

The spatial agglomeration of educated people in Colombia

Juan Tomas Sayago*

Data Mining Lab - Università degli studi di Pavia

Grupo de Investigacion en Economia Regional y Urbana - GIERU

February 12, 2009

Abstract

Colombia's educative system has grown in the last decades, and the access and coverage has reached higher levels but this development has been mainly in bigger cities, and causing an agglomeration of graduates in the tertiary education on these cities and the cities that are around them. The education is one of the pillars of development in the countries and this process can only be fulfilled with the presence of graduates in tertiary education in all the municipalities. The purpose of this article is to test this hypothesis of agglomeration and intend to see which factors that cause this phenomena, such as life quality, income effects and security. Empirical evidence are derived from real data obtained by DANE (Departamento Administrativo Nacional de Estadística) in the Census from 2005 and a violence variable constructed with IEPRI's (INSTITUTO DE ESTUDIOS POLÍTICOS Y RELACIONES INTERNACIONALES) information of political homicide. The understanding of this phenomena would be the only way to stop it and maybe create a different dynamic in the choice of these graduates of what city to work or even live.

keywords: *Education, Education Maps, Spatial Analysis, Moran's Index, Spatial Dependence Models*

*juantomas.sayagomail.com. I want to thank Roger Bivand, Hernan Enriquez, Jorge Aponte, and the ESPE Committee for their comments to this document and Liam Cooke and Matthew Schultz for the help writing this paper. All remaining errors are my own.

1 Introduction

The education system in Colombia has reduced the illiteracy rates and increased the access to education over the last 50 years constituting a major breakthrough. But despite these great accomplishments and advancements, Colombian cities are developing clusters of the quantity of educated people that live in them, these graduates decide to remain in the big cities, even though the access to education, health and quality of life are improving in small cities. Still they don't want to leave the big cities or return to the city where their family live because maybe there may not be enough chance to get a well paid job according to their capacities or it does not seems safe enough or they wont get better living conditions or better opportunities for their children. This way small cities have less educated people that can implement new technologies or projects that could increase the city's opportunities to develop.

In order to compute and observe these clusters data from the 2005 CENSUS done by the DANE¹ will be used. I compute the rate of educated people residing in every municipality that has data available² . Besides it will be shown that there's another cluster present related to education in Colombia that it is the lower rates of graduates tertiary education.

The factors associated with this problem in Colombia, that at first must be thought of are related to social and economic reasons, either because there is a lot of violence in the city where they come from or there is not a good supply of public services or there is not a good job market³ for them to come back from or leave the big cities. The relationship between these concepts will be explored and verified in order to see if there is a statistical relationship between them.

This document is divided in 5 sections besides this one, the second section makes reviews the state of the art related to education and the spatial analysis on this field, the third section analyse the education in the municipalities in Colombia and the 5 classical regions to test the agglomeration hypothesis. The forth section the analysis for the factors related to this agglomeration. The last section concludes this article.

¹DANE: Departamento Administrativo Nacional de Estadística. National Administrative Department of Statistics.

² The methods and the data used to calculate this rates are described on the appendix A.

³Understood as a necessity of their specic capacities or a wage according to their skills.

2 Great advances over the last years

Although literature related to education has been growing during the last 15 years, but the analysis has been restricted to case studies, revision of development, and policy recommendations. However it can be argued that in recent years some authors have tried to implement new ways to analyze this important and problematic sector of colombian society, as will be shown in this section with an overview of recently published documents regarding the educational sector in Colombia.

The implementation of spatial analysis has only been applied once for the education sector; where Jaime Bonet georeferenced data related to the provision for education in the departments, finding different patterns of agglomeration, and concluded that

The dotations for education in Colombia are distributed in such way that the departments with a poor provision are surrounded by departments that have the same conditions and the ones with higher levels are surrounded by prosperous also. (Bonet, 2005)⁴

Bonet also highlights the importance of using information by city. It would also be is important to observe the review to the law made on this article.

Another article with a very interesting approach and related to the tertiary education is done by Mora and Ceballos (2006) found that the the topics that are taught in technical and technological education are not adequate to fit the needs in the region where taught. Considering this one of the main difficulties for some people to remain or go back to their original city. Furthermore it can be said that this education is not accomplishing the objectives for its creation.

For other developments in the education sector its important to review the works presented by Ramirez and Tellez and Ramirez and Salazar in which the history of Colombias education sector is reviewed. Particularly relevant is the work by Ramirez and Tellez as it provides a good historical account of education in the 20th century, underlining the importance of different policies and their effects. It's important to highlight two facts presented by them: first, the recent and quick development in the sector after the 1950's at "an unseen velocity" and even after one slow path a new expansion; and secondly that from the end

⁴Traslation made by the author. Page 36.

of the century the sector was showing important advances.

The development of the educational sector in Colombia can be said to reflect Ramirez and Tellez findings as can be seen by an increase in the access to education pointed out by Iregui, Melo and Ramos(2006). The main problem concerning education in Colombia is more related to efficiency and quality, areas which have not shown constant growth. The same authors have published another article also with an analysis of the efficiency frontier in colombian schools using information from the ICFES analyzing the level of the schools with other factors and concluding that “private schools are more efficient due to a more favourable environment”⁵.

It is also worthwhile to review the results published by the DANE of the 2005 CENSO, which show stable level of literacy in Colombia. It must be said that the percentage of people that attend to school has increased greatly, and continue to remain high running parallel to the demographic increment in population in the last decade.

3 Education agglomerated in Colombian cities

The rate of educated people in Colombia has increased a lot in the last years, because it’s easier to access education and programs have been created to aid the students with low or no income to pay for this education, these politics have developed significantly. Also new universities and technical formation centers have been created all around the country, which has raised the numbers of people educated at the tertiary level. But still the proportion of graduates from the tertiary level is low in almost all municipalities.

The rate of people graduated in the tertiary education is skewed to the left side as can be seen in the kernel density plot and it must be noted in figure 2 that the bigger rates belong to the more populated cities for which the mean rate is 15.7%⁶ . But because this kernel is made using the municipalities’ rate as the measure unit this forms the observed skewness.

⁵Iregui, Melo and Ramos, 2007. Page 21

⁶ Calculations made with data from CENSO 2005 and it is important to remark that they are population based and not by municipality.

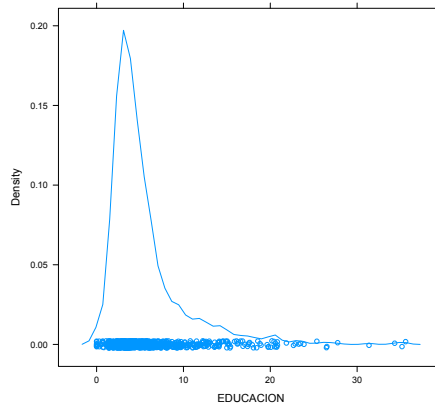


Figure 1: Density kernel of the percentages of tertiary education graduates in Colombia by municipalities. Data: CENSO 2005, DANE.

The existence of spatial relationship between the observations and their neighbors has been widely speiced as the rst law of geography “Everything is related to Everything else, but near things are more related than distant things” stated by Waldo Tobler in the 1970, and as Anselin stated standard econometrics analysis can not be used in presence of spatially correlated data.

The fact that spatial localization matters can’t be doubt because as already seen in figure 2, there appears to be an agglomeration of a bigger percentage in educated people in the municipalities located in the center of the country and specially on the ones that have cities with the highest categories⁷ or with the best opportunities in schooling and public services for the inhabitants, in most cases are the same. Different tests to check spatial agglomeration such as Moran’s Index, Moran’s scatterplot⁸, and the Local Indicators for Spatial Association (LISA)⁹ will be shown.

⁷Highest categories must be understood as the special one, and categories 1, 2 and 3 in most cases. See table 1 for the categories.

⁸These two are the empirical bayesian standarized versions, by populations. For more details see Assuno and Reis, 1999.

⁹Anselin, 1995.

Percentage of people that attended the tertiary education in Colombia

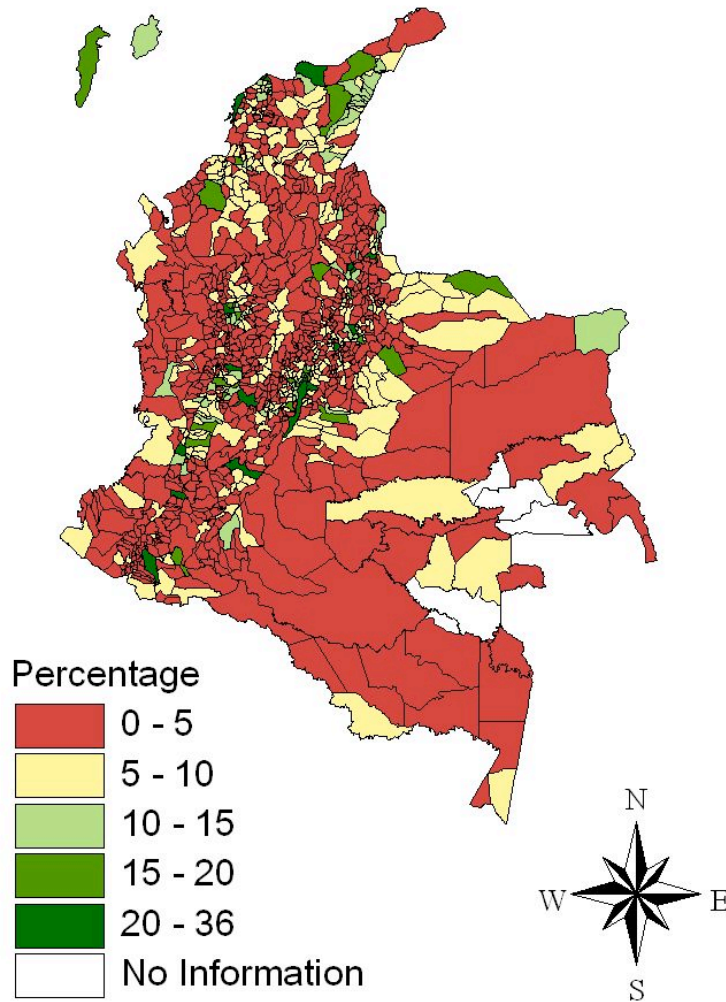


Figure 2: Percentage of tertiary education graduates in Colombia by municipalities.
Data: CENSO 2005, DANE.

Categories	Municipalities	Percentage
1	17	1.5%
2	16	1.4%
3	19	1.7%
4	20	1.8%
5	31	2.8%
6	991	88.6%
E(S)	5	0.4%
NC	20	1.8%
TOTAL	1,119	100%

Source: DANE

The agglomeration of tertiary education graduates in Colombia, as it was said before, seems to need the inclusion of spatial factors; the location of the bigger cities and the cities around those big cities seems to have more graduates than other zones in the country. As can be seen in figure 2 there's an agglomeration of professionals around the richest cities, or maybe the cities with best living conditions, even in Colombia the effects of displacement and the internal conflict violence must be considered. To begin this analysis its important to see if there is spatial autocorrelation in the proportion of graduates in the municipalities in Colombia:

Region	Type of matrix	Moran I
	Queen1	Queen 2
National	0.1516**	0.0983*
Central	0.0974*	0.0154
Oriental	0.1810**	0.0879
Atlantic	0.2576**	0.0648
Pacific	0.1833**	0.0822
A&O	0.1643**	0.0045

** means significant to 5% and * means nearly significant

Source: Data used from CENSO 2005-DANE.

The results from the Moran's Index presented in the above table 2 show us that there is strong spatial correlation eventhough when is considered a second level matrix of contiguity which is also shown in the Moran Scatterplot (figure

3). Finally the Local Indicators of Spatial Association (LISA) on figure 4 show the presence of small clusters around the main capitals and specially the ones with categories special, 1 and 2, but must be highlighted the presence of low rates clusters in both maps which confirms that the municipalities farther from these capitals are the less attractive for these graduates.

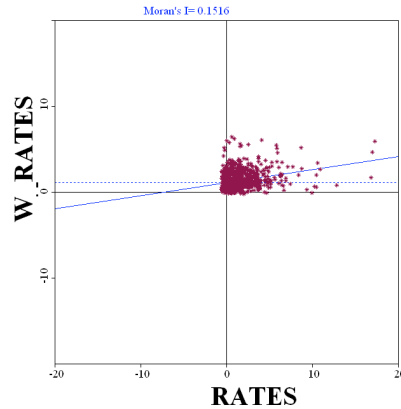


Figure 3: Moran Scatterplot of tertiary education graduates in Colombia. Data: CENSO 2005, DANE.

3.1 Education in Colombian cities grouped by regions

The evaluation of the regions seem to be important to observe the presence of attractors in specific regions, as should be considered that Bogota, Medellin and Cali are not the only nucleus in the country as can be seen mainly in figure 4 from this section, it must be remarked that it is the effect from other capitals and cities near the big capitals which create the cluster of educated people, because they want to avoid the congestions of the bigger cities and search for tranquility with greater contact to nature in their way of living.

Colombia is divided in 5 regions: Central, Oriental, Pacific, Atlantic, and Amazon and Orinoquia¹⁰ this definition of regions is taken like this in order to

¹⁰The regions were formed according to Los Municipios Colombianos hacia los Objetivos de Desarrollo del Milenio. Salud, educación y reducción de la pobreza, 2006 and by discussions with Jorge Aponte.

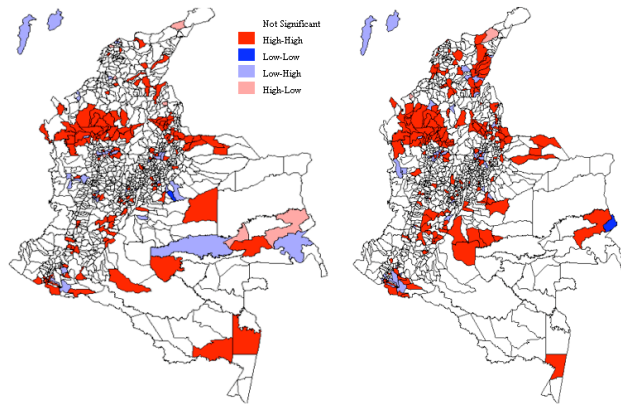


Figure 4: LISA of tertiary education graduates in Colombia by municipalities. On the right the LISA for the Queen matrix level 1 and on the left the level 2. This applies to all the LISA Analysis. Data: CENSO 2005, DANE.

avoid a further discussion on the different concepts of regions in Colombia, a more clear division can be seen in figure 5.

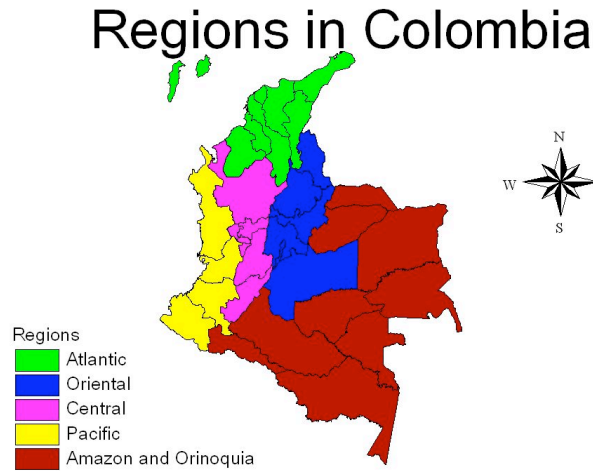


Figure 5: Classification of regions in Colombia. Data: CENSO 2005, DANE.

3.1.1 The Central Region

The Central Region is composed of Antioquia, Caldas, Huila, Quindio, Risaralda and Tolima, where the main capital is Medellin and it is a nucleus of the region where the greatest quantity and proportion¹¹ of graduates of tertiary education is agglomerated as can be seen in figure 6. This seems to show the cluster around the capitals of Antioquia, Risaralda and Caldas. The important effect to be seen is in the part of Antioquia which shares a limit with Choco and Cordoba, where can be seen the percentage of graduates diminishes when they are farther away from Medellin or the other capitals. It should also be highlighted that the amount of cities with larger rates than the national media are few compared to the ones with lower, but this is also an effect already observed on the previous analysis (see figure 6 and the analysis for figure 2).

¹¹It is important to focus on the proportion because of the agglomeration of the population in Colombia is mainly on the big cities.

Percentage of people that attended
the tertiary education in the
Central Region

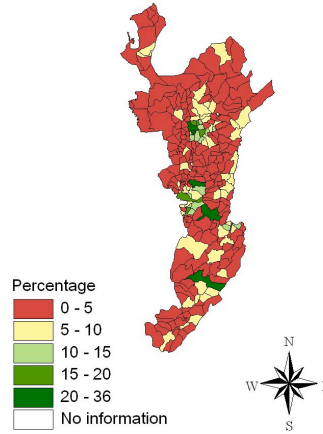


Figure 6: Percentage of tertiary education graduates in the Central Region of Colombia. Data: CENSO 2005, DANE.

The analysis to check for presence of spatial correlation in table 1 shows that there is a significant spatial autocorrelation in the Moran's Index. The results support the evidence of spatial association in the data and also the can be conrmed with the second level matrixes of contiguity. The Results obtained throughout the Moran Scatterplot and the LISA (figure 7 and 8) confirm this and specially the LISA show two clusters of a higher proportion of graduates and three clusters of a lower proportion of graduates.

As can also be seen the region had developed another kind of cluster in the Central Region, the one expected because of the localization of Medellin and another big one around the border to Choco, and Cordoba and Magdalena, as already mentioned, and the spatial autocorrelation appears to weakens when the second level matrixes is included in the analysis and its important to see the small clusters formed around the other important capitals in the region.

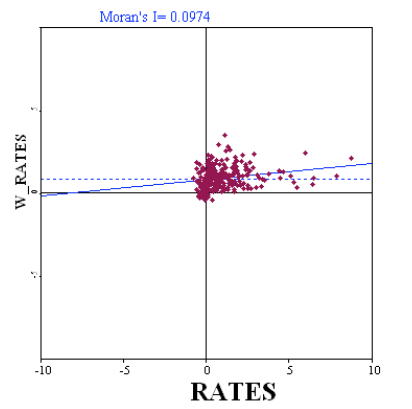


Figure 7: Moran Scatterplot of tertiary education graduates in the Central Region of Colombia. Data: CENSO 2005, DANE.

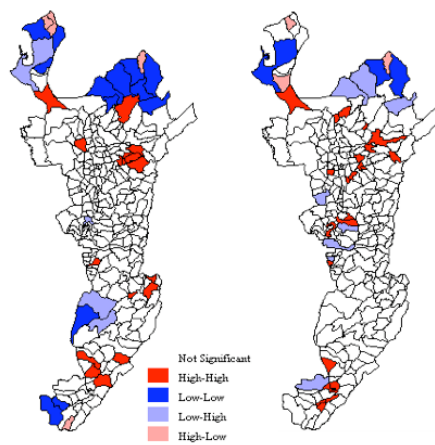


Figure 8: LISA Analysis for the Central region of Colombia. Data: CENSO 2005, DANE.

3.1.2 The Oriental Region

The Oriental Region is composed of Boyaca, Cundinamarca, Meta, Norte de Santander, Santa Fe de Bogota and Santander, and it's important to remember that Bogotá is the capital and the biggest city in the country. Other important cities are Bucaramanga, Cúcuta and Tunja, which can be core for clusters in the region. As should be noted from gure 9 there are higher rates in the municipalities around the important cities. It is also important to note gure 10 where an agglomeration of low level municipalities still surges.

Percentage of people that attended the tertiary education in the Oriental Region

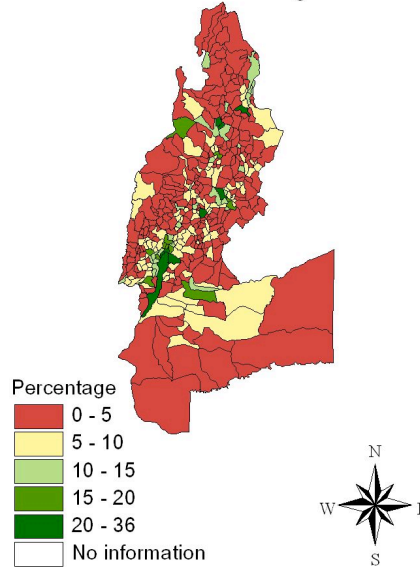


Figure 9: Percentage of tertiary education graduates in the Oriental Region of Colombia by municipalities. Data: CENSO 2005, DANE.

As can be seen in gure 11, there is a cluster around Bogota, and around Bucaramanga y Tunja, that can be seen on the LISA maps. The fact that there are 2 clusters of low levels in the border between Cundinamarca and Boyaca,

and in Meta should be highlighted, although its important to see that they are they are farther away from the main capitals. Also confirmed by the Moran's Index and Scatterplot on table 1 and figure 10.

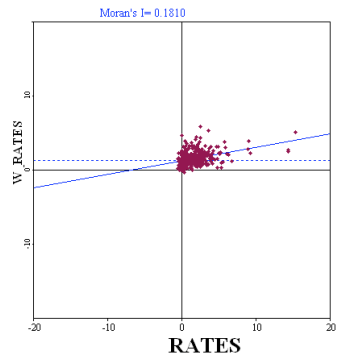


Figure 10: Moran Scatterplot of the tertiary education graduates in the Oriental Region of Colombia. Data: CENSO 2005, DANE.

3.1.3 The Pacific Region

The Pacific Region is composed of Cauca, Choco, Nariño and Valle del Cauca, where the main capital is Cali. We must suppose the cluster created around Cali, the Capital is one of the most important cities in the country which also can be thought of as a pole for development, but also Pasto and Popayan should be considered. Certainly it can be said that Cali is the capital that must be the main attractor in the region.

Different kinds of clusters should be highlighted from figure 9 because there is a high level cluster around Cali and a low level cluster in the border between Cauca and Nariño and a small one in the middle of Choco.

The Moran Index and Moran Scatterplot show interesting results as can be implied from the first one there is a strongly significative spatial correlation but in the second one ther is a spatial correlation but seems not to be very signicative in most cases.

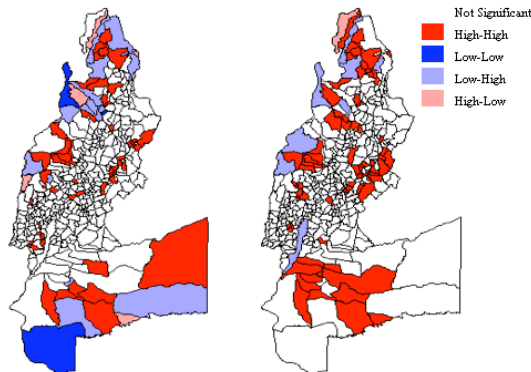


Figure 11: LISA of tertiary education graduates in the Oriental Region of Colombia by municipalities. Data: CENSO 2005, DANE.

The results presented in the Moran Scatterplot give support to the Moran's Index and in figure 14 the LISA map highlights the cluster of lower levels mentioned in the border of Cauca and Nariño and the cluster of higher levels in Valle del Cauca around the main capital Cali.

3.1.4 The Atlantic Region

The Atlantic Region is composed of Atlántico, Bolívar, Cesar, Córdoba, La Guajira, Magdalena and Sucre, where there are three main capitals that are very close to each other Barranquilla, Cartagena, and Santa Marta. It must be highlighted that little clusters can be found in figure 15 and one big cluster of agglomeration for low rates of graduates in the south border of Bolívar, Córdoba, Magdalena and Sucre.

The results from Table 1 show that there is evidence of agglomeration in this region and this result holds for a matrix of contiguity of higher level. Moran Scatterplot analysis supports this analysis and the LISA maps makes the clusters observable (see figures 16 and 17).

Percentage of people that attended the tertiary education in the Pacific Region

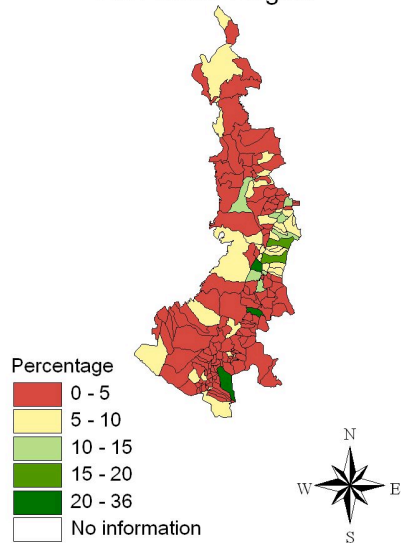


Figure 12: Percentage of tertiary education graduates in the Pacific Region of Colombia by municipalities. Data: CENSO 2005, DANE.

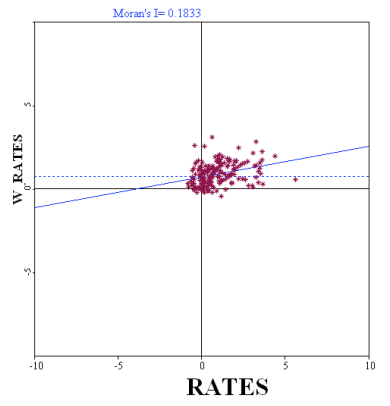


Figure 13: Moran Scatterplot of tertiary education graduates in the Pacific Region of Colombia. Data: CENSO 2005, DANE.

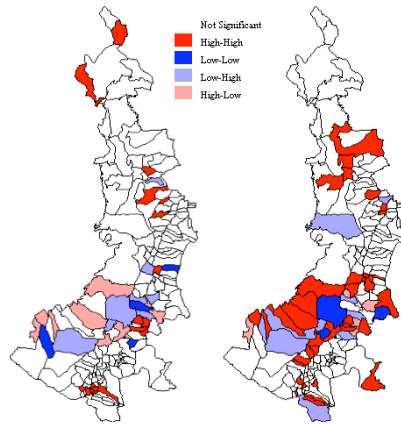


Figure 14: LISA of tertiary education graduates in the Pacific Region of Colombia by municipalities. Data: CENSO 2005, DANE.

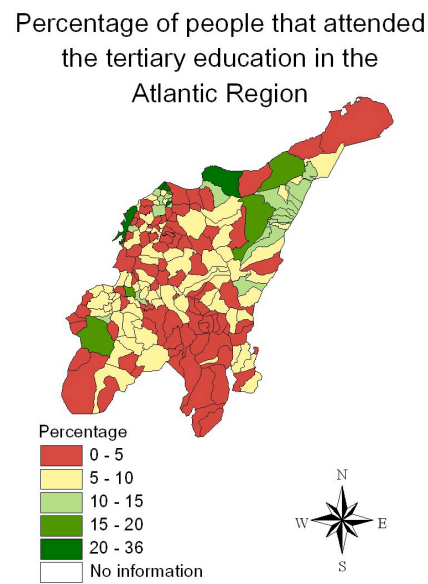


Figure 15: Percentage of tertiary education graduates in the Atlantic Region of Colombia by municipalities. Data: CENSO 2005, DANE.

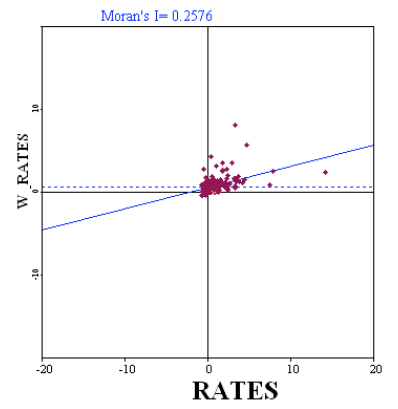


Figure 16: Moran Scatterplot of tertiary education graduates in the Atlantic Region of Colombia. Data: CENSO 2005, DANE.

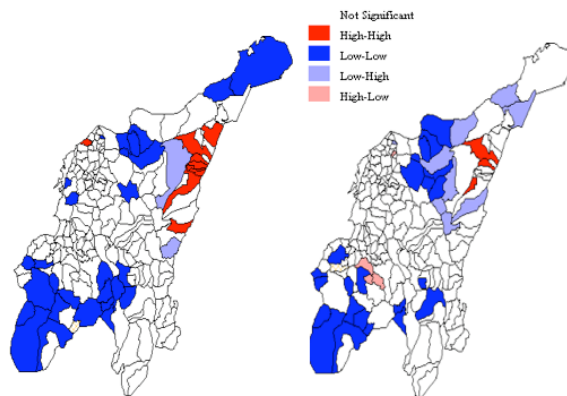


Figure 17: LISA of tertiary education graduates in the Atlantic Region of Colombia by municipalities. Data: CENSO 2005, DANE.

3.1.5 The Amazon and Orinoquia Region

The Amazon and Orinoquia (A & O) Region is composed of Amazonas, Arauca, Caqueta, Casanare, Guainia, Guaviare, Putumayo, Vaupes and Vichada. As is highlighted there is not a big city in the Region that can be an attractor of people and therefore the existence of a cluster of cities with low percentage of graduates of tertiary education, can be seen in figure 18 the higher rates are located mainly in Arauca.

Percentage of people that attended the tertiary education in the Amazon and Orinoquia Region

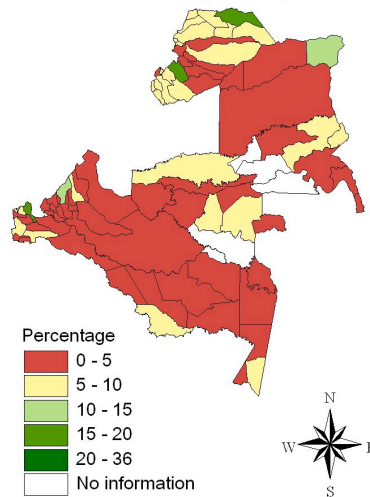


Figure 18: Percentage of tertiary education graduates in the Amazon and Orinoquia Region of Colombia by municipalities. Data: CENSO 2005, DANE.

It should be noted in figure 18 that almost all of municipalities are below the mean and that there is not a city where it is possible to see a high rate unlike the previous sample.

As the data tests for spatial autocorrelation, the results of Moran's Index present statistical evidence of spatial autocorrelation while using first level matrixes of spatial contiguity whilst the second level matrixes seem to be not sig-

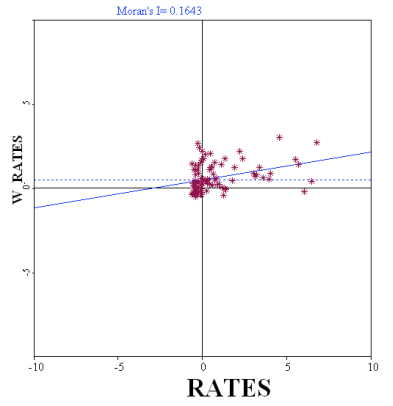


Figure 19: Moran Scatterplot of tertiary education graduates in the Amazon and Orinoquia Colombia. Data: CENSO 2005, DANE.

nificant in both results. Furthermore it's important to observe that the Moran Scatterplot does show the presence of agglomeration supporting the evidence from table 1.

As mentioned before the LISA maps present smaller clusters that seem to be in Arauca and the other clusters shown are too small to be considered as important representations. It is important to remember that this region, does not contain a very developed city so we cannot consider there to be a pole of development or an attractor of educated people.

4 Factors associated to the agglomeration of educated people in Colombia

The factors possibly associated in Colombia as mentioned before are social, economic and political. The analysis to check the spatial correlation of the data tell us that it will be necessary to include a spatial term in the related factors, and the Equation to be considered originally is the following:

$$PTG_i = \beta_0 + \beta_1 Ecv_{ar}_i + \epsilon_i \quad (1)$$

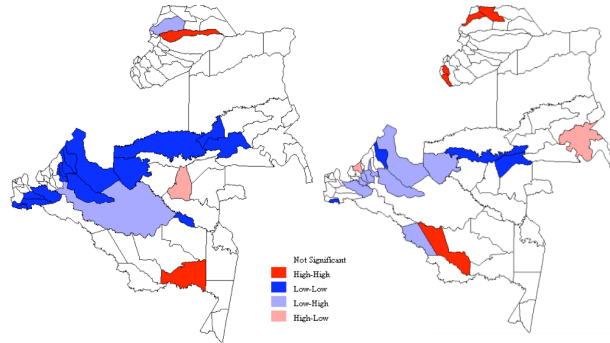


Figure 20: Percentage of tertiary education graduates in Colombia by municipalities.
Data: CENSO 2005, DANE.

$$PTG_i = \beta_0 + \beta_1 Ecv ar_i + \beta_2 Servvar + \beta_3 viov ar + \epsilon_i \quad (2)$$

The variables included are:

PTG_i : The Percentage of University graduates for the city i .

$Ecv ar_i$: The variable considered for economic purposes is the proportion of economic activity times for every thousand persons for the city i and has been dened as $Ecv ar_1$ for the commercial sector, and $Ecv ar_2$ for the service sector¹².

$Servvar_i$:The variable considering the coverage of water supply in the city i .

$Viov ar_i$: The violence Variable accounting for the number of deaths for the years 2000 to 2005 divided by the population in the department and multiplied by 1000¹³.

and for the equations considering spatial data:

¹²The Industrial sector was used without any results to improve, maybe it is related to the problem mentioned by Mora and Ceballos, according to them the tertiary education is focused on the wrong areas and it is not very industry oriented.

¹³I am thankful to Francisco Gutierrez for allowing the use of this information. For more details see the appendix A.

$$PTG_i = \rho W_{PTG_i} + \beta_0 + \beta_1 Ecvar_i + \beta_2 Servvar + \beta_3 viovvar + \epsilon_i \quad (3)$$

$$PTG_i = \beta_0 + \beta_1 Ecvar_i + \beta_2 Servvar + \beta_3 viovvar + u_i \quad (4)$$

where

$$u_i = \lambda W_{u_i} + \epsilon \quad (5)$$

The results are from the OLS Analysis is found in the Appendix B. It can be seen there is important need to include a spatial variable as seen in the Moran Test applied to the residuals in the models with the 3 different economic variables. The results for the analysis using the spatial Lag of the variable and the error lag in the analysis are presented in the following table, where as can be seen and in the appendix C the Lagrange multiplier test show that the best model is the Maximum Likelihood Estimation of the Spatial Error Model and also the results seem to support this model. It must be highlighted that all the variables considered seem to be significant as do the spatial parameter seem to be significant, but this happens in both type of models¹⁴.

Table 2:ML Analysis for PT_i in Colombia						
Variable	Model1 LMlag	Model2 LMlag	Model3 LMlag	Model1 LMErr	Model2 LMErr	Model3 LMErr
intercept	1.716** (0.363)	-1.52** (0.391)	-1.56** (0.326)	-4.4 (0.295)	-0.306 (0.425)	-0.402 (0.420)
EcVar1	0.066 (0.009)	0.0304** (0.008)	-	0.067 (0.008)	0.034** (0.008)	-
EcVar2	-	-	0.101** (0.016)	-	-	0.111** (0.016)
SerVar	-	6.96** (0.539)	7.462** (0.542)	-	8.423** (0.555)	7.925** (0.561)
VioVar	-	-0.857** (0.218)	-0.873** (0.216)	-	-0.866** (0.229)	-0.879** (0.226)
WPUG λ Asimp error	0.493** (0.059)	0.356** (0.055)	0.348** (0.0541)	0.5** (0.061)	0.479** (0.035)	0.487** (0.062)
AIC	6256.6	6055.5	6028.6	6259.3	6033.8	6003.4

** means significant at 5%. The value on top is the coefficient and the p-value

¹⁴For a complete review on the methods applied see Anselin, 1988 and Arbia 2006

in the parenthesis is the standard error.

Its important to observe that the coefficient for the economic variables is significant and the sign shows the expected coefficient of the variable according to what it should be. Doing a remark economic variables and the variable related to the welfare of the people have positive effects and the variable related to violence has a negative one. The previous correlations show us the fact that people tend to go to places where there is less violence and better living conditions and work options. The models are calculated with the queen matrixes with level 2 in order to obtain a better fit of the errors. As can be seen in the models the inclusion of new variables seem to increase the fitting and reduce the AIC, and all the variables have different.

5 Conclusions

A Review of the data proposed in this article should demonstrate the importance of the agglomeration of educated people as a significant determining factor in the development of Colombian cities and the regions in which they are located. It is important not only to consider the distance between the different municipalities but also the specific location to see the real spatial effect and the real geographic association. The first conclusion to be accounted for confirms findings put forward by Bonet: if you have a prosperous neighbor you will be wealthy or at least more developed, and if you have poor neighbors it is probable that you will be poor, and will be the same for educated people.

It is important to consider welfare relationships through the access to public services in order to draw conclusions, because these relationships have been vital in creating agglomerations. Furthermore, this access is also agglomerated and therefore access to the water supply as a necessity to live is crucial for this case.

Violence variables must be accounted for as well: regardless of other factors, an area that is known to have violent reputation strongly discourages people trends moving to or staying in those municipalities are important to consider them because after a violent experience to someone known people are drawn out of the municipality. For future studies it could perhaps be significant to

use a proxy that accounted for in greater historical depth violence variables, to account the people's memory.

Economic variables are indeed related. However, It can be argued that a two-way and mutually reinforcing relationship exists between educated people and developed or wealthy areas. On the one hand, educated people seek better wages, typically found in wealthy areas, while at the same time educated people are an stimulus for economic growth and development¹⁵. This relationship may cast doubts on the results obtained in this article, but should be considered in a comprehensive evaluation of these phenomena.

Because Spatial factors are considered, this type of analysis gives results that the same information and standard econometrics don't. The spatial association of the data as can be seen and as Tobler state it. Regional analysis shows that there are big and small cores to be considered and their presence in the regions is an important factor of socioeconomic relationships.

Furthermore availability of tertiary education in the bigger cities must also accounted as an important factor analyzing this agglomeration and should be clarified that not only the rates are agglomerated but also the population in the cities present agglomeration as was observed by Galvis in Economic topography in Colombia.

It is also necessary to identify that this analysis was done using the municipalities as observation units, and the result of this was to keep hidden any heterogeneity within the municipality which are also very important in urban areas that not visible study. Also another argument that should be considered is the different definitions of region in the results obtained as can be argued Caqueta must be included in the Amazon and Orinoquia Region or the Central Region¹⁶ and other departments are in a different region, however this selection influences vastly the results obtained on the statistical tests.

¹⁵This relationship is considered a stylized fact in economic growth.

¹⁶In the original study it was located in the Central Region.

6 Bibliography

1. Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Norwell, MA: Kluwer.
2. Anselin, Luc. 1993. *Exploratory Spatial Data Analysis and Geographic Information Systems*. Regional Research Institute West Virginia University. Research Paper 9329.
3. Anselin, Luc. 1995. Local Indicators of Spatial Association - LISA. *Geographical Analysis*, Vol. 27. pages. 93-115.
4. Anselin, Luc. 1996. The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association. *Spatial Analytical Perspectives on GIS in Enviromental and Socio-Economic Sciences*.
5. Anselin, Luc. 2003. *An Introduction to Spatial Regression Analysis in R*. Spatial Analysis Laboratory, University of Illinois. Spdep intro R
6. Anselin, Luc. 2005. *Exploring Spatial Data with Geoda: A Workbook*. Center for Spatially Integrated Social Science. Spatial Analysis Laboratory.
7. Anselin, Luc. 2005B. *Spatial Regression Analysis in R: A Workbook*. Center for Spatially Integrated Social Science. Spatial Analysis Laboratory.
8. Arbia, Giuseppe. 2006. *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Springer.
9. Bivand, Roger. 2008. *The spdep Package*.
10. Bonet, Jaime. 2005. Inequidad espacial en la dotación educativa regional en Colombia. *Documentos de Trabajo sobre Economía Regional*. No. 56.
11. DANE. 2005A. *CENSO GENERAL 2005. Cuestionarios*.
12. DANE. 2005B. *CENSO GENERAL 2005. PRINCIPALES INDICADORES DE EDUCACIÓN*.
13. DANE. 2006A. *CENSO GENERAL 2005. Ficha Metodológica*.
14. DANE. 2006B. *CENSO GENERAL 2005. Diseño Conceptual y Metodológico*.
15. Galvis, Luis Armando. 2001. *La topografía Económica de Colombia*. *Documentos de Trabajo sobre Economía Regional*. No. 22.
16. Gutierrez, Francisco. 2006. *Tendencias del Homicidio Político en Colombia 1975- 2004: Una discusión preliminar*. *Nuestra guerra sin nombre*

transformaciones del conicto en Colombia. Eds. Wills M, and Sanchez G. Norma.

17. Iregui, Ana Maria, Ligia Melo and Jorge Ramos, 2006. La educación en Colombia: Análisis del marco normativo y de los indicadores sectoriales. Vol. 9. No.2. Pages 175238.
18. Iregui, Ana Maria, Ligia Melo and Jorge Ramos, 2007. Analisis de ecienza de la educación en Colombia. Revista de Economía del Rosario. Vol. 10. No.1. pages 21-41.
19. Mora, John James y Fernando Ceballos. 2006. Pertinencia de la educación técnica y tecnológica en el mercado laboral de Cali,1994 -2005. Reportes Observatorio de Políticas Públicas. ICESI
20. Moran, Patrick Alfred Pierce. 1948. The interpretation of statistical maps, Journal of the Royal Statistical Society.
21. Ramirez, Maria Teresa y Irene Salazar. 2007. Surgimiento de la educación en la República de Colombia. En que fallamos?. Borradores de Economía. No. 454.
22. Ramirez, Maria Teresa y Juana Tellez, 2006. La educación primaria y secundaria en Colombia en el siglo XX. Borradores de Economía. No. 379.
23. Tobler, Waldo. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography. 1970. Pp. 234-240.

A Data description

The data used in for the recent calculations was taken mostly from the CENSO 2005, the information has been publicated during the last year. The information calculated with the data from DANE is the percentage of people which has a degree in the tertiary level of education, the coverage of the water supply in the city and last the economic variables. The construction of the data for use is presented:

- Percentage of Tertiary Graduates (PTG): The percentage of graduates from the tertiary level education in the city is calculated in the following way:

$$PTG = \frac{Pgl_i}{(TotPop_i)-(Peoples15)} * 100$$

where the data comes from a survey done in the CENSO and Pgl_i is people that graduated at the tertiary level in the city i , $TotPop_i$ is the total population from the survey in the city i , and $Peoples15$ is the population in the survey younger than 15 years old¹⁷

- The Coverage of the water supply in the city (Servvar): Percentage of people that has access to the service of water¹⁸. This data is calculated by the CENSO with a survey also.
- The economic variables are calculated using the number of economic units and the specific sector where they belong the Commercial Sector and the Service Sector.

B OLS Results

The results in table B1 show the need in the models to include a spatial variable in the Analysis.

Table B1: OLS Analysis for PUG_i in Colombia			
Variable	Model1 OLS	Model2 OLS	Model3 OLS
intercept	4.33.01** (0.201)	-0.04 (0.355)	-0.104 (0.355)
EcVar1	0.076 (0.009)	0.034** (0.009)	
EcVar2			0.108** (0.016)
SerVar	-	8.21** (0.531)	7.68** (0.529)
VioVar	-	-0.925** (0.224)	-0.94** (0.221)
$R^2/ps-R^2$	0.05	0.239	0.26
Moran	0.132	0.316	0.325
p-value	0.000	0.000	0.000

¹⁷Its used the 15 years old range because the fact that using 20 years would make that some persons considered in the numerator are not considered in the denominator.

¹⁸Its considered because of the importance and the important necessity of water for a living.

C Lagrange Multiplier Test Statistic

For the better understanding of the test we will describe the models prior execution of the test for the country.

Sp=Spatial Parameter

Model1

$$PUG_i = Sp + \beta_0 + \beta_1 * Ecvar1_i + \epsilon_i \quad (6)$$

where Ecvar1 implies that will be used the variable associated to the Commercial sector.

Model2

$$PUG_i = Sp + \beta_0 + \beta_1 * Ecvar1_i + \beta_2 * Servvar + \beta_3 * Viovar + \epsilon_i \quad (7)$$

Model3

$$PUG_i = Sp + \beta_0 + \beta_1 * Ecvar2_i + \beta_2 * Servvar + \beta_3 * Viovar + \epsilon_i \quad (8)$$

where Ecvar2 implies that will be used the variable associated to the service sector.

As can be seen in the following table both models seem to be significant in the calculation and analysis but now we turn to the error analysis which tends to favor the Spatial Error model because the p-value suggest these models as the likely alternative.

Table C1: Lagrange Multiplier Test			
Variable	Model1	Model2	Model3
LMerr	127.739	105.613	109.206
p-value	0.000	0.000	0.000
RLMerr	0.1963	44.38	51.362
p-value	0.657	0.000	0.000
LMlag	133.13	62.448	59.613
p-value	0.000	0.000	0.000
RLMlag	5.59636	1.216	1.769
p-value	0.018	0.27	0.1835