

Spatial Unemployment Differentials in Colombia

Diferenciales espaciales en la tasa de desempleo en Colombia

Ana María Díaz¹

DOI: 10.13043/DYS.76.3

Abstract

Between 1993 and 2005 Colombian municipalities experienced a polarization process, as municipalities' unemployment rates followed different evolutions that are relative to the National average. This process was been accompanied by the creation of unemployment clusters. This paper uses a spatial Durbin model to explore the influence of various factors in determining differences in regional unemployment rates and proposes a decomposition methodology to quantify how much of the variation in unemployment rate is explained by the variables included in the model and how much is explained by the variables omitted. According to our findings, differences in labor demand, immigration rates, and urbanization are factors behind the observed municipal unemployment disparities. This paper also explores whether different groups of regions will react differently to a labor market impulse.

Key words: Local labor markets, unemployment differential, polarization, clustering, spatial econometrics, spatial Durbin model.

JEL classification: R23, C14, C23, C31.

1 Pontificia Universidad Javeriana, Departamento de Economía, Calle 40 No. 6-23, Ed. Gabriel Giraldo S. J. Piso 8º, Bogotá. E-mail: a.diaze@javeriana.edu.co.

Este artículo fue recibido el 29 de septiembre de 2014, revisado el 2 de junio de 2015 y finalmente aceptado el 5 de noviembre de 2015.

Resumen

Entre 1993 y 2005 la tasa de desempleo de los municipios colombianos experimentó un proceso de polarización. Este proceso estuvo acompañado de la creación de clústeres espaciales de desempleo. En este artículo se utiliza un modelo Durbin espacial para explorar la influencia de diversos factores en la determinación de las diferencias en las tasas de desempleo municipales y propone una metodología de descomposición que permite cuantificar qué parte de la variación en la tasa de desempleo se explica por las variables incluidas en el modelo y qué parte se explica por las variables omitidas. De acuerdo con nuestros resultados, diferencias en la demanda de trabajo, las tasas de inmigración y la urbanización son factores que explican las disparidades municipales de desempleo observadas. Este artículo también explora si los diferentes grupos de regiones reaccionan de manera distinta a un impulso del mercado de trabajo.

Palabras clave: mercados laborales locales, diferencias en desempleo, polarización, econometría espacial, modelo Durbin espacial.

Clasificación JEL: R23, C14, C23, C31.

Introduction

Two of the most evident features of unemployment in Colombia are incidence and persistence over time. Several studies point to structural conditions and rigidities of the labor market as the major causes of these features (Cardenas & Bernal, 2003; Lora & Marquez, 1998). Another interesting but less examined feature of unemployment is its geographical distribution. The aim of this paper is to provide further insights into the geographical patterns of unemployment rates in Colombian municipalities. An exploratory analysis of the unemployment rate in 522 urban areas, using census data from 1993 and 2005, shows that the distribution of unemployment rates was characterized by a polarization process between 1993 and 2005: if the municipal unemployment rate was either very high or very low, relative to the National average, it tended to stay at that level, but if the initial unemployment was close to the average it either increased or decreased. This process has been accompanied by the creation of unemployment clusters; that is to say, close municipalities exhibit similar patterns in unemployment rates.

Why is the geographic distribution of unemployment unequal? Economic theory provides a variety of perspectives on the nature and significance of regional unemployment differentials. Regions with favorable economic and demographic attributes might perform better and experience lower unemployment rates than declining regions. Indeed, regions differ in the industrial composition of their employment; in the age, gender, and skill structure of their populations; and in their levels of urbanization and agglomeration of their economic activity. In the short-term, regional unemployment disparities reflect disparities in the previously mentioned factors. In the long-term regional, differences will gradually erode through labor mobility and/or the firm's relocation. Why do such differences persist? The literature offers, at least, four explanations. First, long-term differentials represent an equilibrium in which factors such as favorable climatic conditions, or an attractive environment encourage people to stay in regions where unemployment rates are high (Marston, 1985). Second, some persistent regional inequalities may reflect labor market rigidities, which restrict mobility (Blanchard, Katz, Hall & Eichengreen, 1992). Third, according to the new economic geography theory, the polarized structure of unemployment rates may reflect the agglomeration of economic activities. The presence of economies of scale that benefit more booming regions, where workers and production are agglomerated, will exhibit lower unemployment rates relative to sparsely populated, peripheral regions (Epifani & Gancia, 2005; Suedekum, 2005; Vom Berge, 2011). The self-reinforcing nature of agglomeration economies, which attracts more workers and firms, translates into a stable core-periphery unemployment gap. Furthermore, as clusters of activity may extend across borders this can result in clusters of high and low unemployment extending across regions (Puga, 2002).

Finally, Bande and Karanassou (2009) use the chain reaction theory to explain the evolution of unemployment differentials across regions. In their model, unemployment is viewed as the outcome of prolonged adjustments to shocks, and since different regions may be exposed to different types of shocks and experience different adjustment processes, unemployment differences persist.

This paper contributes to the empirical unemployment literature by analyzing the effect of certain factors on the dynamics of the geographical distribution of unemployment rates for a developing country. In general terms, the municipal unemployment rate is a reduced form function of factors that affect labor supply and labor demand. These factors can be broadly categorized

as labor market dynamics, non-demographic labor market attributes, human capital, demographic characteristics of the local labor force, and municipal attributes. Although it would be reasonable to assess the contribution of each of these factors from the view that unemployment dynamics are only related with factors within the municipality itself, it is also reasonable to assume that unemployment dynamics in a given municipality are related to the behavior of nearby municipalities due to interdependencies brought about by general equilibrium effects. Recently, in an attempt to bring these effects into the analysis, more and more studies have begun to use what is known as spatial econometrics. Some examples include Molho (1995), Lopez-Bazo, Del Barrio and Artis (2002), Overman and Puga (2002), Aragon *et al.* (2003), Niebuhr (2003), Patacchini and Zenou (2007) and Cracolici, Cuffaro, Nijkamp and delle Scienza (2007). The term "spatial econometrics" is a concept that relates to explanatory regression models that take into consideration that what happens in a particular municipality can also affect events in other nearby municipalities.

Allow me to give a concrete example. Suppose we want to explore the relationship between local human capital and unemployment. We can expect that municipalities with a high proportion of skilled workers experience lower unemployment rates as production shifts towards high-skilled employment, but as a result of human capital externalities nearby, municipalities might also benefit. Spatial econometrics enables us to explore whether human capital has any effect on the unemployment rate of the same municipalities or whether nearby regions are also affected, and, if they are, the extent of the overall impact.

To avoid an ad hoc choice of the specification, this paper uses a spatial Durbin model (SDM). This model is a spatial regression model that includes a spatial lag of the unemployment rate as well as the explanatory variables. The use of the SDM model has several advantages in relation to those models used to analyze regional unemployment differentials that are contained in the previously mentioned literature. One of these advantages is that it allows us to compute fairly simple diagnostics to test this model against more parsimonious alternatives because it nests most of the spatial models (Elhorst, 2010). A second virtue is that it provides consistent parameter estimates even if the true data generating process is a spatial lag or a spatial error model (LeSage & Pace, 2009). Another strength is that it allows us to explore spatial effects for different explanatory variables whilst not imposing prior restrictions on the

magnitude of these effects (Elhorst, 2010). Finally, it allows summary measures of direct, indirect, and total impacts on unemployment rate that arise from changing each explanatory variable in the model, following LeSage and Pace (2009), to be estimated.

The spatial econometric exercise is complemented with a decomposition analysis that quantifies how much of the variation in unemployment rates is explained by the variables included in the model, and how much is explained by the omitted variables. Moreover, it allows us to assess the relative importance of each regressor with respect to its overall effect on the change in municipal unemployment rates.

The final step is to analyze whether different groups of regions will react differently to a labor market impulse. After introducing a shock into the model, namely a policy decision, we estimate four measures: the number of municipalities affected by the system of interactions, the difference between the observed and simulated unemployment rates, a measure of spatial inequality, and a measure of spatial dependence. A comparison of these measures for different scenarios serves to illustrate how a change in a few municipalities can affect other municipalities through the system of interactions and how it can modify the spatial distribution of unemployment rates.

The next section includes a brief literature review. Section II presents the data and describes some underlying trends in Colombian municipal unemployment and potential factors explaining its evolution. Section III, presents the empirical strategy and the results follow in Section IV. The final section offers some concluding remarks.

I. Literature Review

Various theoretical models have explained the existence and persistence of regional disparities in the unemployment rate (Blanchard *et al.*, 1992; Decressin & Fatás, 1995; Elhorst, 2003; Marston, 1985). As stated by Marston (1985), there are two possible explanations. The first one is related to an equilibrium mechanism, while the second is related to a disequilibrium context. According to the first explanation, each region regulates its own equilibrium unemploy-

ment rate, which is determined by local demand and supply side factors.² In the short-term, regional unemployment disparities reflect disparities in these attributes. In the long-term, regional differences will gradually erode through labor mobility up to a point where only compensating differentials between regions remain (Harris & Todaro, 1970; Marston, 1985). Thus, the spatial distribution of unemployment resulting from this interpretation is characterized by constant utility across areas: high unemployment in one area is compensated by some positive factors (e.g., local amenities, climatic conditions, quality of life, local housing conditions, etc.). Marston (1985) claims that, to the extent that unemployment has a natural equilibrium, any policy oriented at reducing regional disparities is useless "since they cannot reduce unemployment anywhere for long".

According to the second view, all regions have a competitive equilibrium unemployment rate (Blanchard *et al.*, 1992). In the short-term, regional inequalities reflect the effect of asymmetric shocks (e.g., a shortage of labor demand in some regions). In the long-term regional differences will eventually level out and disappear through labor migration and/or firm relocation. However, labor market rigidities (e.g., wage bargaining, unions, taxation, welfare state arrangements, and labor laws) may restrict mobility and therefore adverse shocks are not fully absorbed before the regional labor market is hit by new shocks. Thus, the persistence of regional unemployment differentials is determined by the whole history of shocks on the economy. Using this theory, unemployment differentials can be reduced by encouraging flexible labor markets and by reducing structural rigidities (Blanchard *et al.*, 1992).

The equilibrium-disequilibrium views of regional differences in unemployment rates have recently been challenged by what can be observed in the spatial distribution of unemployment rates, in particular by the fact that regions with high (low) unemployment rates are surrounded by other regions with high (low) unemployment rates. Although these patterns are not inconsistent with the equilibrium-disequilibrium views, there is no theoretical causation mechanism that predicts a spatial clustering of unemployment.

2 Demand side factors can be the industry composition of regional production and industrial diversity, while supply side factors relate to attributes of the labor force such as the skill composition and the demographic structure of the workforce.

Models from the new economic geography discipline have attempted to fill this gap. These models posit that the interaction between scale economies and transport costs will create incentives for firms and workers to be concentrated in a certain space. According to Epifani and Gancia (2005), such spatial concentration of economic activity creates core regions, where workers and production are agglomerated; these enjoy lower unemployment than sparsely populated, peripheral regions. Their argument is the following: frictions in the job-matching process lead to equilibrium unemployment, and search costs generate a positive externality of agglomeration in the labor market because agglomeration economies (*i.e.*, productive advantages coming from the spatial concentration of labor and capital) increase firms' core profits and induce new vacancies, thereby lowering unemployment. The opposite happens in the periphery, where the reduction in firms' profits deteriorates the local labor market conditions. The self-reinforcing nature of agglomeration economies, which attracts more workers and firms to the core regions, translates into a core-periphery unemployment gap. Furthermore, clusters of activity may extend across several administrative units, which can result in clusters of high and low unemployment extending across regional borders (Puga, 2002). Consistent with this theory, variables affecting the spatial distribution of economic activity also affect regional disparities in unemployment and might lead to the creation of spatial clusters of high and low unemployment.³

Several empirical studies have analyzed disparities in regional unemployment rates for different countries (*e.g.*, Cracolici *et al.*, 2007; Lopez-Bazo *et al.*, 2002; Molho, 1995; Niebuhr, 2003; Overman & Puga, 2002; Patacchini & Zenou, 2007). In these studies, regional unemployment is related to local characteristics, personal attributes of the local population, local demand variables, and attributes of neighboring regions, which take into account the spatial interaction among regions. These empirical studies have brought to light some interesting facts: a) there are important spatial inequalities in unemployment rates within countries (*e.g.*, UK, Spain, Italy, France, Germany, and Turkey) and also between countries; b) unemployment differences are more pronounced within countries

3 Suedekum (2005) also finds that large core regions will exhibit lower unemployment rates compared to peripheral regions. Moreover, he posits that the core-periphery structure of unemployment resembles the spatial configuration of GDP per capita: low unemployment is centered in the agglomerated area whereas poor regions mostly have high unemployment rates. In other words, regions from the same country, with identical labor market institutions, can evolve very differently, depending on whether they belong to the cluster of central, intermediate or peripheral regions.

than between countries inequalities; c) these differences are highly persistent over time; and d) adjacent regions tend to have similar unemployment rates in comparison to regions located far away, this is unemployment observed at one point in space and is dependent on values observed at other locations.⁴

Persistent regional unemployment inequalities have been explained by spatial differences in labor demand by Molho (1995), Overman and Puga (2002) and Cracolici *et al.* (2007), for the UK, European regions, and Italy respectively. Filiztekin (2009) finds, that for Turkey, not only are there differences in labor demand but there are also regional differences in human capital, which are the sources of observed disparity across regions. Lopez-Bazo *et al.* (2002) and Aragon *et al.* (2003) argue that unequal distribution of amenities is the major cause of spatial inequalities in unemployment rates in Spain and France. Finally, Basile, Girardi and Mantuano (2009) conclude that the excess of labor supply, migration outflows, and spatial proximity determine the polarization of regional unemployment rates.

Spatial dependence of the unemployment rates has been explained by three main factors.

First, data collection of observations associated with spatial units such as countries, states, regions, census tracts do not accurately reflect the nature of the underlying process generating the sample data. Indeed, workers are mobile and can find employment in neighboring areas, thus, unemployment measured on the basis of where people live could exhibit spatial dependence. For example, Patacchini and Zenou (2007), using local UK data, provide evidence of a significant spatial dependence which is mainly explained by commuting flows between local areas. Molho (1995) suggests that spatial dependence arises through migration across regions. Second, the spatial concentration of the variables explains unemployment. For example, regions with favorable economic and demographic conditions may experience lower unemployment rates relative to municipalities with unfavorable conditions. If regions with favorable (or unfavorable) conditions are geographically concentrated this might explain the spatial correlation of unemployment rates (Cracolici *et al.*, 2007). Third, the spatial dependence of unemployment may reflect the agglomera-

4 See for example, Lopez-Bazo *et al.* (2002), Overman and Puga (2002), Niebuhr (2003), Aragon *et al.* (2003), Patacchini and Zenou (2007), Cracolici *et al.* (2007) and Basile *et al.* (2009).

tion of economic activities because the linkages between regions tend to tie together labor supply and demand conditions across nearby areas. This is the conclusion that Overman and Puga (2002) draw from analyzing unemployment clusters in Europe.

There are very few academic analyses trying to explore the regional distribution of unemployment rates in Colombia. Some exceptions are Galvis (2002), Arango (2011), Cardenas, Hernández and Torres (2014), and Merchán (2014). Galvis (2002) finds that the Colombian urban labor market is integrated. Arango (2011) finds important differences across cities for different labor market indicators (e.g., employment, unemployment, wages). Cardenas *et al.* (2014) study the structural determinants of differentials in unemployment rates and labor market performance for 23 Colombian cities. They find that the heterogeneity on unemployment rates can be attributed to labor supply levels, incentives to participate and the age structure of the working age population. Finally, Merchán (2014) studies the determinants of unemployment using longitudinal data for Colombian cities. He finds that the gross domestic product and capital formation in the public sector are the factors that are negatively related to unemployment.

II. Data Description

This study uses Colombian Census data from the Integrated Public Use Micro data Series (IPUMS) for 1993 and 2005.⁵ The IPUMS database is composed of a 10 percent sample of individual records containing information on persons and households. The unit of analysis is the municipality. Individuals are assigned to a municipal area on the basis of IPUMS codes, which are geographical divisions that contain no less than 20,000 inhabitants. Indeed, this code aggregates the information from 1052 municipalities into 532 observations; thus, some observations include more than one municipality. It includes all Colombian municipalities, and corresponds to an intermediate level between the urban and the regional level, which offers both a larger number of observations than standard regions and a reasonable approximation for complete or independent labor market areas. One main advantage of

5 The period between 1993 and 2005 corresponds approximately to an entire business cycle. Thus, to a certain extent, the asymmetries in the municipal response to phases in the cycle are minimized and we can assume that both years are comparable in economic terms.

this data is that it allows us to abstract information from commuting pattern issues, in the sense that areas are large enough that it is very unlikely that individuals commute across them.

A. Outcome Variable

Unemployment rate is the percentage of unemployed people in the working age population. An unemployed individual is defined as someone who is not working and currently available for work during the week in which the Census was undertaken.⁶

B. Covariates

The local unemployment rate is a reduced form function of factors that affects labor supply and labor demand, such as labor market dynamics, non-demographic labor market variables, human capital attributes, demographic characteristics of the local labor force, and municipal attributes. The variables selected to serve as a proxy for these broad categories are the following:⁷

Labor market dynamics: a primary factor determining unemployment differences is employment growth. If a given municipality is creating more employment than the national level, unemployment in that municipality should decrease relatively. However, employment growth at the municipal level may not reduce the unemployment rate. This can occur because a better labor market situation will not only attract jobless workers but also migrants, who may absorb all the new jobs. To control for labor market dynamics, we use a measure of employment growth based on exogenous local labor demand shocks,⁸ and

6 The International Labor Organization (ILO) set guidelines to declare an unemployed individual as someone who is not working, currently available for work, and seeking a job. The ILO introduced modifications regarding this definition by allowing the partial or full relaxation of the active job search requirement in situations in which "the conventional means of seeking work are of limited relevance, where the labor market is largely unorganized or of limited scope, where labor absorption is at the same time inadequate, or where labor force is largely self-employed". Since the Colombian labor market fits this description I do not use the active search requirement in the construction of unemployment rates. Results using standard unemployment definition are robust and available upon request.

7 See Appendix, Table A1.1, for a detailed description of the construction of each variable.

8 This measure is based on the Katz and Murphy (1992) index, which decomposes employment growth into expected share and industry mix components (Partridge and Rickman, 1995; Stevens and Moore,

the ratio of immigrants with working age population. Immigrants are considered as the percentage of the working age population who have moved from one municipality to another in the last five years (as in Basile *et al.*, 2009; Blanchard *et al.*, 1992; Bradley & Taylor, 1997; Molho, 1995).

Non-demographic labor market variables: the diversity of employment opportunities in a municipality may affect the unemployment rate. The more diverse an economy is, the more readily employment reductions in any given sector can be absorbed by other sectors. The greater the industrial diversity the more even the distribution of employees across industries. We measure employment diversity by one minus a two-digit industry Herfindahl index (as in Partridge & Rickman, 1995). Likewise, employment concentrations in particular sectors may have an additional influence on the unemployment rate. Municipalities specializing in declining industries are expected to exhibit higher unemployment rates than those based around growing activities. Consistent with previous analysis (e.g., Basile *et al.*, 2009; Cracolici *et al.*, 2007; Lopez-Bazo *et al.*, 2005; Niebuhr, 2003; Overman & Puga, 2002), we use the employment shares of two main sectors: manufacturing and services.

Human capital variables: to evaluate the effect of human capital on unemployment rates this paper uses the percentage of the working age population who are high school and college graduates. In particular, we should expect this share to be inversely related to the unemployment rate through a composition effect, due to its positive influence on labor demand, and because skilled individuals are geographically more mobile (Manning, 2004; Martin & Morrison, 2003; Mincer, 1991; Saint-Paul, 1996).

Demographic variables: the structure of the population might have important influences on local labor demand and supply (Elhorst, 2003). To control for this effect, we employ the age structure of the population. Additionally, the percentage of females above the age of 15 who are married is included in order to reflect on the possibility that married women withdraw from the labor-force. Similarly, women with young children may be more likely to withdraw

1978). The reason for using this measure instead of a simpler employment growth measure is twofold. First, employment growth perfectly predicts unemployment growth as I am using an extended definition of unemployment in which the job search condition is relaxed. Second, it avoids collinearity with measures of migration and other covariates.

from the labor force. This is controlled by the percentage of women over the age of 15 who are married and have children under the age of 5.

Municipal attributes: we use the population density and the percentage of the municipality's population that lives in urban areas. The standard argument for the inclusion of these variables is that coordination failures between employer and job seekers may be mitigated in urban areas because of the greater diversity of employment opportunities. Recent research has expanded on this by arguing that urban labor markets generate human capital externalities that would not exist in less populated areas (Glaeser & Mare, 2001; Moretti, 2004; Rauch, 1993). In turn, congestion effects may also lead to higher unemployment rates; thus, the relationship between urbanization proxies and unemployment rates is, *a priori* unknown.

C. Spatial Proximity

We measure spatial proximity in terms of contiguity; the neighboring set is therefore defined as the set of municipalities that share a common boundary.⁹ We summarize the possible inter- actions between municipalities using the matrix $W_c = \{\omega_{ij}\}$, where $\omega_{ij} = 1$ if municipalities i and j share a common border and 0 if they do not.

D. Exploratory Evidence

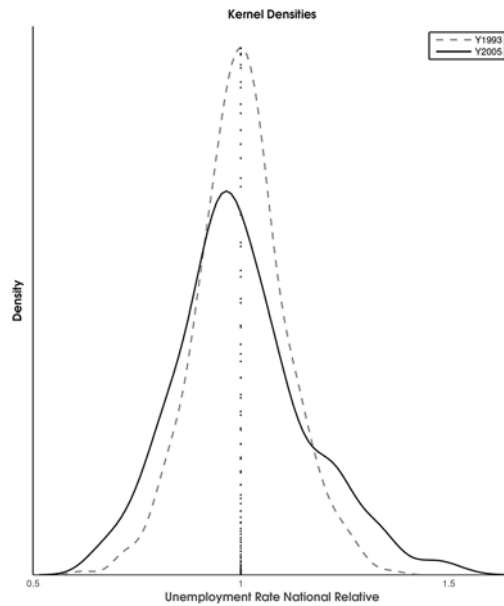
The distribution of unemployment rates in Colombian municipalities has become more uneven. Figure 1 plots the kernel estimates of the density for relative unemployment rates, which are defined as the ratio of the municipal unemployment rate to the national average unemployment rate.¹⁰ The dotted line shows the distribution in 1993, while the solid line shows it in 2005. Note that the line (at 1.0) on the horizontal axis indicates the average unemployment rate. The height of the curve at any point gives the density that any

9 I use a Delaunay triangulation, which is a mesh of non-overlapping triangles created from municipalities' centroids; municipalities associated with triangle nodes that share edges are neighbors. The Spatial Weight Matrix is row normalized.

10 Overman and Puga (2002) use this methodology to evaluate employment clusters across European regions. They argue that using relative unemployment rates helps remove co-movements due to business cycles and trends in the average unemployment rate.

particular municipality will experience that relative rate. It is evident that over time more municipalities have experienced unemployment rates below the average, or above 1.2 times the average, and fewer municipalities have unemployment rates close to national levels.

Figure 1. Municipal Unemployment



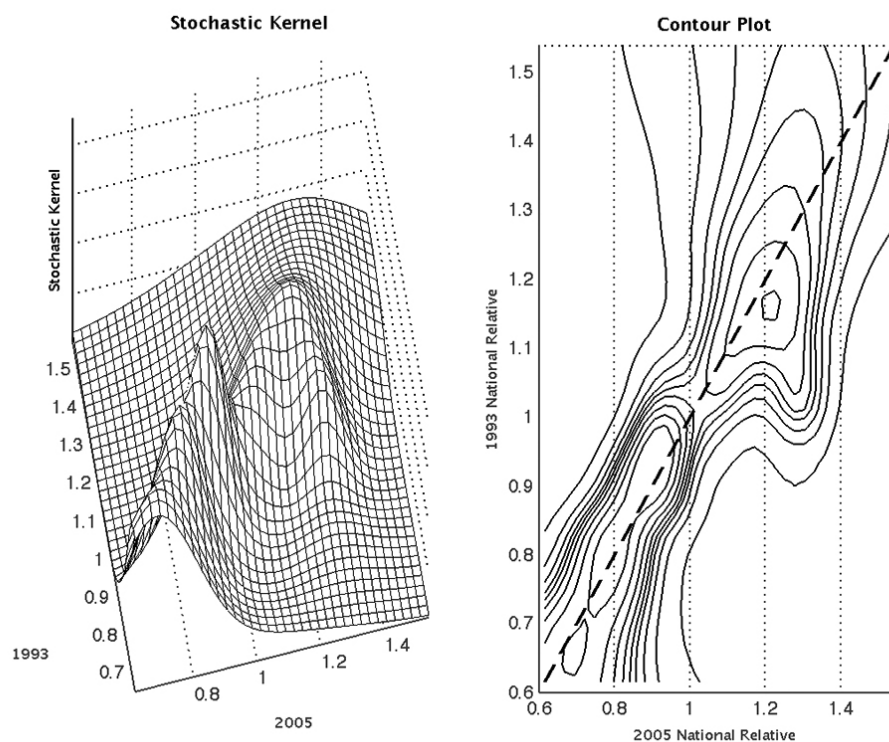
Source: Author's calculations, taken from IPUMS data.

To explore whether the differences in the shapes of these two distributions indicate a structural process, we need to follow the evolution of each municipality's relative unemployment rate over time. To do so we use a stochastic kernel, which provides the likelihood of transiting from one place in the range of values of relative unemployment rates to another (Durlauf & Quah, 1999; Lucas, Prescott & Stokey, 1989).¹¹ Figure 2's left panel plots the transition kernel from the 1993 distribution to the 2005 distribution of the national relative unemployment rates. It provides evidence about the shape of and the mobility within the dynamic distribution. The horizontal axes (for

¹¹ The stochastic kernel is the counterpart of a first-order Markov probability of transition matrix where the number of states tends to infinity. For a formal definition and some properties of stochastic kernels in the study of distribution dynamics, see Durlauf and Quah (1999).

1993 and 2005) show the relative unemployment rates, with 1.0 representing the National average. The vertical axis measures the density function. In terms of the shape, the key issue is to explore whether or not the stochastic kernel has clear peaks. For example, the presence of more than one peak provides evidence of cluster creation. Moreover, polarization is suggested if it were associated with a decline in the middle of the distribution. The plot on the right panel shows a two dimensional contour plot of the three dimensional plot. Lines on the contour plot connect points at the same height on the three-dimensional plot. A 45 degree line is drawn to show where all mass should be concentrated if there was complete persistence in the distribution.

Figure 2. Polarization Evidence from a Stochastic Kernel



Note: Calculations were carried out using Matlab routine developed by Magrini (2007).

The twin-peak nature of Figure 2 confirms that there has been a polarization of unemployment rates. That is, municipalities that had relatively low unem-

ployment rates in 1993 tended to maintain or reduce their relatively low unemployment rate over the next 12 years. Similarly, municipalities that in 1993 exhibited relatively high unemployment rates continued this path until 2005. However, municipalities with intermediate unemployment rates were unlikely to remain in this situation; most experienced their relative rates either increase or decrease.¹²

So far the analysis ignores the spatial distribution of unemployment rates. To explore the role of geography in the unemployment distribution we estimate a Moran's I statistic. This test is a summary measure of spatial correlation that assesses the degree of similarity or dissimilarity of values in spatially close areas¹³. Table 1 shows the estimated Moran's I statistic and its associated significance level for unemployment rates in 1993 and 2005, as well as the difference between these two years.

Results show a high positive spatial correlation for raw unemployment rates. Positive autocorrelation implies that municipalities with relative high (low) unemployment rates are located close to other municipalities with relative high (low) unemployment rates. There is also evidence that the geographic distribution of unemployment in Colombia has become more clustered over time, since the Moran's I increases over time. Thus, we can argue that while municipalities had different unemployment patterns from the national average, they had very similar unemployment outcomes to those of their neighbors, suggesting the creation of unemployment clusters across the territory¹⁴.

To explore whether the determinants of unemployment show similar spatial patterns to the unemployment rates, I estimate Moran's I test for each variable. Table 1 displays the results for each variable using the contiguity matrix as a proxy for spatial proximity. Results confirm that the determinants of

12 Additionally, the concentration of unemployment rates, measured by the Theil coefficient, rose from 0.056 in 1993 to 0.123 in 2005. The Theil index is measured as $TC_t = \sum_{i=1}^N U_{it} \log(U_{it} / WP_{it})$ where U_{it} is the municipal share of unemployment and WP_{it} represents the working population in year t .

13 The Moran's I-Statistic is defined as: $I = \varepsilon' W_c \varepsilon / \varepsilon' \varepsilon$. Where ε represents the residuals from regressing each variable on a constant (i.e., $y_i = \iota + \varepsilon_i$), and W_c is the spatial weight matrix.

14 Figure A1.1 in the Appendix, shows the Local Indicators of Spatial Association for 1993 and 2005. The overall pattern depicts a cluster of high unemployment rates in the north of the country that remains stable over time. The pattern of unemployment rates for the center and south is not uniform across time.

unemployment are positively correlated in space since the Moran's I is significant and different from zero. The difference between both years' values also presents positive autocorrelation.

Table 1. Spatial Autocorrelation of Local Unemployment Rate and Explanatory Variables

Variable	1993		2005		Difference 2005-1993	
	MI	p1	MI	p1	MI	p1
Unemployment	0.43	0.00	0.62	0.00	0.35	0.00
Explanatory Variables						
Local Dynamics						
Employment	0.18	0.00	0.15	0.00	0.17	0.00
Migration	0.37	0.00	0.45	0.00	0.19	0.00
Non demographic labor market						
Ind Divers.	0.09	0.00	0.07	0.01	0.06	0.02
Ind Svs	0.16	0.00	0.07	0.00	0.10	0.00
Ind Manu	0.25	0.00	0.10	0.00	0.12	0.00
Human Capital						
College Share	0.25	0.00	0.32	0.00	0.17	0.00
Demographic						
Age 15 24	0.18	0.00	0.42	0.00	0.23	0.00
Age 55 64	0.15	0.00	0.32	0.00	0.08	0.00
Fem Married	0.40	0.00	0.44	0.00	0.11	0.00
Fem Married with children	0.30	0.00	0.53	0.00	0.12	0.00
Municipality Attributes						
Urban	0.26	0.00	0.39	0.00	0.26	0.00
Pop Density	0.19	0.00	0.20	0.00	0.20	0.00

Note: MI represents Moran's I test, which is calculated as $I = \frac{e'W_e e}{e' e}$. Where e represents the residuals from regressing each variable on a constant, and W_e is the spatial weight matrix. P1 is the p-value based on a standardized z-value that follows a normal distribution.

Source: Author's calculations, based on IPUMS data.

As neighboring municipalities are similar, the similarity of unemployment rates could simply be driven across borders to neighbors. We re-estimate the Moran's I for unemployment rates being conditional on all variables in order to explore their influence on the spatial association of unemployment rates. If clusters of unemployment are only driven by neighboring attributes, then we

should not only observe positive spatial correlation of unemployment determinants but also after conditioning on the entire set of covariates; the spatial correlation for unemployment rates should diminish considerably, and eventually disappear. Table 2 shows the results from this exercise where the first column depicts the unconditional and the second the conditional Moran's I for unemployment rates for each year and their respective difference. After conditioning on these covariates the spatial correlation diminishes, especially in 2005. However, it does not disappear, which suggests that unobservable attributes still affect unemployment clustering.

Table 2. Unconditional and Conditional Moran's I test

	1993		2005		Difference	
	Uncond. (1)	Cond. (2)	Uncond. (1)	Cond. (2)	Uncond. (1)	Cond. (2)
Unemployment	0.43***	0.25***	0.62***	0.26***	0.35***	0.20***

Note: Column (1) shows the results from the unconditional Moran's I Statistic defined as $I = \epsilon' W_c \epsilon / \epsilon' \epsilon$; where ϵ represents the residuals from regressing unemployment rates on a constant and W_c is the spatial weight matrix. Column (2) shows the results from the conditional Moran's I Statistic defined as $I = u' W_c u / u' u$; where u represents the residuals from regressing unemployment rates on a constant and the set of explanatory variables $y_i = \iota + x_i + u_i$ and W_c is the spatial weight matrix. Inference is again based on a standardized z-value that follows a normal distribution.

***p < 0.01, **p < 0.05, *p < 0.1.

Source: Author's calculations, based on IPUMS data.

Interesting stylized facts arise from this explanatory analysis. First, Colombian municipalities have experienced a polarization in their unemployment rates between 1993 and 2005. Second, the unemployment outcomes of individual municipalities have closely followed those of their neighbors, creating clusters of low and high unemployment. Third, the potential determinants of unemployment rates also present a strong spatial correlation. Fourth, the neighbors effect remains strong, even after controlling for similarities in municipal attributes. This suggests that there might still be omitted variables that affect unemployment spatial patterns in Colombian municipalities.

III. Empirical Strategy

This section explains the empirical parametric strategy that will be used to assess the main determinants of the evolution of municipal unemployment rates. In this section, we modify the simple framework proposed by Marston

(1985) to include spatial dynamics. Under Marston's framework we can assume that municipalities are in spatial equilibrium, but are disturbed by a shock, which redistributes labor demand across areas. As a result of the shock, in the short-term, the labor market adjusts by having either changes in the unemployment rate or changes in wages. In the long-term workers might move from high unemployment regions to low unemployment regions (or may enter into non-participation in the workforce); second, firms migrate into high unemployment regions attracted by the large unemployment labor force; third, wages tend to fall due to the excess supply of labor. However, economic and social barriers might impede the adjustment in the short-term, but in the long-term unemployment rates should converge to their spatial equilibrium rates. Thus, we can easily characterize the municipal unemployment rate as having a combination of both equilibrium and disequilibrium components:

$$y_{it} = \sum_{k=1}^K \beta_k X_{k,it} + \alpha_i + \alpha_t + \epsilon_{it} \quad (1)$$

Where i represents municipalities and $t = \{1993, 2005\}$, y_{it} is the log unemployment rate for each municipality i at time t , X_{it} is a matrix of explanatory variables in logs, α_i represents municipal fixed effects, α_t represents the time effects, and ϵ_{it} represents the disturbance term (*i.e.*, disequilibrium component). We can use first-order differences to remove the municipal and fixed time effects and any potential bias arising from it,¹⁵

$$\Delta y_i = \alpha_0 + \sum_{k=1}^K \beta_k \Delta X_{k,i} + \Delta \epsilon_i \quad (2)$$

Where $\Delta y_i = y_{it} - y_{it-1}$, $\Delta X_{k,i} = X_{k,it} - X_{k,it-1}$, $\Delta \epsilon_i = \epsilon_{it} - \epsilon_{it-1}$, and α_0 is the constant term. Expression (2) represents a linear model with independent and identically distributed disturbances. However, the former section showed that disturbances are not independent and identically distributed, even after conditioning on observables. This suggests that there may be spatially correlated time-variant variables that are omitted (*e.g.*, propensity for interregional trade,

15 Colombian municipalities differ widely in geographical location, weather, amenities, and market access. Municipalities where productivity of skilled workers is particularly high, resulting from these attributes, might have lower unemployment rates. Therefore, the first differences model controls for unobserved heterogeneity on the municipal level.

agglomeration economies, and transportation). We can represent the spatial correlation in the following way:

$$\Delta\epsilon_i = \rho W_c \Delta\epsilon_i + \Delta v_i \quad (3)$$

Where ρ is a scalar parameter reflecting the strength of spatial dependence in the process governing the time-variant omitted variable, Δv_i is a vector of disturbances that is assumed to be distributed $N(0, \sigma_v^2 I_n)$, W_c is a $n \times n$ spatial weight matrix, based on contiguity. Therefore, each element of $W_c \Delta\epsilon_i$ represents a linear combination of elements for the unobserved municipal attributes that are associated with neighboring municipalities.

Moreover, observed variables in first differences were also correlated in space. For each explanatory variable, we can represent this correlation like:

$$\Delta X_{k,i} = \phi_k W_c \Delta X_{k,i} + \Delta \omega_{k,i} \quad (4)$$

Where ϕ_k is a scalar parameter, $\Delta \omega_{k,i}$ is a vector of disturbances that is assumed to be distributed $N(0, \sigma_v^2 I_n)$, W_c and W_c is a $n \times n$ spatial weight matrix.

If $\Delta X_{k,i}$ and $\Delta\epsilon_i$ are uncorrelated, the least-squares estimates for β_k in expression (2) are unbiased even if both observed and unobserved variables exhibit spatial dependence. It is very unlikely, however, that these variables are uncorrelated. There might be time-variant demand of supply shocks that commonly affect these variables. Following LeSage and Fischer (2008) and LeSage and Pace (2009), we can express this correlation as,

$$\Delta v_{k,i} = \gamma_k \Delta \omega_{k,i} + \Delta u_{k,i} \quad (5)$$

From equations (2), (3) and (5) we obtain

$$\Delta y_i = \alpha_0 t + \sum_{k=1}^K \beta_k \Delta X_{k,i} + (I_n - \rho W_c)^{-1} \left[\sum_{k=1}^K \gamma_k \Delta \omega_{k,i} + \Delta u_{k,i} \right]$$

By including relation (4), and transforming this equation to obtain i.i.d disturbances, premultiplying both sides of equation by $(I_n - \rho W_c)$, yields

$$\Delta y_i = \beta_0 + \rho W_c \Delta y_i + \sum_{k=1}^K \eta_k \Delta X_{k,i} + \sum_{k=1}^K \eta_{2k} W_c \Delta X_{k,i} + \Delta \mu_i$$

Where

$$\beta_0 = (I_n - \rho W_c) \alpha_0, \eta_{1k} = (\beta_k + \gamma_k), \eta_{2k} = (-\rho \beta_k - \phi_k \gamma_k),$$

and

$$\Delta \mu_i = \sum_{k=1}^K \Delta \mu_{ik},$$

which is distributed $N(0, \sigma_\mu^2)$. In matrix form, this can be written as,

$$\Delta y_i = \beta_0 + \rho W_c \Delta y_i + \eta_1 \Delta X_i + \eta_2 W_c \Delta X_i + \Delta \mu_i \quad (6)$$

This expression represents what has been labeled spatial Durbin model (SDM). This model includes a spatial lag of the dependent variable, $W_c \Delta y_i$, which captures spatial effects working through the dependent variable; ρ is the scalar parameter that reflects spatial dependence, which is expected to be positive in our model. This indicates that unemployment rates are positively related to a linear combination of neighboring unemployment rates, as it was shown in the data description. The model also includes the explanatory variables in differences, ΔX_i , and a spatial lag of the explanatory variable, $W_c \Delta X_i$. Finally, $\Delta \mu_i$ is the error term that is assumed to be $\Delta \mu_i \sim N(0, \sigma_\mu^2 I_n)$. Coefficient estimates on the spatial lag of the explanatory variables capture two types of spatial relationships: spatial effects working through the unemployment rate and spatial effects working through the explanatory variables.¹⁶

This method has some advantages over those previously presented in the literature. First, SDM allow us to consistently estimate the effect of the explanatory variables when endogeneity is induced by omitting a spatially autoregressive variable¹⁷. Second, the model let us quantify the magnitude of spillover effects

16 Estimation was carried out via maximum likelihood using Matlab routines developed by LeSage (1999).

17 Although SDM reduces the problem of omitted variable bias, the results might still be biased and inconsistent due to simultaneity and/or measurement error. The presence of endogenous regressors and plausible measurement errors inevitably leads to 2SLS, which is known to be a consistent estimator. Although not impossible, it is very difficult to find appropriate instruments for each variable on the right hand side. This is left for future research.

arising from both dependent and independent variables. We use the methodology proposed by LeSage and Pace (2009) to estimate summary measures of direct, indirect, and total spatial effects¹⁸.

We can assess the relative importance of each regressor with respect to its overall effect on the change in municipal unemployment rates. Here, we propose a simple decomposition exercise to achieve this goal. To start with, please note that equation (6) can be expressed as:

$$\Delta y_i = (I_n - \hat{\rho} W_c)^{-1} [\hat{\alpha}_0 \iota + \hat{\eta}_1 \Delta X_i + \hat{\eta}_2 W_c \Delta X_i + \Delta \hat{\mu}_i] \quad (8)$$

Where $\hat{\alpha}_0, \hat{\rho}, \hat{\eta}_1, \hat{\eta}_2$ are the maximum likelihood estimates of equation (6), shown in Table 3, and $\Delta \hat{\mu}_i$ is the error in predicting the value of Δy_i , given the value of ΔX_i . Equation (8) can be rewritten as:

$$\Delta y_i = \tilde{\alpha}_0 + (\Delta y_i - \Delta y_i^1) + \Delta \hat{\mu}_i^* \quad (9)$$

Where the first term on the right hand side is the constant term, taking into account the feedback effects of the neighbor's unemployment (i.e., $\tilde{\alpha}_0 = (I_n - \hat{\rho} W_c)^{-1} [\hat{\alpha}_0 \iota]$); The second term is the difference between the observed unemployment growth, Δy_i , and unemployment growth if we assume that none of the explanatory variables changed

18 Partial derivatives from equation (6) take the form of a $n \times n$ matrix:

$$\frac{\partial \Delta y}{\partial \Delta x} = (I_n - \rho W_c)^{-1} (I_n \eta_{1k} + \eta_{2k} W_c) = S_k(W_c) \quad (7)$$

The *average direct effect* is the average of the diagonal elements of the matrix $S_k(W_c)$. This measure summarizes the impact of changes in the i th municipality of variable k using an average across municipalities. For example, if a municipality increases its human capital, the average direct effect accounts for the localized effect and feedback effects, where municipality i affects municipality j and municipality j also affects observation i . The *average indirect effect* is the average of the row-sums of the matrix elements, which corresponds to cross-partial derivatives. This summary impact measure reflects the impacts that take place in other municipalities. It is important to stress that indirect impacts will often exceed the direct impacts because the scalar summary measures cumulative impacts over all regions in the model and not marginal effects. Finally, the *average total effect* is the sum of the direct and indirect impacts.

Survey estimates that the small geographical units will often exhibit high variability. Recent advances in small area estimation incorporating both explicit spatial autocorrelation and empirical likelihood techniques have produced estimates with greater precision. Those methods constitute an interesting topic for future research.

between 1993 and 2005, $\Delta y_i^1 = (I_n - \hat{\rho} W_c)^{-1} [\hat{\alpha}_0 + \Delta \mu_i]$, thus, this term equals to $\Delta y_i - y_i^1 = (I_n - \hat{\rho} W_c)^{-1} [\hat{\eta}_1 \Delta X_i + \hat{\eta}_2 W_c \Delta X_i]$; the last term is $\Delta \mu_i^* = (I_n - \hat{\rho} W_c)^{-1} \Delta \mu_i$.¹⁹

By rearranging terms, dividing both sides by $\Delta \tilde{y}_i = \Delta y_i - \hat{\alpha}_0$, and summing over i , we have

$$1 = \frac{1}{N} \sum_{i=1}^N \left[\frac{\Delta y_i - \Delta y_i^1}{\Delta \tilde{y}_i} \right] + \frac{1}{N} \sum_{i=1}^N \left[\frac{\Delta \mu_i^*}{\Delta \tilde{y}_i} \right] \quad (10)$$

The first part of the right hand side is the part that is explained by the change of the explanatory variables, and the second part represents the change induced by the spatial correlation of the residual component: remember that $\Delta \mu_i^* = (I_n - \hat{\rho} W_c)^{-1} \Delta \mu_i$. Please also note also that having $\Delta \mu_i^*$ rather than $\Delta \mu_i$ in equation (10) allows for the second fraction to be non-zero.

IV. Results

This section reports the empirical investigation's results. First, the results from ordinary least squares are compared to those from the spatial Durbin model. Second, the specification that best fits the data is selected and the results are discussed. Third, a decomposition exercise is proposed to learn the relative importance of different factors. The purpose is to assess how much of the change in unemployment rates is explained by each explanatory variable and their spatial lags, and how much is induced by the spatial correlation of the residual component. While the regression analysis helps us to understand which independent variables are related to the change in unemployment rates, and to explore the forms of these relationships, the decomposition analysis helps us to understand the relative influence of each factor, explaining the outcome variable. Finally, this section presents a simulation exercise in which the new unemployment equilibrium values for each municipality are calculated after a change in a single explanatory variable based on different scenarios.

19 Note that $\Delta \mu_i$ should be uncorrelated across space (parallel to what is assumed about the error component $\Delta \mu_i$), while $\Delta \mu_i^*$ should indeed be correlated across space.

Table 3 compares ordinary least squares results with the spatial Durbin model ones. The first column presents the results from an ordinary least square regression, assuming that disturbances are independent and identically distributed. The second and third columns present the results from the SDM in which spatial lags of both dependent and independent variables are included. The dependent variable is the difference in municipality I's unemployment rate between 1993 and 2005. To be consistent with the exploratory evidence that was described in Section II, I use the same set of explanatory variables in first differences.

Before interpreting the results, let me discuss the best specification selection. Results from both LM SEM and LM SAR tests reject the null hypothesis of no spatial correlation in the model's residuals. These results indicate that OLS residuals, without controlling for the spatial lag of the unemployment rate, (in the SEM model) or controlling for it (in the SAR model), are spatially cor

Table 3. Parameter Estimates from OLS and SDM

	OLS	SDM	
	β_k	η_{1k}	η_{2k}
Local dynamics			
Employment	-0.06 [0.02]***	-0.02 [0.02]*	-0.12 [0.04]***
Migration	-0.14 [0.05]***	-0.08 [0.04]**	-0.20 [0.08]***
Non demographic labor market			
Ind Diversity	0.04 [0.02]**	0.03 [0.02]**	-0.01 [0.02]
Ind Svs	0.01 [0.02]	-0.01 [0.03]	0.02 [0.03]
Ind Manu	-0.04 [0.04]	-0.06 [0.03]*	0.07 [0.06]
Human capital			
College Share	-0.32 [0.08]***	-0.30 [0.07]***	0.16 [0.16]
Demography			
Age 15 24	-0.18 [0.05]***	-0.14 [0.05]***	0.01 [0.00]
Age 55 64	0.15 [0.11]	0.02 [0.08]	0.07 [0.23]

(Continued)

Table 3. Parameter Estimates from OLS and SDM

	OLS β_k	SDM η_{1k} η_{2k}	
Fem married	-0.23 [0.07]***	-0.16 [0.06]***	0.03 [0.14]
Fem married with children	0.32 [0.07]***	0.12 [0.07]*	0.22 [0.08]***
Municipality attributes			
Urbanization	0.05 [0.02]***	0.01 [0.01]	0.05 [0.03]
Pop density	-0.03 [0.10]	-0.06 [0.09]	0.29 [0.19]
Unemployment			
Neighbors' unemployment (ρ)		0.30 [0.06]***	
Constant	0.03 [0.01]***	-0.01 [0.03]	
No. Observations	497	497	
R2 0.21 0.35			
Model Specification Tests			
LM SEM	55.99		
LM SEM p-value	0.00		
LM SAR	70.87		
LM SAR p-value	0.00		
SDM Log Likelihood			972.63
SAR Log Likelihood			954.75
SEM Log Likelihood			948.09

Note: The dependent variable is the change in unemployment rates between 1993 and 2005; the independent variables are in first differences. Standard errors are in brackets. The weight matrix used for SDM takes the form of a binary first-order contiguity matrix, W_c , in which only direct interaction between geographically neighboring regions is allowed for. Two regions are defined as neighbors when they share a common boundary. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: Author's calculations, based on IPUMS data.

related. Ordinary least squares estimates might, therefore, lead to inconsistent and/or inefficient parameter estimates.²⁰

20 OLS estimates will be inconsistent and inefficient if there are omitted variables correlated with independent variables. Moreover, even in the absence of correlation between omitted variables and

It is therefore important to evaluate whether the spatial Durbin model is the best spatial specification. Please remember that the SDM nest most models used in the spatial econometrics literature are the spatial autocorrelation model (SAR) and the spatial error model (SEM). In both cases the best specification is the spatial Durbin model.²¹ In other words, both observed and unobserved explanatory variables exhibit spatial dependence. More over, both observed and unobserved variables are correlated by common spatial correlated shocks. This implies that spatial effects are substantive phenomena rather than random shocks diffusing through the space. Due to this, the preferred specification is SDM.²² As emphasized in the previous section, correct interpretation of the parameter are based on 10.000 sampled raw parameter estimates of the SDM. The weight matrix takes the form of a binary first order contiguity matrix.

If we consider the *average direct impacts*, it is important to notice that they are close to the SDM model coefficient estimates reported in Table 3. Given that the direct impact estimates and the model estimates of the non-spatial lagged variables are similar in most cases, we can conclude that feedback effects between urban areas are negligible. The estimates show interesting features that are consistent with the empirical literature that analyzes unemployment rates in different regions and countries. First, there is evidence that employment growth, migration flows, and the share of employment in the manufacturing sector are negatively related to unemployment growth rates on the municipal level. Second, the evolution of the working age population with high skills is negative related to the unemployment growth rate, and significant. Third, the variables proxying for demographic structure are also correlated with unemployment, especially those for female labor participation. Fourth, urbanization variables seem to be positively related to unemployment growth rates. Finally, the parameters for industry diversity, share of individuals

independent variables, OLS estimates remain unbiased, but are no longer efficient. In the presence of spatial error dependence, standard error estimates will be biased downward, producing Type I errors Anselin (1988).

- 21 The LRT for SAR versus SDM equals to 35.76, which is chi-squared distributed with 12 degrees of freedom, and the associated p-value is 0.001. The LRT for SEM versus SDM is equal to 49.08, which is chi-squared distributed with one degree of freedom, and also has an associated p-value which is very low.
- 22 We also estimated a Bayesian heteroscedastic spatial Durbin model to account for nonconstant variance of the error term. Results, available upon request, are very near to those reported here. This implies that plausible heteroskedasticity is not creating an efficiency problem in this case.

in the agricultural sector, share of elder individuals, and population density are not related to unemployment rates after conditioning on the other variables.

Table 4. Direct, Indirect and Total Impact Estimates

Variables	Direct impact	Indirect impact	Total impact
Employment growth	-0.02 [0.01]**	-0.17 [0.05]***	-0.19 [0.06]***
Migration	-0.09 [0.04]***	-0.31 [0.10]***	-0.40 [0.11]***
Ind Diversity	0.02 [0.02]	-0.01 [0.04]	0.02 [0.05]
Ind Svs	-0.01 [0.02]	0.03 [0.05]	0.03 [0.05]
Ind Manu	-0.06 [0.03]**	0.07 [0.09]	0.02 [0.10]
College Share	-0.30 [0.08]***	0.09 [0.21]	-0.21 [0.23]
Age 15 25	-0.14 [0.05]**	-0.05 [0.14]	-0.18 [0.14]
Age 55 64	0.03 [0.10]	0.09 [0.29]	0.12 [0.32]
Fem married	-0.16 [0.06]***	-0.01 [0.18]	-0.16 [0.20]
Fem married wc	0.13 [0.07]**	0.36 [0.07]***	0.49 [0.09]***
Urbanization	0.02 [0.01]**	0.08 [0.04]***	0.09 [0.04]***
Population density	-0.04 [0.03]	0.07 [0.10]	0.03 [0.07]

Note: Results are based on specification SDM from Table 3. Standard errors, in brackets, are based on 10,000 sampled raw parameter estimates of the SDM. The weight matrix takes the form of a binary first order contiguity matrix.

Source: Author's calculations based in IPUMS data.

The average indirect impacts represent the effect each variable has on unemployment other municipalities: general equilibrium effects. In general terms, the indirect effect is larger (in absolute terms) than the spatial lag coefficient from the SDM model. Second, it is also evident that indirect effects are considerably larger than the mean direct impact. To be able to understand this we need to recall that the scalar summary of the indirect effects measures

the cumulative average impact over space that would result from a change in municipal unemployment rates induced by changes in the explanatory variables. They do not represent marginal impacts. For example, the marginal impact from a one percent change in a single municipality's employment rate on each of the other municipalities' unemployment rates may be small, but cumulatively the impact measures -0.19 percent.

Looking at the results, it is evident that for some variables the local effect dominates: the employment share in manufacturing, the skill composition of the labor force, and some demographic factors. This is because a change in any of these variables has a negative effect on municipal unemployment rates, but their effects are confined to the local labor market. Moreover, employment growth, immigration rates, the percentage of females above the age of 15 who are married and have children under the age of five, and urbanization all have important localized effects (measured by direct effects) and general equilibrium effects (measured by indirect effects). The presence of both direct and indirect effects implies that a municipality-specific change in any of these variables does not only affect the respective local labor market, but instead causes a spillover to neighboring municipalities. The induced change of unemployment in neighboring municipalities causes a spillover to adjacent municipalities, including the municipality where the change took place. According to Molho (1995), this process of spatial adjustments continues until a new equilibrium of regional unemployment is reached.

Total impact estimates measure the sum of the direct and indirect impacts from the previous two columns. From these estimates we can see some surprising results that take into account both direct and indirect impacts. This leads to a total impact that is not significantly different from zero the share of manufacturing, the college share, the share of working age population aged between 15 and 25 years old, and the percentage of females that are married.²³ Moreover, the average total impact for employment growth, migration, the proxy for female labor participation, and urbanization remain significant and their effects are in line with those discussed before.

23 To understand these results it is important to remember that these effects represent the average total impact on a given observation from a change in all municipalities. For example, changing the share of individuals with higher education level in all municipalities has little or no total impact on the unemployment rate of a typical municipality. The intuition here arises from the notion that there are relative advantages in the variables that matter most in terms of reducing unemployment in a given municipality.

To summarize, the findings of this section allow us to understand which of the independent variables are related to the change in unemployment rates, and to explore the forms of these relationships. Differences across municipalities in labor demand, immigration, sectorial specialization, level of education, and urbanization are factors behind observed municipal unemployment disparities. These results are consistent with those of Overman and Puga (2002) and Cracolici *et al.* (2007) for European regions and Italy respectively. The empirical results also make it clear that some characteristics of neighboring municipalities play an important role in determining unemployment rates. For example, municipalities neighboring municipalities with high employment growth were more prone to have better labor outcomes. Immigration seems to play a self-equilibrating role in reducing municipal disparities, as predicted by Burridge and Gordon (1981), Blanchard *et al.* (1992), and Molho (1995). Conversely, municipalities neighboring municipalities with a high share of women married with children under the age of five, and that are highly urbanized are more likely to have higher unemployment rates.

A. Decomposition

Using the information from Table 3 we can evaluate how much of the change in unemployment rates is explained by the explanatory variables and how much is induced by the spatial correlation of the residual component. According to the results, in Table 5, 88.4 percent of the change in unemployment rates is explained by both the explanatory variables and their spatial lags. The unexplained part corresponds to 11.6 percent of the variation in unemployment rates. In addition, we can evaluate the relative importance of each variable and its spatial lag by decomposing the first term of equation (10). To do so, we only need to set the coefficient for each explanatory variable k of interest to zero, estimate the unemployment difference under that scenario (*i.e.*, re-estimate Δy_i^1), and recalculate the first term of equation (4.10).

Table 5 presents the results: The first column shows the percentage of the change in unemployment rates that is explained as a result of local effects of each dependent variable (*i.e.*, Δy_i^1 is evaluated at $\eta_{1k} = 0$), the second column shows the percentage explained by their spatial lags (*i.e.*, Δy_i^1 is evaluated at $\eta_{2k} = 0$), and the third shows how much is explained by each variable (*i.e.*,

Δy_i^1 is evaluated at $\Delta X_k = 0$). It is evident from the results that 22.3 percent of the change in the unemployment rate is explained by the explanatory variables, while 66.1 percent is explained by their spatial lag. The results also indicate that the overall situation is best characterized by several variables, each contributing a certain element, rather than there being a single dominant explanatory variable. However, among these variables, it is clear that employment growth, migration, and urbanization explain 67.7 percent of the variation in unemployment. The largest contribution to unemployment comes from employment growth, which accounts for 32.1 percent, the next largest from migration, which accounts for 23.2 percent, and the smallest from urbanization, the weight of which is 12.3 percent. It can be seen, however, that most of these percentages are explained by these variables' spatial lags.

Table 5. Decomposition

Variables	η_{1k} (%)	η_{2k} (%)	X_k (%)	$\Delta \mu_i^*$ (%)
Total	22.3	66.1	88.4	11.6
Employment growth	3.8	28.4	32.1	
Migration	8.9	14.3	23.2	
Ind Diversity	-6.6	2.3	-4.3	
Ind Svs	-1.8	7.2	5.4	
Ind Manu	5.0	-3.6	1.4	
College Share	-4.7	6.1	1.4	
Age 15 25	9.1	-0.1	8.9	
Age 55 64	1.3	1.3	2.5	
Fem married	10.2	-0.5	9.7	
Fem married wc	-4.1	-2.4	-6.4	
Urbanization	2.4	9.9	12.3	
Population density	-1.1	3.1	2.0	

Note: The first column presents the results of $\Omega_k = \frac{1}{N} \sum_{i=1}^N \left[\frac{\Delta \mu_i^*}{\Delta y_i} \right]$ when Δy_i^1 is evaluated at $\eta_{1k} = 0$, the second column displays the results of Ω_k when Δy_i^1 is evaluated at $\eta_{2k} = 0$, the third column presents the results of Ω_k when Δy_i^1 is evaluated at $X_k = 0$, the last column shows the results from $\frac{1}{N} \sum_{i=1}^N \left[\frac{\Delta \mu_i^*}{\Delta y_i} \right]$

Source: Author's calculations based on IPUMS data.

B. Simulation

We can use the estimated results to calculate the unemployment equilibrium values for each municipality after a change in a single explanatory variable, *i.e.*, the expected values given the model. We create a simple simulation in which each variable is assumed to increase 10 percent in some municipalities, while all the other variables are held constant. We calculate four relevant measures: a) the number of observations that are affected through the system of interactions, b) the difference in the expected unemployment rate under this scenario versus the expected value given the model and the observed data, c) an inequality effect measured by the percentage change in the Theil index, and d) an agglomeration effect calculated as the percentage change in the Morans' I test for unemployment rates. The Table 6 compares these measures for seven different scenarios that differ in the number of tested municipalities: Scenario 1 assumes that the analyzed variable increases 10 percent in all municipalities, Scenario 2 assumes that the change takes place only in the capital of each Department, and Scenarios 3 to 6 compare the results when municipalities are grouped in quartiles according to the initial unemployment rate.²⁴

The first panel in Table 6 presents the results for employment growth. It is evident that if all municipalities experienced a higher than 10 percent increase in the employment growth, then the expected unemployment rate would be 1.9 percentage points lower, the distribution across the space would be 7 percent more unequal, and the spatial correlation for unemployment rates would not change. When we assume that only each department's main cities, 31 observations, experience a change in the employment growth, then 476 municipalities are finally affected through the system of interactions between municipalities, the new equilibrium unemployment is 0.12 percentage points lower, inequality remains at the same level, and unemployment becomes slightly more agglomerated in space.

24 Scenario 3 assumes that the change takes place in the 25 percent of municipalities that exhibited lower unemployment rates in 1993, Scenario 4 modifies those between the 25th and the 50th percentile, Scenario 5 those between the 50th and 75th percentile, while Scenario 6 assumes that it takes place in the 25 percent of municipalities that exhibited higher unemployment rates in 1993.

Table 6. Simulation, Selected Variables

Variable	Scenario	Obs Treated (1)	Obs Affected (2)	Unemp. Change (3)	Pol. Effect (4)	Aggl. Effect (5)
Employment	1. All municipalities	497	497	-1.92	7.38	0
	2. State capital	31	476	-0.12	0	0.17
	3. Lower 25 percent	124	460	-0.47	10.66	12.84
	4. 25th-50th percentile	125	485	-0.49	4.1	6.8
	5. 50th-75th percentile	124	497	-0.48	0.82	1.68
	6. Higher 25 percent	124	488	-0.48	-8.2	-21.15
Migration	1. All municipalities	497	497	-3.97	16.39	0
	2. State capital	31	483	-0.25	1.64	0.17
	3. Lower 25 percent	124	468	-0.98	24.59	24.65
	4. 25th-50th percentile	125	491	-1.01	9.84	13.77
	5. 50th-75th percentile	124	497	-0.99	1.64	5.09
	6. Higher 25 percent	124	489	-0.99	-17.21	-42.23
College Share	1. All municipalities	497	497	-2.11	8.2	0
	2. State capital	31	467	-0.13	4.92	-4.05
	3. Lower 25 percent	124	452	-0.53	22.95	1.04
	4. 25th-50th percentile	125	481	-0.52	8.2	-3.24
	5. 50th-75th percentile	124	497	-0.52	-0.82	1.19
	6. Higher 25 percent	124	476	-0.52	-15.57	-14.38
Urbanization	1. All municipalities	497	497	0.92	-3.28	0
	2. Province capital	31	464	0.06	-0.82	0.03
	3. Lower 25 percent	124	450	0.23	-4.92	-5.87
	4. 25th-50th percentile	125	481	0.24	-2.46	-2.89
	5. 50th-75th percentile	124	497	0.23	-0.82	-0.61
	6. Higher 25 percent	124	475	0.23	4.1	9.46

Note: Column (1) presents the number of municipalities that exhibit a 10 percent increase in the variable analyzed, Column (2) presents the number of municipalities that exhibit a change in their unemployment rates, Column (3) presents the change in unemployment rate, that is $\frac{1}{N} \sum_{i=1}^N [\Delta y(x_k)_i - \Delta y_i]$ where $\Delta y(x_k)_i$ is the simulated unemployment rate and Δy_i is the observed unemployment rate. Column (4) shows the polarization effect, which is the percentage change between the simulated Theil index and that of the observed data (i.e., Theil: 0.12), Column (5) presents the agglomeration effect, which is the percentage change between the simulated spatial correlation test and that of the observed data (i.e., 0.35).

Source: Author's calculations based on IPUMS data.

Particularly interesting results are those from Scenarios 3 to 6 in which we assume that the employment change takes place on a subset of municipalities, which are defined on the basis of the lower, median, and upper quartiles of the initial cross-country unemployment rate distribution. Note that the new equilibrium unemployment rate is the same in all scenarios (0.48 percentage points lower), but important differences in both inequality and agglomeration measures arise. In fact, if the municipalities that did better (*i.e.*, had lower unemployment rates) in 1993 face an additional increase in the employment growth of 10 percent, the distribution of unemployment rates becomes more unequal and more agglomerated. Conversely, if the 10 percent increase in employment growth takes place in those municipalities that did worse (*i.e.*, had higher unemployment rates), the distribution of unemployment rates throughout the country becomes more equal and less agglomerated.

Results for migration and college share are in line with those from employment growth, whereas the results for urbanization work in the opposite direction. In fact, an increase in the urbanization rate of one percentage point in the whole sample increases the unemployment rate by 0.92 percentage points, the distribution of unemployment rates becomes more equal, while the spatial correlation does not change.

This simulation exercise shows some interesting features concerning both the aggregate and spatial distribution of unemployment rates when some of the explanatory variables change. First, policies aimed at increasing employment, migration or educational levels in all municipalities have large effects on the aggregate unemployment rate at the cost of higher spatial polarization. Second, localized interventions can lead to different spatial outcomes depending on the targeted area. Changes in some areas can have differential effects on the spatial distribution of unemployment rates (making them more equal or unequal across space) and in the creation of clusters of municipalities with high and low unemployment rates. In this particular case, increasing employment, migration, and the share of individuals with some college or more in those municipalities that were bad performers in 1993 led to a reduction in aggregate unemployment rates, accompanied by a reduction in both the spatial inequality and the spatial agglomeration.

V. Conclusion

This article increases understanding of the differences in unemployment rates across Colombian municipalities. Using municipal data at an urban level for 1993 and 2005, this paper shows that Colombian municipalities are characterized by diverging unemployment rates; this is a type of polarization process in which municipalities are moving away from the national average. This process has been accompanied by a clustering effect of the unemployment rate since municipalities with high (low) unemployment rates seem to be surrounded by municipalities with high (low) unemployment rates. Moreover, variables that might affect the evolution of unemployment rates exhibit the same spatial patterns. This suggests that the spatial evolution of unemployment rates is the result of different types of municipalities, in terms of economic and socio demographic attributes, that are clustered in space. A simple exploratory analysis confirms that these variables do exert some effect on the spatial evolution of unemployment rates, but even when controlling for similarities in municipal attributes, the neighbors' effect remains strong.

To explore the effect of diverse variables on the evolution of the unemployment rate we use spatial econometric techniques. The approach adopted here uses a unified method for dealing with uncertainty regarding model specification, specifically, the appropriate spatial regression model to be employed. The preferred specification was a spatial Durbin model which allows for two types of spatial interdependencies in the evolution of unemployment rates: spatial effects working through the change in municipal unemployment rates, and the spatial effects working through a set of conditioning variables. Although the results presented in this study are robust, one must be cautious in drawing policy implications. The regression estimated includes variables that might be viewed as simultaneously determined and, therefore, should be interpreted as having a predictive relationship as opposed to a causative one.

The results from this exercise suggest that differences across municipalities in labor demand, immigration, sector specialization, educational attainment, and urbanization are factors behind observed municipal unemployment disparities. The findings also confirm the fact that spatial effects are relevant factors when interpreting municipal disparities in unemployment rates in Colombia. In particular, they show that changes in employment growth, immigration, and urbanization affect not only the local labor market but also their effects spread

into neighboring municipalities. According to the decomposition exercise, these variables explained 67.7 percent of the variation of unemployment. Moreover, these variables' spatial effects account for 52.6 percent of the total variation.

Finally, the simulation exercise illustrates the general equilibrium effects of changing key explanatory variables in different sub-sets of sampled regions. This exercise shows that spatial considerations must be taken into account when using targeted policy to help lift areas out of unemployment. Different unemployment spatial patterns can emerge from interactions between municipalities. In fact, results from this exercise show that uneven spatial distribution of unemployment rates can be reversed by targeting lagging areas instead of leading areas. For example, targeting job creation where unemployment concentrations are high not only helps to reduce aggregate unemployment but also municipal unemployment inequalities. Targeting job creation where unemployment is low leads to larger spatial inequalities while having the same effect on aggregate unemployment.

These results suggest the polarization process could have been mitigated by implementing spatially targeted policies. At the beginning of the 1990s, the unemployment rate in Colombia was low as a response to the macroeconomic stabilization and a more competitive environment induced by structural reforms (Saavedra, 2003). As the economy went into recession in 1999, employment growth stopped, and unemployment rates grew (labor market adjusted through quantities rather than prices, Núñez, 2005). In 2000, the unemployment rate reached 20 percent, and then it fell as a result of modest economic growth and labor market reforms. Indeed, the government started a process of market oriented reforms aiming at increasing employment and spurring macroeconomic stabilization. In 2002 the government introduced a labor reform, which had two main components: i) reducing extra payments for over-time and severance payments, and ii) introducing more flexibility through changes in other labor regulations. These policies did not consider spatial considerations for the unemployment rate. Therefore, the aggregate unemployment rate was reduced but the spatial inequality and the spatial agglomeration were exacerbated.

Acknowledgment

Este artículo se derivó del proyecto PPTA 6484 titulado "Análisis espacial del desempleo en los municipios colombianos".

References

1. Anselin, L. (1988). *Spatial econometrics: Methods and models*. Springer Science & Business Media, New York.
2. Aragon, Y., Haughton, D., Haughton, J., Leconte, E., Malin, E., Ruiz-Gazen, A., & Thomas-Agnan, C. (2003). Explaining the pattern of regional unemployment: The case of the Midi-Pyrénées region. *Papers in Regional Science*, 82(2), 155-174.
3. Arango, L. E. (2011). *Mercado de trabajo de Colombia: suma de partes heterogéneas* (Borradores de Economía 671). Banco de la República.
4. Bande, R., & Karanassou, M. (2009). Labour market flexibility and regional unemployment rate dynamics: Spain 1980-1995. *Papers in Regional Science*, 88(1), 181-207.
5. Basile, R., Girardi, A., & Mantuano, M. (2009). *Regional unemployment traps in Italy: Assessing the evidence* (Working Paper 96). ISAE.
6. Blanchard, O. J., Katz, L. F., Hall, R. E., & Eichengreen, B. (1992). Regional evolutions. *Brookings Papers in Economic Activity*, 1992(1), 1-75.
7. Bradley, S., & Taylor, J. (1997). Unemployment in Europe: A comparative analysis of regional disparities in Germany, Italy and the UK. *Kyklos*, 50, 221-245.
8. Burridge, P., & Gordon, I. (1981). Unemployment in the British metropolitan labour areas. *Oxford Economic Papers*, 33(2), 274-297.
9. Cárdenas, C., Hernández, M. A., & Torres, J. (2014). *An exploratory analysis of heterogeneity on regional labour markets and unemployment rates in Colombia: An MFACT approach* (Borradores de Economía 802). Banco de la República.
10. Cardenas, M., & Bernal, R. (2003). *Determinants of labor demand in Colombia: 1976-1996* (Working Papers 10077). National Bureau of Economic Research, Inc, November.

11. Cracolici, M. F., Cuffaro, M., Nijkamp, P., & delle Scienza, V. (2007). Geographical distribution of unemployment: An analysis of provincial differences in Italy. *Growth and Change*, 38(4), 649-670.
12. Decressin, J., & Fatás, A. (1995). Regional labor market dynamics in Europe. *European Economic Review*, 39(9), 1627-1655.
13. Durlauf, S. N., & Quah, D. T. (1999). The new empirics of economic growth. *Handbook of Macroeconomics*, 1, 235-308.
14. Elhorst, J. P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of Economic Surveys*, 17, 709-748.
15. Elhorst, J. P. (2010). Applied spatial econometrics. *Raising the Bar*, 5(1), 9-28.
16. Epifani, P., & Gancia, G. A. (2005). Trade, migration and regional unemployment. *Regional Science and Urban Economics*, 35(6), 625-644.
17. Filiztekin, A. (2009). Regional unemployment in Turkey. *Papers in Regional Science*, 88(4), 863-878.
18. Galvis, L. A. (2002). *Integración regional de los mercados laborales en Colombia, 1984-2000*. (Documentos de Trabajo sobre Economía Regional 27). Banco de la República.
19. Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, 19(2), 316-342.
20. Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, 60(1), 126-142.
21. Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1), 35-78.

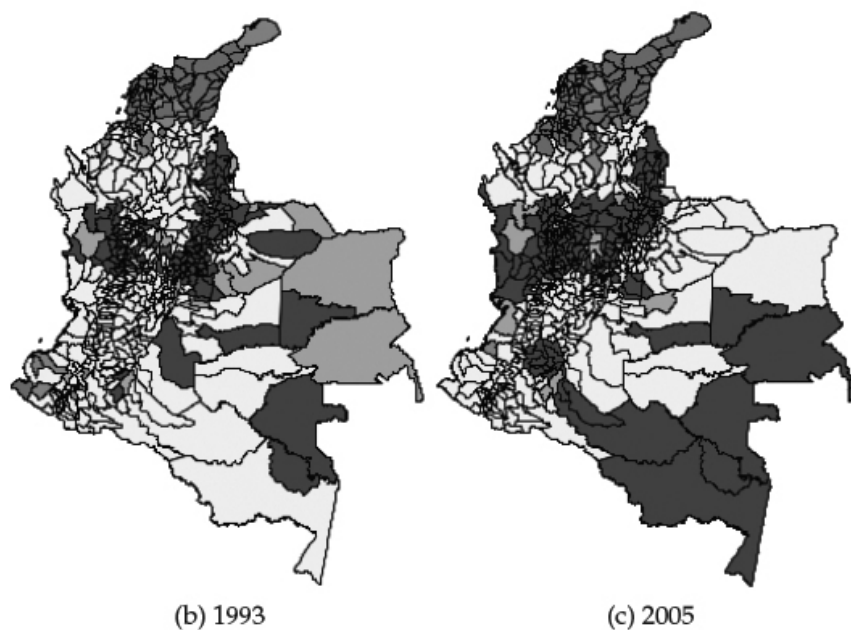
22. LeSage, J. P. (1999). *Spatial econometrics*. Regional Research Institute, West Virginia University, Morgantown, WV.
23. LeSage, J. P., & Fischer, M. M. (2008). Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis*, 3(3), 275-304.
24. LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman & Hall/CRC, Boca Raton, FL.
25. Lopez-Bazo, E., Del Barrio, T., & Artis, M. (2002). The regional distribution of Spanish unemployment: A spatial analysis. *Papers in Regional Science*, 81(3), 365-389.
26. Lopez-Bazo, E., Del Barrio, T., & Artis, M. (2005). Geographical distribution of unemployment in Spain. Regional studies. *The Journal of the Regional Studies Association*, 39(3), 305-318.
27. Lora, E., & Marquez, G. (March 1998). *The employment problem in Latin America: Perceptions and stylized facts* (RES Working Papers 4114). Inter-American Development Bank, Research Department.
28. Lucas, R., Prescott, E. C., & Stokey, N. (1989). *Recursive methods in economic dynamics*. Harvard University Press, Cambridge, MA.
29. Magrini, S. (2007). *Analysing convergence through the distribution dynamics approach: Why and how?* (Working Paper 2007-13). University of Venice, Department of Economics.
30. Manning, A. (2004). We can work it out: The impact of technological change on the demand for low-skill workers. *Scottish Journal of Political Economy*, 51(5), 581-608.
31. Marston, S. T. (1985). Two views of the geographic distribution of unemployment. *The Quarterly Journal of Economics*, 100(1), 57-79.
32. Martin, R., & Morrison, P. (2003). Thinking about the geographies of labour. *Geographies of Labour Market Inequality*, 3-20.

33. Merchán, C. (2014). *Desempleo y ocupación en las ciudades colombianas. Un ejercicio con datos panel* (Technical Report). Departamento Nacional de Planeación.
34. Mincer, J. (1991). *Education and unemployment* (Working Paper 1991). NBER.
35. Molho, I. (1995). Spatial autocorrelation in British unemployment. *Journal of Regional Science*, 35(4), 641-658.
36. Moretti, E. (2004). Human capital externalities in cities. *Handbook of Regional and Urban Economics*, 4, 2243-2291.
37. Niebuhr, A. (2003). Spatial interaction and regional unemployment in Europe. *European Journal of Spatial Development*, 5, 2-24.
38. Núñez, J. (2005). *Success and failure of labor market reform in Colombia* (Technical Report). Banco de la República de Colombia.
39. Overman, H. G., & Puga, D. (2002). Unemployment clusters across Europe's regions and countries. *Economic Policy*, 17(34), 117-147.
40. Partridge, M. D., & Rickman, D. S. (1995). Differences in state unemployment rates: The role of labor and product market structural shifts. *Southern Economic Journal*, 62(1), 89-106.
41. Patacchini, E., & Zenou, Y. (2007). Spatial dependence in local unemployment rates. *Journal of Economic Geography*, 7(2), 169-191.
42. Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography*, 2(4), 373-406.
43. Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, 34(3), 380-400.
44. Saavedra, J. (2003). *Labor markets during the 1990s* (chap. 9, pp. 213-263). Institute for International Economics, Washington, D. C.

45. Saint-Paul, G. (1996). Unemployment and increasing private returns to human capital. *Journal of Public Economics*, 61(1), 1-20.
46. Stevens, B. H., & Moore, C. (1978). *A critical review of the literature on shift-share as a forecasting technique*. Regional Science Research Institute, Amherst, MA.
47. Suedekum, J. (2005). Increasing returns and spatial unemployment disparities. *Papers in Regional Science*, 84(2), 159-181.
48. Vom Berge, P. (2011). Search unemployment and new economic geography. *The Annals of Regional Science*, 50(3), 1-21.

Appendix

Figure A1.1 Local Indicators of Spatial Association



Note: Figure shows the LISA cluster maps. They show the location of significant Local Moran's I statistics, classified by type of spatial association. Grey and dark grey municipalities are indicators of spatial clusters. Grey represents those municipalities with high unemployment rates surrounded by high unemployment rates, and dark grey represents those municipalities with low unemployment rates surrounded by low unemployment rates.

Source: Author's calculations based on IPUMS data.

Table A1.1 Data Description

Variable	Description
Unemployment Rate	The ratio of the non-employed to the working age population. Excludes individuals with physical disabilities to work, persons living from rents, and retired workers.
Employment Growth	The employment growth for municipality m in year t equals $EmpGmt = \left[\frac{\left(\sum gN_{at,t,i} E_{m,t-1,i} \right)}{E_{m,t-1}} \right] - g_{nat,t}$, where $gN_{at,t,i}$ is the national growth rate in industry i , $E_{m,t-1,i}$ is municipality M 's employment in industry i , $E_{m,t}$ is municipality M 's total employment in year $t \in 1$, and $gN_{at,t}$ is the average of national employment growth in year t . The summation is over all two digit sector industries.
Migration Rate	Percentage of the working age population that change of municipality in the last five years.
Industry Diversity	Industry diversity is measured as 1 minus a two digit Herfindahl index, which is $H_c = \sum_{i=1}^N S_{im}^2$, where $S_{im}^2 = \frac{L_{im}}{\sum_{i=1}^N L_{im}}$ and L_{im} is the employment in industry i in municipality m .
Industry Services	Share of labor in services in total employment.
Industry Manu	Share of labor in manufacturing in total employment.
Human Capital	Percentage of the working age population that finished high school and/or college.
Age 15-24 (55-64)	Share of population aged 15-24 (55-64) years in the population in the working age group.
Urbanization	Share of population of the municipality living in the urban area.
Population Density	Ratio population over surface in square kms.

Source: Author's description.

