

Barufi, Ana; Haddad, Eduardo; Nijkamp, Peter

Conference Paper

A comprehensive analysis of the wage curve in Brazil: Non-linearities, urban size, and the spatial dimension

56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Barufi, Ana; Haddad, Eduardo; Nijkamp, Peter (2016) : A comprehensive analysis of the wage curve in Brazil: Non-linearities, urban size, and the spatial dimension, 56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/174643>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A comprehensive analysis of the wage curve in Brazil: non-linearities, urban size, and the spatial dimension

Ana Maria Bonomi Barufi, Eduardo Amaral Haddad, Peter Nijkamp

Abstract

Agglomeration economies have a relevant impact on local labour markets. The interaction of workers and firms in dense urban areas may generate productivity advantages that result in higher wages. City size has an important impact on the relative bargaining power of workers and firms in the labour market. When analysing the relationship of local wages and the business cycle, wage flexibility, measured by the wage curve, is higher in informal sectors in less dense areas in Brazil. Therefore, large agglomerations are supposed to provide a higher bargaining power for workers, as they have further job opportunities. In addition, labour market dualism is an essential ingredient in the evaluation of the wage curve in developing economies. However, this type of analysis should be conducted in the adequate regional level (labour market areas), making it possible to find a relevant impact of city size on the relative bargaining power of workers and firms.

Keywords: wage curve, city size, labour market

1 Introduction

The wage curve represents an empirical negative relation between wages and local unemployment rates at the regional level (BLANCHFLOWER; OSWALD, 1994). This result is expected, as wage flexibility plays a relevant role in adjusting local economies to adverse shocks. Moreover, causality is expected to flow from unemployment size to the level of wages (BLANCHFLOWER; OSWALD, 1995). Many empirical studies have corroborated the original formulation of the wage curve by finding negative and significant relationships in different environments.

Within the context of a developing country, with an expressive informal sector and significant heterogeneity in terms of population density, the wage curve can be an expression of the degree of flexibility in the labour market, given the relative bargaining power of workers and firms and the different costs faced by firms in each sector. City size is one additional factor in this equation, by increasing the number of opportunities for a matching between firms and workers. The investigation of this subject provides an opportunity to further understand how urbanisation economies affect local labour markets, as different city sizes can imply that firms and workers may have more or less precautionary behaviour.

The higher sectorial diversity usually found in bigger cities is one of the factors that contributes to these special conditions, by accommodating uncorrelated random sector-specific shocks and reducing uncertainty (GORDON; MCCANN, 2000). Furthermore, urban areas concentrate diverse and complementary economic activities under intense competition (GLAESER, 2011). In this setting, the interaction of economic agents provides different locational advantages, among which it is possible to highlight the higher probability of a successful matching in the labour market (MARSHALL, 1890). Therefore, not only is the local economic environment supposed to be more resilient to economic shocks, but there is also a broader range of opportunities for firms and workers to become adjusted to adverse economic situations.

Moreover, in the context of developing countries, an additional element becomes relevant, namely the existence of an informal sector and its consequences in labour market outcomes (FIELDS, 2005). In this sense, exploring different aspects of the wage curve for a developing country such as Brazil would provide the opportunity to understand additional effects of agglomeration economies in the labour market and their impacts on different labour market arrangements.

This analysis provides a comparison of the formal and the informal sectors in Brazil, and how their interaction is affected by the intensity of agglomeration. Furthermore, different aspects of the wage

curve in Brazil are investigated, with varying databases, regional aggregation, time periods, and data structure. Finally, an analytical framework based on different bargaining power levels for workers in the formal and the informal sector is proposed, with a direct implication of urban size over the relative power of workers and firms.

This paper is organised as follows: Section 2 presents a literature review on the wage curve, Section 3 explores the empirical literature of the wage curve in different countries and in Brazil in particular, and Section 4 discusses the literature on labour market segmentation and proposes a theoretical framework of analysis. Furthermore, the empirical strategy is presented in Section 5, Section 6 provides a description of the databases and the main results, and Section 7 presents the main conclusions of the paper.

2 The wage curve

The wage curve represents an empirical relation between the level of wages and the unemployment rate in different regions. Considering unemployment on the x-axis and wages on the y-axis, this relation is characterised by a downward sloping curve (BLANCHFLOWER; OSWALD, 1995, 2005). A similar concept but with relevant differences is the Phillips curve, which is concerned with the relationship between the level of unemployment and wage growth. Moreover, while the former has microeconomic foundations and can be discussed in a regional context, the latter is usually considered within a macroeconomic perspective.

A theoretical challenge associated with the wage curve is its regularity between countries and periods, despite the significant institutional differences among them. The wage curve can be interpreted as a measure of the flexibility of the labour market (BLANCHFLOWER; OSWALD, 1995), and several mechanisms have been proposed to explain it, based on imperfect competitive structures:

- Bargaining power: a higher unemployment rate can reduce the bargaining power of workers, as it may be harder for them to find another job in a depressed economic environment. On the other hand, a lower unemployment rate means that firms will have more difficulty finding a replacement that matches their needs. In the first case, workers may lose their ability to demand a higher share of the surplus, while just the opposite occurs in the second case, resulting in higher wages;
- Efficiency wage: firms pay a certain wage aiming to increase labour productivity and workers choose the effort expended at work. In this context, wages influence productivity, but when the unemployment rate is higher, shirking may have a greater penalty if it takes a long time to find a job; consequently, workers will put more effort in the job, even if the payment is low. Then, according to Card (1995), group-specific unemployment rates should be more relevant (for instance, for individuals with different skill levels). Moreover, high and low-amenity regions will be located in different parts of the wage curve;
- Labour contracts: considering the maximisation of the joint welfare of employers and employees, unemployed workers contribute less to the surplus of this combination. Therefore, with a higher wage, it will be interesting to have more individuals employed (leading to a lower unemployment rate);
- Turnover costs: labour markets are perceived as ‘thin’ by workers and the wage curve represents the degree of monopsonistic power of firms, which is generated either by turnover costs or horizontal job differentiation based on geographic location. Then, instead of the perfectly elastic supply curve of the competitive model, firms face an upward-sloping supply curve in the local labour market (MANNING, 2003).

Some of these mechanisms have not necessarily become structured theoretical models, and most importantly, they have not been empirically validated (NIJKAMP; POOT, 2005). An exception to this pattern is the model by Campbell and Orszag (1998) based on the monopsonistic competition structure generated by turnover costs, further confirmed empirically by Morrison, Papps and Poot

(2006). Alternatively, Sato (2000) provided a simple search and matching framework combined with local productivity differentials. Their main conclusion was that regions with higher productivity would have higher wages and lower unemployment rates in equilibrium. These regions also present a higher concentration of people, and consequently, significant congestion costs that ensure the persistence of spatial inequalities.

The wage curve is observed by comparing local labour markets. In this respect, there are two possible ways to see the interaction of spatial units and urban size: on the one hand, neighbouring areas can increase accessibility to job offers; on the other hand, agglomeration size may increase local opportunities and influence the relationship of wages and the unemployment rate. Accessibility to jobs can affect the balance of the bargaining power of workers and employers (LONGHI; NIJKAMP; POOT, 2006), generating a positive effect over the local wage level (wages offered in regions with higher accessibility to jobs should be higher). In addition, the unemployment elasticity to pay is expected to be negatively affected by a higher level of accessibility, meaning that more integrated regions are supposed to show a lower elasticity.

Furthermore, agglomeration economies may increase local productivity, leading to higher wages and lower unemployment rates (Sato, 2000). Therefore, they may be essential to determine the location of different local labour markets in the wage curve. It is also possible that local labour markets are subject to common shocks, with spatially autocorrelated residuals, and spatial proximity can lead to a direct negative effect of unemployment in the neighbouring areas over local wages (LONGHI, NIJKAMP; POOT, 2006).

With these considerations in mind, space and agglomeration size are essential elements in the analysis of the wage curve. The effects of agglomeration are not restricted to the size of the labour market, but also to higher local productivity. These effects can be heterogeneous depending on the degree of formalisation of work relationships. Therefore, formal and informal workers may have their bargaining power affected by local market size and higher local productivity to a different extent. This issue will be further explored in the following sections.

3 Empirical literature

The wage curve has been investigated in the literature as a measure of the flexibility of the labour market in relation to economic downturns, with the unemployment rate synthetizing the business cycle. As mentioned before, many studies have found a significant and negative relationship between wages and the unemployment rate, which is usually explained by market imperfections in the labour market. However, in developing countries there is an additional dimension to be taken into account: the labour market is partially comprised by an informal sector, with more flexible work arrangements.

The empirical literature focused on the analysis of this relationship in a comparison between countries or regions, controlling for individual characteristics and other microeconomic aspects. Nijkamp and Poot (2005) provided a meta-analysis of the coefficients estimated in more than 200 empirical studies, finding that there is a recurrent negative relationship between wages and the local unemployment rate. Correcting for publication and aggregation bias, this elasticity is estimated to be around -0.07.

The most recent advances in the empirical literature have involved a stronger effort to control for worker heterogeneity based on observed and unobserved characteristics. This was made possible by the recent spur of surveys with a longitudinal structure at the individual level, as in Baltagi, Blien and Wolf (2009, 2012) for Germany, and Turunen (1998) for the United States. The latter disaggregated the wage curve by year and worker characteristics, finding very different coefficients depending on the group analysed. In addition, there has been an effort in the literature to control for the possible endogeneity of the unemployment rate (instrumented by its time lag).

The inclusion of agglomeration effects and an explicit spatial dimension has followed the most recent advances in spatial econometrics techniques (BALTAGI; BLIEN; WOLF, 2012; BALTAGI;

ROKICKI, 2014b; ELHORST; BLIEN; WOLF, 2007; LONGHI; NIJKAMP; POOT, 2006). Most studies have found that unemployment in surrounding areas also affects the equilibrium in the local labour market, as workers are supposedly aware of job opportunities in these areas. In the absence of spatial spillovers, the local unemployment elasticity is likely to be overestimated (BALTAGI; ROCICKI, 2014b). Longhi, Nijkamp and Poot (2006) noted that the unemployment elasticity is lower in regions that interact more with other regions, meaning that more isolated areas have a less flexible labour market. Similar studies were conducted by Ramos, Nicodemo and Sanromá. (2015) for Spain, Fingleton and Palombi (2013) in the UK, Baltagi and Rokicki (2014b) for Poland, among others.

As mentioned above, in the context of developing countries, usually a large share of the labour market is unregulated, based on informal relations and does not follow the legislation on the matter. This has important implications for the bargaining relationships in the labour market, as the informal sector represents an additional status apart from formal jobs and unemployment. Therefore, economic shocks may be absorbed in this intermediate sector, with wages varying significantly given a change in the local unemployment rate. In Colombia, Ramos, Duque and Surinach (2010) found that the elasticity of individual wages to the unemployment rate is much higher in the case of informal workers. Baltagi, Baskaya and Hulagu (2013), who estimated different wage curves for the formal and the informal sectors in Turkey, provided a similar result. When controlling for the potential endogeneity of unemployment (FE-2SLS with the lagged value of the unemployment rate as an instrument), they found a coefficient of -0.071 for the formal sector and -0.263 for the informal sector (and -0.107 when both were considered together). Higher wage flexibility was also found for women, young individuals, with low tenure and low education.

Labour market dualism and wage differentials between workers in the formal and informal sectors are very relevant for the degree of wage flexibility in these sectors. However, Menezes-Filho, Mendes and Almeida (2004) showed that when individual observed and unobserved characteristics are taken into account, most of the differences between wages in the formal and the informal sectors in the Brazilian labour market disappear. Controlling for these determinants, the wage level in the informal sector is actually higher than the one in the formal sector (probably as a compensation for providing less benefits). Therefore, there is an indication of a self-selection process of more qualified workers in the formal sector, and any analysis trying to assess labour market outcomes under the perspective of sector dualism should control for it.

Regarding the interaction of the formal and informal sectors, many empirical and theoretical studies have discussed the existence of market segmentation, which can be defined as the existence of a wage gap between jobs in these two sectors. In this sense, controlling for worker, sector and firm observed heterogeneity, a job in the formal sector is supposed to pay a higher salary. With PME data on a more recent period (2003-2008), Corseuil and Foguel (2012) estimated a model for the Brazilian labour market with a similar structure to a wage curve (controlling for individual fixed effects), but included an interaction of the unemployment rate with a dummy variable identifying whether the job is formal or informal. Their interpretation for this result was that the unemployment rate controls for the business cycle and its effect over individual wages varies according to the moment of the economy (expansion or retraction). In an expansion, they find that the formality rate goes up and the wage gap between formal and informal workers increases. Botelho and Ponczek (2011) interpreted this wage gap as a segmentation phenomenon, strongly associated with the strict regulation over certain aspects of labour contracts in Brazil.

Focusing again on the wage curve, most studies in Brazil have focused on a comparison of the level of wage flexibility in the formal and the informal sectors, as well as between rural and urban areas. Barros and Mendonça (1997) applied this methodology to a pseudo-panel combining gender, age groups, educational levels, and metropolitan regions with data from the Monthly Employment Survey (PME-IBGE), from 1982 to 1994. Their conclusion was that the degree of wage flexibility in Brazil was considerably high during that period (a slope of -5), similar to the one observed in industrialised economies, but varied significantly over time (affected by the economic instability). They also found

that a lower wage flexibility for more qualified workers, a similar result found by Reis (2006) – who considered state-level data from 1990 to 1999 from the National Household Survey (PNAD). For a similar period (1989-1993), with data from PNAD, but evaluating the wage curve on macroeconomic terms (the ratio between the growth of real wages and the growth of the unemployment rate), Amadeo and Camargo (1997) found a slope of -10. When adding informal employment variation to unemployment variation, this slope fell to -4.5.

Wage determination in Brazil has changed significantly over the past few decades. According to Carneiro and Henley (1998), at the beginning of the 1980s, the State provided a centralised wage indexation, weakening labour unions. Between the mid-1980s and the beginning of the 1990s, in spite of an increasing spur on the role of unions, extremely high inflation rates led the State to keep pursuing significant control over wage indexation. Finally, in the mid-1990s unions were able to bargain for higher real wages, developing a collective bargaining structure more similar to the one observed in Western Europe. These developments had important implications on the flexibility of wages to economic downturns over time.

Still in the Brazilian context, Souza and Machado (2004) used PNAD data from 1981 to 1999, following a two-stage estimation strategy (individual data in the first stage and macroregion-year data in the second). Their findings suggested that labour markets in urban areas (-0.23) are much more flexible than those in rural areas (-0.06). Silva and Monsueto (2012) considered a similar framework for the period 2002-2009, also with PNAD data. They found an elasticity of wages to unemployment of -0.0474, with an indication that low-education segments are more flexible.

Santolin and Antigo (2009) provided an evaluation of a dynamic wage curve considering averages among wage groups, gender, race, and schooling years, in six metropolitan regions from 1997 to 2005 (with PNAD data). According to them, wage flexibility appeared to be higher for the whole group of workers (-0.15) than for only formal workers (-0.05). Furthermore, the coefficient of the lagged wage was close to zero, indicating that unemployment rates seem to be more related to the level of wages than to their change over time.

A more recent piece of evidence on the wage curve relation in Brazil was presented by Baltagi Rokicki and Souza (2014a), analysing PNAD data from 2002 to 2009. Their estimations controlled for individual observed heterogeneity and state-level fixed effects, instrumenting the unemployment rate by its lag. When the formal and the informal sectors were considered separately, only the latter presented significant results (-0.251, while the coefficient was -0.08 when both sectors were considered together). The authors also compared workers in the formal and the informal sectors with similar probabilities of being formal (below the median), finding that the elasticity of unemployment reached -0.129 and -0.305 respectively. Furthermore, they also found that men were more responsive to the unemployment rate (-0.13), while in the case of women the coefficient was not significant (women were more risk-averse, looking for more stable jobs).

Summing up the results found in the literature, there is an indication of the existence of a wage curve in Brazil, with different degrees of wage flexibility according to educational level, age, gender, size of the labour market and formalisation of the job. Furthermore, over time wage flexibility has decreased, accompanied by an increase in economic predictability and macroeconomic stability of the country, the formalisation of the labour market and an increase of the educational level of the labour force. However, there is room to explore further aspects of the wage curve, in special spatial dependence and agglomeration economies, the control of individual unobserved heterogeneity, and controlling for a possible selection bias of the participation in the labour market.

4 Labour market dualism in Brazil and a theoretical framework

There are pronounced differences in the way individuals interact in the labour market of developing economies according to sector, location, firm size, educational level, among other characteristics. In

this context, the literature on wage dualism has investigated whether similar workers in different sectors (formal and informal) receive different payments (SOARES, 2004). Labour market segmentation can be generated by different mechanisms, among which there are: i) firms' heterogeneity (technology, market power, size, sector), and ii) institutional aspects (unions and/or labour legislation). The latter may lead to segmentation if it imposes effective restrictions on the labour market and if there is an inefficient structure of enforcement to follow these rules, applied only to a share of the agents (PERO, 1992; ULYSSEA, 2006).

According to Estevão and Carvalho Filho (2012), the main characteristics of the formal sector in Brazil were originally established by the Labour Law Consolidation (Consolidação das Leis do Trabalho – CLT), a set of labour laws from 1943. Over the years, some changes have been made to this legislation, especially with the new Constitution, in 1988. A job according to these regulations is characterised in the following way: every worker has a booklet (work card) recording his or her work history; workers and firms must contribute to the social security system and respect certain working conditions related to minimum wages, maximum hours of work and payment for overtime work, pre-paid vacations, pre-paid maternity leave, pre-notification of firing, costs to unjustified demission. Public servants and military are not usually hired under the CLT, but their working relations are based on a contract signed with the government that guarantees similar rights, with more stability.

On the opposite side, informal jobs do not have the benefits and guarantees mentioned above, and can be defined as legal but unregulated activities (ULYSSEA, 2010). Their existence is explained by the significant cost of hiring a formal worker (mostly related to social security contributions) and for the relevant costs of formalising an enterprise. There are different types of informal jobs: high-skilled and low-skilled self-employment (for example, consultants in the first case, and street hawkers in the second case), and individuals hired under informal relations by regulated or unregulated firms. When there is no information on contracts, social security status can be an indicator of informality (HENLEY; REZA ARABSHEIBANI; CARNEIRO, 2009). Another possible characterisation is that it is more difficult to achieve a matching between firms and workers in the informal sector due to the lack of a clear channel to exchange information between economic agents and due to high search costs (EL BADAOU; STROBL; WALSH, 2010).

Agénor (2007) identified three main segments that characterise labour markets in developing countries: rural sector (self-employed individuals and unpaid family workers), informal urban sector (low-skilled self-employed persons and small enterprises, mainly providers of services and other non-tradable goods, with limited access to credit), and formal urban sector (medium and large enterprises producing tradable and non-tradable goods, employing workers with different skills). The informal sector can involve paid labour without a formal contract, subject to high job insecurity, underemployment and low productivity, a few benefits, and high flexibility. On the other hand, employers in the formal sector must provide a number of benefits, and wages can be highly affected by the minimum wage legislation, unions and productivity considerations.

The author continued describing the main institutions and regulations in the labour market of developing countries, which are usually relevant in the differentiation between the formal and the informal sectors. Among these, it is possible to highlight (AGÉNOR, 2007):

- Minimum wage laws: their effects on the labour market will depend on the degree of law enforcement and the frequency of adjustment (encompassing real gains or losses in an inflationary context). Minimum wages may represent a lower bound for wages in the formal sector, mostly affecting payments of unskilled workers.
- Hiring and firing regulations: this sort of legislation aims to protect workers hired in the formal sector, including restrictions for firms to fire workers without a 'proper reason', the requirement of a notification period before dismissal, large severance arrangements, and complicated administrative procedures that increase the cost of layoffs and plant closures.

- Nonwage costs and unemployment benefits: the former includes social security contributions and nonwage benefits (housing, health care, pensions, subsidised transportation, subsidised meals, and family allowances), whose cost is partially shared by firms with workers through lower wages. The cost of unemployment benefits is also shared, and may lead to an increase in the level and the duration of unemployment, apart from discouraging human capital accumulation.
- Indexation practices: may induce a high degree of real wage rigidity, combined with a reduction of the effect of monetary shocks over output and employment. They may also distort the expected effect of policy-induced changes in relative prices. Finally, indexed wages can lead to inflation persistence. Wage rigidity in one sector may coexist with full wage flexibility in another sector – leading to market segmentation.
- Bargaining structures and trade unions: wage bargaining arrangements can be related to centralised collective bargaining, with significant government intervention, possibly resulting in wage segmentation. Trade unions are usually tied with political parties and the state to obtain legislative protection and redistributive policies. If the collective bargaining process is less concentrated, the local labour market will be more flexible.

Therefore, firms aiming to hire a worker respecting labour regulation face a higher cost, in terms of contributions as well as the cost of dismissal. Their adjustment capacity is reduced under adverse economic situations, implying that wage sensitivity to a higher unemployment rate will be lower. Being in the informal sector has the advantages of lowering hiring costs. However, there are some drawbacks related to the risk of being caught, especially for firms that are not respecting labour legislation and self-employed individuals who are not paying the appropriate taxes. Informal labour relations may also lead to difficulties in obtaining credit and participating in certain governmental programmes. Large firms are more likely to get caught for not paying taxes, which implies that the formal sector will usually be associated with large firms (EL BADAOU; STROBL; WALSH, 2010).

In relation to the empirical literature investigating labour market dualism, earlier findings did not include all relevant observed and unobserved heterogeneity and possible selection biases in an appropriate manner (SOARES, 2004). Controlling for worker heterogeneity and firm size seems to tackle most of the formal sector net wage premium in relation to the informal sector (EL BADAOU; STROBL; WALSH, 2010; MENEZES-FILHO; MENDES; ALMEIDA, 2004). Therefore, whenever possible it is important to deploy methods involving selection models, longitudinal databases and instrumental variables, in addition to other techniques that may enhance the control for endogeneity issues and confounding factors.

The characteristics of the Brazilian labour market are typical of a developing country, with significant unemployment and wage dualism (FIELDS, 2005). These two features are essential when modelling the economic relations in the labour market. There is also a certain degree of heterogeneity between jobs in the informal sector, with a special distinction for jobs based on self-employment and wage employment (FIELDS, 1990, 2005). Furthermore, self-employed workers can be divided between low and high-skilled, and it can be hard to distinguish them in the labour market, as both usually provide services in various activity sectors. Depending on the self-employment category, it may even be preferred to a formal job. As a conclusion, the focus here will be on salaried workers¹, and the main difference between formal and informal workers will rely on the existence or not of a contract between workers and firms. While the self-employment category is not considered here, this is not a relevant drawback of this study, as the informal sector can be viewed also as an unregulated microentrepreneurial sector associated with voluntary self-employment (MALONEY, 2004).

The fact that labour markets in developing economies are supposedly segmented, with particular characteristics, imply that their relationships of employment and unemployment are not necessarily

¹ Another specificity of self-employed individuals is that they possibly have a stronger entrepreneurial spirit, something which could differentiate them from wage employees.

the same as the ones observed in advanced economies (AGÉNOR, 2007). Hence, the wage curve relationship may not be similar to what has been found in previous studies for developed economies.

The model discussed in the following section aims to represent the relationships in local labour markets that generate the wage curve under potential labour market dualism. It allows for distinct flexibility degrees observed in different sectors in a dual economy of a developing country. This is achieved by combining turnover costs in an efficiency wage model and a varying bargaining degree. These two mechanisms have an impact in the observed relationship between wages, productivity and unemployment (BOOTH, 2014).

4.1 Theoretical framework: institutional costs, bargaining power and efficiency wages

The previous subsection showed that firms in the formal sector may be subject to “labour codes” that regulate hiring and firing, and impose higher costs to them in terms of higher taxes and benefits that must be payed to workers (FIELDS, 2005). Under these conditions, wages may not adjust completely to the business cycle. The main characteristics of dual labour markets indicate that bargaining power can vary significantly between these sectors. This may be generated by a higher presence of unions in the formal sector, as well as higher turnover costs that can be found in this sector. There is also the risk of a shortage of high-skilled workers necessary for specific tasks in the formal jobs sector, which decrease the propensity of dismissal by firms in this sector. Therefore, one possible way of modelling these differences is to consider that workers in the informal sector have less bargaining power than workers in the formal sector (BLIEN et al., 2013, consider differences in the bargaining power as the key to explain the wage curve observed for different regimes of collective bargaining).

The model developed by Barth et al. (2002) is adapted here to the context of a developing country with a formal and an informal sector. This model combines two mechanisms that generate the relationship of the wage curve: efficiency wage and wage bargaining. The efficiency wage model encompass endogenous turnover costs (CAMPBELL; ORSZAG, 1998), depicted in Equation 1, which are dependent on the probability that the worker quits (negatively related to the unemployment rate). Therefore, under higher unemployment rates, turnover costs are supposed to be lower. Replacement costs per worker (C) are given by the multiplication of hiring costs (H) and the probability that the worker quits (q):

$$C_r = H_r q_r \quad (1)$$

Extending the model proposed by Barth et al. (2002) from a single point economy to a spatial setting, migration is possible and quitting is affected by wages in other local labour markets (MORRISON; PAPPS; POOT, 2006). The elasticity of quits to the expected relative pay is $-\eta$ in the local labour market and $-\mu$ in neighbouring areas. Wages and the unemployment rate of other local labour markets can be approximated by the economy-wide values, and they are weighted by an average migration cost between region r and other regions r' , $\gamma(c_{r'r})$, where $\gamma' < 0$ and $\gamma'' > 0$.

$$q_r = \left(\frac{w_r}{(1 - u_r)\bar{w}_r} \right)^{-\eta} \left(\frac{w_r}{\gamma(c_{r'r})(1 - u_{r'})\bar{w}_{r'}} \right)^{-\mu} \quad (2)$$

Firm's profits are given by $\pi_r = R_r(L_r) - C_r L_r - w_r L_r$. In this equation, $R_r(L_r)$ is the revenue received by firms, C_r is the replacement cost defined in Equation 1. Workers can go on a strike in a bargaining process. In this case, firms receive 0 payoff, while workers may receive a fraction $\varphi \in [0,1]$ of the expected local alternative wage or a fraction of the alternative wage in neighbouring areas ($\theta \in [0,1]$), with $\theta + \varphi < 1$. Then, the equilibrium is achieved through a Nash bargaining framework, which involves the optimisation of the weighted combination of gains of workers and firms:

$$\max N = \max(w_r - \varphi(1 - u_r)\bar{w}_r - \theta\gamma(c_{r'r})(1 - u_{r'})\bar{w}_{r'})^{\beta_r} \pi_r^{1-\beta_r} \quad (3)$$

In this setting, β is the bargaining power of workers. In equilibrium, $w_r = \bar{w}_r = \bar{w}_{r'}$, and all firms are assumed to be identical. Then, following Manning (2011), the equilibrium wage will be given by:

$$w_r = b_r \left(\frac{R_r}{L_r} - C_r \right) + (1 - b_r)(\eta + \mu)C_r \quad (4)$$

Where:

$$0 \leq b_r = \frac{\beta_r}{1 - (1 - \beta_r)(\varphi(1 - u_r) + \theta\gamma(c_{r'r})(1 - u_{r'}))} \quad (5)$$

When $\beta = 0$, the model becomes solely based on the efficiency wage mechanism, with $w_r^* = (\eta + \mu)C_r$, and in this case the elasticity of wages to the unemployment rate is given by:

$$\lambda_r^* = \frac{\partial w_r^*}{\partial u_r} \frac{u_r}{w_r^*} = -\eta \frac{u_r}{1 - u_r} < 0 \quad (6)$$

On the other hand, if $\eta = \mu = 0$ and turnover costs are taken as exogenous, $w_r^B = b_r \left(\frac{R_r}{L_r} - H_r \right)$, the elasticity of wage to unemployment will be:

$$\lambda_r^B = \frac{\partial w_r^B}{\partial u_r} \frac{u_r}{w_r} = -\frac{(1 - \beta_r)\varphi u_r}{1 - (1 - \beta_r)\varphi(1 - u_r)} \leq 0 \quad (7)$$

However, the combination of efficiency wage and wage bargaining mechanisms generate a rent effect of the unemployment rate on bargained wages, whose elasticity is given by:

$$\lambda_r^R = \frac{\partial \left(\frac{R_r}{L_r} - C_r \right)}{\partial u_r} \frac{u_r}{\left(\frac{R_r}{L_r} - C_r \right)} = \eta \frac{u_r}{1 - u_r} \left(\frac{C_r}{\frac{R_r}{L_r} - C_r} \right) > 0 \quad (8)$$

Then, the total elasticity of wages in relation to unemployment will be:

$$\lambda_r = \frac{w_r^*}{w_r} \lambda_r^* + \left(1 - \frac{w_r^*}{w_r} \right) \lambda_r^B + \frac{b_r(1 + \eta) \left(\frac{R_r}{L_r} - C_r \right)}{w_r} \lambda_r^R \quad (9)$$

where $\frac{w_r^*}{w_r} \leq 1$. Taking the derivative of Equation 9 with respect to β_r generates the following relationship:

$$\frac{\partial \lambda_r}{\partial \beta_r} = \underbrace{-(\lambda_r^B - \lambda_r^*)}_{(-)} \frac{\partial \left(\frac{w_r^*}{w_r} \right)}{\partial \beta_r} + \underbrace{\left(1 - \frac{w_r^*}{w_r} \right)}_{(+)} \frac{\partial \lambda_r^B}{\partial \beta_r} + \underbrace{(1 + \eta)}_{(+)} \underbrace{\left(\frac{R_r}{L_r} - C_r \right)}_{(+)} \lambda_r^R \frac{\partial \frac{b}{w_r}}{\partial \beta_r} \quad (10)$$

According to Barth et al. (2002), this result can be interpreted in the following way. Under a pure efficiency wage setting, the introduction of a bargaining mechanism may lead to a higher sensitivity in relation to unemployment, if the efficiency wage effect is small and φ is high. However, as β_r increases, the alternative wage effect embedded in the turnover costs term diminishes, and the rent-sharing goes up, leading λ_r , which is negative, to values closer to zero. Then, if two different groups of workers with distinct levels of bargaining power are compared, those with a higher level of β_r will face less wage flexibility in the labour market in comparison with the other group.

In addition, the unemployment rate in neighbouring areas is inversely related to turnover costs (a higher unemployment rate in neighbouring areas reduces the expected wage associated with a migration decision of the worker, decreasing turnover costs). This term is included in the rent effect, and a higher $u_{r'}$ will decrease wage flexibility to the local unemployment rate. On the other hand, the larger the commuting distance to other regions, the lower the turnover costs, indicating that the

unemployment elasticity of pay will be greater in regions that are more isolated (LONGHI; NIJKAMP; POOT, 2006).

Agglomeration economies are also relevant in the context of the wage curve, as they are related to higher productivity and higher wages, accompanied by lower unemployment rates (SATO, 2000). This is an additional mechanism to generate the wage curve. Therefore, regions with higher urban concentration are expected to be in the left-upper side of the wage curve, while less-agglomerated areas are on the lower-right side. On the other hand, the efficiency wage model presupposes that regional amenities generate spatial differences in the equilibrium of wages and unemployment (CARD, 1995; ELHORST; BLIEN; WOLF, 2007). Then, high-amenity regions are supposed to present lower wages and higher unemployment rates, while low-amenity regions will be in the opposite location of the wage curve (high wages, low unemployment). This issue requires further investigation, but the agglomeration degree is an essential element to generate the wage curve.

In summary, the model developed here indicates that if the main differences between the formal and the informal sectors can be summarised by distinct degrees of bargaining power, where the former is supposed to present a lower wage flexibility to the business cycle than the latter. Accessibility to jobs can also affect the local wage, meaning that local density and the proximity to job opportunities in neighbouring areas may affect the observed wage curve. The next sections will investigate how these elements interact in the Brazilian labour market.

5 Empirical strategy

The model discussed in the previous section indicates that there is a negative relationship between wages and the local level of the unemployment rate. A reduced form of this model can be represented by:

$$w_{r,t} = \gamma_1 u_{r,t} + \sum_k X_{r,t,k} \beta_k + \theta_r + \delta_t + \epsilon_{r,t} \quad (11)$$

$$r = \{1, \dots, R\}, t = \{1, \dots, T\}$$

However, as discussed in Section 3, the empirical literature has started to control for individual heterogeneity in order to obtain a coefficient that actually captures this relationship. Therefore, individual wages at the individual level are related to the unemployment rate at the local level, controlling for observed and unobserved characteristics at both levels whenever possible.

$$w_{i,r,t} = \gamma_1 u_{r,t} + \sum_k X_{i,r,t,k} \beta_k + \theta_r + \delta_t + \epsilon_{i,r,t} \quad (12)$$

$$i = \{1, \dots, N\}, r = \{1, \dots, R\}, t = \{1, \dots, T\}$$

In this initial formulation, $w_{i,r,t}$ is the logarithm of the individual hourly wage and $u_{r,t}$ is the logarithm of the unemployment rate in region r . Moreover, θ_r is the region effect, δ_t is a time effect and $X_{i,r,t}$ is a set of observed characteristics of individual i (age, age², education, sector, occupation, etc.). Whenever there is longitudinal data available at the individual level, an additional term α_i representing the individual fixed effect can be included to control for unobserved individual characteristics fixed in time. The parameter of interest in the equation above will be γ_1 , representing the elasticity of wages against the level of the unemployment rate.

According to the main elements of the theoretical framework discussed before, these models will be estimated for different population groups. Namely, whenever possible, the estimation of the wage curve will be conducted for the formal and informal sectors separately and for different degrees of urban agglomeration. Furthermore, following the argument of Meghir and Whitehouse (1996), the analysis is restricted only to men, as they will most likely stay in the labour market even when facing adverse conditions (low wages, high unemployment, inadequate work conditions).

5.1 Extensions

A few other issues must be investigated in detail to assess the consistency and robustness of the main results. The next subsections will briefly discuss these problems and the strategies developed to avoid them.

5.1.1 Multi-level analysis

The wage curve with controls for individual heterogeneity is characterised by variables in different aggregation levels. If this multilevel nature of the data is not taken into account, there may be a miscalculation of the significance of the effect of aggregated variables over micro-units. Considering the discussion developed by Moulton (1990), individuals in the same labour market may share some components of the variance that are not entirely explained by their observed characteristics, unobserved heterogeneity (constant in time) or by more aggregated variables, such as the unemployment rate (at the area level). In this case, the error term in the equation associated with individual wages, X_{irt} and u_{rt} will be positively correlated across individuals in the same local market and as a consequence the standard error of the unemployment effect will be downwardly biased, generating higher t-statistics (CARD, 1995).

A solution proposed in some studies involves calculating averaged values of all individual-level variables at the area level. Then, the wage curve would be estimated directly at this aggregated level (BALTAGI; BLIEN; WOLF, 2000; BLANCHFLOWER; OSWALD, 1994; SANTOLIN; ANTIGO, 2009). Another possibility is to estimate the wage curve in two stages (BELL; NICKELL; QUINTINI, 2002). In the first stage, the logarithms of individual hourly wages are regressed against individual characteristics and a region-year dummy, which will synthetize the regional characteristics that explain wage differences over time. This first stage can be estimated either with repeated cross-sections or with a panel of individuals (with the control of the unobserved individual heterogeneity constant in time), depending on data availability.

$$w_{i,r,t} = \lambda_{r,t} + \delta_t + \sum_k X_{i,r,t,k} \beta_k + \eta_{i,r,t} \quad (13)$$

Then, the estimated region-year dummy $\hat{\lambda}_{r,t}$ is used in the second stage as a dependent variable, which will be explained by regional characteristics, including the unemployment rate.

$$\hat{\lambda}_{r,t} = \mu_r + \delta_t + \gamma_2 u_{r,t} + v_{r,t} \quad (14)$$

In this case, the parameter of interest in the estimation will be γ_2 , while μ_r and δ_t represent region fixed effects and time dummies.

5.1.2 Endogeneity of the unemployment rate

One of the main concerns of the empirical literature examining the wage curve regards the potential endogeneity of the unemployment rate due to simultaneity. The most common strategy is to consider time lags of this variable as instruments, with the underlying hypothesis that wages in the present do not affect the unemployment rate in the past nor do they share a common causing factor.

5.1.3 Selection bias

Most labour market analyses may be subject to the problem of selection bias, especially when the idea is to obtain estimates that are representative for the whole population and not only for a specific group with available information. In the case of the wage curve, the equation is supposed to represent

the whole population of a working age, while data is observed only for employed individuals (a case of incidental truncation – WOOLDRIDGE, 2002, p. 560).

Within this context, the sample of employed individuals for whom wages are observed may not be randomly selected from the whole population. Card (1995) pointed out that Blanchflower and Oswald (1994) failed to discuss any potential effect of selection bias in the case of the wage curve, but composition bias can affect the elasticity of wages with respect to local unemployment rates. The selection model developed by Heckman (1979) aims to deal with this issue.

It is possible to revise Equation 12, rewriting it in a simplified version:

$$w = \gamma_1 u + X\beta + \varepsilon \quad (15)$$

But the wage is observed only if the person works ($L = 1$). The estimation of the parameters of interest must take into account this additional information in order to correct the limiting distribution of the estimator. The probability of being an employee is given by:

$$L = 1[Z\delta + \xi > 0] \quad (16)$$

in which L represents the condition of being an employee, while Z is a vector of variables associated with this probability. The main assumptions underlying this method are the following: (i) L and Z are always observed for all individuals; (ii) (ε, ξ) is independent of Z with 0 mean; (iii) $\xi \sim N(0,1)$; (iv) $E(\varepsilon|\xi) = \theta\xi$ (linear relationship between the error terms).

Then, the estimating equation can be derived by considering that the element to be estimated is actually $E(w|u, X, L = 1)$. Combining the assumptions above with this statement generates the final model:

$$E(w|u, X, L = 1) = \gamma_1 u + X\beta + \zeta\lambda(Z\delta) \quad (17)$$

where $\lambda(\cdot) \equiv \frac{\phi(\cdot)}{\Phi(\cdot)}$ is the inverse Mills ratio, which represents $E(L = 1|Z) = E(\xi|\xi > -Z\delta) = \lambda(Z\delta)$. A consistent estimator of δ is obtained with a first-stage probit estimation of the selection equation for the whole population. After that, it is possible to consistently estimate γ_1 and β by regressing w against u , X and $\lambda(Z\hat{\delta})$ using the selected sample.

5.1.4 Spatial dependence and neighbouring effects

The model presented in Section 4 highlighted the possibility that the unemployment rate in neighbouring areas could affect wages directly (with a negative relationship) or indirectly, by increasing local wage flexibility. The subject of spatial effects in the wage curve has been significantly explored in the literature, sometimes including the spatial lag of the unemployment rate (LONGHI; NIJKAMP; POOT, 2006) or spatial autocorrelation in the error and in the dependent variable (BALTAGI; BLIEN; WOLF, 2012; FINGLETON; PALOMBI, 2013; RAMOS; NICODEMO; SANROMÁ, 2015; ELHORST; BLIEN; WOLF, 2007; BUETTNER, 1999).

This issue can be investigated in the following way. Departing from Equation 14, which is already at the region level, the error term can present spatial autocorrelation, which is tested using LM test statistics discussed by Elhorst (2014). Then, according to these tests results, two main options are available to account for spatial dependence. The spatial lag model can be specified as:

$$\hat{\lambda}_{r,t} = \mu_r + \delta_t + \rho W \hat{\lambda}_{r,t} + \gamma_2 u_{r,t} + v_{r,t} \quad (18)$$

While the spatial error model is written as:

$$\hat{\lambda}_{r,t} = \mu_r + \delta_t + \gamma_2 u_{r,t} + \varepsilon_{r,t} \quad (19)$$

$$\varepsilon_{r,t} = \varphi W \varepsilon_{r,t} + v_{r,t} \quad (20)$$

where W is the spatial weights matrix based on inverse distances and common local labour market areas (see subsection 6.3.1 for further details), constant in time. The error terms $v_{r,t}$ are supposed to be independently and identically distributed for all r , with zero mean and variance σ^2 . The estimation of these different models follow Elhorst (2014), based on MATLAB routines available online².

5.1.5 Dynamic wage curve

Another possible extension to the basic wage curve model consists of including time-lagged wages in the original regression, obtaining a dynamic wage curve (BALTAGI; BLIEN; WOLF, 2009, 2012). The Phillips curve can be an alternative explanation to the relationship between wages and the unemployment rate (BLANCHARD; KATZ, 1997). In the literature, the estimation of a dynamic wage setting equation aims to test which relationship is more prone to be happening in a certain country. For instance, considering data aggregated at the regional level over time, the wage curve to be estimated can be the following (BELL; NICKELL; QUINTINI, 2002):

$$w_{r,t} = a_r + (1 - \lambda)w_{r,t-1} + \beta u_{r,t} + d_t + \varepsilon_{r,t} \quad (21)$$

In this case, there are three possible outcomes:

- If $(1 - \lambda) = 0$, then $\lambda = 1$, and the wage curve explanation dominates;
- If $(1 - \lambda) = 1$, $\lambda = 0$, and the Phillips curve seems to be more adequate;
- If $0 < \lambda < 1$, there is a dynamic wage curve, in which the impact of past regional changes is prolonged over more than one time period, and the long run elasticity of unemployment can be represented by β/λ .

6 Empirical analysis

The empirical analysis of the wage curve in Brazil is still limited by data availability. Differently from other case studies mentioned in Section 3, in Brazil there is still no longitudinal database with national coverage applied over a long period comparing the formal and the informal sectors. The main databases associated with labour market outcomes are provided by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE) with publicly available microdata:

- Monthly Employment Survey (Pesquisa Mensal de Emprego – PME) – with a moving panel of individuals (who are interviewed over sixteen months), for the period of March 2002 to June 2015, covering only six metropolitan regions (Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, and Porto Alegre);
- National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios – PNAD), with repeated cross-sections for the period 1995-2013, with national coverage and state-level regional disaggregation;
- Demographic Census, a decennial survey consisting of repeated cross sections, for the whole country, allowing for municipal disaggregation.

This basic description shows that each of these databases covers a particular aspect of the labour market. Therefore, the next subsections will develop separate analyses of the wage curve with each of them, namely:

- Subsection 6.1 presents the results of the wage curve in a more aggregated perspective, with the PME database, in which individual heterogeneity is controlled for with the introduction of individual fixed effects and it is also possible to understand the relationship between wage

² Available at <http://www.regroningen.nl/elhorst/software.shtml> and accessed on September 10, 2015.

flexibility and the business cycle; furthermore, results regarding the dynamic wage curve are provided;

- In Subsection 6.2, conclusions are drawn from the PNAD database, covering the whole country for a long period, controlling for state-level unobserved heterogeneity. It also provides a comparison between wage flexibility for different degrees of urbanisation (rural, urban, and metropolitan areas);
- Finally, Subsection 6.3 explores Census data, including Minimum Comparable Areas (MCAs) controls³, neighbourhood effects and the effect of different urban scales over wage flexibility.

All these elements, even if investigated separately, provide a deep understanding of the wage curve in Brazil, with an important contribution to the literature given by the application of different methods to one specific case study. Table A.1 in the Appendix provides a detailed description of all the variables considered, including methods of calculation and data sources.

6.1 The wage curve in a more aggregated context: a comparison between the formal and the informal sectors

The first set of results presented here aim to measure wage flexibility in a more aggregated context, with monthly data, controlling for individual heterogeneity constant over time. Due to data availability, the wage curve is measured only for six metropolitan regions, which account for almost 25% of the population of the country. The next subsection will discuss the main characteristics of this database, while subsection 6.1.2 presents descriptive statistics. Then, subsection 6.1.3 discusses the main results and conclusions of this group of models.

6.1.1 Description of the database – Monthly Employment Survey

The Monthly Employment Survey (Pesquisa Mensal de Emprego – PME) is conducted by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE) providing monthly indicators about demographic, social and economic characteristics of the labour market in six metropolitan regions (Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, and Porto Alegre). The survey underwent a deep methodology review in 2001, adopting the international patterns recommended by the International Labour Organization (ILO). This restriction implies that the period covered here will range from March 2002 to June 2015. However, all regressions will consider information from March 2003 on, in order to allow one-year lag of the unemployment rate to be considered as an instrumental variable.

The surveyed information is restricted to data on the labour market, and it can be analysed with its cross-section dimension or with a moving panel structure (this is the only IBGE survey in longitudinal format). Households are visited for two periods of four consecutive months, eight months apart from each other. Ribas and Soares (2010) provided a methodology and a Stata routine to reduce attrition in the PME panel, identifying individuals within the household and excluding those who have moved or who present inconsistent information over time⁴.

A few steps were conducted to obtain the final database. Only male individuals aged 15 to 59 years old, working as employees in the formal or the informal sector were kept (excluding unemployed and out-of-the economically active population individuals, self-employed individuals, public sector employees, military, employers, and workers in the production for self-consumption). Formality status was defined by the existence or not of a contract under the CLT regulation (see Section 3).

³ See Subsection 4.3 for a longer explanation.

⁴ The Data Zoom website (<http://www.econ.puc-rio.br/datazoom/english/>) provides packages including these algorithms to assess the longitudinal nature of the survey. We are thankful to Data Zoom, developed by the Department of Economics at PUC-Rio, for providing the codes for accessing IBGE microdata.

Furthermore, individuals employed in the public administration sector were also excluded, as well as those with wages equal to zero or with missing information, or with less than 20 or more than 60 working hours per week. Only individuals observed in the second and in the sixth interviews (twelve months apart from each other) were selected to be in the database.

The final database comprised 193,934 observations (96,967 surveyed individuals, observed in two moments in time, representing more than 49 million individuals over the period of more than 12 years when the sampling scheme was taken into account). The distribution of these observations over the years and months is presented in Table A.2 in the Appendix.

The aggregate unemployment rate was calculated for each metropolitan region for each month. In addition, wages were deflated by local inflation indexes (National Consumer Price Index – Índice Nacional de Preços ao Consumidor – INPC – in each metropolitan region). This correction with local inflation indexes aimed to at least partially control for the variation in local living costs over time. The initial differences between each metropolitan region were taken into account by the inclusion of local dummies.

6.1.2 Descriptive statistics

The average unemployment rate for the six metropolitan regions surveyed by PME decreased steadily from 2003 to 2014 and increased in the last few months of the period analysed. On the other hand, wages for the formal and the informal sector presented an upward trend in the period, with the important remark that the formal sector paid higher salaries on average than the informal sector.

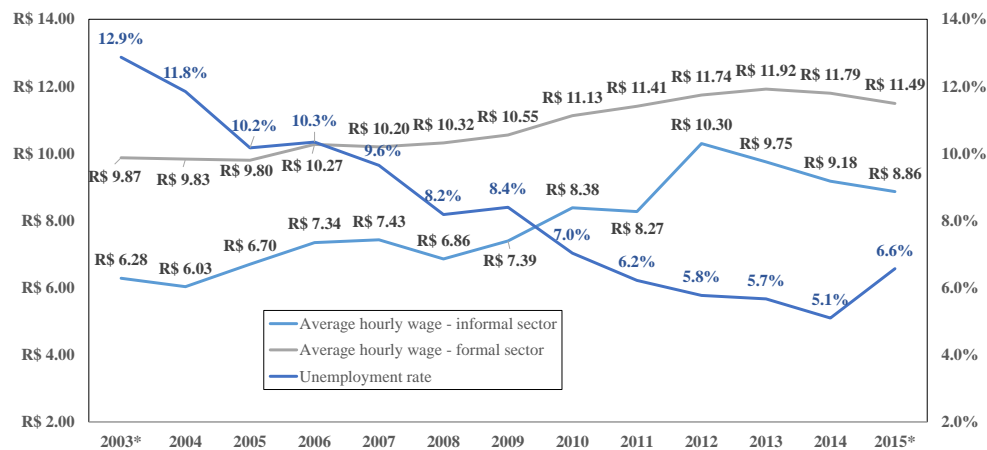


Figure 1 – Average hourly wage in the formal and the informal sectors and unemployment rate, 2003-2015.

* For 2015, the average was calculated only until June. Unemployment rate was calculated for the whole population, while average hourly wages were calculated only for the sample analysed.

Source: PME-IBGE.

Table 1 – Average unemployment rates and hourly wages (both sectors) by metropolitan region, 2003-2015.

	Unemployment rate						Average hourly wage					
	Recife	Salvador	Belo Horizonte	Rio de Janeiro	São Paulo	Porto Alegre	Recife	Salvador	Belo Horizonte	Rio de Janeiro	São Paulo	Porto Alegre
2003*	14.3%	17.0%	11.0%	9.6%	14.4%	9.7%	R\$ 5.32	R\$ 6.99	R\$ 7.98	R\$ 8.74	R\$ 10.74	R\$ 7.95
2004	13.2%	16.4%	10.9%	9.5%	12.9%	8.9%	R\$ 5.41	R\$ 7.26	R\$ 7.68	R\$ 8.13	R\$ 11.04	R\$ 7.86
2005	13.7%	15.9%	9.0%	8.0%	10.5%	7.6%	R\$ 5.36	R\$ 7.35	R\$ 8.04	R\$ 7.87	R\$ 11.31	R\$ 8.17
2006	15.1%	14.0%	8.7%	8.3%	10.8%	8.3%	R\$ 5.50	R\$ 7.55	R\$ 8.28	R\$ 8.59	R\$ 11.80	R\$ 8.49
2007	12.4%	14.1%	7.8%	7.5%	10.4%	7.5%	R\$ 5.77	R\$ 7.22	R\$ 8.31	R\$ 8.90	R\$ 11.60	R\$ 8.48
2008	9.6%	11.8%	6.7%	7.1%	8.7%	6.1%	R\$ 5.67	R\$ 7.73	R\$ 8.90	R\$ 8.91	R\$ 11.44	R\$ 8.55
2009	10.2%	11.7%	6.6%	6.4%	9.5%	5.8%	R\$ 6.51	R\$ 8.19	R\$ 9.12	R\$ 9.29	R\$ 11.58	R\$ 8.78
2010	9.0%	11.3%	5.7%	5.9%	7.3%	4.7%	R\$ 6.48	R\$ 8.00	R\$ 9.68	R\$ 9.80	R\$ 12.47	R\$ 9.42
2011	6.8%	10.0%	5.1%	5.5%	6.4%	4.7%	R\$ 6.95	R\$ 8.32	R\$ 10.11	R\$ 10.26	R\$ 12.48	R\$ 9.89
2012	6.3%	7.5%	4.6%	5.3%	6.2%	4.2%	R\$ 7.64	R\$ 8.67	R\$ 10.82	R\$ 10.93	R\$ 13.05	R\$ 10.05
2013	6.7%	8.3%	4.3%	4.8%	6.2%	3.7%	R\$ 7.37	R\$ 8.45	R\$ 10.62	R\$ 11.41	R\$ 13.03	R\$ 10.54
2014	6.8%	9.4%	3.9%	3.8%	5.2%	3.9%	R\$ 7.77	R\$ 8.87	R\$ 10.64	R\$ 11.95	R\$ 12.39	R\$ 10.65
2015*	8.3%	11.6%	5.3%	5.0%	6.7%	5.3%	R\$ 8.09	R\$ 8.12	R\$ 10.50	R\$ 12.63	R\$ 11.56	R\$ 10.55

Obs.: Unemployment rate was calculated for the whole population, while average hourly wages were calculated only for the sample analysed. For 2003, the average was calculated only from March on, while in 2015, it was calculated only until June.

Source: PME-IBGE.

Table 1 shows that this general trend for wages and unemployment can also be found when each metropolitan region is analysed separately. However, the timing of adjustment may vary between them. Furthermore, in 2015 wages were higher in Rio de Janeiro, São Paulo, Porto Alegre and Belo Horizonte, while higher unemployment rates could be found in Salvador, Recife and São Paulo.

Table 2 – Descriptive statistics of the PME database.

	2003*	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015*
Age	34.3	34.6	34.6	34.6	34.8	35.1	35.5	35.7	35.7	35.9	36.2	36.5	36.8
Metropolitan region													
Recife	4.4%	4.6%	4.5%	4.0%	3.8%	3.4%	3.2%	3.4%	3.8%	3.9%	3.9%	3.8%	4.0%
Salvador	5.3%	5.4%	5.1%	5.1%	5.1%	4.8%	5.2%	4.7%	4.1%	4.5%	4.9%	4.8%	4.4%
Belo Horizonte	10.0%	9.9%	9.2%	9.2%	9.6%	10.1%	10.4%	10.9%	11.0%	10.9%	10.7%	10.1%	9.9%
Rio de Janeiro	24.8%	28.3%	30.0%	30.2%	30.8%	30.2%	28.5%	26.8%	26.3%	26.6%	25.6%	23.2%	21.8%
São Paulo	47.3%	44.0%	43.8%	44.5%	43.5%	43.6%	44.7%	46.2%	46.8%	46.0%	47.4%	50.9%	52.9%
Porto Alegre	8.2%	7.8%	7.4%	7.1%	7.3%	7.8%	7.9%	8.0%	7.9%	8.1%	7.5%	7.3%	7.0%
Student	10.8%	9.9%	9.9%	10.0%	9.9%	9.4%	9.1%	8.4%	8.3%	8.8%	8.4%	7.4%	6.1%
Education level													
Less than 1 year of schooling	2.1%	1.9%	1.7%	1.5%	1.2%	1.2%	1.2%	1.0%	1.2%	1.3%	1.1%	0.8%	0.6%
1 to 3 years of schooling	5.3%	5.3%	4.9%	4.7%	4.6%	4.3%	3.8%	3.2%	2.9%	2.6%	2.2%	2.0%	2.1%
4 to 7 years of schooling	25.1%	25.3%	23.6%	21.8%	21.2%	20.3%	19.9%	18.6%	17.0%	16.4%	15.3%	14.2%	13.3%
8 to 10 years of schooling	22.5%	21.4%	21.4%	21.2%	21.3%	21.3%	20.0%	20.4%	19.9%	18.9%	18.2%	18.7%	16.8%
11 years of schooling or more	45.0%	46.1%	48.3%	50.8%	51.7%	52.8%	55.1%	56.8%	59.0%	60.8%	63.2%	64.3%	67.2%
% of self-reported black or brown	40.3%	43.2%	44.9%	44.1%	45.7%	45.7%	46.2%	46.2%	47.5%	47.7%	47.6%	47.2%	47.2%
% of employees in the formal sector	84.1%	82.6%	82.9%	84.1%	85.0%	85.4%	86.0%	88.0%	88.6%	89.4%	91.1%	92.4%	93.7%
Sector of activity													
Industry	33.5%	32.2%	32.1%	31.5%	30.3%	30.3%	30.6%	30.6%	29.4%	29.0%	28.1%	28.1%	28.6%
Construction	7.0%	7.1%	7.3%	7.4%	7.7%	8.4%	8.6%	8.7%	9.6%	9.4%	9.5%	8.7%	8.3%
Commerce and repair	20.8%	20.9%	20.5%	20.2%	20.7%	20.6%	20.7%	19.7%	19.6%	19.9%	19.7%	19.8%	19.8%
Finance, real estate, serv. to companies	19.6%	19.9%	20.5%	21.2%	21.3%	20.7%	20.0%	20.8%	21.4%	21.4%	21.8%	21.7%	22.1%
Other services	18.6%	19.2%	18.9%	19.1%	19.4%	19.6%	19.6%	19.8%	19.5%	19.9%	20.4%	21.3%	20.7%
Other activities	0.7%	0.6%	0.6%	0.6%	0.6%	0.4%	0.5%	0.4%	0.5%	0.4%	0.5%	0.3%	0.4%

* For 2003, the average was calculated only from March on, while in 2015, it was calculated only until June.

Source: PME-IBGE.

Considering the composition of the database over time (Table 2), it is possible to note an increase in the average age from 2003 to 2015. São Paulo gained importance in the database, as well as higher education levels (especially 11 years of schooling or more, which passed from 45.0% to 67.2% of the population in the period). More importantly, the share of workers in the formal sector⁵ rose from 84.1% to 93.7%. Jobs in the industrial sector lost importance in the total, while the share of workers in the construction and the financial sectors increased over time.

This preliminary analysis indicated that, at least in aggregated terms, there seemed to be a relationship between wages and the local unemployment rate, even if there were only six areas in the database. The following subsection aim to assess whether this relationship still remains after including controls at the individual, time and regional levels.

⁵ Defined as employees in the private sector with a labour contract following the CLT, in opposition to informal workers who do not have this type of contract in their jobs.

6.1.3 Results and preliminary conclusions

The first set of results attempted to control for the issues discussed in subsections 5.1.2 and 5.1.5 (namely, endogeneity of the unemployment rate and the estimation of a dynamic wage curve). This was possible because of the longitudinal nature of the database, with high-frequency information. Other issues such as the selection bias were not treated here because PME focuses mostly on labour market outcomes, and does not detail other dimensions of the individual's life. Spatial dependence was not investigated because the regions available here are disconnected and cover only a very small part of the country. The low number of regions was also the reason why the estimation here was conducted in only one stage (otherwise, the second stage would be based on a very small number of units).

Some controls were included in the estimated models, following the empirical literature. In the case of the models with individual fixed effects (two time periods for each individual separated by one year), controls that did not vary over time for the same individual or at least over a year were dropped (education level, age, month, among others).

The results presented in Table 3 indicated that for the basic model (OLS estimation), a wage curve could be observed only for the informal sector, while the coefficient for the formal sector was not significant. When these two sectors were aggregated, once again it was not possible to obtain a significant effect. However, Model 1 controlled only for observed individual characteristics. When individual fixed effects were included (Model 2), all effects disappeared. Model 3 considered the same estimation as Model 1, but now unemployment rates were instrumented by their time lag. Then, once again, there was an indication of a wage curve relationship only for the informal sector. In Model 4, individual fixed effects were added to the regression, and this generated non-significant coefficients. All these results indicated that the wage curve could be observed only in the case of the informal sector in Brazil, but this result may be generated by the fact that observed individual characteristics were not enough to control for all relevant variability.

Table 3 – Elasticity of wages in relation to the unemployment rate for different estimation strategies and sectors.

Models	Total	Informal	Formal
1) OLS	-0.012 (0.016)	-0.102** (0.045)	-0.007 (0.017)
2) FE	-0.001 (0.010)	0.055 (0.039)	-0.001 (0.010)
3) 2SLS	-0.012 (0.021)	-0.161*** (0.062)	-0.003 (0.023)
4) FE 2SLS	-0.004 (0.013)	0.037 (0.053)	0.003 (0.013)
5) Dynamic OLS	-0.008 (0.007)	-0.058** (0.023)	-0.003 (0.007)
6) Dynamic FE	-0.002 (0.009)	0.021 (0.035)	-0.003 (0.009)
7) Dynamic 2SLS	-0.007 (0.010)	-0.078** (0.032)	0.000 (0.010)
8) Dynamic FE 2SLS	-0.001 (0.012)	-0.005 (0.047)	0.005 (0.012)

Controls: (i) Cross-sectional models: age, age², education attainment, metropolitan region, sector of activity, self-reported black or brown, student, year, month; (ii) Longitudinal models: sector of activity, student, year. Instrument: unemployment rate one year before in the metropolitan region. Dynamic models: the wage of the individual in the previous month was included as an explanatory variable. Numbers in parentheses represent standard errors. All models were estimated with *pweight* in Stata to account for the proportional weight of each observation in the whole population. Standard errors are presented in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Complete estimation results are presented in Tables A.3 and A.4 in the Appendix.

Source: Author's own calculations.

Models 5 to 8 allowed for the possibility of the existence of a dynamic wage curve. When individual fixed effects were included (Models 6 and 8), it was not possible to find any significant results. On the other hand, when only the cross-sectional dimension was taken into account (Models 5 and 7), including the time lag of wages generated coefficients that were half of the size of Models 1 and 3 (in absolute terms) for the informal sector, while there was no significant relationship for the formal sector.

The estimation of dynamic models also allowed for testing whether the relationship between wages and the unemployment rate could be better explained by a Phillips curve. Following the strategy mentioned in subsection 5.1.5, the results presented in Table 4 indicated that when only observed characteristics of the individuals were taken into account, the coefficient of the lagged wage was close to 1 (meaning that the Phillips curve could be a better analytical framework). However, the inclusion of individual fixed effects led to a division by half of these coefficients, signalling that a dynamic wage curve was the more appropriate explanation. The complete regression results can be found in the Appendix (Tables A.3 and A.4).

Table 4 – Elasticity of wages in relation to their time lags for different estimation strategies and sectors.

Models	Total	Informal	Formal
5) Dynamic OLS	0.875*** (0.002)	0.840*** (0.006)	0.877*** (0.002)
6) Dynamic FE	0.422*** (0.003)	0.447*** (0.010)	0.399*** (0.003)
7) Dynamic 2SLS	0.875*** (0.002)	0.840*** (0.006)	0.877*** (0.002)
8) Dynamic FE 2SLS	0.422*** (0.003)	0.447*** (0.010)	0.399*** (0.003)

Standard errors are presented in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Source: Author's own calculations.

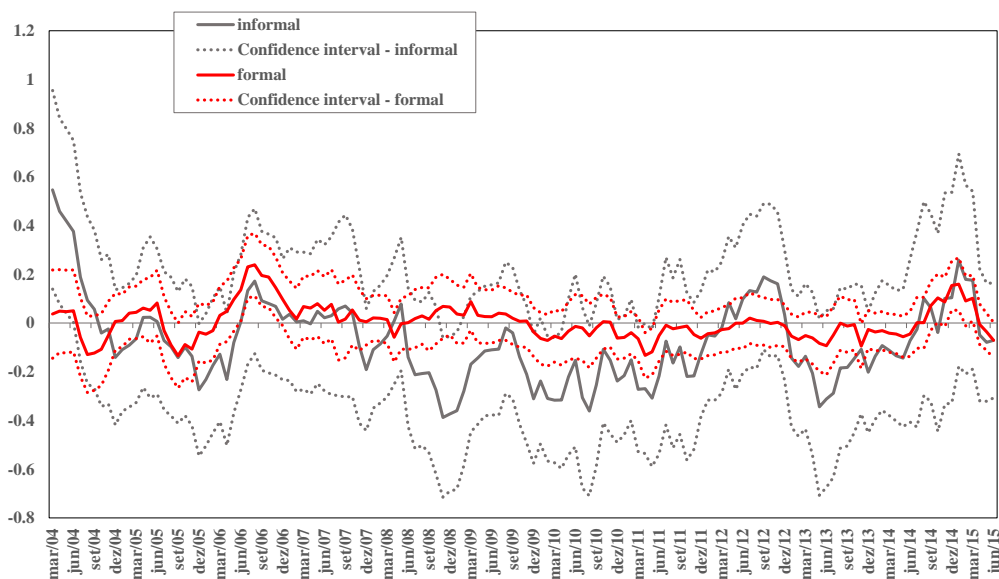


Figure 2 – Coefficients and confidence intervals for OLS regressions, with the sample comprising individuals in 12-month periods ending in each month.

Source: Author's own calculations.

One interesting extension for the models presented in Table 3 is to check how estimations may vary over time. Figure 2 shows the evolution of the coefficients for the formal and the informal sectors (and their confidence intervals) based on the estimation of Model 1. Each observation on the x-axis was obtained by estimating the model for the previous twelve months (moving sample). The confidence intervals showed that elasticities were significantly different from zero only in specific moments for the informal sector. In the case of the formal sector, the coefficient was not different from zero in a consistent way. These results indicated that even when a wage curve was found in aggregated terms, it will not necessarily be constant over time.

In summary, the coefficients for the informal sector seem to match the main results found in the literature for developed economies. A few aspects of this estimation must be highlighted. Firstly, metropolitan regions in a developing country such as Brazil do not necessarily represent the whole country – in general, they are much more diversified than other less dense areas, with better jobs and a labour market with particular characteristics. Moreover, all models including individual fixed effects will compare individuals in only two points in time, reducing within variation, which may explain why no effect was found in these models. These individuals were observed more than once in the database only if they stayed in the same metropolitan region (there were no individuals in the panel who moved during year they are interviewed). Finally, unemployment rates were observed only for six regions, meaning that this variable does not present very much variation between individuals.

6.2 Rural and urban areas and the wage curve

The second set of results discussed here involve the estimation of the wage curve with repeated cross-sections that allow for the control of individual observed characteristics and state-level data. This database presents the advantages of having national coverage, characterising rural and urban areas, and a long time series.

6.2.1 Description of the database – National Household Sample Survey

The National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios –PNAD), conducted by IBGE, investigates, on a yearly basis, demographic and socioeconomic characteristics of the household. Among these characteristics, it provides details about labour supply decisions, education level and income of all sources for the sample of individuals. PNAD is a cross-section survey with state-level representativeness, and has been conducted with a relatively fixed structure ever since the beginning of the 1990s. Every ten years the sampling design is adjusted to the reality presented by the most recent demographic census. In this respect, the survey is not applied during Census years.

PNAD data was analysed for the period from 1996 to 2013 (leaving 1995 as a base year to provide the one-year lag of the unemployment rate as an instrument), after the macroeconomic normalisation and the implementation of the Real as the currency of the country. There are two years in this period in which the survey was not conducted, 2000 and 2010 (Census years). However, the fact that the time lag of the unemployment rate was considered as an instrumental variable requires the calculation of this rate for these missing years. This was done by obtaining the average of the rates observed in the previous and posterior years. Due to this restriction, two different periods of analysis were compared, 1996 to 2013 and 2002 to 2009. The former did not require this imputed unemployment rate, and its results served as a baseline for the main results. Furthermore, rural areas of the Northern states of Rondônia, Acre, Amazonas, Roraima, Pará and Amapá were excluded from the analysis, as they were not surveyed before 2004.

Following the same procedures of subsection 6.1, only male individuals aged 15 to 59 years old were kept in the database. The population of interest comprised employees with positive wages working in the private sector with or without a contract under the CLT regulation. The unemployment rate was

calculated for each state considering all economically active individuals (unemployed and employed), and wages were deflated by the average inflation index of the whole country (National Consumer Price Index – Índice Nacional de Preços ao Consumidor – INPC). It was not possible to correct wages for local components of this index because they were available only for some states.

The results in this subsection sought to understand how the urbanisation level and different sectors of the labour market affect the level of wage flexibility. Furthermore, there was an effort to assess the impact of a possible selection bias in the final results by estimating a Heckman selection model. These selection models were estimated in a specific way for each case. When formal and informal employees in the private sector were considered, the Heckman selection model aimed to control for the probability of being an employee among all male individuals aged 15 to 59 years old. On the other hand, when the model considered only formal or informal workers, the probit model calculated the probability of working in this specific sector.

6.2.2 Descriptive statistics

The analysis must consider not only the characteristics of individuals in the formal and the informal sectors, but also how they compare with individuals outside of these groups (workers in other occupations – domestic workers, self-employed, public employees, employers – unemployed individuals, people out of the labour force). The reason to follow this procedure is that these other groups were part of the comparison group in the Heckman selection model.

Table 5 presents the main characteristics of the database. The first aspect to be highlighted is the fact that the formal sector presented a much higher wage than the informal sector. For the labour force ex-formal and informal, this average was calculated only for those who work. The average age in the informal sector was lower, as well as their tenure at the same company. In the formal sector it was possible to observe higher values of per capital income ex-wages, meaning that these families had a more stable financial condition. In terms of their regional distribution, workers in the informal sector were relatively more concentrated in the Northeast and in the North, while workers in the formal sector were more present in the Southeast and the South.

The education level of workers in the formal sector was higher (2013, 9% of them had 15 years of schooling or more, while only 3.8% of workers in the informal sector had achieved that level). It is noteworthy that between 1996 and 2013, real wages and the educational level of the population increased significantly, while the participation of informal workers in the labour force went down.

Informal jobs were relatively more concentrated in agriculture, construction and services in 2013. At the same time, the formal sector was relatively more present in manufacture, transport and communication, and commerce. A similar pattern was observed for occupational groups: those requiring administrative, technical, artistic, scientific and transportation skills were relatively more frequent amongst jobs in the formal sector. In the informal sector, agriculture, livestock and extractive occupations showed a much higher participation in 2013.

Finally, salaries and the unemployment rate should be compared in each state over time. It is possible to note from Figure 3 that in 1996 the aggregated relationship between wages and the local unemployment rate for the formal and the informal sectors seemed to have a positive sign, while in 2013 their correlation seemed negative (more negative for the informal sector). In fact, Morrison and Poot (1998) identified that aggregated data may not provide any convincing evidence of the wage curve. However, when individual and local characteristics were controlled for, this pattern may change. The next subsection will discuss how wage flexibility varies according to the sector and the local characteristics of the labour market.

Table 5 – Descriptive statistics of the PNAD database.

	1996				2013			
	Labour force ex-formal and informal	Formal or informal	Formal	Informal	Labour force ex-formal and informal	Formal or informal	Formal	Informal
ln(hourly wage)	1.74	1.40	1.68	0.95	2.05	1.79	1.92	1.39
Age	37.27	30.39	31.85	28.11	39.79	33.57	34.25	31.60
Tenure	9.44	4.40	4.78	3.81	10.52	4.83	4.89	4.63
ln(per capita income ex-wage)	4.43	4.42	4.56	4.21	5.87	5.84	5.94	5.56
Region								
North	5.2%	3.7%	2.7%	5.3%	7.3%	5.6%	4.9%	7.8%
Northeast	30.9%	21.4%	14.6%	32.6%	31.1%	22.4%	17.7%	36.6%
Southeast	41.1%	52.4%	59.2%	41.1%	39.9%	47.7%	51.6%	35.8%
South	15.6%	15.8%	18.2%	11.9%	14.0%	15.9%	17.4%	11.6%
Centre-West	7.1%	6.8%	5.3%	9.2%	7.6%	8.3%	8.4%	8.2%
Education level								
Less than 1 year of schooling	14.1%	12.5%	7.0%	21.7%	8.3%	5.7%	4.0%	11.1%
1 to 3 years of schooling	16.1%	15.6%	12.7%	20.5%	7.2%	5.6%	4.0%	10.3%
4 to 7 years of schooling	33.4%	36.5%	36.3%	36.9%	23.2%	20.1%	17.2%	29.2%
8 to 10 years of schooling	16.6%	17.8%	20.9%	12.6%	22.0%	21.3%	20.5%	23.8%
11 to 14 years of schooling	14.2%	14.0%	18.3%	6.9%	29.2%	39.5%	45.3%	21.9%
15 years of schooling or more	5.6%	3.6%	4.9%	1.4%	10.1%	7.7%	9.0%	3.8%
% of self-reported black or brown	44.1%	44.4%	37.8%	55.3%	54.1%	54.5%	51.5%	63.6%
Student	18.0%	11.6%	10.0%	14.2%	19.6%	9.6%	8.4%	13.2%
Sector of activity								
Agriculture	20.3%	18.5%	9.0%	34.1%	11.9%	10.8%	6.2%	24.9%
Manufacture	4.0%	26.5%	34.3%	13.8%	3.4%	22.2%	26.1%	10.3%
Construction	7.3%	11.2%	8.3%	16.1%	10.9%	15.1%	12.8%	22.5%
Other industrial activities	1.4%	1.3%	1.5%	0.8%	0.7%	1.8%	2.2%	0.7%
Commerce	9.1%	14.7%	15.8%	12.8%	7.4%	16.4%	17.5%	13.0%
Services	9.1%	11.9%	12.1%	11.6%	7.9%	9.9%	8.5%	14.2%
Auxiliary services	2.7%	4.0%	4.5%	3.2%	2.7%	7.6%	8.9%	4.0%
Transport and communication	4.0%	6.4%	8.0%	3.7%	4.9%	9.9%	11.0%	6.5%
Social	3.4%	3.2%	3.7%	2.4%	4.2%	3.0%	3.2%	2.4%
Public administration	7.2%	0.0%	0.0%	0.0%	8.5%	0.0%	0.0%	0.0%
Other activities	1.4%	2.2%	2.7%	1.4%	1.1%	3.2%	3.7%	1.6%
Unemployed or out of the labour force	30.1%	0.0%	0.0%	0.0%	36.3%	0.0%	0.0%	0.0%
Occupation group								
Technical, scientific, artistic or similar	4.4%	3.8%	4.7%	2.3%	5.8%	5.1%	5.5%	3.8%
Administrative	8.7%	11.6%	16.0%	4.5%	8.6%	13.8%	16.3%	6.1%
Agriculture, livestock, extractive activities	19.3%	17.7%	8.3%	33.1%	12.0%	10.4%	5.7%	24.8%
Manufacture	13.6%	32.8%	33.3%	32.0%	14.6%	32.7%	32.9%	32.1%
Commerce and auxiliary activities	9.3%	9.5%	9.0%	10.2%	7.2%	13.6%	13.6%	13.3%
Transport, communication, services and other	14.5%	24.6%	28.7%	17.9%	15.5%	24.5%	26.0%	19.8%
Head of the household	56.2%	58.8%	65.0%	48.6%	47.7%	51.4%	53.9%	43.9%
% of people who have children under 15	45.4%	48.5%	53.6%	40.1%	27.9%	35.3%	36.8%	30.6%
Number of observations (unweighted)	59,606	36,862	22,873	13,989	39,752	47,038	35,305	11,733
Number of observations (weighted)	27,544,714	17,499,508	10,895,624	6,603,884	34,974,593	26,908,249	20,270,411	6,637,838

* All calculations were based on a weighted average of the observations (using *pweight* in Stata).

Source: PNAD-IBGE.

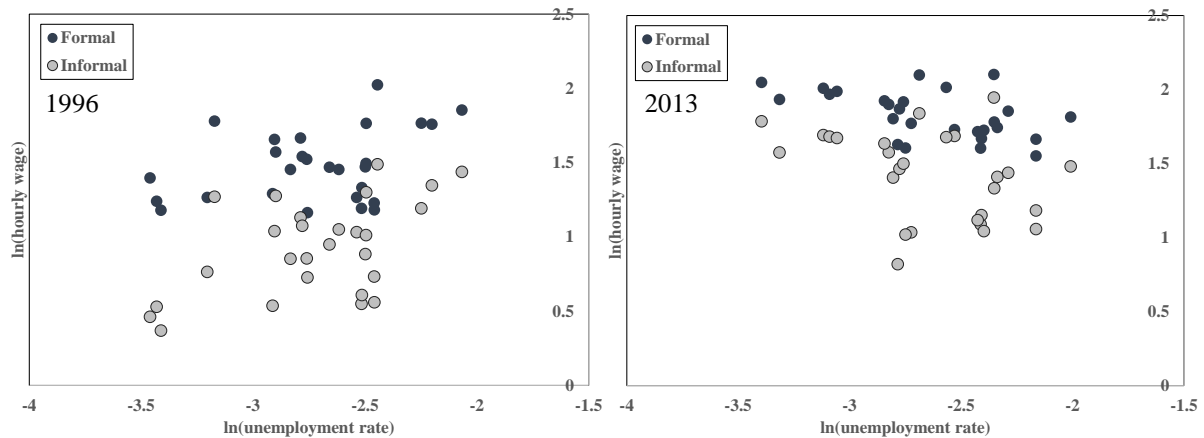


Figure 3 – Logarithm of the average hourly wage for the sample and logarithm of the unemployment rate at the state level, formal and informal sectors, 1996 and 2013.

Source: Author's own calculations.

6.2.3 Results

The first set of results of the wage curve estimated with PNAD data was based on Equation 12, (one-stage regression), with individual and regional-level data. This strategy was adopted here due to the limited number of regions (27 states), which could limit large-sample assumptions for a possible two-stage estimation as proposed in subsection 5.1.1. Furthermore, as discussed in subsection 6.2.1, two different time periods were analysed (1996 to 2013 and 2002 to 2009), to assess the possible varying effects over time, as well as how the results may be possibly affected by the instruments based on an imputation for 2000 and 2010 unemployment rates.

In Table 6, the first three columns present the results for pooled cross-sections OLS estimates of individual wages for the period from 1996 to 2013, controlling for state fixed effects. For the whole country (rural and urban), all coefficients were positive, with a lower value for the informal sector (Models 1a, 2a and 3a). However, then the model was estimated specifically for rural areas (Models 1b, 2b and 3b), only the elasticity of unemployment of the informal sector was significant, and with the expected negative sign. On the other side, for urban areas (Models 1c, 2c and 3c), all elasticities were positive and significant, counteracting the expected result of the wage curve.

Finally, the last three columns in the right hand side considered an IV estimation, with a one-year lag of the unemployment rate as an instrument for the actual unemployment rate in each state. Now, the coefficient for the unemployment rate over wages in the informal sector in the whole country was no longer significant, even though the coefficients for the formal sector and the whole economy were positive (and larger than the ones observed in the OLS estimations). When only individuals living in rural areas were taken into account, the elasticity of the unemployment rate with both sectors analysed simultaneously was negative and slightly significant, while the elasticity of the unemployment rate for wages in the informal sector increased in comparison to the OLS estimation and reached -0.134. Finally, the IV estimation for the urban sector indicated that the elasticity for the informal sector was not significant, while it was positive and significant for the formal sector and for the whole urban labour market. For all models, the F statistic indicated that the lag of the unemployment rate was a relevant instrument for the unemployment rate in t .

Table 6 – Elasticity of individual wages in relation to the state-level unemployment rate, for the whole country, urban and rural areas, and for the whole labour force, formal and informal sectors, 1996 to 2013.

Models	1996 to 2013					
	OLS			IV		
	Total	Informal	Formal	Total	Informal	Formal
Total						
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
ln(unemployment rate)	0.198*** (0.006)	0.052*** (0.010)	0.174*** (0.007)	0.266*** (0.010)	0.025 (0.017)	0.249*** (0.011)
F test instrument				164,515	50,967	106,030
Adjusted R ²	0.564	0.470	0.526	0.564	0.470	0.526
Rural						
	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b	Model 6b
ln(unemployment rate)	-0.005 (0.015)	-0.049** (0.021)	-0.013 (0.019)	-0.051** (0.025)	-0.134*** (0.037)	-0.054* (0.030)
F test instrument				23,161	9,690	11,538
Adjusted R ²	0.443	0.298	0.432	0.442	0.297	0.432
Urban						
	Model 1c	Model 2c	Model 3c	Model 4c	Model 5c	Model 6c
ln(unemployment rate)	0.200*** (0.006)	0.062*** (0.012)	0.182*** (0.007)	0.274*** (0.011)	0.034* (0.020)	0.266*** (0.012)
F test instrument				136,989	38,618	94,736
Adjusted R ²	0.545	0.460	0.523	0.545	0.460	0.522

Controls at the individual level: age, age², tenure, tenure², education level, sector of activity, occupational group, head of the household, has children under 15, self-reported black or brown. Complete tables for the period from 1996 to 2013 are presented in the Appendix. All models were estimated with *pweight* in Stata to account for the proportional weight of each observation in the whole population. Standard errors are presented in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Complete estimation results are presented in Table A.5 of the Appendix.

Source: Author's own calculations.

Table 7 – Elasticity of individual wages in relation to the state-level unemployment rate, for the whole country, urban and rural areas, and for the whole labour force, formal and informal sectors, 2002 to 2009.

Models	2002 to 2009					
	OLS			IV		
	Total	Informal	Formal	Total	Informal	Formal
Rural and Urban						
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
ln(unemployment rate)	0.106*** (0.011)	0.040** (0.020)	0.098*** (0.013)	0.106 (0.077)	-0.680*** (0.182)	0.232*** (0.079)
F test instrument				4,594	1,022	3,271
R ² 2nd stage	0.558	0.456	0.529	0.558	0.450	0.528
Rural						
	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b	Model 6b
ln(unemployment rate)	0.021 (0.030)	0.038 (0.040)	-0.021 (0.036)	-1.500*** (0.473)	-4.661*** (1.781)	-0.499* (0.296)
F test instrument				135	16	163
R ² 2nd stage	0.403	0.267	0.397	0.365	.	0.390
Urban						
	Model 1c	Model 2c	Model 3c	Model 4c	Model 5c	Model 6c
ln(unemployment rate)	0.103*** (0.012)	0.039* (0.022)	0.095*** (0.013)	0.185** (0.079)	-0.447** (0.178)	0.253*** (0.083)
F test instrument				4,302	1,055	3,008
R ² 2nd stage	0.541	0.450	0.526	0.541	0.447	0.526

Controls at the individual level: age, age², tenure, tenure², education level, sector of activity, occupational group, head of the household, has children under 15, self-reported black or brown. Instrument for the unemployment rate at the state level was the local unemployment rate one year before. All models were estimated with *pweight* in Stata to account for the proportional weight of each observation in the whole population. Standard errors are presented in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Complete estimation results can be provided by the author under demand.

Source: Author's own calculations.

When the period of analysis was restricted to 2002 to 2009, there seemed to be more indication of the presence of a wage curve in the relationships of the labour market. In general, elasticities were lower or even became negative, a sign that the wage curve seemed to better express this specific period. This regression covered the same period studied by Baltagi, Rokicki and Souza (2014a). Even though their results seemed to better confirm the expected shape of the wage curve, the main conclusions were very similar to the ones found here. Wage flexibility seemed to be higher in rural areas and in the informal sector. This result was in accordance with the model discussed in Section 4 and added some conclusions to the conclusions obtained in subsection 6.1.

The last set of results in this block refer to the estimation of the wage curve accounting for a possible selection bias. As mentioned before, each model had a different first stage estimation. All of them were based on probit models, but the dependent variable in the selection regression was specific for each case.

Table 8 – Second-stage results of the IV estimation of the elasticity of individual wages in relation to the state-level unemployment rate with a Heckman selection model in the first stage.

Models	1996 to 2013			2002 to 2009		
	Total	Informal	Formal	Total	Informal	Formal
	Total					
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
ln(unemployment rate)	0.296*** (0.011)	-0.010 (0.019)	0.390*** (0.013)	0.227*** (0.087)	-0.763*** (0.205)	0.547*** (0.090)
F test instrument	128,480	138,479	86,812	3,472	1,978	2,682
Adjusted R ²	0.586	0.500	0.533	0.580	0.479	0.537
N	560543	198023	362520	306648	111572	195076
	Rural					
	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b	Model 6b
ln(unemployment rate)	-0.076*** (0.029)	-0.147*** (0.041)	-0.098*** (0.035)	-1.738*** (0.660)	-5.258* (2.869)	-0.419 (0.349)
F test instrument	17,728	31,384	9,856	75	0	125
Adjusted R ²	0.451	0.321	0.443	0.353	.	0.405
N	58774	37626	21148	32069	21039	11030
	Urban					
	Model 1c	Model 2c	Model 3c	Model 4c	Model 5c	Model 6c
ln(unemployment rate)	0.295*** (0.012)	0.010 (0.022)	0.362*** (0.013)	0.241*** (0.087)	-0.567*** (0.196)	0.466*** (0.092)
F test instrument	106,826	100,940	78,588	3,305	2,185	2,511
Adjusted R ²	0.572	0.493	0.541	0.568	0.479	0.545
N	501769	160397	341372	274579	90533	184046

Controls at the individual level: age, age², tenure, tenure², education level, sector of activity, occupational group, head of the household, has children under 15, self-reported black or brown. Instrument for the unemployment rate at the state level was the local unemployment rate one year before. All models were estimated with *pweight* in Stata to account for the proportional weight of each observation in the whole population. Standard errors are presented in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. The complete estimation for 1996-2013 is provided in Table A.6 in the Appendix.

Source: Author's own calculations.

For instance, in the case of the model for the formal sector in rural areas, the selection equation compared individuals in the formal sector against all other workers (including those in the informal sector) and people out of the labour force. Furthermore, in the second stage, all models were estimated with IV (and the instrument was the logarithm of the time lag of the unemployment rate).

The results of Table 8 indicated that wage flexibility in the rural sector seemed to be even stronger when selection bias over observable characteristics was taken into account. In the case of Model 5b, the instrument did not explain much of the unemployment rate, generating problems for the estimation of the second stage of the IV model.

In summary, the analysis of the wage curve with state-level data provided some insights into the differences of wage flexibility between urban and rural areas. In fact, more accessibility to alternative jobs in agglomerated urban areas seemed to reduce the bargaining power of firms, decreasing their capacity to change offered wages according to the business cycle. There was also indication of labour market dualism and a much higher wage flexibility could be found in the informal sector, in accordance with the literature and with previous results discussed here (Section 3).

6.3 City size, spatial dependence, and the wage curve

There is a strong concern of whether state-level regional aggregation is satisfactory for the estimation of the wage curve. It is not clear if the unemployment rate in such large areas can substantially affect the decisions of economic agents at a more disaggregated local level. Furthermore, the state-level unemployment rate does not vary much between individuals, decreasing its explanatory power over the variation of wages. These issues may affect the estimation results. This subsection will be based on a multi-level analysis (see subsection 5.1.1) and will explore the wage curve in a more disaggregated context (municipalities), controlling for urban size, proximity, spatial dependence, and a possible endogeneity of the unemployment rate.

6.3.1 Description of the database – Demographic Census

Every 10 years the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE) conducts the Demographic Census, which covers the whole country with regional disaggregation at the municipal level (or at the neighbourhood level for bigger municipalities). The Demographic Census investigates the main characteristics of individuals and households (education, migration, labour market supply aspects, family composition, living conditions, among others), providing details on the life conditions of the population in each municipality, and serving as a very relevant policy instrument in a country with continental measures such as Brazil.

There is a shorter questionnaire applied to the whole population, while specific individual characteristics are investigated in a longer set of questions applied to a sample. Microdata at the individual level is available for this sample, which can range from 5 to 50 percent of the population depending on the size of each municipality (for instance, the sample size corresponds to 5% of the total population of municipalities with more than 500,000 inhabitants). Even if there are some changes over the years in this survey, the main structure and most methodological aspects have been kept constant over time, mainly in the last three censuses.

The analysis that follows was based on information obtained from the demographic censuses of 1991, 2000 and 2010. Over this period, there has been a significant change in the number of municipalities in Brazil, with the emancipation and consequent foundation of new local administrations. In 1991, there were 4,491 municipalities; in 2000, 5,507; and in 2010, 5,565 municipalities. Any analysis involving data over this time period should take into consideration the fact that the borders of municipalities have been changing, due to merging or splitting processes to create new municipalities (Reis et al., 2011). Therefore, 4,258 Minimum Comparable Areas (MCAs) for the period 1991-2010 were defined keeping constant the borders and areas of each regional unit of analysis. An important detail is that the municipality of reference was based on the place the individual worked (in order to avoid wage sub-estimation in municipalities that aggregate jobs and wage super-estimation in dormitory towns).

A second and bigger level of regional aggregation based on Census data is the Immediate Region of Urban Articulation (Regiões de Influência de Cidades – REGIC areas henceforth), a compound of municipalities that constitutes a labour market area. The criteria to set these local regions were defined by IBGE based on the flow of goods and services (IBGE, 2007) for the 2010 municipal configuration.

Their configuration was harmonised with the 1991-2010 MCAs previously defined, obtaining 478 REGIC areas (from originally 482).⁶ Here, they were included in the concept of neighbourhood. Therefore, MCAs were considered neighbours if they belonged to the same REGIC area, and the strength of this neighbourhood was based on the inverse distance between the centroids of each MCA). For the cases in which only one MCA formed a REGIC area, the neighbourhood comprised the closest MCA, even if it belonged to another REGIC area.

Similarly to the previous procedures to generate the final database, only male individuals aged 15 to 59 years old were kept, employed with or without a formal contract (excluding self-employment, domestic work, workers in the public administration sector, employers, unemployed and individuals out of the economically active population), working from 20 to 60 hours per week. Furthermore, individuals working in the public sector or public administration were also excluded, in addition to those without information about the location of their jobs. Wages in 2000 were deflated to real values of August 2010 by the national INPC index of the period. In addition, a sampling procedure was adopted to generate a computationally manageable database. The original sample in 2000 comprised 2,237,102 interviews, representing 19,381,332 individuals in the population. In 2010, these numbers were 2,706,474 and 26,290,915, respectively.

Finally, the results that follow were based in two-stage regressions (see subsection 5.1.1), with the spatial wage obtained in the first stage being considered as the dependent variable in the second stage. However, there is not enough information to obtain estimates for all MCAs, because the sample does not cover the whole country. Then, when spatial dependence was explicitly considered in the second stage, it was necessary to input values for the missing data, and this was done by calculating the average of the spatial wage among the neighbours⁷.

6.3.2 Descriptive statistics

The main descriptive statistics of the Census database are presented in two parts. Table 9 depicts the distribution of some variables over the sample of individuals for the estimation of the first stage of the wage curve. In general, real wages increased between 2000 and 2010. Over this period, education attainment of the population also went up, and workers were dislocated from agriculture to commerce and services. Manufacture had a relevant participation in the total employment analysed in this database, followed by occupations in commerce and service.

The first stage estimation generates a spatial wage at the MCA-year level, which then is used in the second stage as the dependent variable.

The upper part of Table 10 shows the mean spatial wages in 2000 for each group of MCAs, classified according to their density. It is noticeable that the spatial wage is higher for denser MCAs, in which it is also possible to find higher unemployment rates. However, the comparison of the estimated spatial wages for different sectors because they result from different regressions.

⁶ The definition of the 1991-2010 MCAs and respective REGIC areas can be made available by the author upon request.

⁷ This procedure was adopted for 18 MCAs in 2000 and 3 MCAs in 2010 for the whole labour market (formal and informal), 454 MCAs in 2000 and 170 MCAs in 2010 for the formal sector, and 63 in 2000 and 29 in 2010 for the informal sector.

Table 9 – Descriptive statistics of the variables in the first stage (individual level).

	2000	2010
ln(hourly wage)	2.78	2.98
Age	32.01	33.71
Student	14.0%	12.0%
Self-reported black or brown	44.4%	51.2%
Head of the household	59.6%	50.6%
Education level		
Less than 7 years of schooling	53.6%	37.5%
8 to 10 years of schooling	19.7%	20.7%
11 to 14 years of schooling	21.1%	33.0%
15 years of schooling or more	5.6%	8.7%
Sector of activity		
Agriculture	14.9%	10.9%
Manufacture and construction	34.7%	33.2%
Other industrial activities	2.8%	2.7%
Commerce	18.3%	19.3%
Services	5.4%	6.1%
Auxiliary services	5.9%	7.4%
Transport and communication	7.9%	7.9%
Social services	2.0%	1.9%
Education	3.5%	2.5%
Other activities	4.7%	8.0%
Occupation group		
Leaders	3.8%	3.5%
Scientific, artistic or similar	4.6%	6.3%
Technical level	7.4%	7.2%
Administrative service	8.9%	9.1%
Commerce and service	20.1%	21.7%
Agriculture, livestock, extractive activit	13.6%	9.5%
Manufacture	41.6%	42.7%

Source: Demographic Census.

Table 10. Descriptive statistics of the second stage – spatial wage.

	2000			
	Spatial wage	Spatial wage	Spatial wage	ln(unemployment
	- total	- formal	- informal	rate)
Density of the MCA				
Less than 1	-0.191	-0.219	-0.017	-2.249
1 to less than 2	-0.303	-0.289	-0.144	-2.360
2 to less than 5	-0.376	-0.351	-0.215	-2.274
5 to less than 10	-0.397	-0.376	-0.238	-2.415
10 to less than 20	-0.331	-0.342	-0.188	-2.482
20 to less than 50	-0.271	-0.303	-0.152	-2.228
50 or more	-0.009	-0.107	0.103	-1.860
	2010			
	Spatial wage	Spatial wage	Spatial wage	ln(unemployment
	- total	- formal	- informal	rate)
Density of the MCA				
Less than 1	0.052	0.052	0.172	-2.666
1 to less than 2	-0.070	-0.049	0.055	-2.731
2 to less than 5	-0.119	-0.110	0.005	-2.738
5 to less than 10	-0.145	-0.132	-0.006	-2.859
10 to less than 20	-0.120	-0.134	0.034	-2.942
20 to less than 50	-0.078	-0.126	0.060	-2.811
50 or more	0.094	-0.023	0.246	-2.617
	Spatial wage	Spatial wage	Spatial wage	ln(unemployment
	- total	- formal	- informal	rate)
2000	-0.302	-0.310	-0.157	-2.310
2010	-0.082	-0.104	0.058	-2.809

Source: Author's elaboration.

6.3.3 Results

The main focus of this subsection is in the relationship between the spatial wage and the local unemployment rate. This spatial wage is obtained as the MCA-year effect net of observed individual characteristics. Complete tables with first-stage results are provided in Table A.7 in the Appendix.

The initial results of the second stage are presented in Table 11. First of all, the spatial wage of the informal sector is negatively correlated with the local unemployment rate, while for the total and the formal sectors the results are not significant (Models 1 to 3). When these results are disaggregated by different density levels, the wage elasticity to the local unemployment rate is negative for all density classes in the informal sector, while for the formal sector, whenever this elasticity is negative, it is significant. There is a U shape relationship between wages and the unemployment rate according to density levels. This elasticity is higher in absolute terms for low and high-density MCAs, while it is lower for medium-density MCAs.

As mentioned in Section 5, the local unemployment rate may be endogenous to the wage level. Therefore, Models 7 to 12 provide a tentative instrumental variables estimation, with the 10-year lag of the unemployment rate as an instrument (and the lagged unemployment rate iterated with density groups). Now, all significant coefficients are negative, corroborating the expected shape of the wage curve, but the formal sector is the only one that presents a significant elasticity for all MCAs (Models 7 to 9).

However, there is a potential problem here. Local unobserved characteristics of the labour markets may be affecting the results. Therefore, in Table 12, fixed effects for MCAs are included. These results are very different from the main conclusions obtained from Table 11. In fact, areas with higher population density will show a lower level of wage flexibility, as expected from the model formulated in Section 4. Therefore, the main conclusions obtained from Table 11 were being driven by unobserved local effects.

Finally, the last set of results allows for the existence of neighbourhood effects. As mentioned previously, the empirical literature has started to include not only the spatial lag of the unemployment rate as an explanatory variable, but also to control for spatial dependence. Models 1, 6 and 11 in Table 13 present the basic specification: the spatial wage is explained by local characteristics and the local unemployment rate. In this first estimation, LM tests indicated that the spatial autoregressive model (SAR) seemed to be the most appropriate spatial specification.

Then, the spatial lag of the unemployment rate was included in Models 2, 7 and 12. These results show that the informal sector not only had a higher elasticity for the local unemployment rate, but also for the unemployment rate in the neighbouring areas. Nonetheless, this formulation does not tackle the spatial dependence, which is once again found by LM tests. One last try before estimating the SAR model is to control for unobserved local characteristics, including MCA fixed effects. As expected, the elasticity of wages in relation to local unemployment rates has decreased, but the informal sector still showed the highest coefficients (Models 3, 8 and 13).

Moving to Models 4, 9 and 14, the estimation of a SAR model without local fixed effects indicates that the lag of the dependent variable is significant, and that the unemployment elasticity becomes non-significant (at least for the total labour market). However, the coefficient for the spatial lag of the unemployment rate is still negative and significant. Finally, the last three models (5, 10 and 15) indicate that whenever a SAR model with spatial fixed effects is estimated, the elasticity of the spatial wage to the local unemployment rate is very similar in the local level for the informal and the formal sectors (-0.023 and -0.024, respectively). However, the elasticity for the unemployment rate in the neighbours is much higher for the informal sector (-0.117, against -0.038 in the formal sector).

Table 11 – Regression of the spatial wage against local characteristics, OLS and IV, 2000 and 2010.

	OLS						IV					
	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
ln(unemployment rate)	0.007 (0.005)	-0.028*** (0.006)	-0.006 (0.006)				0.002 (0.020)	-0.021 (0.023)	-0.047* (0.027)			
ln(unemployment rate) * density < 1				-0.052*** (0.008)	-0.084*** (0.009)	-0.060*** (0.010)				-0.098*** (0.023)	-0.122*** (0.027)	-0.141*** (0.031)
ln(unemployment rate) * 1 <= density < 2				-0.011* (0.007)	-0.041*** (0.008)	-0.029*** (0.009)				-0.052** (0.022)	-0.080*** (0.026)	-0.095*** (0.030)
ln(unemployment rate) * 2 <= density < 5				0.010* (0.006)	-0.023*** (0.007)	-0.011 (0.007)				-0.021 (0.022)	-0.047* (0.025)	-0.083*** (0.030)
ln(unemployment rate) * 5 <= density < 10				0.018*** (0.005)	-0.018*** (0.006)	-0.001 (0.007)				-0.011 (0.021)	-0.037 (0.024)	-0.065** (0.028)
ln(unemployment rate) * 10 <= density < 20				0.011** (0.005)	-0.026*** (0.006)	-0.002 (0.006)				-0.027 (0.020)	-0.049** (0.023)	-0.071*** (0.027)
ln(unemployment rate) * 20 <= density < 50				-0.009* (0.005)	-0.043*** (0.007)	-0.011* (0.007)				-0.061*** (0.021)	-0.074*** (0.024)	-0.095*** (0.028)
ln(unemployment rate) * density >= 50				-0.036*** (0.007)	-0.081*** (0.008)	-0.037*** (0.008)				-0.133*** (0.024)	-0.154*** (0.027)	-0.167*** (0.032)
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R ²	0.515	0.440	0.217	0.531	0.455	0.228	0.465	0.381	0.147	0.435	0.354	0.135
F test - instrument: ln(unemployment rate) _{t-1}							700.4	699.6	682.1			
Minimum eigenvalue										91.1	90.1	88.7
N	8,487	8,416	7,885	8,487	8,416	7,885	4,148	4,109	3,726	4,148	4,109	3,726

*Additional controls: constant, percentage of individuals aged 25 or more who completed at least secondary school. Dependent variable is the spatial wage that results from the first stage regression of the logarithm of the individual hourly wage controlled by education attainment, sector of activity, occupation, age, age squared, if the individual is a student, the head of the household, or self-reported black or brown. The local unemployment rate is instrumented by its 10-year time lag. First-stage estimation results are presented in Table A.7 in the Appendix.

Source: Author's elaboration.

Table 12 – Regressions for the spatial wage with MCAs fixed effects.

	FE					
	Total	Informal	Formal	Total	Informal	Formal
ln(unemployment rate)	-0.022*** (0.008)	-0.047*** (0.011)	-0.002 (0.012)			
ln(unemployment rate) * density < 1				-0.029* (0.017)	-0.056** (0.022)	-0.001 (0.025)
ln(unemployment rate) * 1 <= density < 2				-0.032** (0.013)	-0.038** (0.017)	-0.019 (0.020)
ln(unemployment rate) * 2 <= density < 5				-0.036*** (0.011)	-0.048*** (0.014)	-0.040*** (0.015)
ln(unemployment rate) * 5 <= density < 10				-0.039*** (0.009)	-0.061*** (0.012)	-0.014 (0.014)
ln(unemployment rate) * 10 <= density < 20				-0.022** (0.009)	-0.049*** (0.012)	-0.008 (0.013)
ln(unemployment rate) * 20 <= density < 50				-0.008 (0.010)	-0.033** (0.013)	0.015 (0.014)
ln(unemployment rate) * density >= 50				0.037*** (0.013)	-0.023 (0.017)	0.063*** (0.018)
MCA fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
R ² between	0.300	0.058	0.007	0.244	0.059	0.022
R ² within	0.319	0.214	0.303	0.389	0.219	0.307
N	8487	8416	7885	8487	8416	7885

*Additional controls: constant, percentage of individuals aged 25 or more who completed at least secondary school. Dependent variable is the spatial wage that results from the first stage regression of the logarithm of the individual hourly wage controlled by education attainment, sector of activity, occupation, age, age squared, if the individual is a student, the head of the household, or self-reported black or brown. The local unemployment rate is instrumented by its 10-year time lag. First-stage estimation results are presented in Table A.7 in the Appendix.

Source: Author's elaboration.

In sum, the estimation of the wage curve with disaggregated data shows that it is important to include spatial effects in the estimated model. In fact, the theoretical framework discussed in Section 4 had already indicated that whenever another local labour market is sufficiently close to the region analysed, it is possible that workers look for opportunities in these neighbouring areas. This effect is supposed to be even more important here because REGIC areas (local labour markets) are composed by MCAs (the spatial unit of analysis). Therefore, a common shock inside the REGIC area is supposed to affect all MCAs that compose it, with spillover effects between them. This result indicates that there are relevant spatial effects inside a common labour market area. In addition, workers in the informal sector seem to be more affected by these neighbourhood effects.

This analysis is complemented by the results presented in Table 14. In that case, all models include the iteration between the local unemployment rate and the density group of the MCA. Furthermore, the inclusion of spatial effects is done through the iteration of the spatial lag of the unemployment rate and density groups.

The comparison of Models 1, 6 and 11 of Table 14 indicate that there is a higher wage flexibility in the informal sector. The elasticity of the spatial wage in relation to the local unemployment rate follows a U-shaped curve according to the local density. These results corroborate the initial conclusions obtained in Tables 12 and 13. Furthermore, spillover effects seemed to be more relevant for the informal sector, especially when fixed effects at the MCA level are taken into account (Models 3, 8 and 13). A great share of the difference in the wage flexibility between the formal and the informal sectors is captured by the spatial lag of the unemployment rate. This is an indication that the whole labour market area (REGIC area) seems to be more relevant for workers in the informal sector, while workers in the formal sector are usually more influenced by the unemployment rate at the MCA in which they are currently working.

Table 13 – Regression of the spatial wage against local characteristics, OLS, FE and SAR, 2000 and 2010.

	Total					Formal					Informal				
	OLS Model 1	OLS Model 2	FE Model 3	SAR Model 4	SAR - FE Model 5	OLS Model 6	OLS Model 7	FE Model 8	SAR Model 9	SAR - FE Model 10	OLS Model 11	OLS Model 12	FE Model 13	SAR Model 14	SAR - FE Model 15
ln(unemployment rate)	-0.035***	-0.024***	-0.018***	0.005	-0.016*	-0.051***	-0.034***	-0.025***	-0.018***	-0.024**	-0.068***	-0.044***	-0.026***	-0.021***	-0.023**
W*ln(unemployment rate)		-0.026***	-0.102***	-0.033***	-0.074***		-0.039***	-0.059***	-0.024***	-0.038**		-0.052***	-0.149***	-0.032***	-0.117***
Density >= 1 and density < 2	-0.096***	-0.096***				-0.063***	-0.063***				-0.102***	-0.102***			
Density >= 2 and density < 5	-0.154***	-0.154***				-0.123***	-0.122***				-0.155***	-0.155***			
Density >= 5 and density < 10	-0.178***	-0.179***				-0.150***	-0.150***				-0.172***	-0.173***			
Density >= 10 and density < 20	-0.166***	-0.167***				-0.156***	-0.157***				-0.160***	-0.162***			
Density >= 20 and density < 50	-0.111***	-0.111***				-0.142***	-0.143***				-0.115***	-0.116***			
Density >= 50	-0.062***	-0.059***				-0.096***	-0.090***				-0.046**	-0.039*			
W*Dependent variable				0.617***	0.303***				0.431***	0.359***				0.549***	0.193***
Constant	Yes	Yes		Yes		Yes	Yes		Yes		Yes	Yes		Yes	
% people completed at least middle school	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes				Yes	Yes				Yes	Yes			
MCA fixed effects			Yes		Yes			Yes		Yes			Yes		Yes
LM lag	1918.3***	1930.9***	481.9***			1040.5***	1033.9***	746.6***			1093.3***	1079.6***	159.9***		
LM error	494.6***	489.7***	325.0***			317.9***	304.5***	242.2***			338.2***	312.1***	112.8***		
LM lag robust	1426.0***	1443.6***	336.8***			775.7***	783.1***	622.8***			785.0***	795.0***	120.3***		
LM error robust	2.3	2.3	179.9***			53.1***	53.7***	118.4***			29.9***	27.5***	73.2***		
R ²	0.523	0.523	0.318	0.592	0.836	0.221	0.222	0.200	0.300	0.682	0.451	0.453	0.247	0.487	0.784
N	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516	8,516

*Dependent variable is the spatial wage that results from the first stage regression of the logarithm of the individual hourly wage controlled by education attainment, sector of activity, occupation, age, age squared, if the individual is a student, the head of the household, or self-reported black or brown. First-stage estimation results are presented in Table A.7 in the Appendix.

Source: Author's elaboration.

Table 14 – Regression of the spatial wage against local characteristics and the iteration of density groups and the local unemployment rate, OLS, FE and SAR, 2000 and 2010.

	Total					Formal					Informal				
	OLS Model 1	OLS Model 2	FE Model 3	SAR Model 4	SAR - FE Model 5	OLS Model 6	OLS Model 7	FE Model 8	SAR Model 9	SAR - FE Model 10	OLS Model 11	OLS Model 12	FE Model 13	SAR Model 14	SAR - FE Model 15
ln(unemployment)*Density < 1	-0.092***	-0.085***	-0.058**	-0.093***	-0.060*	-0.103***	-0.041	-0.027	-0.045	-0.032	-0.121***	-0.113***	-0.079***	-0.129***	-0.081*
ln(unemployment)*Density >= 1 and density < 2	-0.051***	-0.013	0.033	-0.012	0.033	-0.076***	-0.007	-0.030	-0.016	-0.035	-0.079***	-0.026	0.046*	-0.020	0.047
ln(unemployment)*Density >= 2 and density < 5	-0.028***	-0.017	-0.031**	-0.001	-0.027	-0.055***	-0.057***	-0.094***	-0.050***	-0.089***	-0.059***	-0.045***	-0.050***	-0.034**	-0.047**
ln(unemployment)*Density >= 5 and density < 10	-0.018***	-0.007	-0.038***	0.002	-0.039**	-0.043***	-0.013	-0.014	-0.010	-0.020	-0.052***	-0.020	-0.038***	-0.011	-0.039**
ln(unemployment)*Density >= 10 and density < 20	-0.023***	-0.004	-0.012	0.017*	-0.005	-0.042***	-0.025**	-0.027	-0.045	-0.032	-0.058***	-0.023*	-0.024*	-0.010	-0.018
ln(unemployment)*Density >= 20 and density < 50	-0.041***	-0.049***	-0.046***	-0.001	-0.038*	-0.047***	-0.057***	-0.030	-0.016	-0.035	-0.073***	-0.070***	-0.035*	-0.031**	-0.026
ln(unemployment)*Density >= 50	-0.057***	-0.070***	0.110***	0.011	0.099**	-0.059***	-0.042	-0.094***	-0.050***	-0.089***	-0.102***	-0.124***	0.088**	-0.051*	0.090*
W*ln(unemployment)*Density < 1		-0.020	-0.093***	0.014	-0.057		-0.087***	-0.014	-0.010	-0.020		-0.035	-0.123***	0.015	-0.083
W*ln(unemployment)*Density >= 1 and density < 2		-0.052**	-0.170***	-0.038*	-0.139***		-0.094***	0.000	-0.011	0.009		-0.082***	-0.218***	-0.061**	-0.185***
W*ln(unemployment)*Density >= 2 and density < 5		-0.024	-0.113***	-0.037***	-0.083***		-0.018	-0.060***	-0.030**	-0.056**		-0.041**	-0.130***	-0.033*	-0.097***
W*ln(unemployment)*Density >= 5 and density < 10		-0.024*	-0.106***	-0.031***	-0.073***		-0.051***	0.094**	0.007	0.083		-0.059***	-0.155***	-0.047***	-0.118***
W*ln(unemployment)*Density >= 10 and density < 20		-0.031**	-0.113***	-0.050***	-0.091***		-0.036**	-0.032	-0.053*	-0.006		-0.062***	-0.158***	-0.052***	-0.129***
W*ln(unemployment)*Density >= 20 and density < 50		-0.004	-0.057***	-0.041***	-0.040*		-0.009	-0.066**	-0.057**	-0.034		-0.028*	-0.126***	-0.036**	-0.102***
W*ln(unemployment)*Density >= 50		-0.001	-0.130***	-0.063***	-0.111***		-0.039	-0.020	-0.005	-0.001		-0.008	-0.209***	-0.037	-0.185***
W*Dependent variable				0.610***	0.280***									0.528***	0.187***
Constant	Yes	Yes		Yes		Yes	Yes		Yes		Yes	Yes		Yes	
% people completed at least middle school	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes				Yes	Yes				Yes	Yes			
MCA fixed effects			Yes		Yes			Yes		Yes			Yes		Yes
LM lag	1,936.1	1,948.5	389.0			1,045.2	1,025.5	667.4			800.1	1,086.9	145.9		
LM error	474.3	470.5	396.2			305.6	294.0	255.2			1,100.9	298.2	134.0		
LM lag robust	1,462.8	1,479.7	224.9			786.9	782.6	525.6			25.8	812.6	101.7		
LM error robust	1.0	1.7	232.1			47.3	51.1	113.4			326.6	24.0	89.8		
R ²	0.521	0.522	0.335	0.598	0.838	1045.227	1025.513	667.443	0.000	0.000	0.450	0.453	0.253	0.492	0.785

*Dependent variable is the spatial wage that results from the first stage regression of the logarithm of the individual hourly wage controlled by education attainment, sector of activity, occupation, age, age squared, if the individual is a student, the head of the household, or self-reported black or brown. First-stage estimation results are presented in Table A.7 in the Appendix.

Source: Author's elaboration.

Once again, LM tests indicated that a spatial autoregressive model would be more adequate to deal with spatial dependence. Combining the spatial lag of the dependent variable (spatial wage) with density groups and controlling for local fixed effects, Models 5, 10 and 15 showed that the wage flexibility in relation to the business cycle is much higher in absolute values in the informal sector, for low-density MCAs. In fact, these conclusions meet the predictions of the model discussed in Section 4, in which a worker employed in a less dense labour market has a lower bargaining power. Being surrounded by fewer job opportunities (higher unemployment rates in the neighbouring areas) reduces once again workers' bargaining power, and allow firms to adjust real wages according to the moment of the economy.

The preferred estimates are Models 5, 10 and 15 of Tables 13 and 14. They provide comparisons between the formal and the informal sectors and the whole labour market, control for unobserved local characteristics and try to deal with spatial dependence. From Table 13, the estimated wage elasticity in relation to the unemployment rate is very similar for the formal and the informal sectors (-0.024 and -0.023, respectively), but its spatial lag has a much larger coefficient in the case of the informal sector (-0.117, against -0.038 for the formal sector).

The main conclusions of these results are the following: (i) wage flexibility is higher in less dense local labour markets; (ii) wages in the informal sector are more flexible than in the formal sector; (iii) it is essential to control for unobserved local characteristics in order to obtain the true elasticity of wages to local unemployment rates; (iv) it is important to control for spatial effects when the unit of analysis is sufficiently small; (v) a significant part of the difference between the formal and the informal sectors rely on spatial effects.

7 Conclusion

This paper aimed to provide a comprehensive analysis of the wage curve in Brazil, exploring different databases in order to obtain a deep understanding of the mechanisms that are behind the process of wage bargaining. In order to do that, a theoretical framework combining wage bargaining, institutional costs and efficiency wages is provided, distinguishing the formal and the informal sectors by the relative level of bargaining power of workers employed in each of them.

Many empirical issues are investigated, aiming to understand their influence over the relationship of wages and local unemployment rates. Among them, it is possible to highlight multilevel analysis, endogeneity of the unemployment rate, selection bias, spatial dependence, and the degree to which the wage curve is dynamic or not.

The first set of results examined the wage curve in a more aggregated context, for six metropolitan regions. The informal sector seems be more flexible to economic downturns. This result seems to corroborate other findings in the literature for the wage curve in developing economies. However, the analysis of just a few metropolitan regions of the country is not enough to get a deep understanding of the whole underlying processes. Another conclusion of this set of results is that individual fixed effects should be controlled for whenever possible, as they explain a large share of wage differentials. However, in the particular case of PME, their inclusion do not leave enough variation to capture the wage curve relationship.

Then, the analysis of the wage curve with state-level data from PNAD provided some insights into the differences of wage flexibility between urban and rural areas. In fact, higher accessibility to alternative jobs in agglomerated urban areas seemed to reduce the bargaining power of firms, decreasing their capacity to change offered wages according to the business cycle. There was also an indication of labour market dualism and a much higher wage flexibility could be found in the informal sector, following the literature (Section 3).

Finally, Census data seems to provide the best regional disaggregation to study the wage curve in Brazil. In fact, when local characteristics were adequately controlled for, the findings suggested that

wage flexibility was higher in less dense local labour markets and wages in the informal sector were much more adjustable than those in the formal sector. Furthermore, spatial dependence is a relevant issue to account for all possible effects related to the accessibility of jobs to neighbouring areas and for common shocks at the local labour market level. In this sense, a large part of the difference between the formal and the informal sector is due to the spatial lag of the unemployment rate.

All these results corroborate the main predictions from the model discussed in Section 4. In fact, labour market dualism is an essential ingredient in the evaluation of the wage curve in developing economies. However, this type of analysis should be conducted in the adequate regional level (labour market areas). When this step is followed, it is possible to find a relevant impact of city size on the relative bargaining power of workers and firms.

Appendix

Table A.1 – Definition and source of the main variables considered in the model of the wage curve, PME, PNAD and Census.

PME			(continues)
Variables	Definition	Level	Data source
Hourly wage	Monthly wage received in each month divided by 4.3 times the number of weekly hours in the contract, deflated by the local INPC.	Individual	PME panel
Unemployment rate	Individuals without a job who actively looked for a new position in the previous month over the economically active population.	Metropolitan region	PME
Age	Age at the month of the interview.	Individual	PME panel
Student	Dummy variable that equals one if the individual is currently studying and zero otherwise.	Individual	PME panel
Education attainment	Less than 1 year of schooling, 1 to 3 years of schooling, 4 to 7 years of schooling, 8 to 10 years of schooling and 11 years of schooling or more.	Individual	PME panel
Self reported black or brown	Individuals who identified themselves as black or brown.	Individual	PME panel
Sector of activity	Manufacture, construction, commerce and repair, financial services and real estate, other services, other sectors.	Individual	PME panel
Metropolitan region	Place in which the individual lives: Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo and Porto Alegre.	Individual	PME panel
Month	Month in which the interview was done.	Individual	PME panel
Year	Year in which the interview was done.	Individual	PME panel
PNAD			
Variables	Definition	Level	Data source
Hourly wage	Monthly wage received in October divided by 4.3 times the number of weekly hours in the contract, deflated by the national INPC.	Individual	PNAD microdata
Unemployment rate	Individuals without a job who actively looked for a new position in the previous month over the economically active population.	State	PNAD
Age	Age in October of each year.	Individual	PNAD microdata
Tenure	Number of years in the same job.	Individual	PNAD microdata
Per capita income ex-wage	Household per capita income ex-wage obtained by any household member.	Individual	PNAD microdata
Education attainment	Less than 1 year of schooling, 1 to 3 years of schooling, 4 to 7 years of schooling, 8 to 10 years of schooling, 11 to 14 years of schooling, and 15 years of schooling or more.	Individual	PNAD microdata
Student	Dummy variable that equals one if the individual is currently studying and zero otherwise.	Individual	PNAD microdata
Self reported black or brown	Individuals who identified themselves as black or brown.	Individual	PNAD microdata
Sector of activity	Agriculture, manufacture, construction, other industrial activities, commerce, services, auxiliary services, transport and communication, social services, other activities.	Individual	PNAD microdata
Occupation group	Technical, scientific, artistic or similar; administrative; agriculture, livestock, extractive activities; manufacture; commerce and auxiliary activities; transport, communication, services and others.	Individual	PNAD microdata
Head of the household	Dummy variable that equals one if the individual is the head of the household and zero otherwise.	Individual	PNAD microdata
People who have children under 15	Dummy variable that equals one if the individual has children under 15 and zero otherwise.	Individual	PNAD microdata

Demographic Census		(ends)	
Variables	Definition	Level	Data source
Hourly wage	Monthly wage received in July divided by 4.3 times the number of weekly hours in the contract, deflated by the national INPC.	Individual	Census microdata
Unemployment rate	Individuals without a job who actively looked for a new position in the previous month over the economically active population.	MCA	Census
Density	Total population divided by the area (in km ²).	MCA	Census
Area	Area in km ² .	MCA	Census
Self reported black or brown	Individuals who identified themselves as black or brown.	Individual	Census microdata
Age	Age in July of each year.	Individual	Census microdata
Education attainment	Less than 7 years of schooling, 8 to 10 years of schooling, 11 to 14 years of schooling and 15 years of schooling or more.	Individual	Census microdata
Student	Dummy variable that equals one if the individual is currently studying and zero otherwise.	Individual	Census microdata
Sector of activity	Agriculture, manufacture, construction, other industrial activities, commerce, services, auxiliary services, transport and communication, social services, other activities.	Individual	Census microdata
Occupation group	Leaders; scientific, artistic or similar; technical level; administrative service; commerce and service; agriculture, livestock, extractive activities; manufacture.	Individual	Census microdata

Source: Author's own calculations.

Table A.2 – Number of observations in each month and year of the PME panel database.

	January	February	March	April	May	June	July	August	September	October	November	December
2003			607	1,038	1,050	1,061	1,146	1,097	1,135	1,086	1,228	1,152
2004	1,185	1,104	1,296	1,228	1,280	1,206	1,309	1,152	1,204	1,175	1,336	1,299
2005	1,286	1,137	1,424	1,269	1,339	1,226	1,349	1,276	1,308	1,272	1,359	1,398
2006	1,339	1,190	1,433	1,322	1,377	1,260	1,442	1,407	1,430	1,360	1,443	1,412
2007	1,322	1,246	1,413	1,370	1,367	1,401	1,476	1,430	1,417	1,384	1,551	1,456
2008	1,348	1,289	1,461	1,468	1,326	1,497	1,479	1,396	1,422	1,377	1,529	1,514
2009	1,454	1,365	1,392	1,513	1,430	1,413	1,552	1,382	1,439	1,407	1,472	1,418
2010	1,399	1,325	1,353	1,474	1,454	1,353	1,494	1,353	1,407	1,406	1,542	1,352
2011	1,315	1,267	1,415	1,424	1,393	1,355	1,462	1,357	1,462	1,368	1,590	1,441
2012	1,293	1,285	1,423	1,381	1,401	1,335	1,422	1,314	1,487	1,337	1,487	1,447
2013	1,330	1,289	1,497	1,474	1,515	1,320	1,497	1,436	1,563	1,402	1,523	1,496
2014	1,454	1,306	1,568	1,553	1,456	1,245	828	807	824	744	801	802
2015	757	662	786	770	655	593						

Source: PME panel.

Table A.3 – Complete estimation results of Models 1, 2, 5 and 6 of Tables 3 and 4.

	1) OLS			2) FE			5) DYNAMIC OLS			6) DYNAMIC FE		
	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal
ln(unemployment rate)	-0.012 (0.016)	-0.102** (0.045)	-0.007 (0.017)	-0.001 (0.010)	0.055 (0.039)	-0.001 (0.010)	-0.008 (0.007)	-0.058** (0.023)	-0.003 (0.007)	-0.002 (0.009)	0.021 (0.035)	-0.003 (0.009)
ln(hourly wage) _{t-1}							0.875*** (0.002)	0.840*** (0.006)	0.877*** (0.002)	0.422*** (0.003)	0.447*** (0.010)	0.399*** (0.003)
Student	0.120*** (0.005)	0.077*** (0.011)	0.134*** (0.005)	0.003 (0.005)	-0.007 (0.016)	0.012** (0.005)	0.017*** (0.003)	0.018*** (0.007)	0.018*** (0.003)	0.006 (0.004)	-0.002 (0.015)	0.011** (0.005)
Age	0.075*** (0.001)	0.063*** (0.002)	0.072*** (0.001)				0.008*** (0.000)	0.009*** (0.001)	0.007*** (0.000)			
Age ²	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)				-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)			
Self-reported black or brown	-0.228*** (0.003)	-0.194*** (0.008)	-0.232*** (0.003)				-0.024*** (0.001)	-0.025*** (0.004)	-0.024*** (0.002)			
Education level (less than 1 year of schooling omitted)												
1 to 3 years of schooling	0.061*** (0.011)	0.052* (0.029)	0.063*** (0.012)				0.008 (0.005)	0.021 (0.014)	0.005 (0.006)			
4 to 7 years of schooling	0.181*** (0.010)	0.158*** (0.027)	0.182*** (0.011)				0.021*** (0.005)	0.028** (0.012)	0.019*** (0.005)			
8 to 10 years of schooling	0.327*** (0.010)	0.272*** (0.027)	0.330*** (0.011)				0.043*** (0.005)	0.040*** (0.013)	0.042*** (0.005)			
11 years of schooling or more	0.753*** (0.010)	0.618*** (0.028)	0.759*** (0.011)				0.094*** (0.005)	0.096*** (0.013)	0.092*** (0.005)			
Sector of activity (industry omitted)												
Construction	-0.061*** (0.006)	-0.135*** (0.014)	-0.009 (0.006)	-0.029*** (0.006)	-0.073*** (0.023)	-0.009 (0.007)	-0.011*** (0.002)	-0.028*** (0.007)	-0.003 (0.003)	-0.016*** (0.005)	-0.044** (0.020)	-0.004 (0.006)
Commerce and repair	-0.221*** (0.004)	-0.166*** (0.012)	-0.216*** (0.004)	-0.055*** (0.004)	-0.052*** (0.017)	-0.043*** (0.005)	-0.033*** (0.002)	-0.027*** (0.006)	-0.032*** (0.002)	-0.041*** (0.004)	-0.020 (0.015)	-0.035*** (0.004)
Finance, real estate, serv. to companies	-0.046*** (0.005)	0.130*** (0.016)	-0.059*** (0.005)	-0.022*** (0.005)	0.015 (0.023)	-0.026*** (0.005)	-0.008*** (0.002)	0.017** (0.008)	-0.009*** (0.002)	-0.017*** (0.005)	0.018 (0.020)	-0.022*** (0.005)
Other services	-0.108*** (0.004)	-0.069*** (0.013)	-0.101*** (0.005)	-0.010** (0.005)	0.010 (0.022)	-0.008 (0.006)	-0.015*** (0.002)	-0.014** (0.006)	-0.013*** (0.002)	-0.005 (0.005)	0.016 (0.020)	-0.005 (0.005)
Other activities	-0.274*** (0.018)	-0.285*** (0.034)	-0.232*** (0.021)	-0.067*** (0.017)	-0.088 (0.054)	-0.048** (0.019)	-0.036*** (0.009)	-0.052** (0.022)	-0.026** (0.010)	-0.050*** (0.015)	-0.051 (0.048)	-0.040** (0.018)
Metropolitan region	Yes	Yes	Yes				Yes	Yes	Yes			
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes				Yes	Yes	Yes			
Individual fixed effects				Yes	Yes	Yes				Yes	Yes	Yes
N	193,934	23,553	170,381	193,934	23,553	170,381	193,934	23,553	170,381	193,934	23,553	170,381
Adjusted R ²	0.355	0.386	0.335				0.858	0.832	0.856			
R ² within				0.0221	0.0255	0.0192				0.2002	0.2347	0.1793
R ² between				0.046	0.0824	0.0357				0.8116	0.7522	0.8017

Source: Author's own elaboration.

Table A.4 – Complete estimation results of Models 3, 4, 7 and 8 of Tables 3 and 4.

	3) 2SLS			4) FE-2SLS			7) 2SLS DYNAMIC			8) FE-2SLS DYNAMIC		
	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal	Total	Informal	Formal
ln(unemployment rate)	-0.012 (0.021)	-0.161*** (0.062)	-0.003 (0.023)	-0.004 (0.013)	0.037 (0.053)	0.003 (0.013)	-0.007 (0.010)	-0.078** (0.032)	0.000 (0.010)	-0.001 (0.012)	-0.005 (0.047)	0.005 (0.012)
ln(hourly wage) _{t-1}							0.875*** (0.002)	0.840*** (0.006)	0.877*** (0.002)	0.422*** (0.003)	0.447*** (0.010)	0.399*** (0.003)
Student	0.120*** (0.005)	0.077*** (0.011)	0.134*** (0.005)	0.003 (0.005)	-0.007 (0.016)	0.012** (0.005)	0.017*** (0.003)	0.018*** (0.007)	0.018*** (0.003)	0.006 (0.004)	-0.002 (0.015)	0.011** (0.005)
Age	0.075*** (0.001)	0.063*** (0.002)	0.072*** (0.001)				0.008*** (0.000)	0.009*** (0.001)	0.007*** (0.000)			
Age ²	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)				-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)			
Self-reported black or brown	-0.228*** (0.003)	-0.194*** (0.008)	-0.232*** (0.003)				-0.024*** (0.001)	-0.025*** (0.004)	-0.024*** (0.002)			
Education level (less than 1 year of schooling omitted)												
1 to 3 years of schooling	0.061*** (0.011)	0.052* (0.029)	0.063*** (0.012)				0.008 (0.005)	0.021 (0.014)	0.005 (0.006)			
4 to 7 years of schooling	0.181*** (0.010)	0.158*** (0.027)	0.182*** (0.011)				0.021*** (0.005)	0.027** (0.012)	0.019*** (0.005)			
8 to 10 years of schooling	0.327*** (0.010)	0.272*** (0.027)	0.330*** (0.011)				0.043*** (0.005)	0.040*** (0.013)	0.042*** (0.005)			
11 years of schooling or more	0.753*** (0.010)	0.617*** (0.028)	0.759*** (0.011)				0.094*** (0.005)	0.096*** (0.013)	0.092*** (0.005)			
Sector of activity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Metropolitan region	Yes	Yes	Yes				Yes	Yes	Yes			
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes				Yes	Yes	Yes			
Individual fixed effects				Yes	Yes	Yes				Yes	Yes	Yes
F test instrument (ln(unemp. rate) _{t-12})	199,736	23,828	175,604				199,735	23,817	175,601			
R ² within - 1st stage				0.766	0.762	0.765				0.766	0.762	0.765
R ² between - 1st stage				0.951	0.947	0.950				0.951	0.946	0.950
N	193,934	23,553	170,381	193,934	23,553	170,381	193,934	23,553	170,381	193,934	23,553	170,381
Adjusted R ²	0.355	0.386	0.335				0.858	0.832	0.856			
R ² within				0.022	0.026	0.019				0.200	0.235	0.179
R ² between				0.046	0.084	0.036				0.812	0.753	0.802

Source: Author's own elaboration.

Table A.5 – Complete estimation results of Table 6.

	Total			Rural			Urban		
	1a) Total	2a)	3a) Informal	1b) Total	2b) Formal	2c) Informal	3a) Total	3b) Formal	3c) Informal
ln(unemployment rate)	0.296***	-0.010	0.390***	-0.076***	-0.147***	-0.098***	0.295***	0.010	0.362***
Student	-0.786***	-0.661***	-0.726***	0.684***	-1.584***	0.369***	-0.613***	-0.583***	-0.598***
Age	0.180***	0.014***	0.312***	0.004	-0.009***	-0.043***	0.167***	0.013***	0.274***
Age ²	-0.003***	-0.001***	-0.004***	0.000***	-0.001***	0.001***	-0.003***	-0.000***	-0.004***
Tenure	0.022***	0.013***	0.020***	-0.001	-0.004***	0.008***	0.024***	0.018***	0.021***
Tenure ²	-0.000***	-0.000***	-0.000***	-0.000***	0.000	-0.000***	-0.000***	-0.001***	-0.000***
Self-reported black or brown	0.063***	0.035***	0.046***	-0.358***	0.577***	-0.188***	0.001	-0.008**	-0.007***
Education level (less than 1 year of schooling omitted)									
1 to 3 years of schooling	0.165***	-0.087***	0.458***	0.261***	-0.605***	-0.007	0.196***	-0.003	0.349***
4 to 7 years of schooling	0.287***	-0.242***	0.905***	0.493***	-1.054***	0.024**	0.333***	-0.114***	0.695***
8 to 10 years of schooling	0.543***	-0.344***	1.473***	0.558***	-1.277***	-0.070***	0.551***	-0.208***	1.140***
11 to 14 years of schooling	0.877***	-0.587***	2.080***	0.819***	-2.574***	-0.125***	0.859***	-0.404***	1.663***
15 years of schooling or more	1.003***	-0.180***	1.993***	1.764***	-2.661***	0.817***	1.063***	0.007	1.697***
Sector of activity (Other activities omitted)									
Agriculture	-0.234***	-0.267***	-0.181***	0.133***	0.096***	0.065***	-0.192***	-0.228***	-0.127***
Manufacture	-0.039***	-0.122***	-0.050***	0.158***	0.216***	0.129***	-0.032***	-0.107***	-0.041***
Construction	-0.108***	-0.156***	-0.035***	0.170***	0.093***	0.189***	-0.107***	-0.159***	-0.027***
Other industrial activities	0.075***	-0.121***	0.081***	0.046***	0.059**	-0.019	0.089***	-0.089***	0.090***
Commerce	-0.176***	-0.201***	-0.184***	0.016	0.033	0.014	-0.170***	-0.193***	-0.178***
Services	-0.187***	-0.249***	-0.129***	0.177***	0.162***	0.070***	-0.181***	-0.242***	-0.125***
Auxiliary services	-0.114***	-0.134***	-0.105***	0.200***	0.176***	0.218***	-0.111***	-0.129***	-0.101***
Transport and communication	-0.005	-0.099***	0.023***	0.100**	0.055	0.064	-0.003	-0.095***	0.023***
Social	-0.188***	-0.202***	-0.179***	0.133***	0.193***	0.054	-0.186***	-0.196***	-0.178***
Occupation group (Transport, communication, services and others omitted)									
Technical, scientific, artistic or similar	0.390***	0.296***	0.455***	-0.121**	-0.104	-0.189***	0.382***	0.287***	0.438***
Administrative	0.182***	0.111***	0.203***	-0.540***	-0.388***	-0.462***	0.176***	0.109***	0.194***
Agriculture, livestock, extractive activities	-0.156***	-0.181***	-0.046***	-0.343***	-0.289***	-0.340***	-0.155***	-0.202***	-0.061***
Manufacture	0.040***	-0.025***	0.081***	-0.367***	-0.309***	-0.344***	0.041***	-0.024***	0.081***
Commerce and auxiliary activities	0.040***	-0.033***	0.097***	-0.349***	-0.233***	-0.393***	0.041***	-0.033***	0.094***
Head of the household	0.286***	-0.025***	0.461***	0.011	-0.321***	-0.208***	0.248***	-0.033***	0.386***
People who have children under 15	0.036***	0.027***	0.023***	-0.152***	0.284***	-0.076***	0.015***	-0.001	0.011***
Lambda	2.619***	2.363***	3.233***	-2.185***	4.823***	-1.408***	2.161***	2.273***	2.721***
State	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R ²	0.586	0.500	0.533	0.451	0.321	0.443	0.572	0.493	0.541
N (unweighted)	560,543	198,023	362,520	58,774	37,626	21,148	501,769	160,397	341,372

Source: Author's own elaboration.

Table A.6 – First stage estimation results of Table 7.

	Total (rural and urban)			Rural			Urban		
	Total	Formal	Informal	Total	Formal	Informal	Total	Formal	Informal
Student	-0.435***	-0.303***	-0.337***	-0.500***	-0.415***	-0.361***	-0.421***	-0.273***	-0.337***
Age	0.068***	-0.017***	0.110***	0.022***	-0.011***	0.066***	0.072***	-0.022***	0.112***
Age ²	-0.001***	0.000**	-0.002***	-0.001***	-0.000***	-0.001***	-0.001***	0.000***	-0.002***
Self-reported black or brown	0.075***	0.046***	0.055***	0.213***	0.169***	0.134***	0.049***	0.027***	0.036***
Education level (less than 1 year of schooling omitted)									
1 to 3 years of schooling	0.042***	-0.083***	0.160***	-0.119***	-0.182***	0.091***	0.088***	-0.034***	0.152***
4 to 7 years of schooling	0.045***	-0.215***	0.295***	-0.220***	-0.322***	0.127***	0.102***	-0.151***	0.266***
8 to 10 years of schooling	0.127***	-0.317***	0.491***	-0.205***	-0.399***	0.280***	0.173***	-0.256***	0.441***
11 to 14 years of schooling	0.220***	-0.523***	0.679***	-0.291***	-0.760***	0.427***	0.266***	-0.444***	0.624***
15 years of schooling or more	-0.048***	-0.588***	0.358***	-0.607***	-0.859***	-0.016	0.003	-0.502***	0.317***
Head of the household	0.060***	-0.123***	0.141***	0.095***	-0.127***	0.308***	0.047***	-0.135***	0.125***
People who have children under 15	-0.035***	-0.049***	-0.013***	0.107***	0.058***	0.104***	-0.064***	-0.069***	-0.038***
State	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	0.062	0.054	0.104	0.072	0.064	0.132	0.063	0.051	0.092
N (unweighted)	1,493,518	1,493,518	1,493,518	172,320	172,320	172,320	1,303,187	1,303,187	1,303,187

Source: Author's own elaboration.

Table A.7 – First stage estimation results of Table 11.

	Total	Formal	Informal
Student	0.016*** (0.003)	0.065*** (0.004)	-0.073*** (0.006)
Age	0.055*** (0.001)	0.049*** (0.001)	0.051*** (0.001)
Age ²	-0.001*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)
Self-reported black or brown	-0.100*** (0.002)	-0.105*** (0.002)	-0.084*** (0.004)
Education level (less than 7 years of schooling omitted)			
8 to 10 years of schooling	0.187*** (0.002)	0.174*** (0.003)	0.162*** (0.005)
11 to 14 years of schooling	0.395*** (0.003)	0.371*** (0.003)	0.370*** (0.006)
15 years of schooling or more	0.990*** (0.006)	0.974*** (0.007)	0.919*** (0.016)
Head of the household	0.167*** (0.002)	0.152*** (0.002)	0.167*** (0.004)
Occupational group	Yes	Yes	Yes
Sector of activity	Yes	Yes	Yes
MCA-year dummy	Yes	Yes	Yes
Adjusted R ²	0.572	0.547	0.490
N (unweighted)	459,441	300,280	159,161

Source: Author's own elaboration.

References

- AGÉNOR, P-R. The analytics of segmented labor markets. In: AGÉNOR, P-R.; IZQUIERDO, A.; JENSEN, H.T. (Eds.) **Adjustment Policies, Poverty, and Unemployment: the IMMPA framework**. Malden: Blackwell Publishing, 1st Edition, p. 8-66, 2007.
- AMADEO, E.J.; CAMARGO, J.M. Chapter 7: Brazil: regulation and flexibility in the labor market. In: EDWARD, S.; LUSTIG, N.C. (Eds.) **Labor Markets in Latin America: Combining social protection with market flexibility**. Washington: Brookings Institution Press, p. 201-234, 1997.
- BALTAGI, B.H.; BASKAYA, Y.S.; HULAGU, T. How different are the wage curves for formal and informal workers? Evidence from Turkey. **Papers in Regional Science**, Vol. 92, N. 2, p. 271-284, 2013.
- BALTAGI, B.H.; BLIEN, U.; WOLF, K. The East Germany wage curve 1993-1998, **Economic Letters**, Vol. 69, p. 25-31, 2000.
- BALTAGI, B.H.; BLIEN, U.; WOLF, K. New evidence on the dynamic wage curve for Western Germany: 1980-2004. **Labour Economics**, Vol. 16, p. 47-51, 2009.
- BALTAGI, B.H.; BLIEN, U.; WOLF, K. A dynamic spatial panel data approach to the German wage curve. **Economic Modelling**. Vol. 29, p. 12-21, 2012.
- BALTAGI, B.H.; ROKICKI, B. The spatial Polish wage curve with gender effects: evidence from the Polish Labor Survey. **Regional Science and Urban Economics**, Vol. 49, p. 39-57, 2014b.
- BALTAGI, B.H.; ROKICKI, B.; SOUZA, K.B. The Brazilian wage curve: new evidence from the National Household Survey. **IZA Discussion Paper** N. 8468, 2014a.
- BARROS, R.P.; MENDONÇA, R.S.P. Flexibilidade do mercado de trabalho brasileiro: uma avaliação empírica. **Instituto de Pesquisa Econômica e Aplicada**, Texto para Discussão 452, 1997.
- BARTH, E.; BRATSBERG, B.; NAYLOR, R.A.; RAAUM, O. Explaining variations in the wage curve. Memorandum 2002/03, **University of Oslo**, 2002.

- BELL, B., NICKELL, S., QUINTINI, G. Wage equations, wage curves and all that. **Labour Economics**, Vol. 9, p. 341-360, 2002.
- BLANCHARD, O.; KATZ, L. What do we know and do not know about the natural rate of unemployment. **Journal of Economic Perspectives**, Vol. 11, N. 1, 1997.
- BLANCHFLOWER, D.G.; OSWALD, A.J. **The Wage Curve**. MIT Press, Cambridge, 1994
- BLANCHFLOWER, D.G.; OSWALD, A.J. An introduction to the wage curve. **The Journal of Economic Perspectives**, Vol. 9, N. 3, p. 153-167, 1995
- BLANCHFLOWER, D.G.; OSWALD, A.J. The wage curve reloaded. **NBER Working Paper Series**, Working Paper 11338, 2005.
- BLIEN, U.; DAUTH, W.; SCHANK, T.; SCHNABEL, C. The institutional context of an ‘Empirical Law’: the wage curve under different regimes of collective bargaining. **British Journal of Industrial Relations**, Vol. 51, N. 1, p. 59-79, 2013.
- BOOTH, A. Wage determination and imperfect competition. **IZA Discussion Paper**, N. 8034, 2014.
- BOTELHO, F.; PONCZEK, V. Segmentation in the Brazilian labor market. **Economic Development & Cultural Changes**, Vol. 59, N. 2, p. 438-463, 2011
- BUETTNER, T. The Effect of Unemployment, Aggregate Wages, and Spatial Contiguity on Local Wages: An Investigation with German District Level Data, **Papers in Regional Science**, Vol. 78, p. 47–67, 1999.
- CAMPBELL, C.; ORSZAG, J. M. A model of the wage curve, **Economics Letters**, Vol. 59, p. 119-125, 1998.
- CARD, D. The wage curve: a review. **Journal of Economic Literature**, Vol. 33, N. 2, p. 785-799, 1995.
- CARNEIRO, F.G.; HENLEY, A. Wage determination in Brazil: the growth of union bargaining power and informal employment. **The Journal of Development Studies**, Vol. 34, N. 4, p. 117-138, 1998.
- CORSEUIL, C.H.L.; FOGUEL, M.N. Economic expansion and an increase in labour market formality: a poaching approach. **Revista Brasileira de Economia**, Vol. 66, N. 2, p. 207-224, 2012.
- EL BADAOU, E.; STROBL, E.; WALSH, F. The formal sector wage premium and firm size. **Journal of Development Economics**. Vol. 91, p. 37-47, 2010.
- ELHORST, J.P. **Spatial Econometrics: From cross-sectional data to spatial panels**. Springer Briefs in Regional Science, New York: Springer, 1st Edition, 2014.
- ELHORST, J.P.; BLIEN, U.; WOLF, K. New evidence on the wage curve: a spatial panel approach. **International Regional Science Review**. Vol. 30, N. 2, p. 173-191, 2007.
- ESTEVÃO, M.; CARVALHO-FILHO, I. Institutions, informality, and wage flexibility: evidence from Brazil. **IMF Working Paper**, WP 12/84, 2012.
- FIELDS, G.S. Labour market modelling and the urban informal sector: Theory and evidence [Electronic version]. In: TURNHAM, D.; SALOMÉ, B.; SCHWARZ, A. (Eds.), **The informal sector revisited**. Paris: Organisation for Economic Co-operation and Development, p. 49-69, 1990.
- FIELDS, G.S. A guide to multisector labor market models. **Social Protection Discussion Paper Series**, N. 505, 2005.
- FINGLETON, B.; PALOMBI, S. The wage curve reconsidered: is it truly an ‘empirical law of economics’? **Région et Développement**, N. 38, p. 49-92, 2013.
- GLAESER, E.L. **Triumph of the City: How our greatest invention makes us richer, smarter, greener, healthier, and happier**. London, UK: Macmillan, 2011.
- GORDON, I.R.; MCCANN, P. Industrial clusters: complexes, agglomeration and/or social networks? **Urban Studies**, Vol. 37, p. 513-532, 2000.
- HENLEY, A.; REZA ARABSHEIBANI, G.; CARNEIRO, F.G. On defining and measuring the informal sector: evidence from Brazil. **World Development**, Vol. 37, N. 5, p. 992-1003, 2009.
- LONGHI, S.; NIJKAMP, P.; POOT, J. Spatial heterogeneity and the wage curve revisited. **Journal of Regional Science**, Vol. 46, N. 4, p. 707-731, 2006.
- MALONEY, W.F. Informality revisited. **World Development**, Vol. 32, N. 7, p. 1159-1178, 2004.
- MANNING, A. The real thin theory: monopsony in modern labour markets. **Labour Economics**, Vol. 10, p. 105-131, 2003.

- MANNING, A. Imperfect competition in the labor market. In: CARD, D. ASHENFELTER, O. **Handbook of Labor Economics Vol. 4B**, Amsterdam: North Holland, 1st Edition, p. 972-1041, 2011.
- MARSHALL, A. **The Principles of Economics**. London: Macmillan, 1890.
- MEGHIR, C.; WHITEHOUSE, E. The evolution of wages in the United Kingdom: evidence from micro-data. **Journal of Labor Economics**, Vol. 14, N. 1, p. 1-25, 1996.
- MENEZES-FILHO, N.A.; MENDES, M.; ALMEIDA, E.S. O diferencial de salários formal-informal no Brasil: segmentação ou viés de seleção? **Revista Brasileira de Economia**, Vol. 58, N. 2, p. 235-248, 2004.
- MORRISON, P.S.; PAPPS, K.L.; POOT, J. Wages, employment, labour turnover and the accessibility of local labour markets. **Labour Economics**, Vol. 13, p. 639-663, 2006.
- MORRISON, P.S.; POOT, J. The wage curve and local labour markets in New Zealand. In: MORRISON, P.S. (Ed.), **Labour, Employment and Work in New Zealand**. Victoria University of Wellington, Wellington, 1998.
- MOULTON, B.R. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. **The Review of Economics and Statistics**. Vol. 72, N. 2, p. 334-338, 1990.
- NIJKAMP, P.; POOT, J. The last word on the wage curve? **Journal of Economic Surveys**, Vol. 19, N. 3, p. 421-450, 2005.
- PERO, V.L. A carteira de trabalho no mercado de trabalho metropolitano brasileiro. **Pesquisa e Planejamento Econômico**. Vol. 22, N. 2, p.305-342, 1992.
- RAMOS, R.; DUQUE, J.C.; SURINACH, J. Is the wage curve formal or informal? Evidence for Colombia. **Economics Letters**, Vol. 109, p. 63-65, 2010.
- RAMOS, R.; NICODEMO, C.; SANROMÁ, E. A spatial panel wage curve for Spain. **Letters in Spatial and Resource Sciences**, Vol. 8, N. 2, p. 125-139, 2015.
- REIS, M.C. Os impactos das mudanças na demanda por trabalho qualificado sobre o desemprego por nível de qualificação durante os anos 90 no Brasil. **Revista Brasileira de Economia**, Vol. 60, N. 3, p. 297-319, 2006.
- RIBAS, R.P.; SOARES, S.S.D. O atrito nas pesquisas longitudinais: o caso da Pesquisa Mensal de Emprego (PME/IBGE). **Estudos Econômicos**, Vol. 40, N. 1, p. 213-244, 2010.
- SANTOLIN, R.; ANTIGO, M.F. Curvas de salários dinâmicas: um estudo dos determinantes da histerese do desemprego no Brasil. **Cedeplar / FACE / UFMG**, Texto para Discussão 368, 2009.
- SATO, Y. Search theory and the wage curve. **Economics Letters**, Vol. 66, p. 93-98, 2000.
- SILVA, A.S.P.; MONSUETO, S.E. Flexibilidade do mercado de trabalho: uma análise comparativa entre segmentos socioeconômicos no Brasil (2002-2009). **Proceeding of the 40th Brazilian Economics Meeting, ANPEC – Brazilian Association of Graduate Programs in Economics**, accessed in 15/10/2014, at http://www.anpec.org.br/encontro/2012/inscricao/files_I/i12-351de703e4da2073e562bebd5fd0981.pdf , 2012.
- SOARES, F.V. Some stylized facts of the informal sector in Brazil. **IPEA, Texto para Discussão**, N. 1020, 2004.
- SOUZA, I. V.; MACHADO, A.F. Curva de rendimentos: uma análise no mercado de trabalho urbano e rural no Brasil (1981/99). **Revista de Economia e Sociologia Rural**, Vol. 42, N. 1, p. 35-54, 2004.
- TURUNEN, J. Disaggregated wage curves in the United States: evidence from panel data of young workers, **Applied Economics**, Vol. 30, N. 12, p. 1665-1677, 1998.
- ULYSSEA, G. Informalidade no mercado de trabalho brasileiro: uma resenha da literatura. **Revista de Economia Política**, Vol. 26, N. 4, p. 596-618, 2006.
- ULYSSEA, G. Regulation of entry, labor market institutions and the informal sector. **Journal of Development Economics**, Vol. 91, p. 87-99, 2010.
- WOOLDRIDGE, J.M. **Econometric Analysis of Cross Section and Panel Data**. England: MIT Press, 1st Edition, 2002.