Department of Economics - FEA/USP

New evidence on the wage curve: non-linearities, urban size, and spatial scale in Brazil

Ana Maria Bonomi Barufi
Eduardo A. Haddad
Peter Nijkamp

Working Paper Series Nº 2016-39

DEPARTMENT OF ECONOMICS, FEA-USP
WORKING PAPER Nº 2016-39

New evidence on the wage curve: non-linearities, urban size, and spatial scale in Brazil

Ana Maria Bonomi Barufi (anambb@usp.br)

Eduardo A. Haddad (ehaddad@usp.br)

Peter Nijkamp

Research Group: NEREUS – The University of São Paulo Regional and Urban Economics Lab

Abstract:

Agglomeration economies appear to have a significant 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 plays an important role in the relative bargaining power of workers and firms in the relevant labour market. When analysing the relationship between local wages and the business cycle – wage flexibility, measured by the wage curve –, this influence appears to be higher in informal sectors in less densely populated areas in Brazil. Therefore, large agglomerations are supposed to provide a higher bargaining power for workers, as they have more job opportunities. In addition, labour market dualism is an essential ingredient in the evaluation of the wage curve in developing economies. However, a dual labour market analysis should be conducted at the appropriate 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. Our study aims to shed new theoretical and empirical light on the importance of the wage curve, taking into account various specificities of developing economies. The applied modelling study in Brazil shows that wage flexibility is higher in less dense local labour markets and in the informal sector in relation to the formal sector. Furthermore, it is essential to control for unobserved local characteristics in order to obtain the 'true' elasticity of wages to local unemployment rates, and spatial effects should be accounted for when the unit of analysis is rather small. In this sense, a significant part of the difference between the formal and the informal sectors originates from spatial effects.

Keywords: wage curve; informal sector; bargaining power; agglomeration; rural-urban dichotomy.

JEL Codes: R12; J31; J46.

New evidence on the wage curve: non-linearities, urban size, and spatial scale in Brazil

Ana Maria Bonomi Barufi[* ¥], Eduardo Amaral Haddad[*], Peter Nijkamp[¥ §]

[*] Department of Economics, University of São Paulo, Brazil

[¥] Department of Spatial Economics, VU University Amsterdam

[§] A. Mickiewicz University, Poznán

Resumo

Economias de aglomeração parecem ter um impacto significativo sobre mercados de trabalho locais. A interação de firmas e trabalhadores em áreas urbanas densas pode gerar vantagens de produtividade que resultam em salários mais elevados. O tamanho da cidade tem um papel importante no poder de barganha de trabalhadores e firmas no mercado de trabalho relevante. Ao analisar a relação entre os salários locais e o ciclo de negócios – flexibilidade de salários, medida pela curva de salários –, esta influência parece ser maior no setor informal de áreas menos densas do Brasil. Assim, aglomerações maiores parecem proporcionar um poder de barganha maior para os trabalhadores, já que nelas eles têm acesso a mais oportunidades de trabalho. Adicionalmente, o dualismo no mercado de trabalho é um ingrediente essencial da análise da curva de salários em países em desenvolvimento. Entretanto, uma análise dual do mercado de trabalho deve ser conduzida no nível regional apropriado (mercados de trabalho locais), tornando possível encontrar impactos relevantes do tamanho da cidade no poder de barganha relativo de firmas e trabalhadores. Nosso estudo busca destacar uma nova perspectiva teórica e empírica para a curva de salários, levando em conta as especificidades de países em desenvolvimento. A análise aplicada ao Brasil mostra que a flexibilidade de salários é maior em áreas urbanas menos densas e mais elevada no setor informal quando comparado ao formal. Adicionalmente, é preciso controlar para características locais não observadas para obter a 'verdadeira' elasticidade de salários em relação às taxas de desemprego locais, e efeitos espaciais devem ser considerado quando as áreas de análise são suficientemente pequenas. Nesse caso, uma parte importante da diferença entre os setores formais e informais é originada pelos efeitos espaciais.

Palavras-chave: curva de salários, setor informal, poder de barganha, aglomeração, dicotomia rural-urbana.

Abstract

Agglomeration economies appear to have a significant 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 plays an important role in the relative bargaining power of workers and firms in the relevant labour market. When analysing the relationship between local wages and the business cycle – wage flexibility, measured by the wage curve –, this influence appears to be higher in informal sectors in less densely populated areas in Brazil. Therefore, large agglomerations are supposed to provide a higher bargaining power for workers, as they have more job opportunities. In addition, labour market dualism is an essential ingredient in the evaluation of the wage curve in developing economies. However, a dual labour market analysis should be conducted at the appropriate 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. Our study aims to shed new theoretical and empirical light on the importance of the wage curve, taking into account various specificities of developing economies. The applied modelling study in Brazil shows that wage flexibility is higher in less dense local labour markets and in the informal sector in relation to the formal sector. Furthermore, it is essential to control for unobserved local characteristics in order to obtain the 'true' elasticity of wages to local unemployment rates, and spatial effects should be accounted for when the unit of analysis is rather small. In this sense, a significant part of the difference between the formal and the informal sectors originates from spatial effects.

Keywords: wage curve, informal sector, bargaining power, agglomeration, rural-urban dichotomy

JEL codes: R12, J31, J46

1 Introduction

The wage curve represents an empirical negative relation between wages and local unemployment rates at the regional level (Blanchflower and Oswald, 1994). This result is expected, as wage flexibility plays an important role in the adjustment process of local economies to adverse shocks. Moreover, causality is expected to flow from unemployment size to the level of wages (Blanchflower and Oswald, 1995). Many empirical studies have corroborated the original formulation of the wage curve by finding negative and significant relationships in different socio-economic environments. Within the context of a developing country, with an abundant informal sector and a significant heterogeneity in terms of population density, the wage curve can be an expression of the degree of flexibility in the labour market. City size is one additional factor in this equation, as it increases the number of opportunities for a matching between firms and workers, and it can imply that firms and workers may have a more or less precautionary behaviour in their interactions in the labour market.

Our study provides a comparison of the formal and the informal sectors in Brazil, and how the intensity of agglomeration factors affects them. Different aspects of the wage curve in Brazil are investigated, with various databases, regional aggregation levels, time periods, and data structures. Finally, an analytical framework based on different bargaining power levels for workers in the formal and the informal sector is developed, with a direct implication for the effect of urban size on the relative bargaining power of workers and firms.

2 The wage curve

An empirical relation between the level of wages and the unemployment rate in different regions is represented by the wage curve (Blanchflower and Oswald, 1995, 2005). A theoretical challenge associated with the wage curve is its robust pattern across countries and over time periods, despite the significant institutional differences among them. The wage curve can be interpreted as a measure of the flexibility of the labour market and several mechanisms have been proposed to explain it, based on imperfect competitive structures: bargaining power, efficiency wage, labour contracts, and turnover costs. Some of these mechanisms have not necessarily become structured theoretical models, and most importantly, they have not been empirically validated (Nijkamp and 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 et al. (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 normally observed by comparing local labour markets. Spatial units and urban size can be related in two ways: neighbouring areas can increase accessibility to job offers; and 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 et al., 2006), with a positive effect on the local wage level. In addition, the unemployment elasticity to pay is expected to be negatively affected by a higher level of accessibility. Then, agglomeration economies may increase local productivity, leading to higher wages and lower unemployment rates (Sato, 2000), and they may be essential to determine the location of different local labour markets in the wage curve. Local labour markets can be 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 et al., 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. Such effects can be heterogeneous depending on the degree of formalisation. 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 empirical literature has focused on the analysis of this relationship by comparing countries or regions, controlling for individual characteristics and other relevant microeconomic aspects. Nijkamp and Poot (2005) provided a meta-analysis of the coefficients estimated in more than 200 empirical wage curve 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 have made a stronger effort to control for worker heterogeneity based on observed and unobserved characteristics, (Baltagi et al., 2009, 2012, for Germany, and Turunen,1998, for the United States).

The inclusion of agglomeration effects and an explicit spatial dimension was based on recent advances in spatial econometrics techniques (Baltagi et al., 2012; Baltagi and Rokicki, 2014b; Elhorst et al., 2007; Longhi et al., 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, local unemployment elasticity is likely to be overestimated (Baltagi and Rokicki, 2014b). Longhi, et al. (2006) noted that the unemployment elasticity is lower in regions that interact more with others, meaning that more isolated areas have a less flexible labour market. Similar studies were conducted by Ramos et al. (2015) for Spain, Fingleton and Palombi (2013) in the UK, and Baltagi and Rokicki (2014b) for Poland, among others.

In developing countries, usually a large share of the labour market is unregulated, based on informal relations and does not follow the relevant legislation concerned. This has important implications for the bargaining relationships in the labour market, as the informal sector is distinct from formal jobs and unemployment. Economic shocks may be absorbed in this intermediate sector, with wages varying significantly given a change in the local unemployment rate. In Colombia, Ramos et al. (2010) found that the elasticity of individual wages to the unemployment rate is much higher in the case of informal workers, a similar result as the one found by Baltagi et al. (2013) for Turkey. When controlling for the potential endogeneity of unemployment (FE-2SLS with the lagged value of the unemployment rate as an instrument), Baltagi et al. (2013) 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 et al. (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 vanish. 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 from the perspective of sector dualism should control for it. With data from PME-IBGE (Monthly Employment Survey) on a more recent period (2003-2008), Corseuil and Foguel (2012) estimated a model for the Brazilian labour market with a similar structure as 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. In an economic expansion, they find that the formality rate goes up and the wage gap between formal and informal workers increases.

Wage curve estimation in Brazil has been focusing on the level of wage flexibility in the formal and the informal sectors, and 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 PME-IBGE, from 1982 to 1994. Their conclusion was that the degree of wage flexibility in Brazil was rather 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 a lower wage flexibility for more qualified workers, a similar result as the one obtained by Reis (2006) – who considered state-level data from 1990 to 1999 from 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. 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 only for 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 et al. (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). They also compared workers in the formal and the informal sectors with similar probabilities of being formal (below the median), and the elasticity of unemployment amounted to -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 empirical evidence 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, a formalisation of the labour market and an increase of the educational level of the labour force. Clearly, there is room to explore further aspects of the wage curve, in particular, spatial dependence and agglomeration economies, the control of individual unobserved heterogeneity, and for a possible selection bias of the participation in the labour market.

4 Labour market dualism in Brazil

There are pronounced differences in the way individuals interact in the labour market of developing economies according to sector, location, firm size, educational level, and 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, amongst others: 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). The model discussed in the following section aims to represent the main differences between the formal and the informal sectors 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

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, them as keys to explain the wage curve).

In our study, the model developed by Barth et al. (2002) is adapted for 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 encompasses endogenous turnover costs (Campbell and Orszag, 1998), depicted in (1), which are dependent on the probability that the worker quits (negatively related to the unemployment rate). 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):

$$C_r = H_r q_r \tag{1}$$

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 $\gamma'' > 0$.

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

Firm's profits are given by $\pi_r = R_r(L_r) - C_rL_r - w_rL_r$. In this equation, $R_r(L_r)$ is the revenue received by firms, and C_r is the replacement cost defined in (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:

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

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

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

where:

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

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:

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

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:

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

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 (9) with respect to β_r generates the following relationship:

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

(-) (+) (+) (+)

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 addition, the unemployment rate in neighbouring areas is inversely related to 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 et al., 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 et al., 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.

5 Empirical research strategy

5.1 Model development

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:

$$w_{r,t} = \gamma_1 u_{r,t} + \sum_k X_{r,t,k}\beta_k + \theta_r + \delta_t + \epsilon_{r,t}, \qquad r = \{1, \dots, R\}, t = \{1, \dots, T\} \quad (11)$$

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.

$$w_{i,r,t} = \gamma_1 u_{r,t} + \sum_k X_{i,r,t,k}\beta_k + \theta_r + \delta_t + \varepsilon_{i,r,t}, \quad i = \{1, \dots, N\}, r = \{1, \dots, R\}, t = \{1, \dots, T\} \quad (12)$$

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, age^2, 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).

5.2 Extensions of the analytical framework

5.2.1 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 et al., 2000; Blanchflower and Oswald, 1994; Santolin and Antigo, 2009). Another possibility is to estimate the wage curve in two stages (Bell et al., 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).

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

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.

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

The parameter of interest will be γ_2, while μ_r and δ_t are region fixed effects and time dummies.

5.2.2 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.

5.2.3 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 (12), rewriting it in a simplified version:

$$w = \gamma_1 u + X\beta + \varepsilon \tag{15}$$

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:

$$L = 1[Z\delta + \xi > 0] \tag{16}$$

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:

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

where $\lambda(\cdot) \equiv {}^{\phi(\cdot)}\!/_{\Phi(\cdot)}$ 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.

5.2.4 Spatial dependence and neighbouring effects

The unemployment rate in neighbouring areas could affect wages directly (with a negative relationship) or indirectly, by increasing local wage flexibility. In the literature, such effect was estimated including the spatial lag of the unemployment rate (Longhi et al., 2006) or spatial autocorrelation in the error and in the dependent variable (Baltagi et al., 2012; Fingleton and Palombi, 2013; Ramos et al., 2015; Elhorst et al., 2007; Buettner, 1999). Departing from (14), 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:

$$\hat{\lambda}_{r,t} = \mu_r + \delta_t + \rho W \hat{\lambda}_{r,t} + \gamma_2 u_{r,t} + \nu_{r,t}\ , \tag{18}$$

while the spatial error model is written as:

$$\hat{\lambda}_{r,t} = \mu_r + \delta_t + \gamma_2 u_{r,t} + \epsilon_{r,t} \tag{19}$$

$$\epsilon_{r,t} = \varphi W \epsilon_{r,t} + \nu_{r,t} \tag{20}$$

where W is the spatial weights matrix based on inverse distances and common local labour market areas (see Subsection 6.4.1 for further details), constant in time. The error terms $\nu_{r,t}$ are supposed to be independently and identically distributed for all r, with zero mean and variance σ^2.[1]

5.2.5 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 et al., 2009, 2012). The Phillips curve can be an alternative explanation to the relationship between wages and the unemployment rate (Blanchard and Katz,

[1] The estimation of these models follow Elhorst (2014), based on MATLAB routines available at http://www.regroningen.nl/elhorst/software.shtml and accessed on September 10, 2015.

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 et al., 2002):

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

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: the impact of past regional changes is prolonged over more than one time period, and the long run elasticity of unemployment is given by $^{\beta}/_{\lambda}$.

6 Empirical analysis

6.1 Introduction

The empirical analysis of the wage curve in Brazil is clearly 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.

6.2 Wage curve in a more aggregated context: a comparison between the formal and the informal sectors

Here we aim to measure wage flexibility in a more aggregated context, with monthly data, controlling for individual heterogeneity.

6.2.1 Description of the database – Monthly Employment Survey

This survey is conducted by the Brazilian Institute of Geography and Statistics (IBGE), and here we consider data from March 2003 to June 2015. The moving panel available in the database is analysed based on the methodology provided by Ribas and Soares (2010)[2]. 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. Formality status was defined by the existence or not of a contract under the CLT regulation, and individuals with wages equal to zero or with missing information, or with less than 20 or more than 60 working hours per week were excluded. Only individuals observed in the second and in the sixth interviews (twelve months apart from each other) were selected. 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 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 – INPC – in each metropolitan region.

6.2.2 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. 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. Considering the composition of the database over time[3], 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 sector[4] rose from 84.1% to 93.7%. Jobs in the industrial sector lost importance, while the share of workers in the construction and the financial sectors increased over time.

---

[2] The Data Zoom website (http://www.econ.puc-rio.br/datazoom/english/) provides packages to analyze the PME panel.
[3] Tables with descriptive analysis can be provided by the authors.
[4] 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.

6.2.3 Results and preliminary conclusions

The first set of results attempted to control for the issues discussed in Subsections 5.2.2 and 5.2.5 (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. 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 at least over a year were dropped (education level, age, month, among others).

The results presented in Table 1 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 1 – Elasticity of wages in relation to the unemployment rate for different estimation strategies and sectors.

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

Controls: (i) Cross-sectional models: age, age^2, 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. All models were estimated with *pweight* in Stata. Standard errors are presented in parentheses. Significance levels:* $p<0.10$, ** $p<0.05$, *** $p<0.01$. Complete estimation results can be requested to the authors.
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 5.2.5, the results presented in Table 2 indicated that when only observed characteristics of the individuals were taken into account, the coefficient of the lagged wage was close to one (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 requested to the authors.

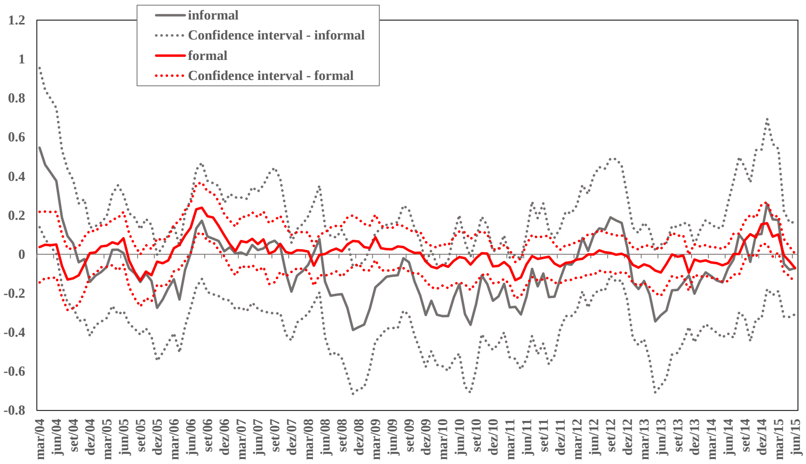
Table 2 – Elasticity of wages in relation to their time lags for different estimation strategies and sectors.

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

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

One interesting extension for the models presented in Table 1 is to check how estimations may vary over time. Figure 1 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.

Figure 1 – 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.

The coefficients for the informal sector seem to match the main results found in the literature for developed economies. However, 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. All models including individual fixed effects will compare them 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. Finally, unemployment rates were observed only for six regions, meaning that this variable does not present very much variation between individuals.

6.3 The rural-urban dichotomy 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.

6.3.1 Description of the database – National Household Sample Survey

PNAD data was analysed for the period from 1996 to 2013, after the macroeconomic normalisation and the implementation of the Real as the currency of the country. The survey was not conducted in 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. Furthermore, rural areas of the states of Rondônia, Acre, Amazonas, Roraima, Pará and Amapá were excluded from the analysis, as they were not surveyed before 2004.

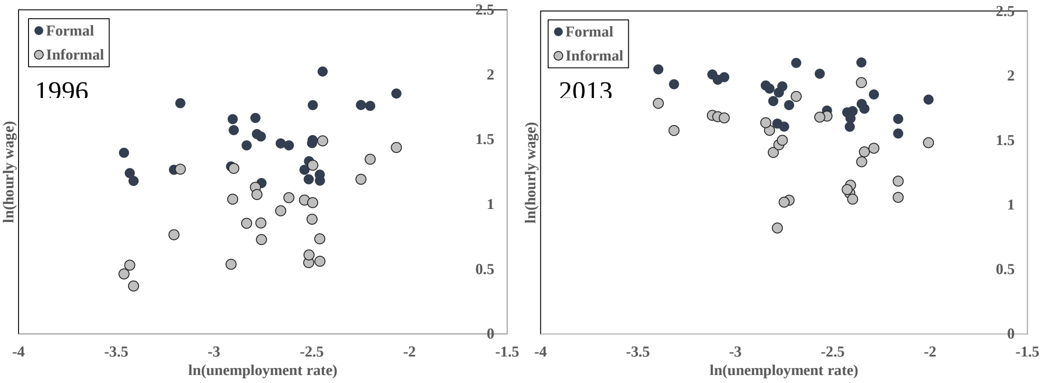
Only male individuals aged 15 to 59 years old were kept in the database. The population of interest comprised employees with positive wages working in the private sector with or without a contract under the CLT regulation. The unemployment rate was calculated for each state considering all economically 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.

6.3.2 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. Regarding the descriptive statistics of the database[5], 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 in the formal sector. In the informal sector, agriculture, livestock and extractive occupations showed a much higher participation in 2013.

Figure 2 – 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.

Finally, salaries and the unemployment rate should be compared in each state over time. It is possible to note from Figure 2 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

[5] Tables with descriptive analysis can be provided by the authors.

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.

6.3.3 Results

The first set of results of the wage curve estimated with PNAD data was based on (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 5.2.1. Furthermore, as discussed in Subsection 6.3.1, 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 3, 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.

Table 3 – 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.

	1996 to 2013					
	OLS			IV		
Models	**Total**	**Informal**	**Formal**	**Total**	**Informal**	**Formal**
			Total			
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
ln(unemployment rate)	0.198***	0.052***	0.174***	0.266***	0.025	0.249***
	(0.006)	(0.010)	(0.007)	(0.010)	(0.017)	(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	-0.049**	-0.013	-0.051**	-0.134***	-0.054*
	(0.015)	(0.021)	(0.019)	(0.025)	(0.037)	(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***	0.062***	0.182***	0.274***	0.034*	0.266***
	(0.006)	(0.012)	(0.007)	(0.011)	(0.020)	(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: age, age^2, tenure, $tenure^2$, 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 can be requested to the authors. All models were estimated with *pweight* in Stata. Significance levels:* p<0.10, ** p<0.05, *** p<0.01.
Source: Author's own calculations.

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.

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 6.2.

Table 4 – 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.

	2002 to 2009					
	OLS			IV		
Models	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***	0.040**	0.098***	0.106	-0.680***	0.232***
	(0.011)	(0.020)	(0.013)	(0.077)	(0.182)	(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	0.038	-0.021	-1.500***	-4.661***	-0.499*
	(0.030)	(0.040)	(0.036)	(0.473)	(1.781)	(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***	0.039*	0.095***	0.185**	-0.447**	0.253***
	(0.012)	(0.022)	(0.013)	(0.079)	(0.178)	(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: age, age^2, tenure, tenure^2, 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. 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 authors.
Source: Author's own calculations.

Table 5 – 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.

	1996 to 2013			2002 to 2009		
Models	Total	Informal	Formal	Total	Informal	Formal
	Total					
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a
ln(unemployment rate)	0.296***	-0.010	0.390***	0.227***	-0.763***	0.547***
	(0.011)	(0.019)	(0.013)	(0.087)	(0.205)	(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***	-0.147***	-0.098***	-1.738***	-5.258*	-0.419
	(0.029)	(0.041)	(0.035)	(0.660)	(2.869)	(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***	0.010	0.362***	0.241***	-0.567***	0.466***
	(0.012)	(0.022)	(0.013)	(0.087)	(0.196)	(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, age^2, tenure, tenure^2, 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. Standard errors are presented in parentheses. Significance levels:* p<0.10, ** p<0.05, *** p<0.01. The complete estimation for 1996-2013 can be requested to the authors.
Source: Author's own calculations.

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. 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. Moreover, 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 5 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).

6.4 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 5.2.1) 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.

6.4.1 Description of the database – Demographic Census

The analysis that follows was based on information obtained from the demographic censuses of 1991, 2000 and 2010. The analysis is done over Immediate Region of Urban Articulation (REGIC areas), that constitutes a labour market area (IBGE, 2007). Their configuration was harmonised with the 1991-2010 MCAs, obtaining 478 REGIC areas.[6] Once again, only male individuals aged 15 to 59 years old, employed with or without a formal contract, working from 20 to 60 hours per week were kept. Furthermore, individuals without information about the location of their jobs were excluded. 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 5.2.1), 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 neighbours[7].

6.4.2 Descriptive statistics

The main descriptive statistics of the Census database show that 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[8]. 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. Then, 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.

[8] Tables with descriptive analysis can be provided by the authors.

Table 6 – Descriptive statistics of the second stage – spatial wage.

	Spatial wage - total	Spatial wage - formal	Spatial wage - informal	ln(unemployment rate)
2000				
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				
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 - total	Spatial wage - formal	Spatial wage - informal	ln(unemployment rate)
2000	-0.302	-0.310	-0.157	-2.310
2010	-0.082	-0.104	0.058	-2.809

Source: Author's elaboration.

6.4.3 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[9]. The initial results of the second stage are presented in Table 7. 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 8, fixed effects for MCAs are included. These results are very different from the main conclusions obtained from Table7. In fact, areas with higher population density will show a lower level of wage flexibility. Therefore, the main conclusions obtained from Table 7 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 9 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. Then, the informal sector not only had a higher elasticity for the local unemployment rate, but also for the unemployment rate in the neighbouring areas. This formulation does not tackle the spatial dependence, according to LM tests. One last try before estimating the SAR model is to control for unobserved local characteristics, with 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, 13).

[9] Complete results can be requested to the authors.

Table 7 – 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	-0.028***	-0.006				0.002	-0.021	-0.047*			
	(0.005)	(0.006)	(0.006)				(0.020)	(0.023)	(0.027)			
ln(unemployment rate) * density < 1				-0.052***	-0.084***	-0.060***				-0.098***	-0.122***	-0.141***
				(0.008)	(0.009)	(0.010)				(0.023)	(0.027)	(0.031)
ln(unemployment rate) * 1 <= density < 2				-0.011*	-0.041***	-0.029***				-0.052**	-0.080***	-0.095***
				(0.007)	(0.008)	(0.009)				(0.022)	(0.026)	(0.030)
ln(unemployment rate) * 2 <= density < 5				0.010*	-0.023***	-0.011				-0.021	-0.047*	-0.083***
				(0.006)	(0.007)	(0.007)				(0.022)	(0.025)	(0.030)
ln(unemployment rate) * 5 <= density < 10				0.018***	-0.018***	-0.001				-0.011	-0.037	-0.065**
				(0.005)	(0.006)	(0.007)				(0.021)	(0.024)	(0.028)
ln(unemployment rate) * 10 <= density < 20				0.011**	-0.026***	-0.002				-0.027	-0.049**	-0.071***
				(0.005)	(0.006)	(0.006)				(0.020)	(0.023)	(0.027)
ln(unemployment rate) * 20 <= density < 50				-0.009*	-0.043***	-0.011*				-0.061***	-0.074***	-0.095***
				(0.005)	(0.007)	(0.007)				(0.021)	(0.024)	(0.028)
ln(unemployment rate) * density >= 50				-0.036***	-0.081***	-0.037***				-0.133***	-0.154***	-0.167***
				(0.007)	(0.008)	(0.008)				(0.024)	(0.027)	(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 - instument: 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

*Controls: constant, % 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 log(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 can be requested.
Source: Author's elaboration.

Table 8 – Regressions for the spatial wage with MCAs fixed effects.

	FE					
	Total	**Informal**	**Formal**	**Total**	**Informal**	**Formal**
ln(unemployment rate)	-0.022***	-0.047***	-0.002			
	(0.008)	(0.011)	(0.012)			
ln(unemployment rate) * density < 1				-0.029*	-0.056**	-0.001
				(0.017)	(0.022)	(0.025)
ln(unemployment rate) * 1 <= density < 2				-0.032**	-0.038**	-0.019
				(0.013)	(0.017)	(0.020)
ln(unemployment rate) * 2 <= density < 5				-0.036***	-0.048***	-0.040***
				(0.011)	(0.014)	(0.015)
ln(unemployment rate) * 5 <= density < 10				-0.039***	-0.061***	-0.014
				(0.009)	(0.012)	(0.014)
ln(unemployment rate) * 10 <= density < 20				-0.022**	-0.049***	-0.008
				(0.009)	(0.012)	(0.013)
ln(unemployment rate) * 20 <= density < 50				-0.008	-0.033**	0.015
				(0.010)	(0.013)	(0.014)
ln(unemployment rate) * density >= 50				0.037***	-0.023	0.063***
				(0.013)	(0.017)	(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

*Controls: constant, % 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 log(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 can be requested.
Source: Author's elaboration.

In 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, but 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). 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).

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 10. 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.

Table 9 – Regression of the spatial wage against local characteristics, OLS, FE and SAR, 2000 and 2010.

	Total					Formal					Informal				
	OLS	OLS	FE	SAR	SAR - FE	OLS	OLS	FE	SAR	SAR - FE	OLS	OLS	FE	SAR	SAR - FE
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	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 can be requested to the authors.
Source: Author's elaboration.

Table 10 – 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	OLS	FE	SAR	SAR - FE	OLS	OLS	FE	SAR	SAR - FE	OLS	OLS	FE	SAR	SAR - FE
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	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***				-0.036***	-0.058**				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 can be requested to the authors.
Source: Author's elaboration.

The comparison of Models 1, 6 and 11 of Table 10 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 8 and 9. 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.

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 9 and 10. 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 9, 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 rather small; (v) a significant part of the difference between the formal and the informal sectors originates from spatial effects.

7 Conclusion

This study 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 so, a theoretical framework combining wage bargaining, institutional costs and efficiency wages was designed, distinguishing the formal and the informal sectors by the relative level of bargaining power of workers employed in each of them. Many empirical issues were investigated, aiming to understand their influence on the relationship of wages and local unemployment rates. Among them, we highlighted 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. In the particular case of PME, the inclusion of individual fixed effects does not leave enough variation to capture the wage curve relationship. State-level data from PNAD provided some new insights into the differences of wage flexibility between urban and rural areas. In fact, a 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 while a much higher wage flexibility could be found in the informal sector, as supported by the literature (Section 3).

Finally, Census data seems to provide the best regional disaggregation to study the wage curve in Brazil. When local characteristics are adequately controlled for, our findings suggest that wage flexibility is higher in less dense local labour markets and wages in the informal sector are 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 empirical results imply that labour market dualism is an essential ingredient in the evaluation of the wage curve in developing economies. When this analysis is done in the appropriate aggregation level, we are able to identify a relevant impact of city size on the relative bargaining power of workers and firms.

References

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.

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.

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.

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.

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.

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.

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.

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/i12-351de703e4da2073e562bebdf5fd0981.pdf , 2012.

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.

WOOLDRIDGE, J.M. **Econometric Analysis of Cross Section and Panel Data**. England: MIT Press, 1st Edition, 2002.