An estimate of spatial wage curve for Brazilian northeast region

Estimativa da curva de salário espacial para a região nordeste brasileira

Carlos Frederico A. Uchoa[1]

Resumo: The chief goal of this paper is to provide an investigation of the Wage Curve spatial heterogeneity for Brazilian northeast labor market. To do so, using the Wage Curve as Theoretical Framework, we apply spatial panel data techniques to compiled data set of 1,793 municipalities over 13 years from RAIS data to obtain the estimates of interest. Our results allow us to conclude there is a statistically significant negative correlation between unemployment and real wage, as also there are spillover effects on wages. In all, our finds confirm our initial assumption and also provide estimates that control for spatial heterogeneity.

Palavras-chave: Wage curve. Unemployment. Panel data. Spatial dependence. Direct and indirect effects.

Abstract: O objetivo deste artigo é investigar da heterogeneidade espacial da Curva de Salário para o mercado de trabalho no nordeste brasileiro. Para tanto, usando a Curva de Salário como ponto de paritda teórico, foram usados modelos dados de painel com efeitos espaciais a um subconjunto de 1.793 municípios ao longo de 13 anos obtidos da RAIS. Os resultados permitem concluir que existe uma correlação negativa e estatisticamente significativa entre o desemprego e o salário real, também que há efeitos estatisticamente significativos nos salários. Além de confirmar nossa suposição inicial da existência de uma Curva Salarial, os resultados fornecem estimativas que controlam os efeitos da heterogeneidade espacial.

Keywords: Curva de salário. Desemprego. Dados em painel. Dependência espacial. Efeitos diretos e indiretos.

JEL codes: E24; C23; R10; R12.

[1]UFBA, Federal University of Bahia, Department of Economics Autor(a) correspondente, *Email:* uchoa@ufba.br .

I Introduction

Through the 2000s Brazilian labor market has seen an important increase of income associated with continuous fall in the unemployment rate. This correlation emphasizes an empirical economic law which has been called the Wage Curve. The law states that there is a negative relationship between real wages and the local unemployment rate. Seminally introduced and empirically evidenced by Blanchflower e Oswald (1994a), the Wage Curve presents characteristics that oppose both Phillips Curve and the model proposed by Harris e Todaro (1970). Mainly because this model inversely relates the real wage level to the unemployment.

Due to theoretical consequences over the last few decades, several authors have investigated the Wage Curve assumptions in different labor markets. Berg e Contreras (2004) concluded that the Wage Curve exists in Chile, while Baltagi, Blien e Wolf (2000) found a similar result for German market. Elhorst, Blien e Wolf (2002), admitting the existence of spatial effects in the model, expanded the analysis provided by Baltagi, Blien e Wolf (2000) and found that wages may not only respond to the regional but also to the national unemployment rate. The North American labor market was analyzed by Blanchard e Katz (1992) and Staiger, Stock e Watson (2001) who obtained results favorable to the Phillips Curve. On the other hand, Hines, Hoynes e Krueger (2001) conclude that the Wage Curve has a more natural theoretical interpretation and better adjusts to the data in that market.

A number of authors have investigated the existence of a wage curve in Brazil. Among these, we highlight the works of Garcia e Fajnzylber (2002), Souza e Machado (2004), Amadeo e Camargo (1997), and Menezes e Alves-Filho (2004). According to these authors, the existence of a Wage Curve implies a high degree of flexibility in the labor market. As result, high unemployment rates are associated with lower levels of real wages.

The main goal of this paper is contributing to debate investigating the flexibility of the Brazilian labor market. The empirical study presented here provides, as far as we know, the first investigation of the wage curve relationship in the northeastern Brazilian labor market. For that, we compiled data set from RAIS data with 1,793 municipalities over the period from 2006 to 2018 years and estimate several spatial panel data (PD) models.

In addition to this introduction, this paper has been divided into five sections. Section II provides a brief presentation of the Wage

Curve model and its expansion to admit the existence of spatial effects. Section III discuss the PD models with spatial effects applied in the empirical analysis. The fourth Section explains how we compile the dataset used in this study. Section V presents the main results while in the last Section we present our final remarks.

II The Wage Curve empirical law

According to Blanchflower e Oswald (1994a) the relationship between income and the unemployment rate is difficult to capture, mainly because cyclical components of unemployment are negatively correlated with income, but structural components are positively correlated. Furthermore, those authors empirically shown that unemployment rate is negatively correlated with real income, which opened a debate about the relationship between real wage and local unemployment rate. Blanchflower e Oswald (1994a) analyzed sixteen different countries and concluded that the Wage Curve adequately describes the relationship between the two variables. The elasticity found was around −0.1, so that doubling the unemployment rate there would be a 10 percent decline in real wages, a result invoked as an empirical law of the economy which was called The Wage Curve.

More specifically, the wage curve assumes an equation in the form

$$w_{irt} = \beta_1 u_{rt} + \beta_k Z_{irt} + \alpha_r + e_{rt} \tag{1}$$

in which w_{irt} it is the real wage of each individual $i = 1, \ldots n$ observed at region $r = 1, \ldots, R$ and period $t = 1, \ldots, T$. α_r is a dummy for each region[2], u_{rt} expresses the unemployment rate in the region r and period t, and β_1 is the coefficient associated with the unemployment rate. Z_{irt} is a set of additional explanatory variables commonly used in the literature, such as characteristics that determine and differentiate the wages, for example gender, race, age, human capital, etc. Card (1995), and $\beta_k = \left(\beta_2, \ldots, \beta_p\right)$ its associated parameters. Finally, e_{rt} is the error term.

Note that in (1) unemployment does not vary at the individual level, which implies that the relevant dimension for estimating purposes does not depend on the individual wage multiplied by the number of periods, but on the number of labor markets multiplied by the number of time periods. As consequence, error term in (1) will be correlated

[2] Alternatively, time dummies could also be included.

across regions and the estimated standard error for the unemployment coefficient would be skewed down (MOULTON, 1990). To work around these problems the usual solution is to aggregate the data at regional level (r), so that we obtain the following specification:

$$w_{rt} = \beta_1 u_{rt} + \beta_k Z_{rt} + \alpha_r + e_{rt}, \tag{2}$$

in which w_{rt}, Z_{rt}, and e_{rt} are average by region r and time t. Note that, since (2) is a log – log specification, estimated parameters are independent of the unit in which the original data were expressed, so that β_1 is the income elasticity with respect to the unemployment rate.

If there is no correlation between the unobserved components in the model 2, the residuals are not correlated with the variables and the Ordinary Least Squares estimator is no longer consistent. However, this is a too restrictive and unrealistic assumption, especially with longitudinal data. To account for individual heterogeneity, we consider that regional differences are captured by the term α_r and apply Random Effects (RE) or Fixed Effects (FE) estimators. The RE estimator assume that unobserved heterogeneity is unrelated to the covariates, but unobserved error has a time-varying component, that is $e_{rt} = \alpha + u_{rt}$ and u_{rt} is an independently identically distributed (i.i.d.) Gaussian error term. FE allows that heterogeneity to be correlated to the covariates, so the individual heterogeneity is captured by the constant (α_r) that varies across individuals but not over time.

An additional issue is that individuals from different regions share similar unobserved characteristics. If this is the case, an increase of wage level in each region might encourage workers in neighboring regions to pressure employers for wage increases. Another possibility is the workforce migration that induces an interdependence between regions. To capture these effects, we consider the existence of cross-sectional dependency in model (2). In this sense, consider the equation (2) rewritten to include a spatial lag term:

$$w_{rt} = \rho \sum_{j=1}^{R} \omega_{ij} u_{jt} + \beta_1 u_{rt} + \beta_k Z_{rt} + \alpha_r + e_{rt} \tag{3}$$

in which ρ is the spatial correlation term and ω_{ij} a typical element of the spatial weight matrix W_{rr}. The next natural step is to discuss certain theoretical aspects that will be addressed before estimating (3), which we will do in the following section.

III Brief review of Spatial Panel Data models

Regression models with spatial dependence are becoming popular in econometric analyzes. The spatial correlation arises because a given event in a certain location can impact neighboring regions, which is related to the idea of relative location, interdependence between geographic regions, and spillover effects.

In order to allow spatial dependence, consider model 3 rewritten in a stacked form, that is let $w = (w_{11}, w_{12}, \ldots, w_{RT})'$, $u = (u_{11}, u_{12}, \ldots, u_{rt})'$, and so on. The FE estimator with a spatially lagged variable can be written as

$$\dot{w} = \rho \left(I_T \otimes W_R\right) \dot{w} + \dot{X}\beta + (\alpha \otimes \iota_T) + \dot{e}, \tag{4}$$

in which $\beta = \left(\beta_1, \ldots, \beta_p\right)$, $X = [u; Z]$, $\dot{w} = w - \bar{w}$, $\dot{X} = X - \bar{X}$, and $\dot{e} = e - \bar{e}$. The superscript '$\dot{}$' denotes that the variable was centered, which means that the group mean, indicated with the superscript '$^-$', was subtracted from each variable to remove the Fixed Effect α. The spatial weights matrix (W), the identity matrix (I) have subscriptions to indicate their respective dimensions, and $\otimes$ denotes the Kronecker's product. Note that $(I_T \otimes W_R)$ is equivalent to multiplying the spatial weights matrix for each T.

It is worthwhile to note that geographical units must remain unchanged, so that territorial limits are constant over time. For example, a new municipality introduces a new geographical unit, reduces the territory of another, and therefore a discontinuity in the spatial weights matrix. In this case, the discontinuity in the neighborhoods would make the analysis more complicated. To avoid this situation, we remove over the analyzed period that were created after 2006.

Along with these considerations, the logarithm of the likelihood function for the FE model takes the form:

$$\ln L = -\frac{RT}{2} \ln \left(2\pi\dot{\sigma}^2\right) + T \ln |I_{RT} - \rho W_R| - \frac{1}{2\dot{\sigma}^2} \left[\dot{w} - \rho \left(I_T \otimes W_R\right) \dot{w} - \dot{X}\beta\right]^2 . \tag{5}$$

Let $\hat{\beta}_O$ and $\hat{\beta}_L$ be, respectively, the parameters estimates $\dot{w}$ and $(I_T \otimes W_R)\dot{w}$ on $\dot{X}$, while $\hat{\dot{e}}_O$ and $\hat{\dot{e}}_L$ the respective residuals. The ρ estimate is obtained through maximization of the concentrated log-likelihood function

$$\ln L = C - \frac{RT}{2} \ln \left[\left(\hat{\dot{e}}_O - \rho \hat{\dot{e}}_L \right)' \left(\hat{\dot{e}}_O - \rho \hat{\dot{e}}_L \right) \right] + T \ln |I_R - \rho W_R| ,$$

in which C is a constant not dependent on ρ.

Maximization should be done numerically since there is no closed solution for the problem. Fortunately, the concentrated likelihood function is concave in ρ, which ensures the uniqueness of solution (ELHORST, 2010b). Once estimated ρ, β and $\dot{\sigma}^2$ estimates are obtained calculating

$$\hat{\beta} = \hat{\beta}_O - \hat{\rho}\hat{\beta}_L$$

and

$$\hat{\dot{\sigma}}^2 = \frac{1}{RT} \left(\hat{\dot{e}}_O - \hat{\rho} \hat{\dot{e}}_L \right)' \left(\hat{\dot{e}}_O - \hat{\rho} \hat{\dot{e}}_L \right) .$$

The associated standard deviations are obtained by replacing the parameters in the asymptotic variance matrix. For that, consider the covariance matrix estimator given by

$$\widehat{\text{Avar}} \left(\dot{\beta}, \rho, \dot{\sigma}^2 \right) =$$
$$\begin{bmatrix} \frac{1}{\dot{\sigma}^2} \dot{X}' \dot{X} & & \\ \frac{1}{\dot{\sigma}^2} \dot{X}' \left(I_T \otimes W_n \right) \dot{X} \dot{\beta} & T \text{tr} \left(\tilde{W} \tilde{W} + \tilde{W}' \tilde{W} \right) + \frac{1}{\alpha^2} \dot{\beta}' \dot{X}' \tilde{W}' \tilde{W} \dot{X} \dot{\beta} & \\ 0 & \frac{T}{\dot{\sigma}^2} \text{tr} \left(\tilde{W} \right) & \frac{RT}{2\dot{\sigma}^2} \end{bmatrix}^{-1}$$

in which $\tilde{W} = [I_{RT} - \rho (I_T \otimes W_R)]^{-1}$.

Elhorst (2010a) argues that the centering procedure produces biased estimates of $\dot{\sigma}^2$. So, we adopt the transformation proposed by the author to correct the bias, which is of the form $\frac{T}{T-1}\hat{\dot{\sigma}}^2$.

On the other hand, if regional heterogeneity is treated as random, the log-likelihood function for RE model becomes

$$\ln L = -\frac{RT}{2} \ln \left(2\pi \ddot{\sigma}^2 \right) + T \ln |I_{RT} - \rho W_R| - \frac{1}{2\ddot{\sigma}^2} \left[\ddot{w} - \rho \left(I_T \otimes W_R \right) \ddot{w} - \ddot{X}\beta \right]^2 , \tag{6}$$

in which the superscript '¨' now denotes the transformations $\ddot{w} = w - (1 - \theta)\,\bar{w}$ and $\ddot{X} = X - (1 - \theta)\,\bar{X}$. If θ is known, maximum likelihood estimate of ρ is obtained through an analogous procedure to the one used for FE. So, θ estimate is obtained by maximizing the concentrated log-likelihood function

$$\ln L = -\frac{RT}{2} \ln \left[e(\theta)' \, e(\theta) \right] + \frac{R}{2} \ln \theta^2,$$

with respect to θ.

Here, $u(\theta) = w - (1-\theta)\bar{w} - \rho \left[(I_T \otimes W_R) w - (1-\theta)(I_T \otimes W_R) \bar{w} \right] - \left(X - (1-\theta)\bar{X} \right) \beta$. Note that the procedure includes the parameter θ which is estimated alternatively and together with ρ, β, and $\ddot{\sigma}^2$ until the convergence criterion is reached for all of the unknowns.

The asymptotic covariance matrix for RE model takes the form

$$\widehat{\text{Avar}}\left(\beta,\rho,\theta,\ddot{\sigma}^2\right) =$$
$$\begin{bmatrix} \frac{1}{\ddot{\sigma}^2}\ddot{X}'\ddot{X} & & & \\ \frac{1}{\ddot{\sigma}^2}\ddot{X}'(I_T \otimes W_R)\ddot{X}\beta & T\operatorname{tr}\left(\tilde{W}\tilde{W} + \tilde{W}'\tilde{W}\right) + \frac{1}{\sigma^2}\beta'\dot{X}'\tilde{W}'\tilde{W}\dot{X}\beta & & \\ 0 & -\frac{1}{\ddot{\sigma}^2}\operatorname{tr}\left(\tilde{W}\right) & R\left(1+\frac{1}{\theta^2}\right) & \frac{RT}{2\ddot{\sigma}^2} \\ 0 & \frac{T}{\ddot{\sigma}^2}\operatorname{tr}\left(\tilde{W}\right) & -\frac{R}{\ddot{\sigma}^2} & \frac{RT}{2\ddot{\sigma}^2} \end{bmatrix}^{-1}.$$

Note that both models can be slightly modified to include spatially lagged covariates or spatial dependence in the error terms. In the first we get the spatial Durbin model by adding a covariate of the form $(I_T \otimes W_R) X$. In the second we set $\epsilon = \lambda We$ to get the spatial error model (SEM). Estimators for both cases are obtained through a slightly modifications of expressions already presented and do not significantly change the results above (LESAGE; PACE, 2009).

An important feature of spatial models is the possibility of calculating direct and Indirect Impacts. Once the time period is fixed, the impacts of a given explanatory variable on the dependent variable spatially lagged or not is obtained, with notation abuse, through following matrix of partial derivatives:

$$\left[\frac{\partial w}{\partial x_1}, \frac{\partial w}{\partial x_2}, \ldots, \frac{\partial w}{\partial x_k} \right] = \left[I_R - \rho W_R \right]^{-1} (\beta_k I_R) \tag{7}$$

in which β_k is a typical element of vector $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$.

The elements of diagonal capture the influence of x_k, in the spatial unit r, on response. On the other hand, elements outside the diagonal are the influence, in the other spatial units than r, of x_k on response.

According to LeSage e Pace (2009), the average of main diagonal on the right-hand side of (7) is the Direct Impact. The average of sum of columns outside the main diagonal the Indirect Impact. Furthermore, Elhorst (2010a) notes that, as the spatial units remains constant over time, the impact measure is not time dependent. In this case, the calculation is similar to those of cross-sectional models.

Although the direct and indirect effects can be easily calculated, it is not so easy to test the validity of these estimates, The standard errors cannot be obtained in the usual way, but, fortunately, LeSage e Pace (2009) proposed a Monte Carlo method to simulate the distribution of the impacts. The parameter distribution is obtained *a-posteriori*, through Bayesian Markov Chain Monte Carlo (MCMC) simulation, which produce valid hypothesis tests as demonstrated by Gelfand e Smith (1990). Nevertheless, there are two possible ways to conduct the simulation. One is to use the impact matrix itself, but this approach has the disadvantage to require, at each step, to invert a matrix with size equal to the sample. The alternative method consists in replace the matrix by its trace, which is employed in this paper.

IV Dataset

To construct the dataset, we had to resort to a hard compilation procedure, for which the primary source is the Annual Relation of Social Information (RAIS). The RAIS is an administrative register maintained by the Ministry of Labor, which can be seen as an annual census of formal employment. We should note that the use of RAIS is not immune to criticism, one of the most common is that RAIS only collect information about formal employment. However, in view of the options available, using RAIS is the only way to build a data set at municipal level.

Considering model (3), we set the dependent variable as the real average wage per worker at the municipality level ($wage_{rt}$) deflated by the IPCA. The explanatory variable of interest is the unemployment rate ($unemp_{rt}$), defined as population legally employed divided by the working age population, that is those aged 15 to 64. Full sample is composed of 1,793 Brazilian municipalities over the period 2006 to 2018, making a total of 23,309 observations.

We also include some control variables to avoid a possible omission bias. These variables aim to capture human capital, demographics, and other labor market characteristics. In this sense, to measure the human capital, we calculate the average human capital level ($educ_{rt}$), which is a good proxy for human capital because it represents stock of human capital. For that, we calculate the total years of schooling of all employees in the municipality and then we divide it by the number of workers. In addition to the human capital, we also include the following control variables: the percentage of male ($male_{rt}$), the percentage of white ($race_{rt}$), the average age of workers legally employed (age_{rt}), and a full set of time dummies (γ_t).

The spatial weight matrix was built using adjacency criteria. As such, units are considered to be neighbors if they share a common boundary. It is worthwhile to note that the spatial weight matrix, which specifies connectivity across geographic units, must be specified a priori and cannot be estimated within the model.

Despite the discussion about the impact of this measure on the estimates, LeSage e Pace (2010) showed with a thorough investigation that there is no evidence to support this fact. The resulting matrix has been transformed in the conventional way so that the sum of the lines is equal to 1 (see LeSage e Pace, 2009). The links between geographical units (municipalities) are shown in Figure 1.

V Main results

The regression model has the following specification:

$$\begin{aligned}\log(wage_{rt}) = \rho\,(I_T \otimes W_R)\log(wage_{rt}) + unemp_{rt} + \log(educ_{rt}) + \\ male_{rt} + race_{rt} + age_{rt} + \gamma_t + \varepsilon_{rt},\end{aligned} \tag{8}$$

in which the error term (ε) may be spatially auto-correlated or not and γ_t is a full set of time dummies.

As a start point, we estimate both the FE and RE models to decide which model to accept. These results are shown in Table 1 in which column (1) shows parameter estimates by RE, column (2) Spatial RE, column (3) FE, and column (4) Spatial FE.

All estimated models report a statistically negative coefficient for the unemployment rate. The coefficients associated with human capital, gender, and age are also all significant, while the coefficient for race is positive but not significantly different from zero. Wald tests ($\chi^2(\gamma_t)$) strongly rejects the null hypothesis that time dummies are jointly equal to zero. Finally, Pesaran CD tests strongly reject the null of no cross-sectional dependence.

To decide which model best fits the data we report the Akaike information criterion (AIC), calculated as $2k + RT\log(ssr/RT)$, and the

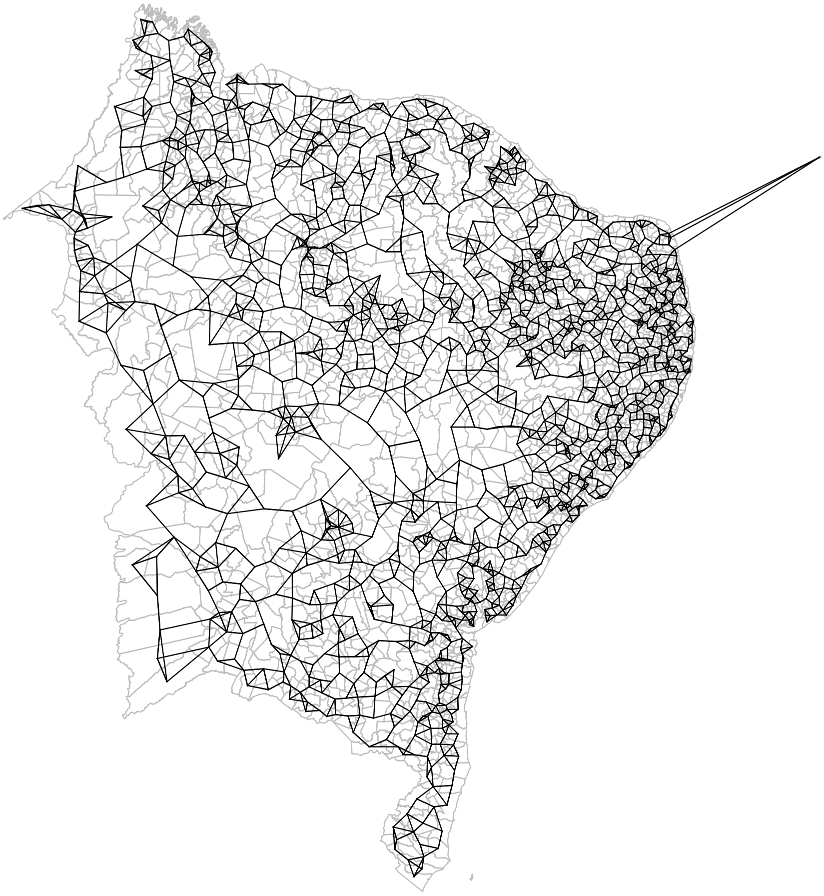

Figura 1: Brazilian Northeast Region: municipalities and linkages between units.

Hausman test. The test results indicate that it is not possible to ignore the presence of significant effects. So, we conclude that the FE estimator is preferable to the RE estimator which is inconsistent. These results are expected and are in line with the observations of Elhorst (2010b). The results shown in In addition, Table 1 allow us to conclude that, between the two FE alternatives, model (4) best fits the data.

In Table 3 we present results for 6 different Spatial FE specifications. Columns (1) and (3) are Spatial FE models and columns (2) and (4) spatial Durbin FE models, which include the spatial lag of unemployment ($unemp_{r-i}$). Once again, as expected, all models exhibit negative and significant coefficients for unemployment rate. The coefficients associated to spatial lag (ρ) and spatial error (λ) are also significant in all specifications.

Overall, the results in Table 3 indicate that neighborhoods positively influence the wages in Brazilian municipalities, so that there are 'high-high' and 'low-low' clusters. This positive spatial correlation suggests that an increase of real wage in a neighboring tends to produce an increase in the wage level in the reference municipality.

The Akaike information criterion favors model (1), so we conclude that this is the model that produces the best fit. We also note that the Durbin model the spatial lag of unemployment ($unemp_{r-i}$) not statistically different from zero. Models (3) and (4) contains the estimates for spatial error term (λ). Note that the spatial error correlation is negative in all models, suggesting the existence of spatial autocorrelation of the errors.

Unemployment negatively impacts wages so that higher unemployment rates are associated with lower real wage averages. Moreover, in all models the estimated coefficients for unemployment (≈ -0.21) are close to those obtained by Blanchflower e Oswald (1994a). Reported estimates for human capital, gender, and age are statistically significant, but race is not. Among these, we highlight the human capital and the percentage of male workers. The first is soundly different from zero and varies between 0.20 and 0.23. The second, also positive and statistically different from zero, indicating that male workers receive, on average, higher wages than women.

Next step is to assess the Direct, Indirect, and Total Impacts on real wage due to changes in one of the covariates. Direct Impact measure the impact of a change in the covariate on response. For example, if the average real wage increases in a neighboring municipality, a positive impact on the wages is expected since $\rho > 0$. Total Impact results from the sum of direct and indirect effects. This is a measure of the

Tabela 1: Comparison of spatial and non spatial, FE- and RE-models. Dependent variable is the log real wage.

Model	(1) *RE*	(2) *SRE*	(3) *FE*	(4) *SFE*
Constant	5.30^{***}	4.44^{***}		
	(0.03)	(0.03)		
ρ		0.14^{***}		0.11^{***}
		(0.01)		(0.01)
λ		1.52^{***}		
		(0.06)		
unemp	-0.25^{***}	-0.25^{***}	-0.21^{***}	-0.21^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
log(*educ*)	0.25^{***}	0.25^{***}	0.23^{***}	0.23^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
male	0.44^{***}	0.42^{***}	0.42^{***}	0.41^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
race	0.01	0.01	−0.00	−0.00
	(0.00)	(0.00)	(0.00)	(0.00)
age	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
$\chi^2\,(\gamma_t)$	$94{,}375.8^{***}$	$68{,}951.9^{***}$	$91{,}308.7^{***}$	$8{,}928.3^{***}$
	(0.00)	(0.00)	(0.00)	(0.00)
Hausman-test		152.6		152.2
		(0.00)		(0.00)
nT	23,309	23,309	23,309	23,309

Source: Author(s), own elaboration. **Note:** Estimated standard errors in parentheses are cluster robust at the municipal level and *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. $\alpha \otimes \iota_T\ (\chi^2)$ is testing the joint hypothesis that the full set of time dummies is not jointly different from zero with 12 degrees of freedom. Pesaran CD report tests for cross sectional dependence in panel models.

total cumulative impact resulting from an increase in wages on neighboring. Indirect Impact is calculated as the total minus the Direct Impact, which is a measure of the impact of a change in the average wage in other municipalities on the wages in the municipality under consideration.

The Direct, Indirect, and Total Impacts for our preferred model, which is model (4) according to the AIC criterion, are exhibit in Table 3. The estimated Direct impacts are higher than the reported in the regression model, which occurs due to the positive feedback from the impacts passing through neighboring municipalities and returning.

One can note that unemployment, human capital, gender, and age exhibit impacts significantly different from zero. Average age has a positive impact on wages, which seems intuitively plausible, mainly because it is a proxy for worker's experience. On the other hand, the estimated effects for race are all remarkably close to zero and still statistically significant, which suggests that is not a determinant factor to the wage level in the municipality. However, it is worthwhile to note that this result should not be viewed as an absence of racial income gap.

The Indirect Impact of unemployment rate on wages are also negative and smaller than Direct Impact. This result suggest that increase of wages in units experiencing reduction in unemployment. The Indirect Impact from unemployment in nearby municipalities are something like ten times those of Direct, which suggests a small to moderate spillover.

Total Impact, consisting of equal parts of Direct and Indirect impacts, is also negative. Same analysis can be done for the other variables. Both percentage of male workers and average age increase real wages. A further study could assess this issue in detail, since it seems a good starting point for discussion and further research of economic inequality by gender.

Overall, the consequences of these results are important and show that the average real wage is strongly correlated with the unemployment rate. As pointed by Blanchflower e Oswald (1994b), these findings are consistent with research showing that the wage curve casts doubt on some of the most important ideas in macroeconomics, labor economics, and regional economics. One source of weakness in this study which could have affected our measurements is the absence of a causality analysis, an intriguing one which could be usefully explored in further research.

Tabela 2: Summary of Panel Data Regression Results for Fixed Effects models. Dependent variable is the log real wage.

	(1)	(2)	(3)	(4)
ρ	0.11^{***}	0.11^{***}	0.34^{***}	0.34^{***}
	(0.01)	(0.01)	(0.02)	(0.02)
λ			-0.26^{***}	-0.27^{***}
			(0.02)	(0.02)
unemp	-0.21^{***}	-0.21^{***}	-0.20^{***}	-0.20^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
$unemp_{r-1}$		0.00		0.03
		(0.02)		(0.02)
$\log(educ)$	0.23^{***}	0.23^{***}	0.20^{***}	0.20^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
male	0.41^{***}	0.41^{***}	0.39^{***}	0.39^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
race	−0.00	−0.00	−0.00	−0.00
	(0.01)	(0.01)	(0.01)	(0.01)
age	0.01^{***}	0.01^{***}	0.01^{***}	0.01^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
$\chi^2\,(\gamma_t)$	$8{,}928.3^{***}$	$89{,}594.8^{***}$	$85{,}869.4^{***}$	945.0^{***}
	(0.00)	(0.00)	(0.00)	(0.00)
nT	23,309	23,309	23,309	23,309
AIC	−104,905.9	−104,903.8	−104,711.3	−104,698.4

Source: Author(s), own elaboration. **Note:** Dependent variable is Log of real wage $\log(w)$. Estimated standard errors in parentheses are cluster robust at the municipal level and *, **, *** indicates significance at the 90%, 95%, and 99% level, respectively. $\alpha \otimes \iota_T\,(\chi^2)$ is testing the joint hypothesis that the full set of time dummies is not jointly different from zero with 12 degrees of freedom.

Tabela 3: Direct (1), indirect (2), and total (3) effects measures for model (1). Dependent variable is the log real wage.

	(1)	(2)	(3)
unemp	-0.214^{***}	-0.026^{***}	-0.240^{***}
	(0.014)	(0.003)	(0.015)
log(*educ*)	0.226^{***}	0.028^{***}	0.253^{***}
	(0.008)	(0.002)	(0.009)
male	0.415^{***}	0.051^{***}	0.465^{***}
	(0.013)	(0.004)	(0.014)
race	0.000	0.000	0.000
	(0.008)	(0.001)	(0.009)
age	0.011^{***}	0.001^{***}	0.012^{***}
	(0.001)	(0.000)	(0.001)

Source: Author(s), own elaboration. **Note:** Estimated standard errors in parentheses are cluster robust at the municipal level and * , ** , *** indicates significance at the 90%, 95%, and 99% level, respectively.

VI Concluding remarks

This paper aimed to estimate the Wage Curve with spatial effects for the Brazilian Northeast region. The main contribution was to investigate this empirical law through spatial panel modes. As far as we can see, there is no other study in Brazil with the same subject, database, and technique. In this sense, prior to this study there was a shortage of spatio-temporal analysis of association between wages and unemployment.

In summary, one of the more significant findings to emerge from this study is spatial dependence is an important determinant and cannot be ignored. In particular, as is widely know, if spatial autocorrelation is neglected the estimates can be biased, type I error increase or the relationships become reversed. So, according to our results, this is an important issue and deserves the attention of researchers.

About public policies, there is strong evidence that higher unemployment rates are associated with lower real wages. Our results also confirm the existence of a wage Curve in the region, bringing to the fore all the consequences registered in previous studies. Taken together, these findings support the recommendation that spatial interdependence when formulating employment and income policies. This information can be used to develop targeted interventions aimed at improving legal employment.

Referências

AMADEO, E. J.; CAMARGO, J. M. Brazil, regulation and flexibility in the labor market. In: EDWARDS, S.; LUSTIG, N. (Ed.). *Labor Markets in Latin America: Combining Social Protection with Market Flexibility*. Washington, DC: Brookings Institution Press, 1997.

BALTAGI, B. H.; BLIEN, U.; WOLF, K. The east german wage curve 1993-1998. *Economics Letters*, v. 69, n. 1, p. 25–31, 2000.

BERG, J. M.; CONTRERAS, D. Political-economic regime and the wage curve: Evidence from chile, 1957-96. *International Review of Applied Economics*, v. 18, n. 2, p. 151–165, 2004.

BLANCHARD, O. J.; KATZ, L. F. Regional evolutions. *Brookings Papers on Economic Activity*, v. 1, p. 1–75, 1992.

BLANCHFLOWER, D. G.; OSWALD, A. J. Estimating a wage curve for britain: 1973-90. *Economic Journal*, v. 104, n. 426, p. 1025–1043, 1994.

BLANCHFLOWER, D. G.; OSWALD, A. J. *The Wage Curve*. 1. ed. Cambridge, MA: MIT Press, 1994.

CARD, D. E. *The Wage Curve: A Review*. [S.l.], 1995.

ELHORST, J. P. Matlab software for spatial panels. In: *IVth World Conference of the Spatial Econometrics Association (SEA)*. Chicago, IL: [s.n.], 2010. Disponível em: <http://www.regroningen.nl/elhorst/>.

ELHORST, J. P. Spatial panel data models. In: FISCHER, M.; GETIS, A. (Ed.). *Handbook of Applied Spatial Analysis*. New York, NY: Springer, 2010. C.2, cap. 3, p. 377–407.

ELHORST, J. P.; BLIEN, U.; WOLF, K. *A spatial panel approach to the east German wage curve*. [S.l.], 2002.

GARCIA, L. L.; FAJNZYLBER, P. A. Curva de salário para o brasil: Uma análise microeconométrica a partir dos dados da pnad de 1992 a 1999. In: ANPEC. *XXX Encontro Nacional de Economia*. [S.l.], 2002.

GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, American Statistical Association, v. 85, n. 410, p. 398–409, 1990.

HARRIS, J. R.; TODARO, M. P. Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, v. 60, p. 126–142, 1970.

HINES, J. R.; HOYNES, H. W.; KRUEGER, A. B. *Another Look at Whether a Rising Tide Lifts All Boats*. [S.l.], 2001. (Working Paper Series, 8412). Disponível em: <http://www.nber.org/papers/w8412>.

LESAGE, J. P.; PACE, R. K. *Introduction to Spatial Econometrics*. Boca Raton, FL: CRC Press, 2009.

LESAGE, J. P.; PACE, R. K. *The Biggest Myth in Spatial Econometrics*. [S.l.], 2010.

MENEZES, W.; ALVES-FILHO, O. A curva de salário para a rms. In: ANPEC REGIONAL. *IX Encontro Regional de Economia do Nordeste*. Fortaleza, CE, 2004.

MOULTON, B. R. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, v. 72, p. 334–338, 1990.

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*, v. 42, p. 35–54, 01 2004.

STAIGER, D.; STOCK, J. H.; WATSON, M. W. *Prices, Wages and the U.S. NAIRU in the 1990s*. [S.l.], 2001. (Working Paper Series, 8320). Disponível em: <http://www.nber.org/papers/w8320>.