

Unionization and wages: a quantile analysis¹

Gustavo Saraiva Frio, Marco Túlio Aniceto França
and Davi Salomão Canavesi Farias Eskenazi

Abstract

The aim of this article is to analyse the wage effects of unionization in a sample of the Brazilian population in 2015. Using data from the National Household Sample Survey (PNAD) and its supplement on labour relations and unionization, econometric techniques (ordinary least squares, propensity score matching and unconditional quantile regressions) are used to test the hypotheses that being unionized affects wages, and whether there is an additional benefit in the case of unions that participate more actively with their members. The results show that there is a positive relationship between unionization and average wages. It was also found that an increase in total unionization seems to reduce the pay of workers at the lower end of the distribution, but raise wages for the majority of the population. In addition, unions that participated in an agreement in the last 365 days generated average gains of 4.3%.

Keywords

Trade unions, labour relations, wages, economic analysis, econometric models, Brazil

JEL classification

J31, J51, J83

Authors

Gustavo Saraiva Frio holds a PhD in Economics from the Pontifical Catholic University of Rio Grande do Sul (PUCRS), Brazil, and a Master's degree in Applied Economics from the Federal University of Pelotas (UFPEL), Brazil. Email: gustavo.frio@gmail.com.

Marco Túlio Aniceto França is Professor of the Graduate Program in Development Economics at the Pontifical Catholic University of Rio Grande do Sul (PUCRS), Brazil, and holds a PhD in Economics from the Federal University of Paraná (UFPR), Brazil. Email: marco.franca@puers.br.

Davi Salomão Canavesi Farias Eskenazi has a Master's degree in Economics from the Pontifical Catholic University of Rio Grande do Sul (PUCRS), Brazil. Email: eskenazi.davi@gmail.com.

¹ The authors are grateful to the Higher Level Personnel Training Coordination Unit (CAPES) and the National Council for Scientific and Technological Development (CNPq) for funding this research. Any errors or omissions are the authors' responsibility.

I. Introduction

In 2016, the Brazilian government sent a draft law to the national congress, making several changes to the structure of the *Reforma Trabalhista* [Labour reform] instrument. Of these changes, withdrawal of the mandatory nature of the union contribution proved controversial. The law required Brazilian workers to donate one day of service per year to the union of their job category. In 2016, about R\$3,532,487,702 was collected in this way, corresponding to 0.06% of that year's gross domestic product (GDP). According to Menezes-Filho and others (2002), the unions, which have government authorization to act, have exclusive jurisdiction to negotiate category wages, including those of non-unionized workers, who also have to pay the contribution. According to these authors, this amount is distributed as follows: 15% for the state union federation, 5% for the national union federation, 20% for the Special Employment and Wage Account (CEES) and the remaining 60% for the union itself.

In July 2017, there were 16,625 unions registered in Brazil, between workers' unions (68.63%) and employers' associations (31.37%). However, this large number does not necessarily represent the members in question. According to Campos (2014), Brazil had a unionization rate of 17%, very similar to many of the other countries researched by that author, including Chile and Japan. Campos (2014) showed that, civil servants have the highest unionization rate among urban workers (36.8%), while unregistered employees in the private sector have the lowest rate. The author also notes that the Brazilian model differs from that of other countries because union entities can speak for the category as a whole, even if there are non-affiliated workers. According to Cardoso (2014), the membership rate varied between 22.5% and 18.1% between 1988 and 2012.

Thus, against this backdrop, unions will have to seek new forms of financing as from 2018, when the law comes into force, either by charging higher dues to their members or by attracting new members. When the "union tax" ends, these entities will have to be more competitive and participatory. The expected natural effect will be the end of inefficient unions and the growth of those that succeed in serving their members in a satisfactory manner, either by negotiating better wages, or by improving the work environment or offering other services that attract them.

Globally, many articles have reported that unionized workers are more likely to have job protection and to receive social benefits and pensions (Ebbinghaus, Göbel and Koos, 2011; Lu, Tao and Wang, 2010; Western and Rosenfeld, 2011). There are also studies showing that workers receive a wage premium from union membership (Borjas, 1979; Gyourko and Tracy, 1988; Mellow, 1982; Mishel, 2012; Yao and Zhong, 2013). In the case of Brazil, Campos and Moura (2017) analyse the probability that unionized workers receive additional benefits; and they use conditional quantile regressions to show that these workers earn more, mainly along the wage distribution.

In view of the above, two hypotheses are tested using the 2015 National Household Sample Survey (PNAD) and its supplement on labour relations and unionization of the same year — *Pesquisa Nacional por Amostra de Domicílios: aspectos das relações trabalhistas e sindicalização* [National Household Sample Survey: aspects of labour relations and unionization] (IBGE, 2015). The methodology of unconditional quantile regressions (regression of the recentred influence function (RIF)) is used. (Firpo, Fortin, and Lemieux, 2009), along with ordinary least squares (OLS) weighted by propensity score matching, to test the following hypotheses:

1. Hypothesis 1: being unionized increases average wages

According to the international literature on the subject (Bloch and Kuskin, 1978; Borjas, 1979; Gyourko and Tracy, 1988; Yao and Zhong, 2013; Yilmaz and San, 2017), unions have a positive effect on wages. The methodology employed comprises the Oaxaca-Blinder decomposition on the mean (ordinary least

squares); OLS weighted by propensity score matching, and recentred influence functions (RIF) regression, because the model does not presume the distribution of the dependent variable and makes it possible to visualize the results in the different income percentiles. The data source is the 2015 PNAD. The analysis of Campos and Moura (2017) was elaborated further by controlling for variables representing occupation and branch of activity, which made the results more robust. The RIF regression has the advantage of analysing the effect of an increase in the total number of unionized workers on quantile wages, which affords a broader view than conditional quantile regressions, which only make a local analysis.

2. Hypothesis 2: the degree to which individuals are involved in union activities affects wages directly

Where differentials exist between the wages of unionized and non-unionized workers, the analysis will verify whether the degree of unionization affects wages. Using information from the PNAD 2015 supplement on labour relations and unionization, and the doubly robust OLS and propensity score matching models, the hypothesis to be tested will compare only those who are unionized. On average, the results of the models (OLS and propensity score matching (PSM)) show that unionization is associated with an average gain of between 7% and 7.3%. The results of the PSM model were corroborated by the Rosenbaum bounds test, which did not detect the presence of unobservable characteristics that could affect either the probability of being unionized or the outcomes. Across the wage distribution, the RIF regression model shows that an increase in total unionization in Brazil would generate an increase in income for most workers, except for those in the lowest decile, who are estimated to lose.

With respect to the degree of participation in the union, and the use made of it, the results do not show statistically significant gains relative to members who neither make use of their union nor participate in it. However, on average, workers who reported that their unions participated in a dispute or wage negotiation process in the last 365 days have 4.3% higher wages than their peers who did not report any such action by their union.

This study contributes to the literature by verifying the effect of unionization on wages in Brazil. It makes it possible to evaluate not only the wage gains for unionized workers, but also how the degree of unionization (measured by the individual's participation in meetings and his/her use of the services offered by the union) affects the earnings of the respective members. Moreover, unconditional quantile regressions find that the effect of unionization is increasing along the distribution, but is negative for affiliates in the lower tail. Campos and Moura (2017) used conditional quantile regressions to show the individual's gain from joining a union. The data are interesting, as unions in Brazil defend the job category and not just the member. Thus, "free-rider" behaviour can be expected. However, the unions display higher wages than their peers, both in the middle and in the upper tail and median part of the distribution.

This article is divided into six sections, including this introduction. Section II makes a brief review of the literature, and section III provides contextual background on Brazilian union groupings. Section IV describes the data source and the methods used in the study. Section V reports the results obtained, and the sixth and last section offers final remarks.

II. Literature review

This section makes a brief review of the national and international literature on the role played by unions in wages. In the case of Brazil, only Campos and Moura (2017) estimate unionization wage premia directly, both at the mean and across the wage distribution.

There is a vast international literature on the effects of unionization on workers' wages, focused mainly on the United States and China. In the case of the United States, Bloch and Kuskin (1978) use OLS and data from the 1973 U.S. Census Bureau Current Population Survey (CPS) to estimate the wages of union and non-union workers. The authors conclude that the wage structure differs between the two groups. Using data from the National Longitudinal Surveys from 1969 to 1971, Duncan and Leigh (1980) respond to the authors through OLS and Heckman models (to correct sample selection bias). The authors conclude that if selection bias is not corrected, the effect of unionization on wages may be underestimated.

Also with respect to the United States, Borjas (1979) uses selection bias correction (Heckman, 1979) and the OLS model to estimate unionization wage premia. The data are taken from the National Longitudinal Survey of Older Men; and the results show that the reservation wage of unionized individuals is lower than that of their non-unionized peers, while unionized workers earn a premium from membership.

Gyourko and Tracy (1988) and Mellow (1982) estimate the effects of unionization on wages by using OLS and correcting for selection bias. The data source was the U.S. Census Bureau's Current Population Survey (from 1977 and 1979, respectively). The results show that unionized workers are consistently paid more than their non-unionized peers. Correcting for selection bias, Gyourko and Tracy (1988) find that the wage premium obtained from union membership is higher among public sector workers than among those of the private sector.

More recently, Mishel (2012) combines data from the U.S. Census Bureau's Current Population Survey with descriptive statistics to assess the effect of unionization. The author finds that union membership dropped from 26.7% in 1973 to 13.1% in 2011. Moreover, this decline affected the pay of men at the median of the distribution and increased the pay gap between those in managerial and technical positions. Western and Rosenfeld (2011) use the same data (1973–2007) and the OLS model with variance decomposition. The results show an inverse relationship between the unionization rate and wage inequality.

Firpo, Fortin and Lemieux (2009) employed the methods of unconditional quantile regressions (RIF) and conditional quantile regressions to test for differences between union and non-union workers along the income distribution. The data source was the Outgoing Rotation Group supplement to the U.S. Census Bureau's Current Population Survey from 1983 to 1985. The results of the conditional quantile regressions show that being unionized increases wages for men across the wage distribution. Moreover, RIF regressions show that an increase in unionization increases wages in the median and first income decile, but reduces them for those in the bottom decile.

In the case of China, Lu, Tao and Wang (2010) and Yao and Zhong (2013) test the effects of unionization on wages and other variables. The paper by Lu, Tao and Wang (2010) used OLS and instrumental variable methods, with data taken from the 2006 Private Enterprise Survey, to assess the effect of unionization on several variables. The results show that being unionized affects labour productivity positively and increases the likelihood that workers receive housing assistance, maternity leave, and other benefits. However, the results on wages show no significant differences between unionized workers and their non-unionized peers.

The results reported by Yao and Zhong (2013), using seemingly unrelated regressions (SUR) and data from 1,268 Chinese firms in 2006, show that unionization increases median workers' wages and reduces hours worked. It also increases the likelihood that workers have pension coverage. These results remain robust even after adding controls for education, employment and immigration.

Using a multilevel logit model and data from the European Social Survey, Ebbinghaus, Göbel, and Koos (2011) investigate the determinants of unionization in 19 European countries. Controlling for microeconomic and macroeconomic variables, the authors find that the expectation of being protected by unemployment insurance is a strong determinant of unionization. In the case of Romania, using data

from the National Institute of Statistics and data analysis, Stan and Erne (2016) test the hypothesis that migration increases unions' wage bargaining power in the Romanian health sector. The results show no direct effects between migration, labour demand and wage increases through union agreements. Yilmaz and San (2017) use data from the 2004 and 2008 Turkish National Household Surveys and Oaxaca-Blinder decomposition and quantile regressions. Their results show that, on average, unionized individuals earn more, and these results hold across quantiles, indicating larger gains among unionized workers in the lower tail of the income distribution.

In the case of Brazil, several authors, such as Becker (2015), Jacinto and Rodeghiero (2015) and Cruz and others (2016), use being unionized as one of the explanatory variables in wage equations, always using PNAD data. In general, the results report positive effects of unionization on wages. However, in the work of Cruz and others (2016), unionization shows no significant effect on wages in some regions. Gonçalves and Machado (2004) find a wage premium for members — the authors consider unionization as an institutional characteristic— decreasing from 25.38% in 1992 to 19.71% in 2001.

Campos and Moura (2017) combine data from the 2015 PNAD with quantile regressions and income to compare earnings between unionized and non-unionized workers. They use binomial logistic regression to identify the determinants of access to indirect remuneration (transportation, food and health subsidies). The results are significant and indicate that unionized workers have a higher chance (odds ratio) of access to benefits. With respect to earnings, unionized workers earn more both at the mean and in the quantiles, and their earnings increase as they move through the distribution to the right.

III. Labour unions in Brazil

This section provides details of trade unions in Brazil, based on data from the Ministry of Labour and Employment (MTE). Data from PNAD are also used to demonstrate the trend in the number of unionized workers in 2002–2015.

In July 2017, Brazil had 16,625 unions, divided between workers' unions (11,409) and employers' associations (5,216), representing 68.63% and 31.37% of total workers, respectively.² Table 1 shows the distribution of unions between rural and urban areas.

Table 1
Brazil: distribution of unions by zone, July 2017

Type of union	Urban zones	Percentages	Rural zones	Percentages	Total
Employer unions	3 595	68.92	1 621	31.08	5 216
Worker unions	8 482	74.34	2 927	25.66	11 409
Total	12 077		4 548		16 625

Source: Prepared by the authors, on the basis of Ministry of Labour and Employment of Brazil.

As shown in table 1, most unions, whether employer or worker, are concentrated in urban areas.

Using PNAD data, table 2 shows the proportion of workers who are union members, using the sample weights provided by the Brazilian Institute of Geography and Statistics (IBGE).

² Data for July 2017 [online] <http://www3.mte.gov.br/cnes/default.asp>.

Table 2
Brazil: union members by gender, 2002–2015
(Percentages)

Year	Unionized	Unionized men	Unionized women
2002	15.89	17.17	14.16
2003	16.70	17.64	15.44
2004	17.09	18.21	15.60
2005	17.38	18.39	16.06
2006	17.62	18.76	16.16
2007	16.99	17.87	15.86
2008	17.49	18.47	16.24
2009	17.07	18.09	15.77
2011	16.72	17.36	15.89
2012	16.19	16.83	15.35
2013	15.59	16.15	14.86
2014	16.17	16.28	16.03
2015	18.84	19.13	18.47
Mean	16.90	17.72	15.84

Source: Prepared by the authors on the basis of the National Household Survey (PNAD).

Note: A difference in means test showed that men are more likely to be unionized than women — a result that is statistically significant at 1%. The National Household Sample Survey does not have data for 2010, because the Brazilian census was held in that year.

Between 2002 and 2006, the proportion of workers who were unionized each year increased from 15.89% to 17.62%. After 2006, the rates fall slightly until 2013, and then recover between 2013 and 2014. In 2015, the first year of a more profound national crisis, the percentage of unionized workers increased substantially. In terms of gender, it is found that, in general, men are more unionized than women.

In the case of men, unionization rates varied between 17.17% and 18.87% in 2002–2011, and then slipped to 16.15% in 2014. Accompanying the total, the proportion grew again in 2015, to reach the peak of the historical series, at 19.13%. The proportion of women who are unionized varied between 14.16% and 16.16% in 2002–2014. As in the case of men, there was a sharp rise in 2015, to 18.47%.

Table 3 shows how union members are distributed between urban and rural employees, self-employed workers (*autónomos*), casual own-account workers (*avulsos*), independent professionals (*profissionais liberais*) and other types of unions.³

Until 2009, the proportion of urban employees who were unionized trended in a narrow range between 56.82% and roughly 60%. Subsequently, there was a sharp drop followed by a rebound in 2015, when the rate returned to 57.74%. In 2011, for example, 50.70% of unionized workers were urban employees. The proportion of rural employees who were union members varied between 28% and 30% until 2008, before starting a decline which pushed the rate down to 24.35% in 2015.

The proportion of self-employed workers (*autónomos*) and own-account casual workers (*avulsos*) who are union members decreased throughout the historical series presented in the sample: in the case of the self-employed, the rate almost halved from 2.61% to 1.39%. The same phenomenon occurred in the case of own-account workers, whose unionization rate decreased from 0.41% in 2007 to 0.20% in 2015. The rate also decreased in the independent professionals group, from 5.33% to 2.52% at the end of the historical series represented in the sample.

³ In Brazil, a distinction is made between self-employed workers, own-account or casual workers, and independent professionals. Self-employed workers (*autónomos*) exercise a professional activity individually and are paid for it. Own account workers (*avulsos*), who may or may not be unionized, provide urban or rural services without an employment contract to several firms that hire labour through mandated intermediaries. Independent professionals have technical or academic training in a specific area and exercise their activities independently, but these are regulated and legalized.

Table 3
Brazil: distribution of unionized workers by type of organization, 2002–2015
(Percentages)

Year	Urban employees	Rural employees	Self-employed workers	Own-account workers	Independent professionals	Other unions
2002	59.83	28.07	2.61	0.36	5.33	3.79
2003	58.43	28.78	2.50	0.38	5.04	4.87
2004	58.82	29.43	2.28	0.29	4.57	4.61
2005	59.97	30.15	2.12	0.32	3.60	3.85
2006	59.41	29.90	1.92	0.38	3.93	4.45
2007	56.82	29.70	2.04	0.41	3.23	7.81
2008	59.51	28.89	1.83	0.38	3.14	6.24
2009	59.66	27.54	1.61	0.20	2.69	8.29
2011	50.70	27.74	1.63	0.34	2.66	16.93
2012	54.58	25.76	1.67	0.23	2.55	15.22
2013	53.37	26.31	1.67	0.33	2.46	15.85
2014	54.06	26.31	1.60	0.29	2.48	15.25
2015	57.74	24.35	1.39	0.20	2.52	13.80
Average	57.15	27.92	1.91	0.32	3.40	9.30

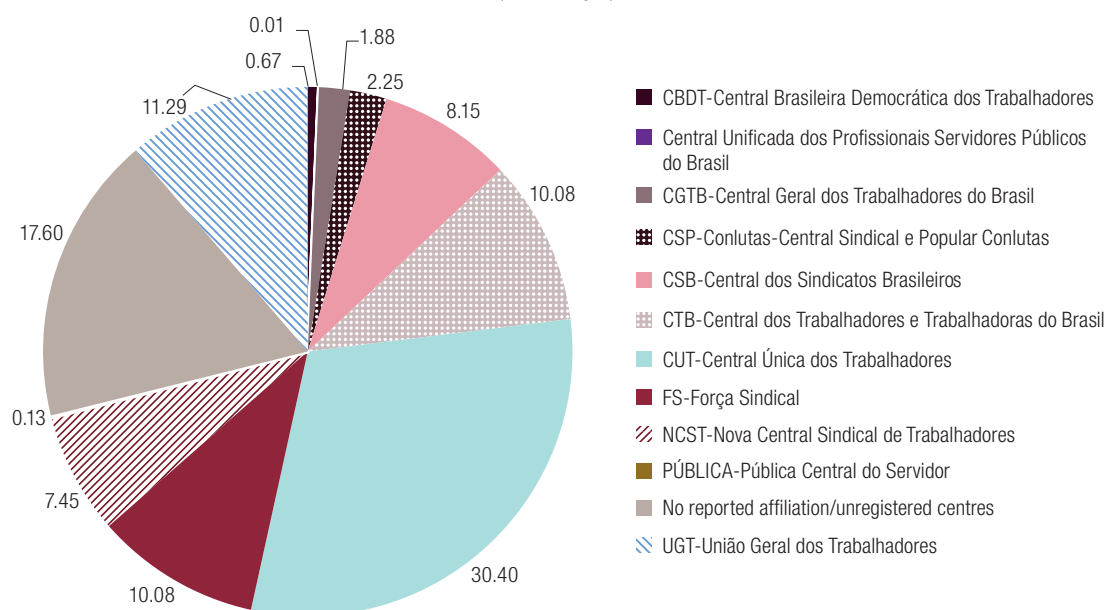
Source: Prepared by the authors on the basis of the National Household Survey (PNAD).

Note: The National Household Sample Survey does not have data for 2010, because the Brazilian census was held in that year.

Given that there were proportional reductions in several groups, the proportion of professionals who reported belonging to another type of union rose by about 4.5 times, from 3.79% to 16.93%, between 2002 and 2011. It is worth noting that, although the percentage is lower, 13.8% of individuals declaring themselves unionized in 2015 belonged to this group.

The Ministry of Labour and Employment publishes the number of affiliated entities in each union federation, as well as the number of members. The data in figure 1 refer to 2013.

Figure 1
Brazil: union membership by trade union federation, 2013
(Percentages)



Source: Prepared by the authors, on the basis of Ministry of Labour and Employment of Brazil.

In 2013, some union federations, such as Central Brasileira Democrática dos Trabalhadores (CBDT) or PÚBLICA, had fewer than 100,000 members. In contrast others had over 1 million, including Central Única de Trabalhadores (CUT), Força Sindical (FS), Central de Sindicatos do Brasil (CSB) and Central de Trabalhadores e Trabalhadores do Brasil (CTB).

Of the approximately 13 million affiliated workers, 30.40% are members of CUT; but 17.6% of workers have union membership that is not registered with the government or affiliated to any federation. In other words, nearly one in every two workers is either a member of CUT, or not a member of any federation, or else is affiliated to a federation that is not registered. The General Workers Union (*União Geral dos Trabalhadores* – UGT), FS and the CTB are also strong federations, accounting for approximately 31.5% of total union members.

Table 4 displays the foregoing data, by federative unit; in other words, the number of workers affiliated to unions in each state or the Federal District, or in unions that are not attached to a specific federative unit. The states with the largest number of labour unions are Minas Gerais, São Paulo and Rio Grande do Sul, with more than 1,000 each. The states with the fewest are Acre and Roraima, with 37 each. Roraima is the state with the fewest affiliates. São Paulo, the most populous state, also has the largest number of union members, at close to 3 million.

Table 4
Brazil: unions and unionized workers by federative unit, 2013

Federative unit	Unionized	Unions	Federative unit	Unionized	Unions
-	487	1	PA	261 073	290
AC	14 387	37	PB	265 227	338
AL	150 334	192	PE	1 013 362	394
AM	107 326	148	PI	400 566	284
AP	24 913	44	PR	747 638	825
BA	864 051	558	RJ	766 588	546
CE	480 061	427	RN	204 333	241
DF	395 269	136	RO	100 553	112
ES	336 316	229	RR	12 935	37
GO	206 223	341	RS	953 184	1 056
MA	409 830	335	SC	562 895	717
MG	1 000 504	1 163	SE	144 395	197
MS	201 372	297	SP	2 952 909	1 670
MT	128 070	258	TO	50 510	74
Total	12 755 446	10 947			

Source: Prepared by the authors, on the basis of Ministry of Labour and Employment of Brazil.

Note: Acre (AC); Alagoas (AL); Amazonas (AM); Amapá (AP); Bahia (BA); Ceará (CE), Distrito Federal (DF) Espírito Santo (ES); Goiás (GO); Maranhão (MA); Mato Grosso (MT); Mato Grosso do Sul (MS); Minas Gerais (MG); Pará (PA); Paraíba (PB); Paraná (PR); Pernambuco (PE); Piauí (PI); Roraima (RR); Rondônia (RO); Rio de Janeiro (RJ); Rio Grande do Norte (RN); Rio Grande do Sul (RS); Santa Catarina (SC); São Paulo (SP); Sergipe (SE); Tocantins (TO).

IV. Data and models

This section is divided into three subsections, dealing respectively with: the source data, descriptive statistics and the dependent and independent variables; the theoretical models, describing the theoretical basis behind the economic models; and the empirical models, namely unconditional quantile regression models, OLS with Oaxaca-Blinder and propensity score matching.

1. Data

The aim of the National Household Survey (PNAD), conducted by the Brazilian Geographical and Statistical Institute (IBGE), is to collect socioeconomic and demographic data. The survey is generally conducted annually (except for years in which a demographic census is being held), with the aim of gathering information on the profile of the inhabitants of selected households (sex, age, education, work and income, in addition to household characteristics); and, depending on demand, to collect data on migration and fertility, among other matters.

In 2015, IBGE collaborated with the Ministry of Labour and Employment and the International Labour Organization (ILO), to produce the PNAD supplement on aspects of labour relations and unionization (*Pesquisa Nacional por Amostra de Domicílios: Aspectos das Relações Trabalhistas e Sindicalização*) (IBGE, 2015). Unlike the annual survey, which targets workers aged 10 years or older, this survey studied workers aged 16 years or older and examined the situation of unionized workers in the inference period, generating data on union density in Brazil.

Table 5 reports the descriptive statistics on the variables. The dependent variable is the wage per hour worked in the week. The main variable of interest is binary and measures the effect of the person being unionized. When only those who work are considered, the total number of union members drops by almost 1.2 percentage points from 18.84% to 17.65%.

Table 5
Descriptive statistics

Variable	PNAD 2015		SUPPLEMENT ^a	
	Mean	Standard deviation	Mean	Standard deviation
Hourly wage	13.76	63.68	17.28	68.94
Education level	9.08	4.12	9.66	4.24
Experience	23.04	13.30	24.89	12.88
In-firm experience	7.11	8.48	8.97	9.69
Married	62.93%	0.48	70.82%	0.45
Head of household	48.45%	0.50	56.62%	0.50
Metropolitan area	33.47%	0.47	32.51%	0.47
Urban area	87.99%	0.32	84.72%	0.36
White	46.35%	0.50	49.89%	0.50
Formal	49.42%	0.50	69.56%	0.46
Male	59.23%	0.50	64.15%	0.48
Unionized	17.25%	0.38	-	-
Made use of the union	-	-	19.67%	0.40
Participated in the union	-	-	14.3%	0.35
Active union	-	-	12.25%	0.33

Source: National Household Survey (PNAD), 2015.

Note: The wage variable is expressed in reais per hour; the variables education level, experience and in-firm experience are expressed in years; and the other variables are expressed in percentages.

^a Part of the supplement *Pesquisa Nacional por Amostra de Domicílios: Aspectos das Relações Trabalhistas e Sindicalização*, in which workers report being unionized.

The control variables are level of education (individuals have an average of 8.5 years of schooling), experience in the labour market (average of 23.2 years) and in-firm experience (average of 7.62 years). These variables are directly related to productivity and are consonant with the income equation (Mincer, 1974).

The independent variables are linked to family and place of work: being head of household, being married, living in a metropolitan region or an urban area. These variables are used in the literature because they are directly associated with wage variations (Cirino and Lima, 2016; Hoffmann and Ney, 2004; Jacinto and Rodeghiero, 2015; Reis and Ramos, 2011; Santos and others 2010).

Dummy variables were also included to represent the type of employment, in each state, each type of occupation listed in the Brazilian Classification of Occupations (CBO) and according to the type of occupation from the National Classification of Economic Activities (CNAE). These variables are important because they separate the wage effect from the fact that employees and employers have different wages, as well as being able to be in different branches of occupation and activity. In addition, the state is considered since, as noted above, unionization patterns vary between them (Lu, Tao and Wang, 2010; Teixeira and Menezes-Filho, 2012).

The variables of interest for the degree of unionization hypothesis (hypothesis 2) are: *Used the union* (V90882), which is a dummy variable referring to the use (or non-use) of services such as legal assistance, medical-dental agreements, among others; *Participated in the union* (V90884), which is also a binary variable identifying persons who reported having participated in union activities (lectures, meetings, among others); and lastly, *Active union* (V90887), which is a binary variable in which the respondent indicates whether the union took action of any kind (negotiation or collective dispute) in the 365 days prior to the interview. It should be noted, however, that the part of the PNAD supplement on labour relations and unionization that was used refers to persons who declared themselves to be unionized.

The cross sections made in the PNAD database refer to age, occupation position and work activity group (both for the main job). Only persons aged 18 to 65 years, who are not military or civil servants and do not belong to civil service activity groups, were considered. For the supplement, in addition to the PNAD cross-sections, only unionized workers were considered, in order to understand the effects of participation, utilization and having an active union on the average earnings of unionized workers.

2. Theoretical model⁴

The wage differential that a worker gains for being unionized is called the wage gap. The difference in earnings between being unionized and non-unionized workers can be calculated using the following equation:

$$D_i = \frac{S_S^i - S_N^i}{S_N^i} \quad (1)$$

where D is the wage differential, S is the wage, the superscript i corresponds to the person and the subscripts S and N refer, respectively, to being unionized and not being unionized.

This differential refers to a single worker. The average pay differential obtained from being a member can be calculated by summing the differences between participating and not participating, so that the average difference is given by:

$$G = \frac{\sum_i^k D_i}{k} \quad (2)$$

where G is the average gain from being unionized and k is the number of individuals.

However, this difference in equations (1) and (2) does not occur because a person is or is not unionized; and production variables must be controlled for in the calculation. Hence, the average wage gap between persons of similar production characteristics but different unionization status is given by the following equation:

$$D = \frac{\bar{S}_S - \bar{S}_N}{\bar{S}_N} \quad (3)$$

The average wage differential arising from union membership is calculated as a percentage.

⁴ Based on Borjas (2009).

3. Empirical models

The models used here have two stages: estimation of wages using the income equation proposed by Mincer (1974) —with the doubly robust propensity score matching method and the Rosenbaum bounds sensitivity test— and estimation of the income equation in income percentiles, according to the RIF regression model of Firpo, Fortin and Lemieux (2009), in addition to decomposing the differential at the mean with the Oaxaca-Blinder model.

(a) Ordinary least squares

The OLS model is used to estimate the Mincerian wage equation, which consists of using the natural logarithm of hourly pay regressed against the variable of interest (union membership), along with productivity variables (education, labour market experience and firm experience) and other variables that affect wages (marital status, household responsibility, gender, race), plus a random error term (μ):

$$\ln y_i = \alpha + \beta d_sind_i + \theta X_i + \mu \quad (4)$$

where α is the intercept, β is the coefficient on each of the independent variables, X is a vector of covariates, and d_sind is the binary variable that takes value 1 when the person is unionized.

(b) Oaxaca-Blinder

The Oaxaca-Blinder decomposition model arises from the work of Blinder (1973) and Oaxaca (1973). The decomposition creates a counterfactual from the separate estimation of the earnings equations for union and non-union members. The equation is presented in the following form:

$$Y_{NS} - Y_S = X_{NS}(\beta_S - \beta_{NS}) + (X_S - X_{NS})\beta_S \quad (5)$$

where Y is the natural logarithm of hourly income, X is a vector of covariates and β represents the parameters associated with the values of X . The subscripts NS and S represent non-union and union workers, respectively.

The left side of the equality represents the union/non-union wage differential. On the right side, the first term is attributed to discrimination —a price effect— and the second term is the differential explained by observable components —the characteristic effect— (Crespo and Reis, 2004). In other words, the inter-group wage differential is decomposed into two parts: the part attributed to observable characteristics and the part considered in the literature as discrimination.

(c) Propensity score matching

Although the OLS model controls for observable characteristics that affect wages, union members may have observable and unobservable characteristics that differ from those of their non-unionized peers. Thus, a probability model is employed in which the dependent variable is *Being unionized*. Propensity score matching is a way of matching individuals according to observable characteristics, who have similar probabilities of being unionized relative to those who are not.

Two assumptions are considered for the model proposed by Rosenbaum and Rubin (1983). The first is that individuals whose probabilities are within a common range ($0 \leq p(T = 1|X) \leq 1$) are compared; and, when controlling for observable characteristics, being unionized or not

becomes random $Y_i(0), Y_i(1) \perp T_i | X$. Thus, the mean effect of being unionized or average effect on the treatment group (ATT) will be given by the difference of the mean scores found for each group: $ATT = E[Y_i(1)|T_i = 1, X = x] - E[Y_i(0)|T_i = 0, X = x]$, where $E[Y_i(1)|T_i = 1, X = x]$ is the population mean for unionized workers and $E[Y_i(0)|T_i = 0, X = x]$ is the population mean for others.

The first step in the estimation of the propensity score matching involves estimating a probability model, probit, given by:

$$P(C_i = 1 | Z) = \phi(Y_{\beta} + Z_i' \alpha) = \phi(S_{\beta} \delta) \quad (6)$$

where the cumulative density function is represented by Φ ; and S_i is a binary variable that takes the value 1 if the person is unionized. The index function is given by $Z_i' \alpha$, where Z_i is the vector of explanatory variables, which affect the unionization decision.

Having estimated the probit model, the next step is to include the estimated propensity score, $\hat{p}(x_i)$ in the regression, which can be done by imputation or reweighting. However, $\hat{p}(x_i)$ may exhibit non-positive values; so the doubly robust methodology, as highlighted by Hirano, Imbens and Ridder (2003), generates additional gains in robustness, owing to the elimination of the effect of omitted variables, and reduces the correlation between omitted and included variables. The use of the estimated propensity score is more efficient than the true $p(x_i)$. It should be noted that the propensity score is sensitive to the matching estimator, and that the functional form is sufficiently sensitive to consider the overlap hypothesis.

(d) Sensitivity

In addition to good matching, a test must be made for omitted variables that affect both the treatment group (being unionized) and the outcome variable (natural log of wages) at the same time, which may cause bias. In this way, "omitted variable bias" can be tested using Rosenbaum's bounds (Rosenbaum, 2002).

According to Rosenbaum (2002), the probability of participation (in this case, unionization) of an individual i is given by:

$$\pi_i = Pr(T_i = 1 | X_i) = F(\beta X_i + \gamma \varepsilon_i) \quad (7)$$

where π_i is the probability that the person is unionized, given a set of observable variables, X_i ; and γ is the effect of the unobservable error on the person in the treatment group ($T_i = 1$). In a case where the bias is non-existent, $\gamma = 0$, the probability of being in the treatment group is given by the observable variables alone.

In a case where the function $F(\cdot)$ has a logistic distribution and two matched individuals, i and j , with relative probabilities of being in the treatment group given by $\frac{\pi_i}{1-\pi_i}$ and $\frac{\pi_j}{1-\pi_j}$, respectively, the ratio of these probabilities is given by:

$$\frac{\frac{\pi_i}{1-\pi_i}}{\frac{\pi_j}{1-\pi_j}} = \frac{e^{(\beta X_i + \gamma \varepsilon_i)}}{e^{(\beta X_j + \gamma \varepsilon_j)}} = e^{\gamma(\varepsilon_i - \varepsilon_j)} \quad (8)$$

Equation (8) represents the ratio of the probabilities, showing that, when the omitted variables do not affect the probability of being in the treatment group, $\gamma = 0$, the ratio will be equal to 1. When this value is different from 1, there is an important variable that is omitted. In short, Rosenbaum's bounds consist of understanding how the omitted variables affect the differential given by $\varepsilon_i - \varepsilon_j$.

(e) RIF regression

The model developed by Firpo, Fortin and Lemieux (2009) is known as RIF regression, in which the dependent variable becomes an influence function (IF). The model assumes that a linear function of the independent variables can be used to model the conditional expectation of the $RIF(Y;v)$:

$$E[RIF(Y;v) | X] = X\gamma + \varepsilon \quad (9)$$

where, by means of OLS, it is possible to estimate the parameters γ .

In the quantile case, the $RIF(Y; Q_\tau)$ is an equality with $Q_\tau + IF(Y, Q_\tau)$ and can be written as follows:

$$RIF(y; Q_\tau) = Q_\tau + \frac{\tau - 1\{y \leq Q_\tau\}}{f_Y(Q_\tau)} \quad (10)$$

where $f_Y(\cdot)$ is the density of the distribution, at the margin, of Y ; Q_τ is the τ -quantile population of the distribution of Y , unconditional; and $1\{\cdot\}$ is an indicator function.

Thus, the following equation is estimated computationally:

$$RIF(\widehat{y}; \widehat{Q}_\tau) = \widehat{Q}_\tau + \frac{\tau - 1\{y \leq Q_\tau\}}{\widehat{f}_Y(Q_\tau)} \quad (11)$$

V. Results

Table 6 displays the results of the effect of unionization on wages. The estimation procedure was divided into two parts: OLS and doubly robust propensity score matching. The results are quite similar, which makes initial model more robust.

Table 6
Income equations

Models	Ordinary least squares	Propensity score matching
	(1)	(2)
Unionized	0.073*** (0.005)	0.07*** (0.005)
Education level	0.037*** (0.001)	0.043*** (0.001)
Experience	0.003*** (0.000)	0.004*** (0.000)
In-firm experience	0.008*** (0.000)	0.01*** (0.000)
Married	0.095*** (0.004)	0.07*** (0.006)
Head of household	0.071*** (0.004)	0.074*** (0.006)
Metropolitan area	0.148*** (0.004)	0.156*** (0.006)
Urban area	0.106*** (0.008)	0.099*** (0.011)
White	0.074*** (0.004)	0.07*** (0.006)

Table 6 (concluded)

Models	Ordinary least squares	Propensity score matching
	(1)	(2)
Formal	0.092*** (0.005)	0.098*** (0.011)
Male	0.135*** (0.005)	0.146*** (0.007)
Constant	1.345*** (0.406)	2.26 (169.39)
D_UF	YES	YES
D_EMPLOYMENT_TYPE	YES	YES
D_CNAE	YES	YES
D_CBO	YES	YES
No. of observations	120 090	119 870
Adjusted R ²	0.397	0.482

Source: Prepared by the authors.

Note: * significant at 1%; ** significant at 5%; *** significant at 10%. D_UF is a dummy variable for each federative unit; D_EMPLOYMENT_TYPE represents dummy variables for employer, self-employed and employee; D_CNAE refers to a dummy variable for each branch of activity of the National Classification of Economic Activities (CNAE), and D_CBO refers to a dummy variable for each occupation of the Brazilian Classification of Occupations (CBO).

The results of the production variables, such as level of education, experience and specific experience are related directly to workers' earnings. Education accounts for an average wage increase of between 3.7% and 4.3% per year of study. Labour market experience (EXP) increases wages by 0.3% to 0.4% per additional year. On average, the effect of in-firm experience is greater, between 0.8% and 1% per additional year.

The main variable of interest is the effect of being unionized on wages. As there are controls for the effects of type of work (employee, employer, self-employed), effects related to CBO occupations, activity groups and federative unit, it is possible to control for the effects of all unions on wages. Thus, responding to hypothesis 1, the effect of being unionized turned proved positive, ranging from 7% to 7.3%. Cruz and others (2016) found estimated effects of unionization on earnings of between 9.2% and 11.03% in Brazil's regions in 2006. Jacinto and Rodeghiero (2015) found that unionization had an average effect of 16% on workers' wages in the Porto Alegre metropolitan region. In the case of China, Yao and Zhong found a unionization wage premium of between 8.7% and 12.6%. The effect estimated by Campos and Moura (2017) was 8.4%; and this study estimates the wage premium at between 7% and 7.3% for the unionized group. This is lower than the premia estimated by Gonçalves and Machado (2004), of 25.38% and 19.71% in 1992 and 2001, respectively.

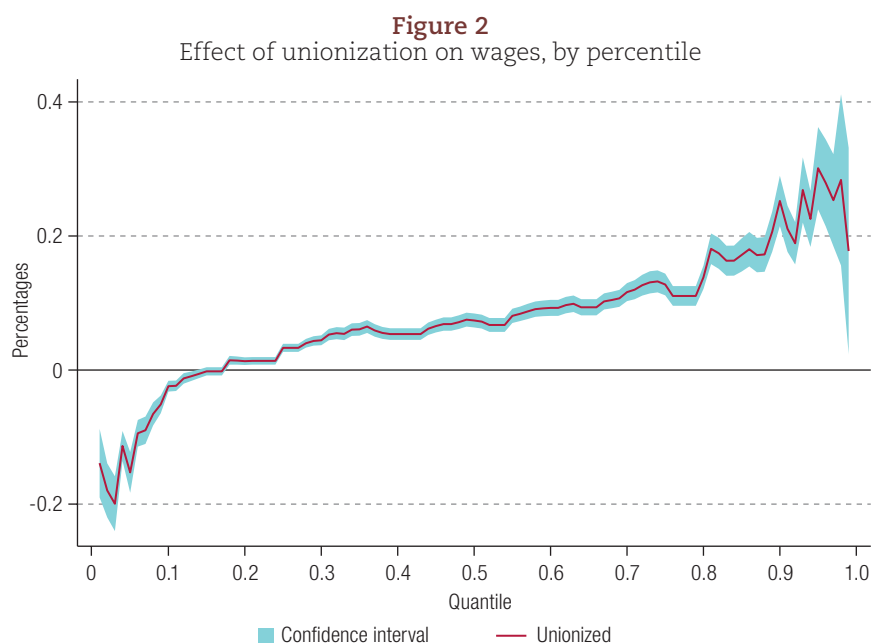
The effect of having a spouse varies between 7% and 9.5% in the models estimated. Being head of household has a negligible effect, just 0.3 percentage points (between 7.1% and 7.4%). The variables representing living in a metropolitan region or an urban area, self-declaring as white, and being male increase people's wages by: 14.8–15.6%; 9.9%–10.6%; 7%–7.4%; and 13.5%–14.6% in the two models, respectively. Individuals working in the formal sector earn between 9.2% and 9.8% more than their peers in the informal sector.

Table A1.1 in annex A1 separates the wage differential between the groups into two components: explained and unexplained — the Oaxaca-Blinder model. The total differential is 28.72% in favour of union members. This is almost entirely explained by production factors (21.41%), such as education level, experience and the other explanatory variables of the OLS and propensity score-matching models. The unionization wage premium is the component explained solely and exclusively by union membership, at 7.31%. In other words, roughly 75% of the wage differential between the groups is observable in the characteristics of the individuals.

Annex A1 presents some important results that corroborate the effect found in table 6 by robust propensity score matching. First, figure A1.1 reports the kernel density before and after matching. While the groups (union and non-union workers) are very different before matching, afterwards they become very similar. Table A1.2 reports the test of means of the variables used in the model, before and after matching. This shows that, before the matching procedure, most of the variables were statistically different between the groups. However, after matching, only two variables remained statistically different between groups.

For the results to be reliable, a sensitivity test needs to be performed. The Rosenbaum bounds omitted variable bias test is reported in table A1.3. The result shows that, at a 1% confidence level, it is possible to reject the null hypothesis that there is an omitted variable that simultaneously affects the treatment (unionization) and the outcome (natural log of hourly wage).

Figure 2 shows the effect of unionization on wages at each income percentile. Estimates were made using the non-conditional quantile (RIF) regression model. The differential varies between approximately 15% against unionized workers and 25% in favour of unionized workers. As shown in table 7, unionization has a positive effect on mean wages.



Source: Prepared by the authors.

Table 7
Income equation for unionized workers

Variables	Ordinary least squares
Made use of the union	0.011 (0.013)
Participated in the union	-0.000 (0.015)
Active union	0.043*** (0.012)
Education level	0.043*** (0.002)
Experience	0.003*** (0.000)

Table 7 (concluded)

Variables	Ordinary least squares
In-firm experience	0.011*** (0.001)
Married	0.058*** (0.01)
Head of household	0.063*** (0.01)
Metropolitan area	0.175*** (0.012)
Urban area	0.118*** (0.02)
White	0.067*** (0.01)
Formal	0.085*** (0.021)
Male	0.168*** (0.013)
Constant	1.993*** (0.074)
D_UF	YES
D_EMPLOYMENT_TYPE	YES
D_CNAE	YES
D_CBO	YES
No. of observations	19 150
R ² -Adjusted	0.521

Source: Prepared by the authors.

Note: * significant at 1%; ** significant at 5%; *** significant at 10%. D_UF is a dummy variable representing each federative unit; D_EMPLOYMENT_TYPE represents dummy variables for employer, self-employed and employee; D_CNAE is a dummy variable representing each branch of activity of the National Classification of Economic Activities (CNAE), and D_CBO is a dummy variable for each occupation of the Brazilian Classification of Occupations (CBO).

Figure 2 shows that up to about the tenth percentile, union members earn less than non-union members. From that point on, the trajectory increases gradually at each percentile, with small oscillations. The results also show that there is a substantial increase in union earnings between the seventy-fifth and eighty-fifth percentile.

Contrary to findings in studies targeted on the United States (Firpo, Fortin and Lemieux, 2009), the greatest protection is given to the wealthiest and not to the lower paid (who are even harmed). The interpretation of the RIF regression model is a higher level of analysis, not just local as in the case of conditional quantile regressions. As shown in figure 2, increasing the number of unionized workers reduces the wages of the lowest paid (first 10 percentiles). Thereafter, earnings rise incrementally until the end of the distribution, where the gains exceed 20%. Firpo, Fortin and Lemieux (2009) find gains of 0.195 for the first decile and losses of 0.135 in the last decile, associated with an increase in the unionization rate. For Brazil, using conditional quantile regressions, Campos and Moura (2017) find that the gains from being unionized increase between the start and the end of the wage distribution. However, the results are analysed by focusing on the income quantile and, in general, the gains across the distribution are statistically identical to the gain at the mean. The result of the unconditional quantile regression, shown in figure 2, indicates the extent to which an increase in unionization affects wages in general across the quantiles. In this case, in the lowest quantiles wages fall, indicating that unions defending those workers' interests are inefficient from the standpoint of financial gains.

Table 7 presents the result of the wage equation among unionized workers. The sample consists of 19,150 individuals. The equation was estimated with a fairly large number of explanatory variables, which partially justifies the adjusted R^2 of above 50%. The aim is to test hypothesis 2, in other words to verify the effect of the degree of unionization on wages. The degree of unionization is measured by the answers given by individual respondents.

The variables representing being a union participant or making use of the union are not significant in the wage equation. However, the mean wage of unionized workers who declare that their union participated in a dispute or collective agreement in the most recent period, is 4.3% higher than that of unionized workers whose union did not participate in any agreement.

Although unionized workers have a higher level of schooling than the overall sample, the gain per year of study is higher, at 4.3%. Experience generates gains of 0.3% per additional year, while the gain from in-firm experience is 1.1% per year. The other controls explaining wages are also significant: being married increases unionized workers' wages by 5.8%, while being a head of household raises them by 6.2%.

Binary variables related to place of residence show gains for people in urban areas and metropolitan regions of 12.6% and 17.7%, respectively. Two variables related to personal characteristics, being male and self-reporting as white, increase wages by 16.8% and 6.7%, respectively. On average, the wages of unionized workers in the formal sector are 8.5% higher than those of their informal-sector peers.

VI. Final remarks

There are roughly 17,000 unions in Brazil and about 17% of people are union members; of these, 73% live in urban areas and represent 68% of all unionized workers.

The initial motivation of this paper was to test the hypothesis that unions affect wages and that the effects are likely to differ across the wage distribution. An additional aim was to test whether the degree of unionization affects wages.

To answer these questions, microdata from the 2015 PNAD and its supplement on labour relations and unionization were used. The methodology employed consisted of OLS models (with Oaxaca-Blinder decomposition), propensity score matching and RIF regression.

The results show that, on average, unionized workers in the sample earn between 7.3% and 7.5% more than their non-unionized peers. These findings are consistent with those reported in the literature: being unionized has a positive effect on wages. When income percentiles are considered, the gain is greater towards the upper tail of the distribution. Nonetheless, higher rates of unionization generally cause a reduction in the wages of the lowest paid workers (up to the tenth percentile).

Hypothesis 2, that the degree of unionization affects wage levels among unionized workers is not confirmed. However, if the unionized worker declares that his or her union participated in collective agreements or disputes in the 365 days prior to the survey (active union), a wage difference of about 4.3% is observed.

As a caveat, it is impossible to capture the indirect effect of collective agreements that affect both unionized and non-unionized individuals. This effect should make the wages of the two groups very similar, but the results do not bear this out.

A closer analysis shows that, although unions have undergone various transformations in the last four decades, as institutions they remain strong, since union membership has a direct effect on the income of unionized workers. It is worth noting that these results weaken a device that has been in place since the 2017 labour reform of individual bargaining between employers and employees. Although it

was impossible to prove income discrimination according to the individual's degree of participation in the union, the wages of unionized workers who reported that their unions had been active in negotiations in the last 365 days were 4.3% higher, on average, than those of their peers whose unions were less active. Moreover, differences in union effects across the distribution may reflect the fact that different occupational categories have different bargaining powers when negotiating for better wages. These differences may also be signalling a decline in the number of small unions, which aggregate within the larger trade union federations.

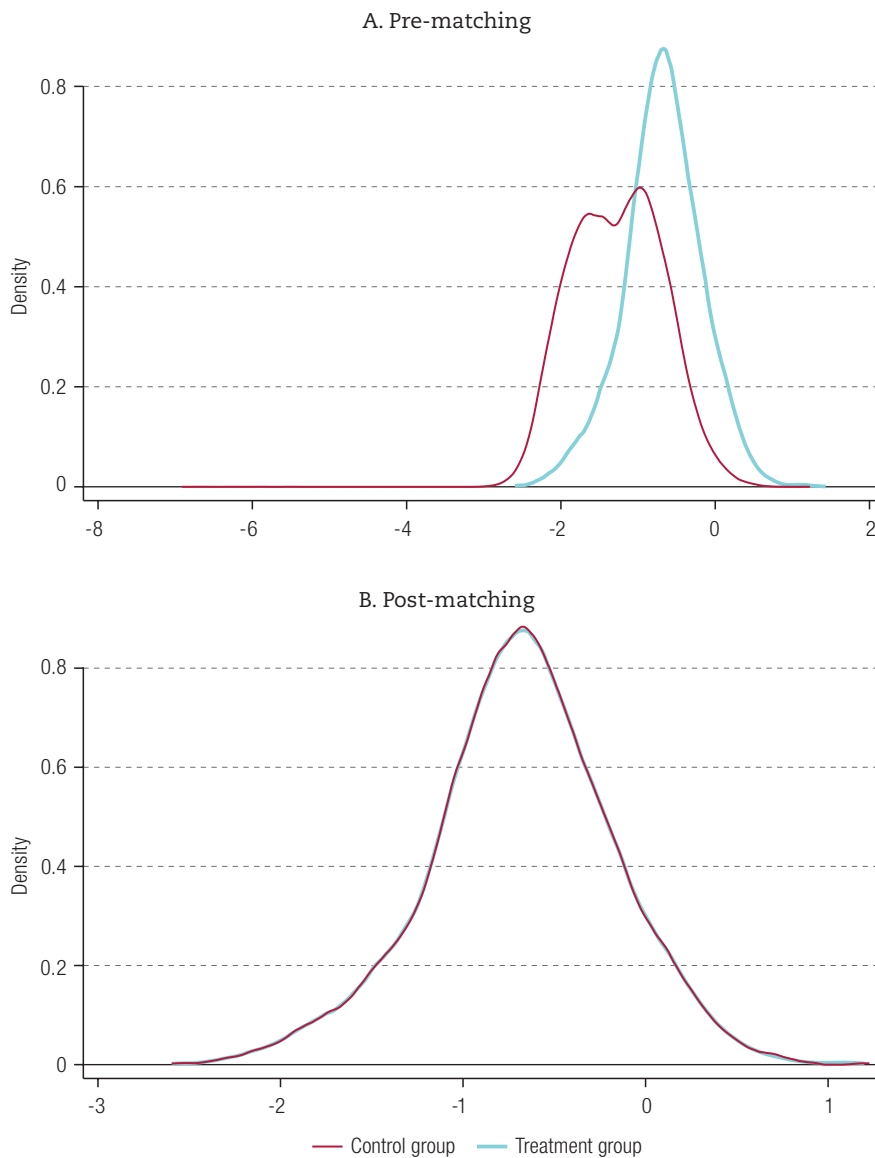
Bibliography

- Becker, K. L. (2015), "Concentração do emprego industrial, sindicatos e salários: uma aplicação para o caso da indústria brasileira de metalurgia", *Revista Brasileira de Estudos Regionais e Urbanos*, vol. 4, No. 1.
- Blinder, A. S. (1973), "Wage discrimination: reduced form and structural estimates", *Journal of Human Resources*.
- Bloch, F. E. and M. S. Kuskis (1978), "Unions and wages: wage determination in the union and nonunion sectors", *ILR Review*, vol. 31, No. 2.
- Borjas, G. (2009), *Economia do Trabalho*, Porto Alegre, AMGH Editora.
- (1979), "Job satisfaction, wages, and unions", *Journal of Human Resources*.
- Campos, A. G. (2014), "Dilemas do trabalho: sindicatos no Brasil hoje", *Texto para Discussão*.
- Campos, A. G. and A. C. Moura (2017), "Diferenças de remuneração entre trabalhadores sindicalizados e não sindicalizados: evidências sobre o mercado de trabalho brasileiro", *Texto para Discussão*, Brasília, Institute of Applied Economic Research (IPEA).
- Cardoso, A. (2014), "Os sindicatos no Brasil", Institute of Applied Economic Research (IPEA) [online] http://repositorio.ipea.gov.br/bitstream/11058/3765/1/bmt56_nt01_sindicatos_brasil.pdf.
- Cirino, J. F. and J. E. de Lima (2016), "Diferenças de rendimento entre as Regiões Metropolitanas de Belo Horizonte e Salvador: uma discussão a partir da decomposição de Oaxaca-Blinder", *Revista Econômica do Nordeste*, vol. 43, No. 2.
- Crespo, A. R. V. and M. C. Reis (2004), "Decomposição do componente de discriminação na desigualdade de rendimentos entre raças nos efeitos idade, período e coorte", *Anais do XXXII Encontro Nacional de Economia*, Natal, National Association of Postgraduate Programs in Economics (ANPEC).
- Cruz, M. S. Da and others (2016), "Determinantes das desigualdades salariais nas macrorregiões brasileiras: uma análise para o período 2001–2006", *Revista Econômica do Nordeste*, vol. 42, No. 3.
- Duncan, G. M. and D. E. Leigh (1980), "Wage determination in the union and nonunion sectors: a sample selectivity approach", *ILR Review*, vol. 34, No. 1.
- Ebbinghaus, B., C. Göbel and S. Koos (2011), "Social capital, 'Ghent' and workplace contexts matter: Comparing union membership in Europe", *European Journal of Industrial Relations*, vol. 17, No. 2.
- Firpo, S., N. M. Fortin and T. Lemieux (2009), "Unconditional quantile regressions", *Econometrica*, vol. 77, No. 3.
- Gonçalves, É. and A. F. Machado (2004), "Análise da composição do rendimento do trabalhador empregado no Brasil: uma abordagem de oferta, demanda e instituições (1992–2001)", *Economia*, vol. 5, No. 1.
- Gyourko, J. and J. Tracy (1988), "An analysis of public- and private-sector wages allowing for endogenous choices of both government and union status", *Journal of Labour Economics*, vol. 6, No. 2.
- Heckman, J. J. (1979), "Sample selection bias as a specification error", *Econometrica*.
- Hirano, K., G. W. Imbens and G. Ridder (2003), "Efficient estimation of average treatment effects using the estimated propensity score", *Econometrica*, vol. 71, No. 4.
- Hoffmann, R. and M. G. Ney (2004), "Desigualdade, escolaridade e rendimentos na agricultura, indústria e serviços, de 1992 a 2002", *Economia e Sociedade*, vol. 13, No. 2.
- IBGE (Brazilian Geographical and Statistical Institute) (2015), *Pesquisa Nacional por Amostra de Domicílios: aspectos das relações trabalhistas e sindicalização*, Brasília.
- Jacinto, P. de A. and L. B. Rodeghiero (2015), "Retorno em escolaridade: um estudo para a região Metropolitana de Porto Alegre", *Revista Brasileira de Estudos Regionais e Urbanos*, vol. 6, No. 1.
- Lu, Y., Z. Tao and Y. Wang (2010), "Union effects on performance and employment relations: evidence from China", *China Economic Review*, vol. 21, No. 1.
- Mellow, W. (1982), "Employer size and wages", *The Review of Economics and Statistics*.

- Menezes-Filho, N. A. and others (2002), "Unions and the economic performance of Brazilian establishments", *Research Network Working papers*, No. R-464, Inter-American Development Bank (IDB).
- Mincer, J. (1974), "Schooling, experience, and earnings", *Human Behavior & Social Institutions*, No. 2.
- Mishel, L. (2012), "Unions, inequality, and faltering middle-class wages", *Issue Brief*, vol. 342.
- Oaxaca, R. (1973), "Male-female wage differentials in urban labour markets", *International Economic Review*, vol. 14, No. 3.
- Reis, M. C. and L. Ramos (2011), "Escolaridade dos pais, desempenho no mercado de trabalho e desigualdade de rendimentos", *Revista Brasileira de Economia*, vol. 65, No. 2.
- Rosenbaum, P. R. (2002), *Observational Studies*, Springer.
- Rosenbaum, P. R. and D. B. Rubin (1983), "The central role of the propensity score in observational studies for causal effects", *Biometrika*, vol. 70, No. 1.
- Santos, G. C. and others (2010), "Mercado de trabalho e rendimento no meio rural brasileiro", *Economia Aplicada*, vol. 14, No. 3.
- Stan, S. and R. Erne (2016), "Is migration from Central and Eastern Europe an opportunity for trade unions to demand higher wages? Evidence from the Romanian health sector", *European Journal of Industrial Relations*, vol. 22, No. 2.
- Teixeira, W. M. and N. A. Menezes-Filho (2012), "Estimando o retorno à educação do Brasil considerando a legislação educacional brasileira como um instrumento", *Revista de Economia Política*, vol. 32, No. 3.
- Western, B. and J. Rosenfeld (2011), "Unions, norms, and the rise in US wage inequality", *American Sociological Review*, vol. 76, No. 4.
- Yao, Y. and N. Zhong (2013), "Unions and workers' welfare in Chinese firms", *Journal of Labour Economics*, vol. 31, No. 3.
- Yilmaz, E. and S. San (2017), "Wage gap and dispersion in a partially unionized structure in Turkey", *Empirical Economics*, vol. 52, No. 2.

Annex A1

Figure A1.1
Kernel density before and after matching



Source: Prepared by the authors.

Table A1.1
Oaxaca-Blinder Decomposition

	Coefficient
Non-unionized	1.9763*** (0.002)
Unionized	2.2635*** (0.006)
Difference	-0.2872*** (0.007)
Explained	-0.2141*** (0.004)
Not explained	-0.0731*** (0.005)

Source: Prepared by the authors.

Note: Robust standard errors in parentheses. Significant at: 1%***, 5%** , 10%*.

Table A1.2
Means of variables before and after matching

Variable	Before			After		
	Unionized	Non-unionized	t	Unionized	Non-unionized	t
Education level	10.44	9.27	***	10.44	10.44	
Experience	23.29	22.01	***	23.28	23.26	
In-firm experience	4.46	6.26	***	7.45	7.21	***
Married	0.69	0.60	***	0.69	0.69	
Head of household	0.54	0.45	***	0.54	0.54	
Metropolitan area	0.46	0.43	***	0.46	0.46	
Urban area	0.94	0.94		0.94	0.93	***
White	0.47	0.43	***	0.47	0.78	
Formal	0.78	0.46	***	0.78	0.78	
Male	0.61	0.55	***	0.61	0.61	

Source: Prepared by the authors.

Table A1.3
Rosenbaum bounds test for bias in matching (2002)

Rosenbaum bounds	p-value	Rosenbaum bounds	p-value	Rosenbaum bounds	p-value
1.00	0.00	1.20	0.00	1.40	0.00
1.05	0.00	1.25	0.00	1.45	0.00
1.10	0.00	1.30	0.00	1.50	0.00
1.15	0.00	1.35	0.00		

Source: Prepared by the authors.