Three equations are estimated, using two different samples. First, we run an IV estimation using a cross-section sample for year 2009. We next run a pooled-OLS estimation using a longitudinal sample. Finally, we add to union literature by using the latter sample to estimate the union premium via fixed effects. Variance decomposition is performed post-estimation.

Our first stage results show that individual union history is relevant in the union sorting process. Firm size is also significant, as are other firm characteristics. Observable characteristics of the individual seem not to play an important role. Regarding the second stage, union status proves to be non-significant in the cross-section estimation. Our pooled estimation shows that the union premium is approximately 21%. In this case, union status explains only 2% of income inequality. We think that it is due to the low Chilean union density rates. Lastly, once we control for unobservable characteristics, union premium once again loses significance. Therefore, union status does not affect (this measure of) income inequality. These results lead us to think that it is not the union institution that leads to higher wages, but individual non-observable characteristics of union members.

Given this result, we analyse international evidence of union and collective negotiation structure for OECD countries and compare it to the Chilean case. Chilean congress is currently processing a labour reform, which is also analysed. We find that Chile is well behind OECD standards in terms of centralization level and percentage of collective negotiation, union density and worker replacement. The current labour reform changes some of these elements, but still leaves Chile lagging behind.

The paper is organized as follows: section 2 provides an institutional framework related to unions in the Chilean economy. Section 3 gives an econometric framework in which we describe how to estimate the wage premium as well as its effect on income distribution. Section 4 develops the steps contained in the estimation procedure, while data and variables are described in section 5. Section 6 presents our empirical findings and is followed by a discussion section. Section 8 concludes.

# 2 The Chilean Labour Market and Union Legislation

Over the past twenty years, Chile has been one of the twenty five countries with the largest economic growth in the world. This ranks Chile second among OECD countries and first in South America. Currently, its GNP is around US$ 22,000 per capita. Chile classifies as a high income level country.

The current Chilean labour legislation has its origins in 1979 with the labour plan implemented by Minister of Labour Jose Piñera during Augusto Pinochet's dictatorship. It deregulated and flexibilized the labour market, minimizing unions and collective bargaining. The reform was based on the ideas of Milton Friedman and Frederik Von Hayek.

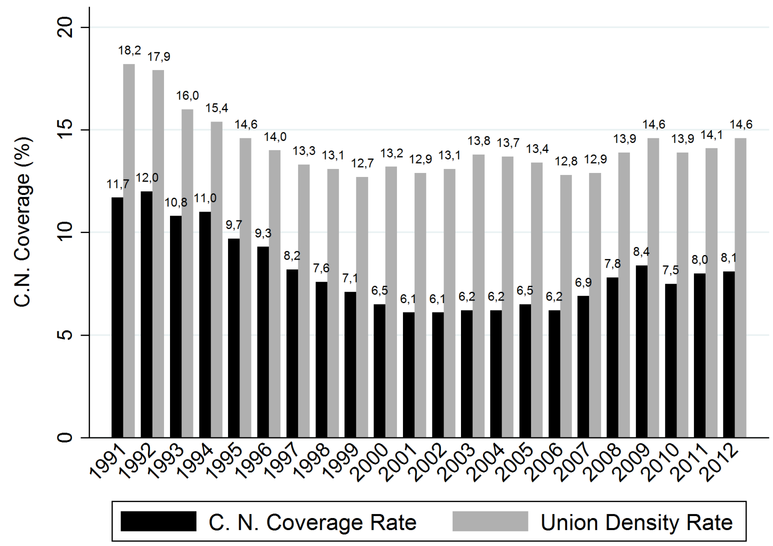
The labour plan of 1979 had four main points: non-paralysing strike, union depoliticization, union decentralization and organizational parallelism. These four points were expressed in two laws. One restricted union density and another restricted the collective negotiation process. They have remained untouched with the governments that followed Pinochet's dictatorship.

Unions were outlayed in 1973, when the military coup took place. At that time, trade union density was 33%. The implementation of the 1979 labour plan made unions legal again. However, their political power was weakened compared to pre-dictatorship scenario.

In 1990, with the return to democracy, the so-called *"Concertacion"* implemented another labour reform. This reform marginally strengthened the position of unions, maintaining the structure of the labour plan of 1979.

By 1991, trade union density reached 18,2%. In the following years, trade union density has tended to fall. Nowadays it is 15.3%. Meanwhile, only 8% of workers participate in a collective bargaining process. Figure 1 shows union density and collective negotiation coverage rates since the return to democracy.

**Figure 1:** Effective Collective Negotiation Coverage and Union Density Rates




Source: Durán (2013)

Labour legislation allows external worker replacement during a strike. Firms can subdivide unions so big firms can face hundreds of unions instead of only one. Union negotiation takes place only on firm-level. In other countries unions can negotiate on an industry-level or even on a national-level[1]. Also, when the collective negotiation process is done, its benefits are extended to all the workers of the firm, not only to the unionized ones. The latter encourages workers not to be part of a union since they can receive the same benefits without paying union fees.

Legislation also imposes entry barriers to union participation. All workers with temporary contracts are immediately excluded. Also, workers that join the union after the collective bargaining process already started are excluded from the resultant benefits. For public sector employees, union participation is forbidden by the law. Nonetheless, in practice they do conduct collective negotiations through the National Federation of Public Sector Employees (ANEF). Additionally, unions are not the only worker's organization inside the firms. Organizational parallelism allows firms to face multiple organizations, which weakens the union and also disaggregates the topics that unions can face.

# 3 Econometric Framework

We follow the latent equation approach found in Landerretche et al. (2013) to estimate the union effect on wages[2]. This approach assumes that each worker, independent of his/her union status, has a latent union wage and a latent non-union wage. Therefore, its the difference between these two latent wages that a correct estimation of the union premium should capture. Since a worker will never be in both the union and non-union group at the same time, data will only show one wage - union status pair per period. The observed wage of the individual is as following:

[1]See section 7 for a more detailed analysis of international situation.
[2]Note that these models are for panel data. However, they can be applied to cross-section data if $t$ is ignored.

$$w_{i,t} = s_{i,t} \cdot w^1_{i,t} + (1 - s_{i,t}) \cdot w^0_{i,t} \quad (1)$$

Where $w^s_{i,t}$ is the log of latent wage of individual $i$ at time $t$ for union status $s$. Union status $s_{i,t}$ equals 1 when the worker is unionized and 0 if not. Regrouping terms:

$$w_{i,t} = w^0_{i,t} + s_{i,t} \cdot (w^1_{i,t} - w^0_{i,t}) \quad (2)$$

In (2), $w^1_{i,t} - w^0_{i,t}$ corresponds to the individual union wage premium and cannot be observed. Assuming a latent wage consisting of an observable and an unobservable omponent, our estimating equation is the following[3]:

$$w_{i,t} = x_{i,t} \cdot \beta + s_{i,t} \cdot \gamma + \varepsilon_{i,t} \quad (3)$$

Observable characteristics can be found in $x_{i,t}$, while $\beta$ is a vector of parameters that describes the effect of each characteristic on wages. All unobservable components are gathered in $\varepsilon_{i,t}$. Finally, any consistent estimator of $\gamma$ would reveal the union effect on wages.

It cannot be ignored that union status is not assigned randomly. Therefore, the mean difference of wages between union and non-union sectors is not the union wage effect. It is an endogenous variable as the selection process into each sector is a *choice*. Thus, it must be modelled. The union status equation is given by:

$$\begin{aligned} y_{i,t} = z_{i,t} \cdot \delta + s_{i,t-1} \cdot \theta + b_i + v_{i,t} \qquad \forall \;\; t = 1, ..., T \wedge i =, ..., N \\ s_{i,t} = 1 \;\; if \;\; y_{i,t} \geq 0 \\ s_{i,t} = 0 \;\; if \;\; y_{i,t} < 0 \end{aligned} \quad (4)$$

where $y_{i,t}$ is an unobserved latent variable which governs the union status $s_{i,t}$ of worker $i$ at time $t$, $z_{i,t}$ a vector of exogenous observable characteristics, $\delta$ its coefficient vector and $\theta$ the coefficient vector of lagged union status $s_{i,t-1}$.

For the income distribution analysis, we will use the variance of the log of income as a measure of inequality. We will then decompose this variance as presented in Fields (2002) to identify the effect that unions have on income distribution.

This methodology is based on Mincer's (1974) model. Ignoring $t$ and $i$ for simplicity, equation 3 can be rewritten as following:

$$w = \sum \beta_j \cdot b_j \quad (5)$$

where $b_j$ represents all $j$ independent variables included in our income equation (including union status). According to Mood, Graybill & Boes's (1974) theorem:

$$Cov(\sum \beta_j \cdot b_j, w) = \sum Cov(\beta_j \cdot b_j, w) \quad (6)$$

[3]For a more detailed derivation of the equation see Landerretche et al. (2013)

The term on the left is the covariance of $w$ with itself, then:

$$\sigma^2(w) = \sum Cov(\beta_j \cdot b_j, w) \tag{7}$$

Dividing the above expression by $\sigma^2(w)$:

$$100\% = \sum \frac{Cov(\beta_j \cdot b_j, w)}{\sigma^2(w)} = \sum S_j \tag{8}$$

We know that:

$$Corr(\beta_j \cdot b_j, w) = Cov(\beta_j \cdot b_j, w)/(\sigma_{\beta_j b_j} \cdot \sigma_w)$$

Therefore:

$$S_j = \frac{\beta_j \cdot \sigma(b_j) \cdot Corr(b_j, w)}{\sigma_w} \tag{9}$$

where $S_j$ represents the share in which each explicative variable affects inequality. Further, given that $R^2$ is the fraction of the log-variance that is explained by all variables together:

$$P_j = \frac{S_j}{R^2} \tag{10}$$

where $P_j$ is the fraction that each variable $j$ affects the share of inequality explained by the included variables.

The main disadvantage of this approach is that it imposes a parametric form to the income generating process. It does, however, present two important advantages. It allows the effect of each variable on income inequality to be isolated and quantified. Therefore, it allows us to identify the effect of unions in particular. Also, given the log-linear model, the measure of inequality that's used becomes irrelevant[4].

# 4 Estimation Procedure

Throughout this paper we will estimate a variety of regressions as we correct different estimation problems. We will first use our cross-section sample to estimate equation 3 by ordinary least squared. Then, we will estimate it through an instrumental variables approach. Next we will use our longitudinal sample for a pooled and fixed effects estimation. Finally, we proceed with the variance decomposition analysis. Details of the estimations are given in this section.

As mentioned above, first we will estimate equation 3 by OLS using our cross-section sample. However, our income equation presents the endogeneity problem detailed in the previous section. We cannot estimate it by OLS, as $\gamma$ wouldn't reflect the causal effect of being unionized on wages. It would include the effect of unobservable differences between the union and non-union sector.

Therefore, we will use an instrumental variables approach to correct this endogeneity. Our instruments will be the lagged union statuses of the individual. We will use a joint significance test of the excluded instruments (those included in the first stage but not in the second) to test the validity of our instruments. However, we will not use the typical two stage least squared (2SLS) procedure. To take into account the binary nature of our endogenous variable, we will estimate

[4]See Fields (2002) for a demonstration.

the selection equation by maximum likelihood[5].

This, in turn, will bring us the *forbidden regression* problem. It states that mimicking 2SLS with probit will not produce consistent estimate as neither the conditional expectations operator nor the linear projection carry through nonlinear functions[6]. We will use a three stage procedure proposed by the same author and found in Adams et al. (2009) to correct this problem. In the first stage, we estimate - via probit - the determinants of union status (including the instruments). In the second stage, we use the predicted values from the previous stage in a new union status equation. This time, however, it will be estimated via OLS and will not include the instrumental variables. The third stage is estimated as usual.

Given our rich longitudinal data, we will use it to try to estimate the union premium correctly. For our first stage we detect evidence of Heckman's (1981) *initial conditions* problem. It states that a lagged variable (in this case union status) cannot be assumed exogenous, as it is correlated with the individual-specific unobserved component. To correct this problem we will use Wooldridge's Conditional Maximum Likelihood (CML) estimator found in Landerretche et al. (2013). The procedure detailed in the above paragraph will also be applied. Both methods are detailed in the Appendix.

Once we obtain the fitted valued from our first (second) stage, we estimate our income equation. We initially estimate it for a pooled sample and then through a fixed-effects estimation. Non corrected benchmark equations are also estimated. Using the estimated parameters we can calculate the effect that each one has on income distribution, following Fields (2002). These effects are captured in the $S_j$ and $P_j$ terms explained in section 3.

A challenge for future investigations would be to model employment status as well as union status.

# 5 Data and Variables

## 5.1 Data

For this investigation, we used data from the Social Protection Survey (Encuesta de Protección Social or EPS). This survey has four editions published: 2002, 2004, 2006 and 2009. These have been run by the Microdata Center of the University of Chile and the University of Pennsylvania.

The 2002 version was applied to individuals affiliated to the pension system. Respondents described their labour history for the 1980-2002 period. For the following versions, labour history for the period since the last survey was described. Information provided includes individual and firm characteristics. For the 2002-2006 versions, the total sample is composed of roughly 16,500 people and 14,500 for the 2009 edition.

## 5.2 Sample Construction

We constructed two samples for the paper: a cross-section using the 2009 edition of the survey, and a four-period balanced panel using all four editions.

[5]`PROBIT and XTPROBIT` commands in Stata.

[6]See Woolridge (2002) section 15.7.3 for proof.

We first built a raw database for each year using the last labour observation of the interviewee. Individuals who presented unemployment/inactivity spells were dropped, as were non-wage earners (in both the public and private sector). Respondents who were older than 65 were also dropped.

For the cross-section database, we exploited the longitudinal nature of our data by extracting the union status of (all three) previous periods in time. The final sample consists of 2,290 observations. Our panel sample is smaller. It consists of 1,261 individuals for three periods. From the 2002 survey we only extract union status. The main cause of the difference is that, in this case, individuals must fulfill the above conditions in all editions of the EPS.

### 5.3 Variables & Descriptive Statistics

Variables included to explain wages are those found in classic mincer equation literature. Schooling is included as following: $\alpha_1 \cdot schooling + \alpha_2 \cdot d_1 \cdot (sch - 8) + \alpha_3 \cdot d_2 \cdot (sch - 12)$ where $d_1 = 1$ if $sch > 8$ and $d_2 = 1$ if $sch > 12$. This approach allows us to capture nonlinearities in the returns to schooling[7]. We also include gender(=1 if individual is male), potential experience($exp = age - sch - 6$) and its square, firm size and industry, occupation[8] and a public sector dummy. For union status determinants the same variables were used, and lagged union status were used as instrumental variables.

[7]The return to each year of elementary education is $\alpha_1$, to secondary education education $\alpha_1 + \alpha_2$ and to higher education $\alpha_1 + \alpha_2 + \alpha_3$.

[8]Occupations according to ISCO-88 and industry to ISIC Rev. 2.

**Table 1:** Descriptive Statistics: EPS 02-09

| Variables | Cross-section Union | Cross-section Non-Union | Panel Union | Panel Non-Union |
|---|---|---|---|---|
| Log Wage/hour | 7.481 | 7.216 | 7.499 | 7.237 |
| | (0.636) | (0.660) | (0.622) | (0.638) |
| Male = 1 | 0.615 | 0.683 | 0.616 | 0.690 |
| | (0.487) | (0.465) | (0.487) | (0.463) |
| Schooling (years) | 12.336 | 11.001 | 12.466 | 11.118 |
| | (3.588) | (4.005) | (3.724) | (3.973) |
| Experience | 27.319 | 27.508 | 27.571 | 27.404 |
| | (10.267) | (11.220) | (10.326) | (11.415) |
| Experience Squared | 851.556 | 882.467 | 866.491 | 881.122 |
| | (597.288) | (670.024) | (604.642) | (680.644) |
| Public Sector Worker = 1 | 0.365 | 0.116 | 0.395 | 0.122 |
| | (0.482) | (0.320) | (0.490) | (0.328) |
| *Firm Size* | | | | |
| 0-9 | 0.035 | 0.243 | 0.031 | 0.216 |
| | (0.184) | (0.429) | (0.174) | (0.411) |
| 10-49 | 0.143 | 0.298 | 0.170 | 0.322 |
| | (0.350) | (0.457) | (0.377) | (0.468) |
| 50-199 | 0.219 | 0.210 | 0.244 | 0.237 |
| | (0.414) | (0.408) | (0.430) | (0.425) |
| 200+ | 0.604 | 0.249 | 0.554 | 0.226 |
| | (0.489) | (0.433) | (0.498) | (0.418) |
| *Occupation* | | | | |
| High command | 0.010 | 0.019 | 0.017 | 0.015 |
| | (0.097) | (0.135) | (0.130) | (0.123) |
| Professionals | 0.135 | 0.093 | 0.153 | 0.105 |
| | (0.342) | (0.291) | (0.361) | (0.306) |
| Technicians | 0.158 | 0.087 | 0.159 | 0.092 |
| | (0.365) | (0.282) | (0.366) | (0.290) |
| Clerks | 0.143 | 0.146 | 0.131 | 0.158 |
| | (0.350) | (0.353) | (0.338) | (0.365) |
| Service workers | 0.127 | 0.124 | 0.122 | 0.124 |
| | (0.333) | (0.329) | (0.328) | (0.330) |
| Skilled agri. and fish. | 0.027 | 0.053 | 0.023 | 0.057 |
| | (0.162) | (0.224) | (0.149) | (0.232) |
| Craft | 0.108 | 0.154 | 0.105 | 0.146 |
| | (0.310) | (0.361) | (0.307) | (0.354) |
| Plant and machine opps. | 0.179 | 0.133 | 0.170 | 0.129 |
| | (0.384) | (0.340) | (0.377) | (0.335) |
| Unskilled | 0.114 | 0.192 | 0.119 | 0.173 |
| | (0.318) | (0.394) | (0.325) | (0.378) |
| *Industry* | | | | |
| Agriculture | 0.049 | 0.133 | 0.037 | 0.138 |
| | (0.216) | (0.339) | (0.189) | (0.345) |
| Mining | 0.060 | 0.011 | 0.051 | 0.009 |
| | (0.238) | (0.104) | (0.221) | (0.093) |
| Manufacturing | 0.130 | 0.138 | 0.145 | 0.158 |
| | (0.337) | (0.345) | (0.352) | (0.365) |
| Utilities | 0.014 | 0.008 | 0.011 | 0.010 |
| | (0.119) | (0.092) | (0.106) | (0.099) |
| Construction | 0.041 | 0.115 | 0.037 | 0.094 |
| | (0.199) | (0.319) | (0.189) | (0.291) |
| Wholesale, retail and hotels | 0.106 | 0.175 | 0.094 | 0.172 |
| | (0.308) | (0.380) | (0.292) | (0.377) |
| Transport and communications | 0.081 | 0.078 | 0.063 | 0.078 |
| | (0.273) | (0.268) | (0.242) | (0.268) |
| Financial intermediation | 0.076 | 0.087 | 0.068 | 0.088 |
| | (0.265) | (0.283) | (0.252) | (0.283) |
| Personal services | 0.442 | 0.254 | 0.494 | 0.254 |
| | (0.497) | (0.436) | (0.501) | (0.436) |
| Unionization Rate 2002 | | | 16.97 | |
| Unionization Rate 2004 | | | 21.41 | |
| Unionization Rate 2006 | | | 26.64 | |
| Unionization Rate 2009 | | | 27.91 | |
| Observations | 631 | 1,659 | 352 | 909 |

For panel, only information from 2009 was used.
SD in parentheses

The above table presents the descriptive statistics of each sample. Both samples have similar mean values. As expected because of our sample selection, the unionization rates of our samples are higher than those of the full sample and the official numbers. We also see that unionization rates of our sample are on the rise.

At first glance, we can see that the union sector presents higher wages. It also shows less males and more public sector workers. Additionally, there is a clear positive relation between firm size and unionization. Mining has a larger weight within the union sector than the non-union sector, as does personal services. Agriculture, construction and wholesale sectors present the opposite phenomenon. We expect these variables to be relevant in our first stage.

# 6 Empirical Results

In this section we present our results. Stage 2 of the forbidden regression correction procedure is omitted. The second stage presented in this section corresponds to our income equations.

## 6.1 First Stage

The following table shows the results of our first stage. Columns (1) and (3) show our benchmark equations for the cross-section and panel data estimation, respectively while Columns (2) and (4) show our corrected estimations for both samples. The latter include lagged union status as additional regressors and, for the panel data sample, corrections for the *initial conditions problem*.

**Table 2:** First Stage Results: Probit Estimation of Union Status

| | Cross Section | | Panel Data | |
|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) |
| Union Status 02 | | 0.280*** | | 1.034*** |
| | | (0.0986) | | (0.157) |
| Union Status 04 | | 0.672*** | | |
| | | (0.0928) | | |
| Union Status 06 | | 0.856*** | | |
| | | (0.0832) | | |
| Union Lag | | | | 0.540*** |
| | | | | (0.124) |
| Male = 1 | -0.0142 | -0.0710 | -0.0265 | 0.00752 |
| | (0.0762) | (0.0829) | (0.118) | (0.0960) |
| Experience | 0.0311** | 0.0193 | 0.0437** | 0.0772** |
| | (0.0145) | (0.0155) | (0.0189) | (0.0379) |
| Experience Squared | -0.000438* | -0.000370 | -0.000343 | -0.000681 |
| | (0.000256) | (0.000274) | (0.000355) | (0.000648) |
| Elementary | -0.0169 | -0.0270 | 0.0221 | -0.0375 |
| | (0.0360) | (0.0385) | (0.0519) | (0.0443) |
| Secondary | 0.0683 | 0.0475 | 0.107 | 0.0744 |
| | (0.0515) | (0.0552) | (0.0727) | (0.0618) |
| Higher Education | -0.0393 | -0.00814 | -0.120** | -0.0182 |
| | (0.0357) | (0.0385) | (0.0483) | (0.0429) |
| Public Sector Worker = 1 | 0.653*** | 0.410*** | 0.795*** | 0.0876 |
| | (0.0977) | (0.108) | (0.122) | (0.162) |
| 10-49 | 0.564*** | 0.500*** | 0.618*** | 0.380** |
| | (0.128) | (0.138) | (0.149) | (0.175) |
| 50-199 | 1.058*** | 0.928*** | 1.188*** | 0.768*** |
| | (0.127) | (0.137) | (0.153) | (0.182) |
| 200+ | 1.530*** | 1.305*** | 1.728*** | 0.986*** |
| | (0.121) | (0.130) | (0.149) | (0.177) |
| Mining | 1.239*** | 1.242*** | 1.122*** | 0.221 |
| | (0.247) | (0.265) | (0.336) | (0.431) |
| Manufacturing | 0.416** | 0.287 | 0.652*** | 0.0726 |
| | (0.172) | (0.191) | (0.220) | (0.283) |
| Utilities | 0.678** | 0.590* | 0.693 | 0.262 |
| | (0.316) | (0.350) | (0.461) | (0.589) |
| Construction | -0.179 | -0.0434 | 0.240 | 0.387 |
| | (0.196) | (0.211) | (0.258) | (0.341) |
| Wholesale, retail and hotels | 0.250 | 0.300 | 0.399* | 0.120 |
| | (0.177) | (0.194) | (0.231) | (0.295) |
| Transport and communications | 0.355* | 0.345* | 0.448* | -0.199 |
| | (0.186) | (0.203) | (0.253) | (0.337) |
| Financial intermediation | 0.401** | 0.487** | 0.484* | 0.162 |
| | (0.187) | (0.204) | (0.256) | (0.340) |
| Personal services | 0.661*** | 0.540*** | 1.045*** | 0.264 |
| | (0.166) | (0.184) | (0.225) | (0.314) |
| Professionals | 0.459 | 0.350 | -0.0914 | -0.455 |
| | (0.310) | (0.336) | (0.328) | (0.373) |
| Technicians | 0.806** | 0.600* | 0.224 | -0.367 |
| | (0.320) | (0.348) | (0.337) | (0.379) |
| Clerks | 0.621* | 0.422 | -0.138 | -0.631 |
| | (0.324) | (0.351) | (0.339) | (0.386) |
| Service workers | 0.848** | 0.588 | 0.135 | -0.508 |
| | (0.332) | (0.359) | (0.349) | (0.403) |
| Skilled agri. and fish. | 1.119*** | 0.969** | 0.440 | 0.263 |
| | (0.381) | (0.415) | (0.420) | (0.497) |
| Craft | 0.980*** | 0.846** | 0.347 | -0.107 |
| | (0.339) | (0.367) | (0.361) | (0.431) |
| Plant and machine opps. | 1.052*** | 0.853** | 0.370 | -0.451 |
| | (0.334) | (0.362) | (0.358) | (0.433) |
| Unskilled | 0.771** | 0.705* | -0.0468 | -0.572 |
| | (0.334) | (0.362) | (0.359) | (0.419) |
| Constant | -3.456*** | -3.060*** | -4.500*** | -4.213*** |
| | (0.481) | (0.520) | (0.606) | (0.708) |
| Observations | 2,290 | 2,290 | 3,783 | 3,783 |
| Number of individuals | | | 1,261 | 1,261 |
| IV $F$-statistic | | 105.2 | | 117.6 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
EPS 02-09

First off, as a weak instruments test we use a common rule of thumb for models with one endogenous regressor: the $F$-statistic of a joint test whether all excluded instruments are significant should be bigger than 10. In both estimations our instruments pass this test.

We will begin looking at the cross-section estimation. Perhaps the most interesting finding is the statistical significance and magnitude of the lagged union statuses, which are larger the closer they are to the present. This might lead to think that a public policy devoted to promoting union participation today may have long term effects. Aside from that, firm characteristics such as if it belongs to the public sector and its size have a significant and positive relation with belonging to the union sector. Mining and personal services also prove to be significant. These results were expected. Craftsmen, plant and machine operators, skilled agricultural and fishing and unskilled workers are occupations that increase the probability of being unionized. These variables are robust to both estimations. Experience and its square, manufacturing, clerks and service workers lose significance once lagged union status is included in the estimation. Personal characteristics do not seem to play a role in the union sorting process.

Next we will analyse our panel data results. Once again, lagged union status is highly significant and positive, as is the union status of the initial period. The latter phenomenon speaks of the importance of the initial conditions problem. In this estimation industry and occupational characteristics lose significance. Experience becomes significant in this estimation. Firm size proves to be robust to both samples and estimation procedures.

## 6.2 Second Stage

**Table 3:** Second Stage Results: IV Estimation of the Union Premium

| | Cross Section | | Panel: Pooled | | Panel: Fixed Effects | |
|---|---|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
| Union | 0.0606** | | 0.0808*** | | 0.0662*** | |
| | (0.0253) | | (0.0207) | | (0.0235) | |
| Union (First Stage) | | 0.0835 | | 0.208*** | | -0.173 |
| | | (0.0551) | | (0.0554) | | (0.192) |
| Male = 1 | 0.154*** | 0.155*** | 0.175*** | 0.174*** | | |
| | (0.0247) | (0.0247) | (0.0196) | (0.0196) | | |
| Experience | 0.0172*** | 0.0171*** | 0.0184*** | 0.0175*** | 0.0784*** | 0.0832*** |
| | (0.00425) | (0.00429) | (0.00294) | (0.00296) | (0.00650) | (0.00786) |
| Experience Squared | -0.000188*** | -0.000186** | -0.000169*** | -0.000160*** | -0.000216** | -0.000253** |
| | (7.20e-05) | (7.24e-05) | (5.43e-05) | (5.44e-05) | (0.000109) | (0.000116) |
| Elementary | 0.00158 | 0.00168 | 0.0211*** | 0.0209*** | 0.0544*** | 0.0540*** |
| | (0.00836) | (0.00833) | (0.00692) | (0.00689) | (0.00885) | (0.00898) |
| Secondary | 0.0662*** | 0.0658*** | 0.0465*** | 0.0443*** | 0.0205 | 0.0238 |
| | (0.0144) | (0.0144) | (0.0106) | (0.0106) | (0.0177) | (0.0179) |
| Higher Education | 0.0199 | 0.0201 | 0.0107 | 0.0127 | 0.00713 | 0.00607 |
| | (0.0128) | (0.0128) | (0.0105) | (0.0105) | (0.0153) | (0.0153) |
| Public Sector Worker = 1 | -0.0745** | -0.0797** | -0.0898*** | -0.120*** | -0.0196 | -0.0107 |
| | (0.0356) | (0.0363) | (0.0276) | (0.0301) | (0.0395) | (0.0392) |
| 10-49 | 0.126*** | 0.124*** | 0.113*** | 0.103*** | -0.00910 | -0.00155 |
| | (0.0287) | (0.0289) | (0.0222) | (0.0225) | (0.0247) | (0.0255) |
| 50-199 | 0.212*** | 0.207*** | 0.200*** | 0.178*** | 0.0227 | 0.0463 |
| | (0.0302) | (0.0320) | (0.0236) | (0.0253) | (0.0255) | (0.0314) |
| 200+ | 0.264*** | 0.255*** | 0.223*** | 0.181*** | 0.0302 | 0.0651* |
| | (0.0331) | (0.0386) | (0.0246) | (0.0299) | (0.0279) | (0.0377) |
| Mining | 0.527*** | 0.519*** | 0.446*** | 0.417*** | 0.0884 | 0.0894 |
| | (0.0733) | (0.0764) | (0.0627) | (0.0637) | (0.0863) | (0.0853) |
| Manufacturing | 0.118*** | 0.116*** | 0.121*** | 0.109*** | 0.00351 | -0.00498 |
| | (0.0394) | (0.0397) | (0.0306) | (0.0311) | (0.0337) | (0.0353) |
| Utilities | 0.191** | 0.187** | 0.166** | 0.154* | -0.0361 | -0.0292 |
| | (0.0843) | (0.0858) | (0.0795) | (0.0799) | (0.0831) | (0.0823) |
| Construction | 0.166*** | 0.167*** | 0.147*** | 0.151*** | 0.0532 | 0.0617 |
| | (0.0412) | (0.0414) | (0.0378) | (0.0377) | (0.0497) | (0.0498) |
| Wholesale, retail and hotels | 0.104*** | 0.104** | 0.0412 | 0.0372 | -0.0174 | -0.0231 |
| | (0.0404) | (0.0404) | (0.0314) | (0.0315) | (0.0363) | (0.0369) |
| Transport and communications | 0.150*** | 0.148*** | 0.135*** | 0.128*** | -0.0525 | -0.0721 |
| | (0.0479) | (0.0480) | (0.0427) | (0.0428) | (0.0568) | (0.0594) |
| Financial intermediation | 0.267*** | 0.266*** | 0.197*** | 0.195*** | -0.0457 | -0.0513 |
| | (0.0453) | (0.0453) | (0.0386) | (0.0387) | (0.0451) | (0.0453) |
| Personal services | 0.0899** | 0.0867** | 0.0227 | 0.00451 | -0.0147 | -0.0181 |
| | (0.0389) | (0.0396) | (0.0311) | (0.0322) | (0.0447) | (0.0451) |
| Professionals | -0.241** | -0.242** | -0.299*** | -0.302*** | -0.0532 | -0.0757 |
| | (0.0986) | (0.0985) | (0.0759) | (0.0761) | (0.0712) | (0.0714) |
| Technicians | -0.471*** | -0.475*** | -0.509*** | -0.524*** | -0.100 | -0.116 |
| | (0.106) | (0.107) | (0.0776) | (0.0777) | (0.0772) | (0.0772) |
| Clerks | -0.758*** | -0.761*** | -0.760*** | -0.766*** | -0.173** | -0.202** |
| | (0.108) | (0.108) | (0.0772) | (0.0774) | (0.0802) | (0.0815) |
| Service workers | -0.909*** | -0.913*** | -0.955*** | -0.969*** | -0.209** | -0.234*** |
| | (0.110) | (0.111) | (0.0788) | (0.0791) | (0.0817) | (0.0812) |
| Skilled agri. and fish. | -0.941*** | -0.946*** | -1.074*** | -1.086*** | -0.214** | -0.217** |
| | (0.114) | (0.114) | (0.0817) | (0.0819) | (0.0896) | (0.0892) |
| Craft | -0.814*** | -0.818*** | -0.963*** | -0.979*** | -0.170* | -0.180* |
| | (0.111) | (0.112) | (0.0802) | (0.0805) | (0.0926) | (0.0917) |
| Plant and machine opps. | -0.858*** | -0.863*** | -0.974*** | -0.994*** | -0.149* | -0.172** |
| | (0.110) | (0.111) | (0.0791) | (0.0796) | (0.0872) | (0.0864) |
| Unskilled | -0.989*** | -0.992*** | -1.101*** | -1.112*** | -0.198** | -0.225*** |
| | (0.109) | (0.109) | (0.0788) | (0.0790) | (0.0855) | (0.0848) |
| Constant | 7.039*** | 7.046*** | 6.845*** | 6.886*** | 4.761*** | 4.726*** |
| | (0.138) | (0.140) | (0.103) | (0.103) | (0.144) | (0.147) |
| Observations | 2,290 | 2,290 | 3,783 | 3,783 | 3,783 | 3,783 |
| R-squared | 0.524 | 0.523 | 0.500 | 0.500 | 0.235 | 0.233 |
| Number of individuals | | | | | 1,261 | 1,261 |
| $S_j$ | 0.0073 | 0 | 0.0098 | 0.0219 | 0.0081 | 0 |
| $P_j$ | 0.0139 | 0 | 0.0197 | 0.0439 | 0.0342 | 0 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
EPS 02-09

The above table shows the results for our second stage estimations. Once again we include benchmark estimations, shown in columns (1),(3) and (5).

When estimating by (uncorrected) OLS, the union premium is of 6%. The corrected OLS estimation is shown in column (2). It shows that once we model the union sorting process, union status is no longer significant. Gender, experience, secondary education and firm size are significant and positive. Public sector has a negative effect on income. Industry and occupational characteristics are also significant. All variables are robust to both estimations, most showing slightly lower values in column (2).

To further explore the union effect on wages, we run a pooled-OLS estimation. It differs from the previous estimation because it exploits the longitudinal characteristic of our data in its first stage. The results of the corrected estimation can be found in column (4). In this estimation, the union premium is 21% and highly significant. This result is interesting because of its similarity to the one obtained by Landerretche et al. (2013) in their pooled estimation. Apart from wholesale, retail and hotels and personal services, which lose significance, and elementary education (becomes significant) the variables that were significant in estimation (2) remain significant. Once again, all variables significant in the benchmark estimation are robust to the inclusion of predicted union status, showing lower values in the corrected version.

Our last estimation is shown in column (6). It uses the fitted values obtained from the same first stage as the pooled estimation, but is estimated via fixed effects. Therefore, it allows us to control for unobservable characteristics (that are constant in time). Union status proves to be non significant in this estimation. Experience and its square and elementary education are significant and positive. Firm size becomes (barely) significant in those cases where there are more than 200 workers once the predicted union status is included. Some occupational categories are also significant. Repeating the pattern of the estimations detailed above, all parameters are robust to the included corrections, showing lower values in the latter version. The exception is experience, which becomes larger in magnitude.

### 6.3 Variance Decomposition

The last two rows of table 3 show the $S_j$ and $P_j$ corresponding to union status. Since union status does not have an effect on wages in estimations (2) and (6), it doesn't have an effect on income distribution. In our pooled estimation, where the union premium is 21%, it explains slightly over 2% of inequality, and around 4,4% of the inequality explained by the included variables.

## 7 Discussion

Our results show that the wage gap between union and non-union workers is due to unobserved components and not the union institution. Therefore, it is not unreasonable to think that the structure of the Chilean labour market has played a part in syndicalism decline.

As mentioned in section 2, the current labour plan banned collective bargaining on industry and national level and excluded workers with temporary contracts as well as the new members of the union. It also admitted worker replacement during strike and generated multiple workers organizations besides the union. The latter with the purpose of restricting unionization rates and the collective bargaining process itself.

This structure leads us to think that in Chile workers have a very low scope of action. Naturally, in this scenario, it is quite difficult for unions to have a significant wage gap relative to the non-union sector, and therefore, to have a significant impact on inequality.

The chart below shows comparative evidence for 23 OECD countries. It considers degree of centralization of the collective negotiation and percentage of workers involved in it. It also shows union density rates and the case of external replacement during strike. It can be seen that Chile and Japan are the only countries where collective negotiation takes place only on firm level. The rest of the countries allows negotiation on industry level, and 7 of them allows it on a national level. Regarding the percentage of workers who participate in the collective negotiation, Chile takes last place with only 8%. On the other hand, only 4 countries allow external replacement during strike. As we mentioned above, Chile is one of them.

**Table 4:** Level and Percentage of Collective Negotiation, OECD Countries

| Country | National | Industry | Firm | C.N. % | Union Density (%) | Replacement |
|---|---|---|---|---|---|---|
| Austria | | XXX | X | 95 | 27,4 | No |
| Belgium | XXX | XX | X | 96 | 55 | No |
| Czech Republic | | XX | XXX | 38 | 13,4 | No |
| Denmark | | XX | XXX | 80 | 66,8 | No |
| Finland | | XXX | X | 91 | 68,6 | No |
| France | X | XX | XXX | 98 | 7,7 | No |
| Germany | | XXX | XX | 59 | 17,7 | No |
| Greece | X | XXX | X | 65 | 21,3 | No |
| Hungary | X | X | XXX | 33 | 10,6 | No |
| Ireland | XXX | X | XX | 44 | 29,6 | Yes |
| Italy | | XXX | X | 80 | 36,9 | No |
| Netherlands | | XXX | X | 81 | 17,6 | No |
| Portugal | | XXX | X | 25 | 20,5 | No |
| Slovakia | | XX | XXX | 35 | 16,8 | No |
| Spain | X | XXX | X | 71 | 17,5 | No |
| Sweden | | XX | XXX | 88 | 67,7 | No |
| UK | | X | XXX | 29 | 25,4 | No |
| Norway | X | XXX | X | 70 | 53,5 | - |
| Australia | | X | XXX | 40 | 17 | - |
| Canada | | X | XXX | 32 | 27,2 | Yes |
| Japan | | | XXX | 20 | 17,8 | - |
| USA | | X | XXX | 14 | 10,8 | Yes |
| Chile | | | XXX | 8 | 15,3 | Yes |

X: indicates existence; XX: semi-dominant level; XXX: dominant level.
Source: Fundación Sol (2015).

If we look at Norway, where trade union density reaches 53,5% several features stand out. Unions can negotiate on national level and replacement of workers on strike is forbidden. There is strong evidence supporting a significant effect of unions on wages. Barth et al. (2000) find a positive effect of workplace trade union density on the level of the individual's pay in establishments covered by collective agreements.

In addition, Blanchflower (1996) makes a comparison between the United States and 18 other countries members of the OECD. He finds a significant wage gap between unionized and non-unionized workers in the US as well as for the rest of the countries of the sample. He also finds that unions did better in countries with centralized as opposed to decentralized wage setting systems.

Chilean government is currently working on a labour reform. It's main purpose is to fully obey 87th and 98th ILO conventions. This is directly associated to giving more freedom to unions, as well as expanding collective bargaining coverage.

Nonetheless, the project can hardly remove the main ideas of the 1979 Labour Plan. The right to strike is still not acknowledged. The case of worker replacement is now an ambiguous situation, therefore, not completely eradicated. Collective negotiation remains untouched and workers can only negotiate on firm level. The reform binds unions to providing minimum services to the firm even during strike.

It can be inferred, hence, that this labour reform does not achieve to give more freedom to the unions since it doesn't allow the collective negotiation to be developed outside the firm, and it still excessively regulates negotiation procedures. These ideas are opposed to 87th and 98th ILO conventions.

# 8 Conclusion

In this paper we attempt to estimate the effect of trade unions on inequality in Chile. We estimate the union premium via instrumental variables, using history of the individual as an instrument. Income distribution analysis is performed through a variance decomposition approach.

A series of conclusions can be obtained from our estimations. For the union sorting process the union history of the individual proves to be relevant. Individual characteristics seem not to be important. On the contrary, firm characteristics are significant, in particular firm size. An OLS estimation of the union premium shows that it is of around 6%. However, once we model union status, it loses significance. If non-observable characteristics are included in the first stage but not the second stage, unionization has an effect of 21% on wages. This result is similar to the one found in Landerretche et al. (2013). Nonetheless, its effect on income distribution is marginal. A fixed effects estimation of the union premium shows that it is not significant. Therefore, even though union status has an effect on wages, it is not due to the union institution, but to individual non-observable characteristics of union members.

This paper also analyses Chilean labour market legislation. It appears to be flexible and deregulated. Section 7 compared Chile a sample of OECD countries in different areas regarding union structure. Chile lags behind OECD countries in terms of labour legislation, even if the labor reform in process is implemented. Further research should point towards replicating this study once Chilean labour legislation satisfies OECD standards.

There are 3 main contributions that can be obtained from this study. First, the union sorting process presents parameters robust to different samples and estimation procedures which adds empirical value to the idea that union history of individual is relevant in the union sorting process. Second, the use of fixed effects in our estimation concludes that the wage differential between unionized and non-unionized workers is due to individual non-observable characteristics of union members and not because of the union institution. This reinforces the idea that unions have low scope of action in the current Chilean economy. Third, this paper not only measures the union premium, but also its effect on income distribution. This is new in the literature and relevant to Chile, who performs poorly in terms of labor legislation and income distribution relative to OECD standards.

# 9 Appendix

## 9.1 Appendix A: Woolridge Procedure for Forbidden Regression

In this appendix we detail forbidden regression correcting procedure found in Woolridge (2002) (p. 623, procedure 18.1). Assume equation 3 as main equation and $s_{i,t}$ endogenous dummy variable.

(a) Estimate the binary response model $P(s = 1|x, z) = G(x, z; \gamma)$ by maximum likelihood.

(b) Obtain the fitted probabilities, $\hat{G}_i$.

(c) Estimate equation 3 by 2SLS using instruments 1, $\hat{G}_i$ and $x_i$.

## 9.2 Appendix B: Woolridge's CML Estimator [9]

CML Estimator is used in this paper to correct the initial conditions problem. It assumes that, in the union status equation, the individual-specific unobservable component can be approximated using the following distribution:

$$b_i|d_{i0}, W_i \sim N(\lambda_0 + \lambda_1 \cdot s_{i0} + \lambda' W_i, \sigma_c^2) \tag{11}$$

$W_i$ is the row of all non-redundant explanatory variables in all time periods (or their means). Therefore, $b_i$ can be replaced by:

$$b_i = \lambda_0 + \lambda_1 \cdot s_{i0} + \lambda' W_i + c_i \tag{12}$$

where $c_i|(d_{i0}, W_i) \sim N(0, \sigma_c^2$. Now we can say that $s_i t$ follows a probit model with transition probability:

$$Pr[s_{it}|s_{it-1}] = \Phi[(2 \cdot s_{it} - 1) \cdot (z'_{it}\gamma + \delta \cdot s_{it} + \lambda_0 + \lambda_1 \cdot s_{i0} + \lambda' W_i + c_i)] \tag{13}$$

Therefore, Woolridge's CML estimator turns out to be the standard random effects probit estimator but with $s_{i0}$ and $W_i$ included as additional regressors.

# References

Adams, R., Almeida, H., and Ferreira, D. (2009). Understanding the Relationship Between Founder-CEOs and Firm Performance. *Journal of Empirical Finance.*

Ashenfelter, O. and Johnson, G. (1972). Unionism, Relative Wages, and Labour Quality in U.S. Manufacturing Industries. *International Economic Review*, 13(3):488–508.

Barth, E., Raaum, O., and Naylor, R. (2000). Union Wage Effects: Does Membership Matter? *The Manchester School*, 68(3):259–275.

Blanchflower, D. (1996). The Role and Influence of Trade Unions in the OECD.

Contreras, D. and Gallegos, S. (2011). Desigualdad Salarial en América Latina: una década de cambios. *Revista CEPAL.*

Duncan, G. and Leigh, D. (1985). The Endogeneity of Union Status: an Empirical Test. *Journal of Labour Economics*, 3(3):385–402.

[9] Procedure details in appendix B from Landerretche et al. (2013).

Durán, G. (2013). Panorama Sindical y de la Negociación Colectiva en el Chile de los US$ 22.655. *Revista de Derecho y Seguridad Social*, 3:85–96.

Durán, G. and Kremerman, M. (2015). Sindicatos y Negociación Colectiva: Panorama Estadístico Nacional y Evidencia Comparada.

Fields, G. (2002). Accounting for Income Inequality and its Change: A New Method, With Application to the Distribution of Earnings in the United States.

Fundación Sol (2015). Reforma Laboral ¿Pone fin al plan laboral de la dictadura o lo consolida?

Heckman, J. (1979). Sample Selection as a Specification Error. *Econometrica*, 47(1):153–161.

Hirsch, B. (2004). Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic. *Journal of Labour Research*, 15(2).

Landerretche, O., Lillo, N., and Puentes, E. (2013). The Union Effect on Wages in Chile: A Two-Stage Approach Using Panel Data. *Labour*, 27(2):164–191.

Lee, L.-F. (1978). Unionism and Wage Rates: a Simultaneous Equations Model with Qualitative and Limited Dependent Variables. *International Economic Review*, 19(2):415–433.

López, R., Gutiérrez, P., and Figueroa, E. (2013). La "Parte del León": Nuevas Estimaciones de la Participación de los Súper Ricos en el Ingreso de Chile.

Ruiz-Tagle, J. and Sehnbruch, K. (2015). More but not better jobs in Chile? The fundamental importance of open-ended contracts. *International Labor Review*, 154(2).

Stock, J. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models*.

Woolridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*.

Woolridge, J. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models With Unobserved Heterogeneity. *Journal of Applied Econometrics*.