

Rosero, Grace Carolina Guevara

Conference Paper

Growth agglomeration effects in spatially interdependent Latin American regions

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Rosero, Grace Carolina Guevara (2015) : Growth agglomeration effects in spatially interdependent Latin American regions, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124693>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

Growth agglomeration effects in spatially interdependent Latin American regions

Carolina Guevara
University of Lyon, Lyon
CNRS, GATE Lyon Saint-Etienne, Ecully

September 7, 2015

1 Introduction

Identifying the mechanisms of agglomeration is a key element towards understanding how agglomeration impacts on growth. The theoretical literature analyzing the link between agglomeration and growth has been developed through the association of different concepts and findings from several theories. Among them, important findings derive from the conjunction of elements of growth theory and elements from economic geography theory. Based on the New Economic Geography, growth models can predict how firms choose their location to produce and how the resulting distribution of economic activities affects the rate of technological progress (Martin and Ottaviano, 1999). Other theoretical advances have been developed in urban economics. This theoretical approach enables the determination of the mechanisms of city formation and how human capital accumulation leads to increasing the city size and, in turn, how it promotes technological progress. A distinction between developed and developing countries is also feasible to the extent that later advances in urban theory recognize the underlying characteristics of economies.

Those theoretical approaches have demonstrated a strong positive correlation between agglomeration and growth. Indeed, such a link is bidirectional which is determined by forward (demand) and backward (cost) linkages (Martin and Ottaviano, 1999, 2001; Baldwin and Forslid, 2000). However, the positive relationship may change due to different factors. The common explanation of reduced effects or even negative effects of agglomeration on growth is related with congestion effects. Further concentration of activities entails high housing costs, pollution, traffic and high

transport and commuting costs which produces dispersion of economic agents. Another aspect, less studied in the literature, lies on the role of the level of development of countries in explaining the agglomeration-growth relationship. Little attention has been dedicated to such a factor (Venables, 2005; Duranton, 2007) most likely due to the implicit assumption that the experience of developed countries can be translated to that of developing ones. This is what Chauvin et al. (2013) called the knowledge mismatch which recognizes the importance of focusing research on developing countries. The fact that the economic conditions vary across the stages of development might matter for the effects of agglomeration. Indeed, the empirical evidence provided by Williamson (1965) shows that the degree of agglomeration is higher in early stages of development than in later stages. Accordingly, the effects of agglomeration on growth are more pronounced in the former stages than in the latter stages (Brühlhart and Sbergami, 2009). Thus, powerful¹ mechanisms seem to occur in developing countries that have not yet been studied. The factors that cause the differences have not been widely explored. Broadening the study of mechanisms in developing countries will lead to a better understanding of agglomeration effects.

This paper then investigates how agglomeration impacts on growth and whether the mechanisms of agglomeration are related with the level of development. To answer these questions, we use sub-national data of Latin American countries.

The remainder of this paper is organized as follows. In section 2, the context of spatial disparity of Latin American countries is described. In section 3 we present the data, the strategy estimation and the results. Section 4 concludes.

2 The context of inequality in Latin America

Lately, Latin America has experienced a continued economic expansion. During the last decade, the average growth rate² has increased from 2.4% in 2000, year in which several countries faced both internal and external crisis, to 5.8% in 2005 and by the end of the decade, the average growth rate has reached 6.5% in 2010, reflecting one

¹As Venables (2005) qualified the mechanisms of mega cities in developing countries that keep attracting more firms.

²These statistics were obtained using data from the Economic Commission for Latin America and the Caribbean (ECLAC). They correspond to the 10 biggest countries of South America, Mexico and Panama.

of the best signs of recovery after the world recession in 2008. The growth patterns differ across countries. During the period 2004-2010³, these countries recorded an annual growth rate of 5.5% in average. Countries with higher annual growth rates than the average were Panama (8.13%), Argentina (7.44%), Uruguay (7.11%), Peru (6.83%) and Venezuela (6.83%). Countries with lower annual growth rates were Paraguay (5%), Colombia (4.7%), Ecuador (4.65%), Bolivia (4.51%), Brazil (4.47%), Chile (4%) and Mexico (2.32%).

Within countries, the economic scenario is very heterogeneous. Few domestic regions concentrate most of the economic activity. Overall, capital regions alone or together with no more than two regions account for most of the Gross Domestic Product (GDP) of each country. For instance, the capital region of Peru, alone represented 52% of the total GDP; in Chile, the capital region accounted for 48%; in Ecuador only two regions recorded 54%; and in Colombia three main regions produced 52% of the total GDP.⁴ In general, spatial concentration at regional level is a common pattern in Latin America.

Such a spatial configuration has been the result of a sequence of historic events. It was first formed in concordance with Indian settlements (Aztec, Maya and Inca empires). Then, the colonizing countries (mainly Spain and Portugal) favored specific regions which become political and economic centers (Atienza and Aroca, 2012; Massiris-Cabeza et al., 2012). Globalization was another key factor that reinforced spatial concentration in regions where sources of comparative advantage of countries were located. Along time, those regions with comparative advantage have developed certain conditions, urban infrastructure among them. Thus, the process of urbanization was different across regions within countries. The primacy of politically and economically favored regions was enhanced. As shown in Figure 1, Panama, Paraguay, Uruguay, Argentina, Chile and Peru concentrate more than 40% of their total urban population in their primate regions⁵. In other countries, urban population has been concentrated in the main region and secondary regions, for example Bolivia, Ecuador, Honduras, Colombia, Venezuela and Brazil.⁶

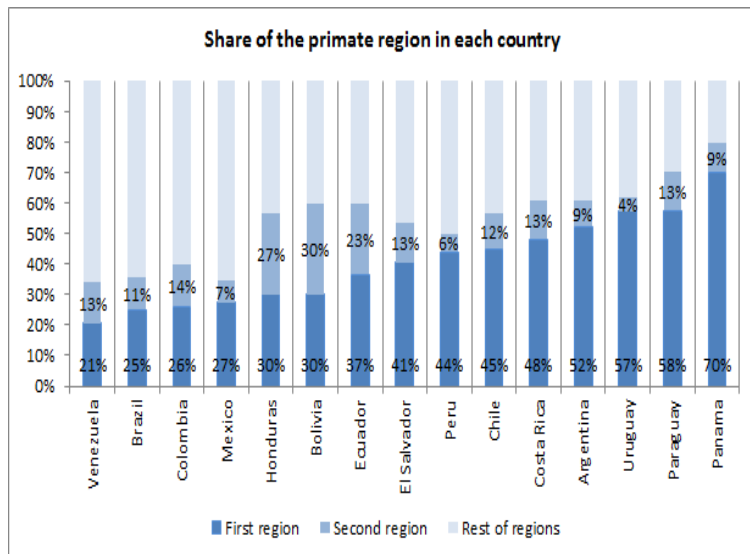
³The average on countries is taken since 2004 in order to avoid years of crisis in some of them.

⁴See appendix A.1, table 9 for the distribution of GDP at regional level within Latin American countries.

⁵For more details about this information, see appendix B

⁶In Bolivia, there exists a system with two main regions: La Paz and Santa Cruz. Ecuador has two main urban centers: Guayas and Pichincha. Honduras also has two main regions: Francisco Morazan and Cortes. In Colombia, the two main regions are Cundinamarca and Antioquia. Venezuela has a two-region urban system with the State of Miranda whose main city is Caracas and the State of Zulia whose main city is Maracaibo. Finally, Brazil was initially a urban system

Figure 1: Percentage of urban population in primate regions; 2000



Source: ECLAC database on National Population Censuses.

It is also argued that the trade policy of Import Substitution Industrialization⁷ (ISI acronym is spanish) played an important role in reinforcing regional concentration and primacy within countries. The reason is that the increase of industrial activity in main cities has attracted great flows of internal migration from rural areas towards urban areas (Atienza and Aroca, 2012).

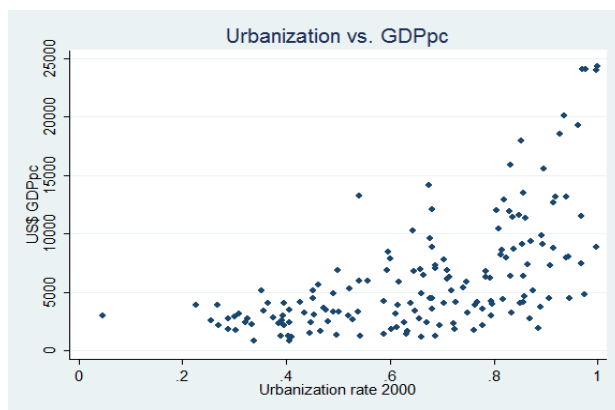
The dynamics of urbanization have involved internal migration and labor informality. The improvement of agricultural techniques led to a surplus of labor force, specially young population, which had to move from rural areas towards urban areas. Besides, given that favored regions accounted for good provision of public services and provided better standards of living, more population located in those regions. However, not all migrants improved their standard of living. The migrants with low levels of education were employed in the informal labor market in order to survive in big cities. Such a situation together with social issues promoted an informal

with three-main-regions: Sao Paolo, Rio de Janeiro and Minas-Gerais. However, Rio de Janeiro and Minas-Gerais have decreased their share of the total urban population.

⁷In the decades of the fifties and sixties, most of Latin American governments adopted the ISI policy which consisted in replacing foreign industrialized imports with domestic production.

housing market, the so-called invasions, where a lot of persons settled in without property titles. All things considered, spatial disparities have been accentuated not only in the economic aspect but also in the social aspect. Thus, provision of services, access to health, education and technology vary from one region to another. The gap between high income and low income regions is large. In economic terms, at a country level, the GDP per capita of the wealthiest region is 8 times as large as the poorest region.⁸ In social terms, some regions within countries have more than 60% of their population in poverty situation.⁹ Regions with high levels of urbanization, commonly the aforementioned favored regions, provide high levels of wealth. As we can observe in Figure 2, the correlation between urbanization rate and GDP per capita is positive (the correlation coefficient is 0.37 and it is significant) at regional level.

Figure 2: Urbanization and GDPpc



Source: ECLAC database on National Population Censuses and National Accounts of each country.

The analysis presented throughout this section reveals a high degree of spatial concentration within Latin American countries. Agglomerated regions have higher standards of living. Thus, those regions are likely to enjoy the benefits of agglomeration. The next section 3 is devoted to determine to what extent agglomeration affects growth of Latin American regions.

⁸For more details, see Table10 of appendix A.1.

⁹For information about regional poverty of countries, see Table 11 of Appendix A.1.

3 Estimation of the effects of agglomeration in Latin American regions' growth

Based on the stylized facts presented in previous section 2, the study of agglomeration effects in Latin American regions merits special attention. Spatial disparities are present not only in economic terms but also in social terms. Despite this, empirical literature is scarce. The main reason is the lack of consolidated data. The present study surpasses such issue, being the first attempt to analyze the impact of agglomeration on Latin American regions' growth.

This study focuses on data at regional level for two reasons. First, the heterogeneity between regions can be taken into account. Second, the study of agglomeration economies entails geographical interactions which are more evident at a rather small spatial scale. The geographical proximity between economic agents promotes more immediate and dynamic connections. Thus, externalities of agglomeration are more likely to be observed within regions. Spatial interactions between regions would be more intense than between countries. The reason is that regions within countries have common characteristics that enable them to interact more easily with each other than with regions of other countries.

3.1 Data Issues and Data Sources

Since data at the sub-national level is limited in some Latin American countries¹⁰, we use a regional database which accounts for information of 166 regions of 8 countries: Argentina, Bolivia, Chile, Colombia, Ecuador, Mexico, Peru and Panama. The inclusion of more variables reduces the sample to 87 regions of three countries: Colombia, Ecuador and Mexico. Hereinafter, we refer to the first database as *A* and to the second database as *B*.

In spite of the fact that the availability of temporal data differs across countries, we could obtain a panel database for the period 2000-2009. Since the process of agglomeration takes time, we believe that the annual variation is not enough to explain growth (Gardiner et al., 2010; Bosker, 2007; Brülhart and Sbergami, 2009). Hence, we use a three year-period database: 2001-2003, 2004-2006 and 2007-2009.

¹⁰The countries that do not account for regional information of National Accounts are Costa Rica, El Salvador, Guatemala, Haiti, Honduras, Nicaragua and Venezuela.

All data was collected from National Statistical Institutes of each country, developing a feedback relationship with public servants. In particular we obtained regional series for real GDP, population by age, surface area, education level of population and government expenditure.

The second source of information is the regional institution ECLAC (Economic Commission for Latin American countries) from which we obtained the Gross Domestic Product (GDP) at the regional and industrial level,¹¹ and urbanization rates at regional level. The database of urbanization contains information until 2000. We updated it with information on recent population censuses of countries. We impute an intermediate value of urban population between the last census and the recent census using an annual growth rate. The rationale behind this computation is that urbanization is an increasing process over time. This allows capturing the effect of urbanization in the dynamic economic growth rate, using three year period data.

Given that national currencies vary across Latin American countries, we standardize the measure of nominal GDP to real GDP using the Purchasing power parity (PPP) conversion factor of 2000 of the World Bank, International Comparison Database.¹² The geographical coordinates were obtained from the GeoHack system of Wikitech.

One data issue we face is the difference in the procedures measuring urban population in each country. Overall, there are three general methods to establish the urban character: i) localities with more than 2000 inhabitants, ii) definition (i) combined with the provision of public services and urban conditions, and iii) capital cities constitute urban areas; the areas outside capitals are considered rural. We consider that the variety of procedures constitutes a robustness test since if a pattern is observed despite the diversity of measures, the result would be more persuasive.

¹¹Data on industries is used to compute the industrial specialization index of each region.

¹²The PPP factor measures the number of units of a country's currency required to buy the same amount of goods and services in the domestic market as U.S. dollar would buy in the United States. Thus, it allows transforming the GDP in national currencies to dollars in real terms, which are comparable.

3.2 Estimation Strategy

Using both databases, we estimate the impact of agglomeration on the regional growth of Latin America. Database *A* allows estimating a parsimonious model of agglomeration. Database *B* allows estimating an extended model including more variables. The general specification of the model is:

$$\Delta y_{i,t+k}/y_{i,t} = \alpha + \psi \ln y_{i,t} + \gamma \text{agglom}_{i,t} + X_{it}\beta + \mu_i + \xi_t + \varepsilon_{it} \quad (1)$$

where $\Delta y_{i,t+k}/y_{i,t}$ is the growth rate¹³ of per capita GDP of region *i* between time *t* and *t* + 3¹⁴, $y_{i,t}$ is the initial per capita GDP of region *i* and its associated coefficient estimate ψ is generally interpreted as the conditional convergence parameter in panel models (Arbia and Piras, 2005). If the coefficient estimate is negative, convergence exists¹⁵. agglom_{it} is the level of agglomeration of region *i* at time *t*. It is proxied by two indicators: urbanization rate and population density. The vector X_{it} includes additional explanatory variables, such as the level of education of region *i* at time *t* which is measured by gross school enrollment ratio in tertiary education; the role of government in region *i* at time *t* measured by the share of public investment in regional GDP¹⁶; and industrial specialization¹⁷ of regions which is computed using the formula: $\text{spe}_i = \max_j \frac{GDP_{ij}/GDP_i}{GDP_j/GDP}$ where *i* refers to the region and *j* refers to the industry. It is calculated at each time *t*. μ_i represents the non observed regional-

¹³It is worth noting that we do not use GDP in levels because of two reasons. First, we would deviate from the theoretical literature that focuses on the effect of agglomeration on growth and not in economic performance. Second, the relationship between urbanization and GDP in levels is potentially endogenous.

¹⁴We use the logarithmic formula to calculate the growth rate. $\Delta y_{i,t+k}/y_{i,t} \cong \ln y_{i,t+k} - \ln y_{it}$

¹⁵It is worth noting that only three periods are taken into account. Then, the results of convergence do not show a long-run equilibrium. It is interpreted as the convergence of regions to their own steady states.

¹⁶It is the annual average of three years in each period.

¹⁷We compute the indicator of specialization based on seven aggregated sectors: 1. the primary sector which comprises Agriculture, forestry and fishing and Mining and quarrying, 2. the manufacturing sector, 3. the market services sector which comprises: Arts, entertainment and recreation, Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use, Accommodation and food service activities, Professional, scientific and technical activities, Information and communication, Transportation and storage, Real estate activities, Financial and insurance activities which comprises Transportation and storage, 4. public services which comprises: Public administration and defence; compulsory social security, Education and Human health and social work activities, 5. Wholesale and retail trade; repair of motor vehicles and motorcycles, 6. Electricity, gas, steam and air conditioning supply and Water supply; sewerage, waste management and remediation activities and 7. Construction.

specific effects, ξ_t represents the time-specific effects¹⁸ and ε_{it} is the idiosyncratic error term which is independent and identically distributed with mean zero and constant variance. The parsimonious model estimates the parameters α , ψ and γ . The extended model includes the parameters in β .

Since the level of agglomeration, $agglom_{i,t}$, is likely to be correlated with regional characteristics μ_i such as geography or historic factors of development (commonly not observed), the assumption of uncorrelated errors in Ordinary Least Squares estimation is violated. Then, the OLS estimator will be both biased and inconsistent. In order to obtain consistent and efficient estimates, a panel model with fixed effects is adequate. It allows controlling for unobserved heterogeneity.¹⁹

Another issue when estimating the effect of agglomeration on growth is the endogeneity from reverse causality, highlighted in Section 1. In the literature, it has been well documented that the relationship between the two is bidirectional (Martin and Ottaviano, 1999, 2001; Baldwin and Forslid, 2000). On one hand, geographic concentration of economic activity increases growth through the expansion of innovation activity, local input linkages and other positive externalities of agglomeration. On the other hand, it is also likely that higher growth in one region induces firms to locate in that region because it has the highest expenditure level and the highest demand of differentiated products. The method of instrumental variables is appropriate to treat reverse causality. It consists in finding a third set of variables z which has to be relevant (correlated with *agglom*) and valid (uncorrelated with the error term of the regression of y on *agglom*) (Combes and Lafourcade, 2012). The use of this method will depend on the exogeneity test of agglomeration.

Now, we turn to the expected effects of our variables. The impact of agglomeration, measured by urbanization and population density, is ambiguous. As outlined in the literature review, concentrated regions are appropriate to produce positive ag-

¹⁸By including time fixed effects for periods 2001-2003 and 2007-2009, we control the effects of internal and external crisis. Between 1999 and 2002, external and internal crisis have occurred: the Asian crisis in 1998, the Brazilian crisis in 1999 (Brazil is not included in our analysis), the Argentinean crisis in 2000-2001 and the Ecuadorian crisis in 1999. In 2008, world crisis took place.

¹⁹In the case where u_{it} is assumed to be fixed and the remainder disturbances are stochastic with ε_{it} independent and identically distributed with mean zero and a constant variance, the fixed effects model is an appropriate specification. The second case is to assume that u_{it} is random where the individual effect is characterized as random. In order to choose between fixed effects model or random effects model, the Hausman test will be used. (For more details about the Hausman (1978) test, see Baltagi (1995))

glomeration economies (externalities that come from sharing, matching and learning mechanisms (Duranton and Puga, 2003)) that boost economic growth. Then, a positive sign is expected. Nevertheless, negative effects could also emerge. A very high concentration of economic activities may generate traffic jam, high housing prices and high commuting costs. These congestion effects might overshadow growth effects of agglomeration. Then, the resulting sign of the effect of agglomeration would be negative. In the wake of Williamson (1965) and Brülhart and Sbergami (2009), agglomeration benefits would prevail at early stages of development. Hence, we expect to observe positive effects of agglomeration on growth of Latin American regions.

The expected effects of the variables contained in vector X_{it} are the following. First, we consider education of population. Back to the endogenous growth model of (Lucas, 1988), human capital has been widely considered a key factor for growth. The investment (education) in human capital not only increases the productivity of workers but also the productivity in the whole economy through the human capital accumulation. In models linking the NEG and endogenous growth models, the accumulation of knowledge and human capital increases the innovation activity (new varieties of the differentiated good) which, in turn, induces growth (Baldwin and Forslid, 2000; Martin and Ottaviano, 1999, 2001; Fujita and Thisse, 2003; Martin, 1999). Therefore, the effect of the level of education on growth is expected to be positive.

Second, the role of government is introduced. Since governmental investment comprises infrastructure, it represents physical capital, widely argued to influence the technological change in growth models (Lucas, 1988). As the role of governments involves the improvement of living conditions of population, public investment also includes social investment. Unfortunately, we cannot differentiate capital formation from social investment, so that the interpretation of this variable implies the combination of the two.

Third, in growth theory, specialization in one good during the process of learning-by-doing generates increasing returns and productivity growth. This is in concordance with Marshallian externalities which state that the concentration of one industry in a given region facilitates knowledge spillovers between firms and induces growth of that industry and of that region (Glaeser et al., 1992). But if the sector of specialization does not bear potential technological progress, then the economic performance would be poor (Lucas, 1988; Grossman and Helpman, 1991; Romer, 1987). Besides, the effects of specialization on growth depends on the volatility of the sector (Estevadeordal and Volpe-Martincus, 2009). More volatile sectors lead to lower long-term growth rates (Fatás, 2002). Thus, the expected effect of specialization is

ambiguous.

3.2.1 Spatial data issues

Last but not least, when using spatial data, a general issue is the interaction between spatial units. In our case, the spatial dependence is likely because one region's growth can affect neighboring regions' growth. As we describe in Section 2, the centers of agglomeration can gain benefits from surrounding regions. In empirical literature, this issue has been considered (Ciccone and Hall, 1995; Ciccone, 2002; Crozet and Koenig, 2008; Bosker, 2007), however not systematically. Since the spatial dependence violates the independence assumption in the panel model, we are required to test it. We use the Cross Dependence (CD) test proposed by Pesaran (2004)²⁰ for panel data models and the test proposed by Moran (1948), the so-called Moran's I test²¹.

If evidence of interaction effects among regions exists, we deal with this issue by means of spatial econometrics. The strategy is to use spatial weight matrices representing the relation between regions. We presume that spatial dependence occurs in growth of regions. For that reason, we consider a model specification with a spatial lag of the dependent variable as follows.

$$\Delta y_{i,t+k}/y_{i,t} = \alpha + \psi \ln y_{i,t} + \gamma \text{agglom}_{it} + \rho \sum_{j=1}^N w_{ij} \Delta y_{j,t+k}/y_{j,t} + X_{it}\beta + u_{it} \quad (2)$$

where ρ is the spatial dependence coefficient associated with the spatial lag of the dependent variable, w_{ij} are the elements of the spatial weight matrix and u_{it} is the error term.²² Commonly, the spatial weight matrix is row-standardized so that the spatial lag contains the arithmetic averages of neighbors (Anselin, 1988). In our case, the spatial lag of growth rate represents the average growth rate of neighboring regions. In the literature, many ways of specifying the spatial dependence have

²⁰The Pesaran's CD test is based on average of pair-wise correlation coefficients of the OLS residuals from the individual regressions in the panel, and it is used to test for cross section dependence. Through the investigation of the small sample properties of the test, it is shown that the test is appropriate for small samples. Furthermore, the Pesaran's CD test does not require a priori specification of the spatial matrix. For details of the test statistic, see Pesaran (2004)

²¹The Moran's I is a global index of spatial correlation that indicates the degree of similarity between geographical units. The index ranges between -1 (dissimilar values are closer) and 1 (similar values are closer).

²²In order to know whether this specification using the spatial lag of the dependent variable is appropriate, the spatial dependence after the regression has to be ruled out.

been considered. In our model, three specifications of the spatial weight matrix are tested: $k=1$ nearest neighbors-based matrix, distance weight matrix and gabriel-method based weight matrix. These weight matrix specifications are explained in more detail later in Section 3.4.3.

The methodology used is the Maximum Likelihood Estimation because it provides minimum variance unbiased estimators. Given that our database is not large, the computational problems that this method could cause are ruled out.²³

All things considered, a specific-to-general approach is used. We start with a non-spatial panel model and then we test whether the model needs to be extended with spatial interaction effects.

3.3 Results

3.3.1 Non-Spatial Panel models

All things considered, the specific-to-general approach is used. We start with a non-spatial panel model and then we test whether or not the model needs to be extended with spatial interaction effects.

3.4 Results

3.4.1 Panel models

Before presenting the results of panel models, we show the temporal and cross-sectional variations of agglomeration and growth in Table 1. The three period growth rate varies along time and across regions. Urbanization slightly varies along time but strongly across regions. This reflects the spatial disparity within countries. Population density has a similar pattern as urbanization. The consideration of a panel of three periods is important because temporary effects of economic events can be modeled. Put differently, the effect of urbanization can be identified after controlling for time fixed effects, for example, crisis.

²³NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>, 15/01/2014

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Temporal variation	Cross variation
GDPpc growth rate	.04	0.08	.07	.05
urbanization rate	.67	0.21	.013	.21
population density	217.93	1280.37	15.04	1282.87

In Table 2, we present²⁴ the fixed effects panel model estimations using two measures of agglomeration: urbanization in columns (1, 3) and population density in column (2, 4). We show the results for both samples A and B.

²⁴Outlier observations are eliminated. They are Campeche (Mexico), Buenos Aires (Argentina), CABA (Argentina) and El Beni (Bolivia).

Table 2: Non Spatial Panel models using sample A and sample B

	(1)	(2)	(3)	(4)
	urb A	Popd A	urb B	Popd B
urb	0.634 (2.883)***		1.067 (2.456)**	
ln Population density		0.172* (1.912)		0.111 (0.882)
initial GDPpc	-0.160 (-3.772)***	-0.149 (-3.457)***	-0.262 (-3.251)***	-0.219 (-2.542)**
yearcris09	-0.104 (-15.126)***	-0.111 (-13.720)***	-0.106 (-12.419)***	-0.105 (-11.938)***
yearcris01	-0.144 (-9.571)***	-0.142 (-10.155)***	-0.137 (-5.174)***	-0.145 (-6.045)***
Constant	1.014 (3.383)***	0.782 (2.511)**	1.654 (2.981)***	1.535 (3.032)***
N observations	486	486	258	258
N regions	162	162	86	86
Region-specific effects	Y	Y	Y	Y
F	133.0	136.0	110.1	98.52
p-value F	0.000	0.000	0.000	0.000
R^2	0.563	0.560	0.599	0.584
Hausman test	28.99	38.67	29.77	25.08
p-value	0.000	0.000	0.000	0.000

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < 0.01$

According to the Hausman test (bottom of Table 2), the assumption that non-observable individual effects are not correlated with explanatory variables does not hold true and fixed effects model is preferred. The standard errors of the coefficients are robust to heteroskedasticity.

All estimations show a positive and significant effect of agglomeration on economic growth after controlling for region-specific effects and time-specific effects. Both proxies of agglomeration yield to similar results in terms of the coefficient sign; but the effect is not significant enough when using population density. One interpretation of this result might be that urbanization rate is a better measure of agglomeration because it ensures a structure with appropriate conditions to generate agglomeration externalities. Since urbanization is potentially endogenous as mentioned before, we test whether it can be treated as exogenous in our growth model. The instrument that we use is the altitude²⁵. Geological variables are commonly used to correct the endogeneity issue of agglomeration (Combes et al., 2008). According to the difference-in-Sargan statistic²⁶, we cannot reject the null hypothesis that urbanization rate is exogenous (sample A: C-statistic=0.009, p-value=0.925; sample B: C-statistic=0.10, p-value=0.747). Based on these results, we proceed to use urbanization rate as our measure of agglomeration in future estimations.

The coefficient estimate of the GDP per capita at the initial year of each period shows the beta convergence conditional on the level of agglomeration. The speed of convergence²⁷ is about 6% annually when using urbanization and 5% when using population density. These results are larger compared to the Serra et al. (2006)'s estimates of convergence which range between 1% and 2%. They conditioned the convergence rate only with regional characteristics. Given that, regions reach their steady states much faster if the level of agglomeration is taken into consideration. One explanation could be the larger growth effects of agglomeration in low income regions. This presumption will be tested later.

Finally, the assumption that economic growth outputs across regions are independent might not hold true because regions are very likely to be connected to each other, specially within countries. Then, we test the spatial dependence in the

²⁵We use the average altitude of the capital city in each region.

²⁶Under the null hypothesis, the suspected variable can be treated as exogenous. The test statistic is distributed as chi-squared with degrees of freedom equal to the number of regressors tested.

²⁷The speed of convergence is calculated as: $s = -\ln(\psi + 1)/k$ where k is equal to 3 (3 periods) in our case. (Bosker, 2007)

model using the Cross Dependence (CD) test proposed by Pesaran (2004) for panel data. It is a general test based on average pair-wise correlation coefficients of the OLS residuals. Therefore, a priori spatial matrix imposing a relationship pattern is not required. However, back to Tobler’s rule of geography, it is widely argued that close units are more related than distant units. Then, a spatial matrix using distance between regions is convenient for the analysis of spatial dependence. Such a matrix is required to perform the Moran’s I test²⁸. It is a global index of spatial autocorrelation that compute the degree of similarity between regions weighted by their proximity. In order to check the robustness of the spatial dependence test, we conducted both, the Pesaran’s CD test and the Moran’s test. For the latter, we use three spatial matrices: 1. the $k = 1$ -nearest neighbors (W k1), 2. the neighborhood based on Gabriel method (W g) and 3. distance weight matrix (W d). Now, we provide basic concepts of such matrices, leaving a detailed description to Section 3.4.3.²⁹ The first matrix implies that each region has only one neighbor. The second matrix implies more links of neighborhood than the former. The third matrix relates all regions with each other according to the distance between them.

In Table 3, the CD Pesaran’s test shows that outputs of regions are spatially correlated in both samples. For both samples A and B, the Moran’s I reports that global spatial correlation exists, whatever the spatial matrix used. We address this issue by estimating our specification 2 using spatial panel techniques in Section 3.4.3.

²⁸The Moran’s I statistic is computed as in Anselin (1995): $I = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N \sum_{j=1}^N w_{ij}}$.

²⁹Other spatial configurations are included in the analysis of spatial correlation between Latin American regions in Section 3.4.3.

Table 3: Spatial Dependence tests

Test	Statistic value	p-value
Sample A		
CD Pesaran's test	15.025	0.0000
Moran's I (W k1)	period1=0.422	period1=0.000
	period2= 0.438	period2=0.000
	period3=0.437	period3=0.000
Moran's I (W g)	period1=0.43	period1=0.000
	period2= 0.399	period2=0.000
	period3= 0.439	period3=0.000
Moran's I (W d)	period1=0.149	period1=0.000
	period2=0.146	period2=0.000
	period3= 0.168	period3=0.000
Sample B		
CD Pesaran's test	7.023	0.0000
Moran's I (W k1)	period1=0.223	period1=0.048
	period2=0.288	period2=0.016
	period3=0.348	period3=0.005
Moran's I (W g)	period1=0.235	period1=0.006
	period2=0.181	period2=0.024
	period3=0.225	period3=0.008
Moran's I (W d)	period1=0.053	period1=0.030
	period2=0.043	period2=0.056
	period3=0.049	period3=0.039

3.4.2 Non-Spatial Panel Models, Williamson's hypothesis

Now, we turn to the analysis regarding the role of the level of development of regions in the agglomeration-growth relationship. Based on the literature (Williamson, 1965; Henderson, 2000; Davis and Henderson, 2003; Brülhart and Sbergami, 2009)(Henderson, 2003), we presume that the effect of urbanization on growth might be different between regions according to their level of development. The intuition is the following. Since transport and communication infrastructure is scarce at low levels of development, activities concentrate in the region that provides better conditions (Brülhart and Sbergami, 2009). Hence, it is more likely to benefit from agglomeration externalities. Conversely, at high levels of development, better transport connection would promote dispersion. Therefore, agglomeration effects could vanish

in such a dispersed environment.

Since this study considers the heterogeneity of regions within countries, we are able to distinguish the effects at regional level. Thus, we avoid attributing an effect to a whole country.

In order to test the Williamson's hypothesis, the econometric strategy is to run various simulations³⁰ using sub-samples built out of cutoffs according to the level of development (the initial per capita GDP). Put differently, we look at the changes of the effect of urbanization on growth at different levels of development. Through this exercise, the threshold at which the effects of urbanization on growth reach their maximum could be identified. Such a threshold is US\$5700 of per capita income.³¹ Likewise, we attempt to determine the threshold at which urbanization effects become negative. According to the simulations, negative effects seem to appear at US\$10,500 of per capita income. This value is close to the threshold identified by Brühlhart and Sbergami (2009) at US\$10,000 in a worldwide analysis. The results of simulations are shown in Table 13 of Appendix C.

Based on our threshold of US\$5700, we split the sample A into more developed regions and less developed regions. The group of high-developed regions contain 29 regions from Mexico (total 31); 12 from Chile (total 13); 12 from Argentina (total 22), 8 from Colombia (total 33), 1 from Peru (total 24) and 2 from Panama (total 9). The rest of regions of those countries and Bolivian (total 8) and Ecuadorian (total 21) regions are classified into the group of less developed regions.³² The classification of regions is shown in Appendix C.

Table 4 reports the Fixed Effects panel models for low-developed regions (column 1) and high-developed regions (column 2). Moreover, we present the results related to the threshold of negative effects (column 3).

³⁰An alternative econometric strategy would be to introduce an interaction term between urbanization and the level of development. However, such an interaction term implies high collinearity with urbanization rate and the estimates are no longer efficient. In order to avoid multicollinearity, the database is divided in sub-samples according to the initial per capita GDP of regions. Then, we run the regression using a sub-sample of regions that have an income level less/higher than a certain value and we look at the effect of urbanization for those regions.

³¹It is noteworthy that Henderson (2000) found that primacy generates negative effects on growth. The annual growth rate losses from excessive primacy increases until a level of income of US\$4900.

³²The total of regions corresponds to the regions of study without outliers.

Table 4: Panel Models testing the Williamson's hypothesis

	(1)	(2)	(3)
	Low-dev.	High-dev.	neg
urb	0.683 (2.365)**	0.229 (0.771)	-0.0471 (-0.084)
initial GDPpc	-0.174 (-3.274)***	-0.103 (-2.293)**	-0.280 (-5.715)***
yearcris01	-0.169 (-7.815)***	-0.110 (-6.843)***	-0.107 (-4.409)***
yearcris09	-0.0921 (-8.187)***	-0.109 (-12.748)***	-0.1000 (-7.409)***
Constant	1.087 (3.111)***	0.826 (1.991)**	2.794 (3.674)***
N observations	288	198	75
N regions	96	66	25
Region-specific effects	Y	Y	Y
F	45.24	107.5	94.00
p-value F	0.000	0.000	0.000
R^2	0.521	0.717	0.776

t statistics in parentheses

* $p < .10$, ** $p < .1$, *** $p < 0.01$

Firstly, we find a positive effect of agglomeration on regional economic growth for low developed regions. Regarding the group of high developed regions, the coefficient is positive but not statistically significant. This result confirms the Williamson (1965) hypothesis. The coefficient estimates of *urb* are indeed statistically different between the two groups ($t=-0.938$, $p\text{-value}=0.007$)³³. Hence, most of Bolivian, Ecuadorian, Colombian, Peruvian and Panamanian regions enjoy increasing benefits of urbanization. Conversely, most of Mexican, Chilean and Argentinean regions face decreasing benefits of agglomeration, still positive but not significant.

Furthermore, there are indications that negative effects might emerge at higher levels of income than US\$10,500 (column 3). Notwithstanding, the negative effect of urbanization is weak and not significant. Therefore, Latin American regions do not yet face strong negative effects.

Moreover, it seems that low-developed regions have a higher rate of convergence (6%) than high-developed countries (3%)³⁴. However, the difference between the coefficient estimates of *initial GDPpc* is not significant ($t\text{-statistic}=-1.5$ $p\text{-value}=0.135$).

In Section 3.4.3, we present the results when taking into account the potential spatial dependence using spatial panel models. Section 3.4.5 presents model estimations including omitted variables that influence the dependent variable.

3.4.3 Spatial panel models

Upon finding evidence of significant cross section spatial dependence in previous models according to the CD test and Moran's I test, shown in Table 3, we estimate the model using spatial panel data techniques.

In the literature, only few studies analyze the spatial interdependence between Latin American countries. By examining the knowledge spillovers and their channels of diffusion in South America, Guevara and Autant-Bernard (2015) show that the level of productivity in one country generates indirect effects in other countries' productivity. They state that the spatial dependence is driven by both physical proximity and trade intensity between countries. Ramírez and Loboguerrero (2002) show that spatial dependence is significant in a worldwide regression. Notwithstanding, the results for Latin America show no significance of the spatial dependence term.

³³The statistical test is shown in Table 14 of Appendix C

³⁴The formula of the speed of convergence is: $s = -\ln(\psi + 1)/k$ where k is equal to 3 (3 periods) in our case. (Bosker, 2007)

Blanco (2011) focuses on the role of spatial interdependence in the Foreign Direct Investment in Latin America. Her results show that the spatial correlation is not significant.

Given the lack of concensus about the presence of spatial correlation between Latin American countries, this section attempts to deepen the understanding of spatial relationships between these countries and their domestic regions. In our case, it is crucial to look at the intensity of spatial interconnections as they could influence the effects of agglomeration.

In order to study geographic interactions between Latin American regions, spatial connections between countries have to be defined. To this end, we present a brief recap of the integration agreements context in Latin America. Since 1960s important integration agreements have been signed between Latin American countries. The initial agreements were the Latin American Free Trade Association (LAFTA) which included South American countries and Mexico; and Central American Common Market (CACM) which included Central American countries. However, the outcome in terms of integration was ineffective and other agreements between subgroups of countries were signed (ECLAC, 2015). For instance, the agreement of the Andean Community (CAN acronym in spanish) held in 1969 by Ecuador, Bolivia, Peru and Venezuela; and the Southern Common Market (MERCOSUR acronym in spanish) held in 1991 between Brazil, Uruguay, Paraguay and Argentina. Lately, the Union of South American Nations (UNASUR acronym in spanish) has been held in 2011. Given the numerous modifications and agreements, the integration of all Latin American countries is still in progress.

Considering those historical patterns of integration, two spatial configurations between Latin American regions can be defined:

- A: High integration between countries implying international connections of their domestic regions. In this configuration, the link of country x and country y immediately implies relationships between domestic regions of both countries.
- B: Low integration between countries implying weak international connections of regions. In this configuration, the link between domestic regions in the same country is high. But the link of domestic regions of different countries is lower.

The differentiation between the levels of integration will shed light on the mechanisms of spatial interactions between regions and the impact of a given configuration of integration on growth.

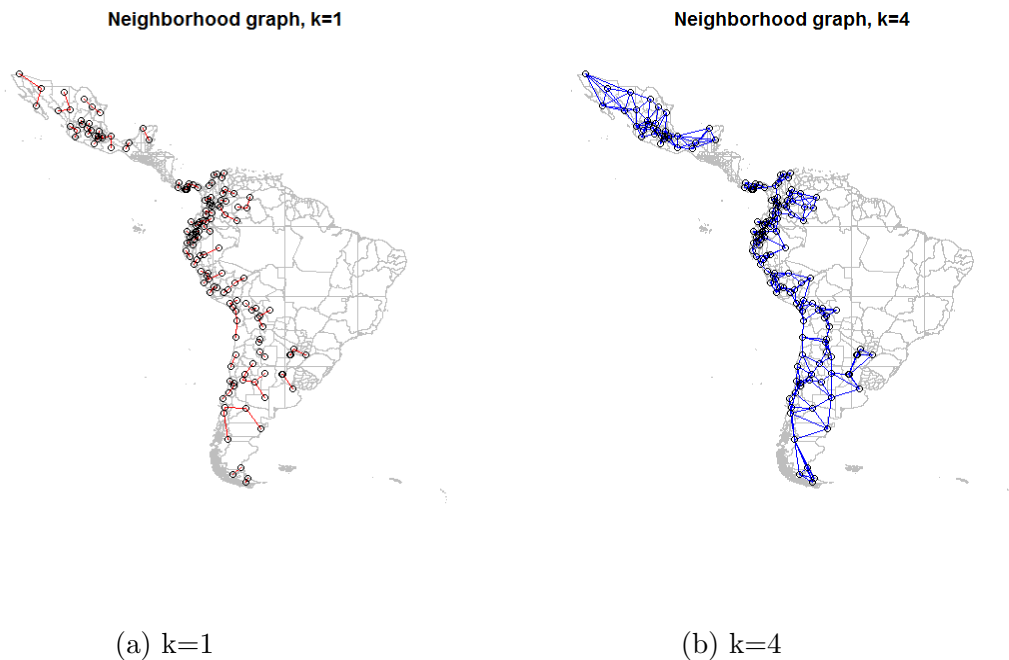
It is worth noting that pairs of regions of the same country tend to establish a stronger connection than pairs of regions of different countries due to national border barriers. These are the so-called border effects (McCallum, 1995). Besides tariffs, another factor determining border effects in Latin America is the difference between national currencies.

Regarding the construction of the spatial weight matrix, we account for geographical data of coordinates of capitals of each region collected from a tool of geo-localization, so-called GeoHack Wikimedia Toolserver Wiki. We use these geographical coordinates to set different neighborhood configurations. To construct the weight matrices, we eliminate islands (Galapagos-Ecuador and San Andres-Colombia).

In this study, we apply the k -nearest neighbors method and the Gabriel method³⁵. We use the inverse of row standardized matrices as suggested by Anselin (1988). The k -nearest neighbors criterion implies that region i is considered as neighbor of region j if their distance is equal, or less than equal, to the minimum possible distance can that be found between region i with all other regions. This definition ensures that each spatial unit has the same number k neighbors. For instance, allowing $k = 1$ implies that each region has one neighbor; allowing $k = 4$ implies that each region has four neighbors. The resulting graphs using such a methodology are shown in Figure 3. In order to connect all regions with each other, we set the number of neighbors to the maximum a region can have. In the case of sample A , there are 160 regions. Then, each region can have at most 159 neighbors. Then, we set $k = 159$.

³⁵This method was proposed by Gabriel and Sokal (1969). Mathematically, two points i and j are connected if the square of the distance between them is less than the sum of the squared distance between each of these points and any other point k .

Figure 3: k-nearest neighbors configuration



In case A of high integration, we use the spatial weight matrix setting k to the maximum. The elements w_{ij} of such a matrix represent the distance between region i and j where every region i is connected to all other regions ($i \neq j$). This spatial weight matrix is noted as (W d).

In case B of low integration, we use the nearest neighbors $k = 1$ spatial weight matrix. It is noted as (W k1). We also define a spatial weight matrix to illustrate pure border effects by assuming that regions from different countries do not connect at all. To do so, the connections in (W d) of regions of country x with those of country y , $x \neq y$ are set to 0. The resulting matrix (W beH) connects all domestic regions within countries but does not connect international regions. The high intensity of connections of domestic regions is represented by H in (W beH).

Using the Gabriel's method, we construct a spatial weight matrix noted as (W g). Such a matrix connects countries through their frontier regions. Thus, it also represents the case A of high integration but to a lesser extent than the spatial matrix (W d). The connectivity graph is shown in Figure 4.

We also use this method to construct another weight matrix of pure border effects. To do so, we proceed as in (W d). The resulting weight matrix (W beL) reflects less number of connections of domestic regions within countries than in (W beH). Low intensity of domestic connections is represented by L in (W beL).

Figure 4: Gabriel method graph



Elaboration: by author

To sum up, the spatial weight matrices that reflect high integration are (W d) interpreted as quasi-complete integration; and (W g) interpreted as moderate integration. The spatial matrices that reflect low integration are (W k1) interpreted as the configuration of lowest integration; (W beH) interpreted as pure border effects with high level of domestic connections; and (W beL) interpreted as pure border effects with moderate level of domestic connections. It is worth mentioning that all matrices illustrate the border effects definition because the intensity of connections decreases with distance between regions, i.e., the further a region is from a given neighbor, the less intense their relation of neighborhood will be.

In order to take into account the spatial interdependence between regions, we es-

estimate a Spatial Autoregressive Model by Maximum Likelihood. The spatial panel model includes a spatially-lagged dependent variable capturing interaction effects between regions. Region-specific effects and time-specific effects are also included.

Before commenting the results, we present the test of spatial dependence. According to the Moran’s I test, the null hypothesis of no spatial correlation is not rejected for most of periods t (see Table 5). However, when using the spatial weight matrix (W k1), spatial correlation remains in periods p1 and p3. In the case of matrices (W g) and (W beL), spatial correlation has been considerably reduced. There is absence of spatial dependence in the residuals after the estimations using the distance weight matrix (W d).

Table 5: Moran’s I test for SAR models Sample A

	(W d)	(W k1)	(W g)	(W beH)	(W beL)
	I (p-value)	I (p-value)	I (p-value)	I (p-value)	I (p-value)
p1	-0.014 (0.337)	-0.195 (0.027)	-0.069 (0.183)	-0.023 (0.339)	-0.056 (0.248)
p2	-0.003 (0.421)	-0.078 (0.233)	-0.008 (0.488)	0.007 (0.369)	-0.005 (0.492)
p3	-0.016 (0.303)	-0.266 (0.004)	-0.175 (0.008)	-0.051 (0.132)	-0.175 (0.011)

In order to interpret the estimates, it is worth noting that the presence of the spatial autoregressive parameter in our model produces indirect effects that come from geographical proximity. Thus, the total effect of the changes in the explanatory variables has two components: the direct effect or local effect and the indirect effect or spatial effect. These effects come from the matrix of elasticities Ξ . In the case of the estimate coefficient of agglomeration, we have the following matrix.

$$\Xi_Y = \hat{\gamma}(I - \hat{\rho}W)^{-1}$$

The direct effect of agglomeration in region i on its economic growth is represented by the i , i th element from the diagonal of matrix Ξ . The indirect effect of agglomeration of region j on economic growth of region i is represented by the i , j th element of such a matrix.

In Table 6, we present the results using the aforementioned spatial weight matrices. In column (1), the distance matrix (W d) representing the configuration of high integration is used, in column (2), the matrix based on the Gabriel method

(W g) representing the moderate configuration, in column (3), the matrix based on k nearest neighbor with k=1 (W k1) representing the configuration of low integration. The results using matrices reflecting pure border effects are also shown: border effects with high intensity of intraconnection (W beH) is used in column (a) and border effects with low intensity of intraconnection (W beL) is used in column (b). Both direct and indirect effects are presented.

Table 6: SAR Model with Parsimonious specification-Database A

	(1)	(2)	(3)	(a)	(b)
	(W d)	(W g)	(W k1)	(W beH)	(W beL)
Direct					
urb	0.598 (3.332)***	0.594 (3.274)***	0.641 (3.508)***	0.580 (3.224)***	0.572 (3.159)***
initial GDPpc	-0.167 (-5.779)***	-0.161 (-5.510)***	-0.165 (-5.621)***	-0.165 (-5.690)***	-0.159 (-5.486)***
yearcris01	-0.120 (-12.389)***	-0.124 (-13.035)***	-0.129 (-13.635)***	-0.121 (-12.761)***	-0.123 (-12.965)***
yearcris09	-0.0818 (-10.546)***	-0.0865 (-11.476)***	-0.0934 (-12.656)***	-0.0763 (-9.521)***	-0.0845 (-11.200)***
Spatial					
ρ	0.403 (6.054)***	0.225 (5.797)***	0.134 (4.677)***	0.610 (6.478)***	0.245 (6.324)***
Variance					
σ_e^2	0.00173 (15.481)***	0.00172 (15.383)***	0.00176 (15.404)***	0.00171 (15.448)***	0.00169 (15.365)***
Indirect					
urb	0.446 (2.573)**	0.166 (2.845)***	0.0975 (2.892)***	0.261 (2.708)***	0.162 (2.832)***
initial GDPpc	-0.125 (-3.256)***	-0.0448 (-3.989)***	-0.0251 (-3.817)***	-0.0741 (-3.693)***	-0.0451 (-4.145)***
Total					
urb	1.043 (3.170)***	0.760 (3.262)***	0.738 (3.496)***	0.840 (3.181)***	0.734 (3.155)***
initial GDPpc	-0.292 (-4.935)***	-0.205 (-5.457)***	-0.190 (-5.591)***	-0.239 (-5.396)***	-0.204 (-5.450)***
N observations	480	480	480	480	480
N regions	160	160	160	160	160
Region-specific effects	Y	Y	Y	Y	Y
R^2	0.585	0.584	0.581	0.587	0.585
Hausman test	16.40	16.67	17.88	17.74	17.45
p-value	0.00580	0.00516	0.00310	0.00330	0.00372

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < 0.01$

The comments of the results are as follows. Firstly, we observe that Latin American regions have a positive and highly significant effect of urbanization in their economic growth after controlling for spatial dependence. The coefficient estimate of the direct effect is stable across all spatial configurations. It is slightly lower than the coefficient estimate of the panel model (see Table 2). The reason is that a part of the total effect is attributed to spatial interactions between regions.

The spatial effects of urbanization are significant. It indicates that urbanization in one region brings benefits for growth of its neighbors. It could be indirectly related to knowledge spillovers among regions. Urban conditions in one region would allow knowledge diffusion towards neighboring urbanized regions thank to geographical proximity. Thus, economic growth of close regions is positively affected.

The highest indirect effect of urbanization is observed when using the distance weight matrix (W_d). The indirect effect of urbanization is small when using the $k=1$ nearest neighbors-based matrix (W_{k1}). This intuitive result was expected given that the spatial configuration of (W_d) implies higher integration than the spatial configuration of (W_{k1}). Regarding spatial configurations with border effects, the indirect effects of urbanization are larger in the framework of high intensity of internal connections (W_{beH}) than in the framework of low intensity of internal connections. The more interconnected the regions, the higher the spatial growth effects of urbanization.

The total effect of the coefficient of convergence is close to the panel model's estimate. The contribution of indirect effects coming from geographical proximity between regions to convergence is small but significant. Besides, those spatial effects vary depending on the spatial configuration used. This implies that proximity between regions matters for the process of convergence as stated by Serra et al. (2006).

The spatial autocorrelation coefficient is statistically positive and significant for all models. It assess the extent to which one region's growth affects the economic growth of other regions. The degree of spatial autocorrelation increases with the level of integration that spatial weight matrices reflect. For instance, the value of ρ is high when using the distance weight matrix (W_d), which exhibits a quasi-complete integration and it is low when using the $k = 1$ nearest neighbor matrix (W_{k1}), which exhibits low integration between Latin American regions.

Regarding the analysis of border effects spatial configurations, interesting in-

sights are drawn. By eliminating the connections of regions from different countries and keeping the internal connections (W beH), the spatial correlation coefficient ρ is large, even larger than the coefficient estimate of (W d) where all regions are connected. It means that the interdependence between regions within one country is much higher than between regions of different countries. Similarly, the spatial autocorrelation coefficient of (W beL) is higher than that of (W g).³⁶ Based on these results, we confirm that border effects are significant in Latin America. The aspects that could explain this result are related to the diversity of national currencies across countries and to the scarce infrastructure across frontiers.³⁷

The model that fit better the data is that using (W g) which links South America with Mexico. Such a configuration seems to be appropriate because in trade terms Mexico is one of the ten largest export destination countries of South America accounting for 3% of total exports; and it is one of the ten largest import origin countries representing 3.8% in total imports (Guevara and Jarrin, 2011). On the contrary, the distance weight matrix (W d) that shows a strong relationship might not explain the actual integration context of Latin American countries. Even though the relationship between countries within Latin America has improved, it does not mean that there exists a complete integration. Border regions might be connected but it is hard to think that all regions within a country can be connected with all other regions within another country. All things considered, we rely on the results of the model using the (W g) spatial weight matrix.

3.4.4 Spatial Panel Model, Williamson's hypothesis

Now, we estimate Spatial Autoregressive models for each group of regions (Table 7): low-developed regions (column 1) and high-developed regions (column 2).

Regarding the spatial correlation parameter, it is significant for both high-

³⁶Recall that (W beL) is based on the spatial configuration of (W g). For that reason, we compare the results between these two matrices.

³⁷Only recently, UNASUR has proposed an Initiative for the Integration of South America regarding infrastructure (IIRSA). The planing of such a project started in 2000 and its implementation in 2005. Hence, the outcomes of said initiative are not tangible yet. In this regard, one direction of future research is the assessment of the implementation of infrastructure projects proposed by IIRSA in modifying the strength of border effects. To this end, spatial weight matrices using actual flows of trade or road density would be appropriate instead of spatial weight matrices based on geographical distance used here.

Table 7: Spatial panel models for low-developed and high-developed regions

	(1)	(2)
	(W g) Low-dev	(W g) High-dev
<hr/>		
Direct		
urb	0.682 (2.462)**	0.350 (1.286)
initial GDPpc	-0.177 (-4.955)***	-0.102 (-3.034)***
yearcris01	-0.142 (-10.574)***	-0.0815 (-8.180)***
yearcris09	-0.0849 (-6.136)***	-0.0810 (-10.002)***
<hr/>		
Spatial		
ρ	0.235 (4.536)***	0.287 (6.074)***
<hr/>		
Variance		
σ_e^2	0.00214 (11.904)***	0.000766 (9.712)***
<hr/>		
Indirect		
urb	0.201 (2.359)**	0.143 (1.412)
initial GDPpc	-0.0513 (-3.124)***	-0.0406 (-2.407)**
<hr/>		
Total		
urb	0.897 (2.845)***	0.504 (1.487)
initial GDPpc	-0.230 (-4.512)***	-0.145 (-2.687)***
<hr/>		
N observations	288	192
N regions	96	64
Region-specific effects	Y	Y
R^2	0.552	0.753
Hausman test	12.69	7.583
p-value	0.0264	0.181

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < 0.01$

developed and low-developed regions. It seems to be higher in high-developed regions than in low-developed regions. However, the difference between both coefficient estimates is not statistically significant (t-statistic = -0.744391226, p-value > 0.05).

The results confirm the claim that positive agglomeration effects on economic growth increase up to a certain threshold of development level. The total effect of urbanization for low-developed regions is significant at low stages of development (0.68) whereas at later stages, agglomeration does not impact on economic growth.

The conditions of development might produce specific mechanisms that provoke such differences. One of those conditions could be the urbanization rate itself. At low income levels, the pace of urbanization is rapid (Henderson, 2003). In our sample, low-developed regions record an average growth rate of urbanization of 5.8% between 2001 and 2007,³⁸ whereas high-developed regions have a lower average growth rate of urbanization of 2.3% in the same period. The difference between those means is statistically significant (t statistic= 13.64, p-value=0.000).

We also analyze other variables, omitted in the model, that account for conditions of development of each country. Such omitted variables are most likely to be part of fixed-specific effects in the models as they represent characteristics of development that are not assumed to change in a short period of time, abruptly. Recall that according to our threshold, most of regions of Mexico, Chile and Argentina are classified in the group of high-developed regions of Latin America. In terms of education, Mexico and Chile stand as countries with the highest shares of working population with master's degree.³⁹ The countries with the highest ratios of research and development expenditure over GDP are Argentina (0.48%) and Mexico (0.43%) followed by Chile (0.43%) and Ecuador (0.39%). Regarding transportation infrastructure, Panama (34.6%), Mexico (32.8%) and Argentina (29.4%) record the highest percentages of road paved.⁴⁰

In short, Mexico, Chile and Argentina are commonly the countries with the best indicators in terms of development. The high values of those macroeconomic variables implicitly reflect the enormous development of their internal markets. A

³⁸These statistics corresponds to sample A divided into groups of developed and developing regions.

³⁹The indicators are 0.087% for Mexico and 0.11% for Chile in 2009. The information is obtained from the database of Network of Indicators of Science and Technology (RICYT acronym in spanish).

⁴⁰The information corresponds to 1999 and the source of the data is the World Development Indicators of the World Bank.

process of rapid agglomeration surely took place in those markets which, in turn, induced further agglomeration. The progress is such that one can reasonably ask whether the benefits of agglomeration might be reaching their limits. According to our results, most of regions of these countries face non-significant effects. Conversely, it seems that the other countries with low levels of development are in a certain stage in which their embryony local infrastructure, markets, human capital and other development conditions promote an accelerated process of agglomeration which, in turn, produces positive effects for growth.

3.4.5 Extended Spatial Panel Model

Until now, this study has introduced key variables of our specification 2 but some explanatory variables still miss in the model. The following extended model includes a set of observable control variables (X_i in our specifications 1 and 2). It is estimated only with sample B due to the availability of additional variables. To recall, sample B contains information on 85 regions⁴¹ of Colombia, Ecuador and Mexico over the period 2001-2009.

In concordance with growth theory, we include the enrollment ratio in tertiary education, public investment and level of industrial specialization. The first variable measuring the mass of educated population could entail collinearity with urbanization since access to education would improve with urbanization. The willingness of families to invest in education would increase in urban areas where schools and universities facilities are located. Indeed, the correlation coefficient between education and urbanization is 0.6320, which is significant⁴². Besides, we think that the enrollment ratio in tertiary education could be an alternative measure of agglomeration bearing benefits for growth. Thus, it could be conceived as one of the Marshallian agglomeration economies: the local pool of skilled labor (Marshall, 1890). For that reason, we present in Table 8 the estimations with both variables urbanization and tertiary education in column 1, tertiary education only in column 2, and urbanization only in column 3.

⁴¹Sample B initially has 87 regions but 2 island regions (Galapagos from Ecuador and San Andres from Colombia) and 1 outlier region (Campeche from Mexico) are eliminated.

⁴²The correlation matrix between variables included in the extended model is presented in Appendix D.

Table 8: Spatial Autoregressive Extended Models

	(1)	(2)	(3)
	(W g) urb-educ	(W g) educ	(W g) urb
Direct			
urb	0.343 (1.272)		0.719 (2.715)***
initial GDPpc	-0.375 (-7.065)***	-0.370 (-9.065)***	-0.253 (-6.051)***
yearcris01	-0.0911 (-6.981)***	-0.0941 (-7.166)***	-0.107 (-9.400)***
yearcris09	-0.0860 (-9.431)***	-0.0841 (-7.997)***	-0.0772 (-7.538)***
inv_GDP	0.193 (3.368)***	0.188 (3.118)***	0.241 (4.043)***
tertiary_educ	0.506 (3.831)***	0.561 (4.881)***	
IER	-0.0453 (-3.120)***	-0.0431 (-3.055)**	-0.0480 (-3.366)***
Spatial			
ρ	0.304 (6.161)***	0.306 (6.203)***	0.298 (5.905)***
Variance			
σ_e^2	0.00116 (11.077)***	0.00117 (11.074)***	0.00124 (11.077)***
Indirect			
urb	0.139 (1.165)		0.286 (2.380)**
initial GDPpc	-0.152 (-3.736)**	-0.155 (-4.021)***	-0.101 (-3.765)**
inv_GDP	0.0771 (2.799)**	0.0786 (2.599)**	0.0955 (3.299)***
tertiary_educ	0.205 (2.810)**	0.236 (3.329)**	
IER	-0.0184 (-2.586)***	-0.0179 (-2.724)***	-0.0192 (-2.913)***
Total			
urb	0.482 (1.252)		1.005 (2.707)***
initial GDPpc	-0.526 (-6.541)***	-0.525 (-7.477)***	-0.354 (-5.790)***
inv_GDP	0.270 (3.412)***	0.266 (3.070)***	0.336 (4.092)***
tertiary_educ	0.711 (3.712)***	0.796 (4.592)***	
IER	-0.0637 (-3.095)***	-0.0610 (-3.058)**	-0.0672 (-3.374)***
N observations	252	252	252
N regions	84	84	84
Region-specific effects	Y	Y	Y
R^2	0.692	0.688	0.681
Hausman test	26.78	25.37	22.77
p-value	0.000771	0.000652	0.00187

t statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < 0.01$

The urbanization rate is no longer significant after controlling for more variables and spatial autocorrelation (see column 1 of Table 8) On the contrary, the enrollment ratio in tertiary education⁴³ is highly significant in models (1) and (2). This could indicate that skilled labor pool might be a stronger measure of agglomeration than urbanization. When excluding the variable of education (see column 3 of table 8), the urbanization rate becomes significant.

Moreover, the effects of educated labor force spill across regional borders. The mass of educated workers in one region impacts on its neighbors' growth. Those spatial effects represent the so-called knowledge spillovers⁴⁴. The intuition regarding this positive growth effect of education on neighboring regions is the following. When a new product, resulting from research and development investment, is created in one region, neighboring regions would have access to it thank to geographical proximity. Thus, the economic growth of close regions is positively affected.

The public investment has a positive impact. It fosters economic growth of regions through the improvement of transport connectivity and the creation of opportunities for population. Public investment in basic services, such as education and health improves living conditions which allow working-age population to access to the labor market which, in turn, increases productivity and economic growth. On the other hand, investment in communication and transportation infrastructure also induces to higher growth by increasing productivity.

Regarding spatial effects of public investment, they are significant and account for almost 45% of the total effect. Given that the indivisible facilities funded by public investment cannot be used in more than one region, we presume that significant spatial effects come from social investment such as education and health.

The concentration of production in few sectors seems to be detrimental for regional growth. The negative effect could be due to the nature of the sector of specialization. As mentioned before, if the sector in which one region specializes does not entail potential technological progress, its effect on growth would be negligible. Besides, if the main sector is volatile, external shocks can significantly reduce its performance. In this sample, many regions specialize in the primary sector, manufacturing and public services and very few regions specialize in services industry. Then, the negative average effect may come from the first three sectors. We presume

⁴³We also use the number of schooling years of labor force to measure skilled labor force but the estimation shows a lower effect. Based on that, the mass of skilled labor force matters more than the level of education of workers.

⁴⁴Previously, we have stated that spatial effects of urbanization are indirectly explained by knowledge spillovers.

that the primary sector is very likely to bear potential negative effects due to its high volatility.

By bringing together the effects of human capital and public investment, the speed of convergence increases. Now, the regions of the three countries converge at rates of 16% when using skilled labor force and 9% when using urbanization. This could be due to the presence of spatial effects between regions. As mentioned before, a region that spills over its knowledge will affect its neighbor's growth through spatial effects. Likewise, there exist positive spatial effects of public investment. Thus, regions would converge more rapidly to their steady states.

4 Conclusions

In this study, we seek to answer three questions regarding agglomeration effects on economic growth of Latin America at regional level. The first question consists in establishing whether the level of agglomeration of regions in these developing countries influences their growth. The second question investigates to what extent the effects of agglomeration on growth change with the level of development of regions. Finally, the third question attempts to determine the degree of spatial interdependence between regions. To answer these questions, this study analyzes a panel data of regions of eight countries over the last decade by means of panel and spatial panel data methodologies.

Some conclusions can be drawn from this empirical investigation. One of them is the clear evidence that agglomeration is of significant importance for the economic growth of Latin American regions. However, the effects vary across regions. As Williamson (1965) argued, the key element of the differences is the level of development. According to our results, at low stages of development less than 5,700 dollars per capita income, the effects of urbanization are magnified. Then, the effect of urbanization decreases until a threshold of 10,500 dollars per capita income, at which negative effects might appear. However, those negative effects are weak and not significant, suggesting that Latin American regions do not yet face strong negative effects.

The reasons that explain the non-monotonic effect of agglomeration on growth are twofold. On one hand, low-developed regions enjoy a rapid pace of urbaniza-

tion so that benefits are more important there than in high-developed regions where urbanization slows down (Henderson, 2000; Davis and Henderson, 2003) (Henderson, 2003). On the other hand, changes of agglomeration effects on growth are due to the level of development inherent to regions. At higher stages of development, development features, such as the level of education and R&D investment can act as mechanisms that produce over-agglomeration, causing less pronounced marginal effects on growth.

Additionally, Latin American regions are spatially interdependent. Therefore, indirect effects on regional growth due to geographical proximity exist. Those effects depend on the level of integration between countries and between regions within countries. The results suggest that the international integration of Latin American countries is beneficial to obtain spatial benefits of agglomeration. Nevertheless, spatial effects are even more important when there is only intra-national integration with null international integration. This finding points out the strong border effects in Latin America. Lastly, the regions of some Latin American countries (sample B) enjoy significant effects of human capital which diffuse across space. Likewise, public investment produces positive direct and indirect spatial effects. Those elements together make regions to converge more rapidly to their steady states.

The current study contributes to existing literature in the following aspects. First of all, it responds to the claim of extending the knowledge base about the effects of agglomeration for developing countries by looking at sub-national regions in Latin America. Therein, we have been able to distinguish the specific effect of regions within countries instead of attributing a general effect to a whole country. Given that, we could test the Williamson's hypothesis taking into account the heterogeneity of regions within countries in terms of their levels of development. By considering the spatial dimension, we have also been able to quantify the marginal effects of agglomeration coming from geographical proximity. It provides interesting insights about the intensity of spatial effects of agglomeration depending on the level of integration in Latin America, which involves international and intra-national connections.

Despite this contribution, some limitations of the study have to be underlined. The major limitation stems from the little production of statistics at sub-national level in Latin American countries regarding some variables that are commonly introduced in growth models. This issue was partly solved in the extended model using sample B. Likewise, this study is constrained by the lack of data of some Latin American countries as Brazil, Venezuela and the Central American countries.

Furthermore, careful interpretation of the results is requested since the temporal dimension of the sample is rather small. In addition, the definition of agglomeration using urbanization is another limitation since it does not take into account the industrial composition of regions. The consideration of such an element is a complementary topic to be investigated. The fact that developing countries typically have predominant primary activities would have influence on agglomeration externalities. Lastly, regional integration defined by geographical distance is limited. Although spatial weight matrices based on distance illustrate general spatial configurations, they cannot shed light on the evolution of spatial patterns. Using spatial weight matrices based on trade flows or road density would provide promising results regarding the evolution of economic integration. This analysis constitutes one direction for future research. In particular, an assessment of the effect of inter-regional infrastructure projects on economic integration, recently undertaken in Latin America, could be conducted.

A Appendices

A.1 Economic Concentration within LAC

Table 9 shows the three regions of each country with high values of Gross Domestic Product of 2009. The row total indicates the total of regions of analysis for each country and the total GDP. Overall, the capital regions of countries generates most of the national GDP.

Table 9: LAC: share of regions in the total GDP, 2009

Country	Number region	Name region	GDP(US\$billion) 2009	Share of total GDP
Argentina	1	Buenos Aires	150	0.333
Argentina	2	Buenos Aires Autonomous city	110	0.244
Argentina	3	Catamarca	5.5	0.012
Total	24		450	1
Bolivia	1	Santa Cruz	4.5	0.284
Bolivia	2	La Paz	3.8	0.24
Bolivia	3	Cochabamba	2.5	0.158
Total	9		15.85	1
Chile	1	Santiago Metropolis	97	0.482
Chile	2	Bio bio	20	0.099
Chile	3	Valparaiso	18	0.089
Total	13		201.3	1
Colombia	1	Bogota Capital District	89	0.264
Colombia	2	Antioquia	51	0.151
Colombia	3	Valle	37	0.11
Total	33		337.54	1
Ecuador	1	Guayas	21	0.3
Ecuador	2	Pichincha	17	0.243
Ecuador	3	Manabi	5.7	0.082
Total	22		69.88	1
Mexico	1	Mexico District	240	0.184
Mexico	2	Mexico State	120	0.092
Mexico	3	Nuevo Leon	100	0.077
Total	31		1301.3	1
Peru	1	Lima	62	0.52
Peru	2	Arequipa	6.8	0.057
Peru	3	La libertad	5.8	0.049
Total	24		119.18	1
Panama	1	Panama	25	0.699
Panama	2	Colon	4.7	0.131
Panama	3	Chiriqui	2.4	0.067
Total	9		35.74	1

Source: Institutes of Statistics of each country. Elaboration: by author.

Table 10 contains the regions with the maximum and minimum GDP per capita of each country. The gap between them is generally high, except for Bolivia.

Table 10: LAC: gap between high income and low income regions by country, 2009

Country	US\$ GDPpc (max)	US\$ GDPpc (min)	Gap between regions (times)
Argentina	35162.55	3961.746	8.88
Bolivia	3122.311	1170.295	2.67
Chile	24134.7	5485.635	4.4
Colombia	23280.88	2137.815	10.89
Ecuador	16626.52	1884.339	8.82
Mexico	57132.73	5186.45	11.02
Peru	9943.319	1350.976	7.36
Panama	19652.26	2775.971	7.08
Mean			7.64

Source: Institutes of Statistics of each country. Elaboration: by author.

Table 11 reports the incidence of poverty of regions by country.

Table 11: Incidence of Poverty in regions within LAC

Year	Country	Poverty measure	Name region	Poorest region	Name region	Least poor region
2010	Mexico	Income poverty	Chiapas	0.78	Nuevo Leon	0.21
2009	Chile	Income poverty	Arauca	0.27	Antofagasta	0.08
2010	Colombia	Income poverty	Choco	0.65	Bogota	0.15
2010	Peru	Income poverty	Huancavelica	0.66	Madre de Dios	0.09
2008	Ecuador	Income poverty	Bolivar	0.66	Pichincha	0.19
2001	Argentina	Unsatisfied Basic Needs	Formosa	0.34	Buenos Aires	0.08
2001	Bolivia	Unsatisfied Basic Needs	Potosi	0.8	Santa Cruz	0.38

Source: Institutes of Statistics of each country. Elaboration: by author.

B Primacy in LAC

Table 12 records the primacy rates of Latin American countries between 1950 and 2000.

Table 12: Primacy in Latin American countries 1950-2000

Country	Primate region	1950	1960	1970	1980	1990	2000
Panama	Panama	0.6027	0.6484	0.6712	0.6780	0.6815	0.7018
Paraguay	Central	0.5338	0.5592	0.5938	0.5813	0.5699	0.5770
Uruguay	Montevideo-Canelones		0.6358	0.6130	0.6017	0.5892	0.5728
Argentina	Buenos Aires	0.6071	0.5996	0.5952	0.5625	0.5269	0.5232
Chile	Metropolitana	0.4135	0.4269	0.4433	0.4457	0.4555	0.4488
Peru	Lima	0.3238	0.4166	0.4411	0.4493	0.4411	0.4398
Ecuador	Guayas	0.3159	0.3561	0.3545	0.3501	0.3588	0.3656
Bolivia	La Paz	0.4004		0.3621		0.3231	0.3005
Mexico	Mexico (Federal District)	0.2964	0.3049	0.3191	0.3350	0.2847	0.2734
Colombia	Cundinamarca	0.2959	0.2186	0.2411	0.2463	0.2483	0.260
Brazil	Sao Pablo	0.2558	0.2562	0.2741	0.2760	0.2641	0.2508
Venezuela	Miranda (Federal District)	0.3024	0.3137	0.3124	0.2753	0.2491	0.2071

Source: ECLAC database based on National Population Censuses. Elaboration: by author.

C Definition of the threshold of development

Threshold Estimation

In order to determine the threshold of the level of development at which the effects of urbanization are maximized, we run some simulations. We split the database by income levels. For instance, at an income level of US\$4000, we keep regions under such a value and we estimate the model. We look at the parameter estimates and so on. According to the results, the maximum effects of urbanization on growth are reached between US\$5500 and US\$6000 of per capita income. Table 13 shows the results when setting the parameter of the level of development lower than US\$5000, lower than US\$5700, lower than \$6000 and lower than US\$7500. The maximum coefficient is reached at US\$5700 of per capita income. Likewise, we identify the threshold at which the coefficient estimate of urbanization becomes negative.

Table 13: Simulations, level of development threshold

	(1)	(2)	(3)	(4)	(5)
	< 5000	< 5700	< 6000	< 7500	> 10500
urb	0.640 (2.132)*	0.683 (2.365)*	0.671 (2.366)*	0.648 (2.438)*	-0.0471 (-0.084)
initial GDPpc	-0.170 (-3.091)**	-0.174 (-3.274)**	-0.171 (-3.261)**	-0.161 (-3.255)**	-0.280 (-5.715)***
yearcris01	-0.176 (-7.804)***	-0.169 (-7.815)***	-0.169 (-7.849)***	-0.160 (-8.106)***	-0.107 (-4.409)**
yearcris09	-0.0915 (-7.451)***	-0.0921 (-8.187)***	-0.0937 (-8.664)***	-0.0979 (-10.308)***	-0.1000 (-7.409)***
Constant	1.074 (2.960)**	1.087 (3.111)**	1.070 (3.124)**	1.001 (3.035)**	2.794 (3.674)**
N observations	270	288	300	354	75
N regions	90	96	100	118	25
F	41.31	45.24	48.47	73.82	94.00
p-value F	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.529	0.521	0.522	0.525	0.776

t statistics in parentheses

* $p < .05$, ** $p < .01$, *** $p < 0.0001$

In order to test whether the differences in the effects of urbanization between the regions with lower levels of income and the regions with higher levels of income than the threshold are significant, we estimate a pooled interaction regression equation. Therein, we introduce an interaction variable between the regressor urbanization and a dummy variable indicating the group of development (low developed and high developed region), $levelHdev * urb$. Table 14 presents the regression for low developed regions (column 1), the regression for high developed regions (column 2) and the regression testing the difference of the estimates of urbanization urb between the two (column 3). The estimation of the interaction term indicates that the differences between both type of regions is significant.

Table 14: Test of the differences of effects of urbanization by level of development

	(1)	(2)	(3)
	Low dev.	High dev.	difference test
urb	0.683 (2.365)*	0.229 (0.771)	
$levelHdev * urb$			-0.939 (-2.719)**
initial GDPpc	-0.174 (-3.274)**	-0.103 (-2.293)*	-0.164 (-3.817)**
yearcris01	-0.169 (-7.815)***	-0.110 (-6.843)***	-0.146 (-9.657)***
yearcris09	-0.0921 (-8.187)***	-0.109 (-12.748)***	-0.0998 (-14.164)***
Constant	1.087 (3.111)**	0.826 (1.991)	1.213 (3.723)**
N observations	288	198	486
N regions	96	66	162
R^2	0.521	0.717	0.569

t statistics in parentheses

* $p < .05$, ** $p < .01$, *** $p < 0.0001$

Based on previous results, less developed regions and high developed regions are classified as follows.

Table 15: LAC: Regions by level of development

Low-developed regions			High-developed regions		
Region	Country	GDPpc 2001	Region	Country	GDPpc 2001
Chiapas	Mexico	5636.268	Neuquen	Argentina	35032.38
Azuay	Ecuador	5319.271	Casanare	Colombia	28147.75
Oaxaca	Mexico	5185.019	Antofagas	Chile	24097.25
Pichincha	Ecuador	5158.448	Tierra de	Argentina	24057.04
Vichada	Colombia	5136.355	Distrito	Mexico	24004.46
Entre rio	Argentina	5125.451	Nuevo leo	Mexico	20082.19
Boyaca	Colombia	4940.763	Santa cru	Argentina	19279.38
Santiago	Argentina	4907.089	Magallane	Chile	18543.26
Lima	Peru	4810.598	Quintana	Mexico	15912.08
San juan	Argentina	4659.29	Coahuila	Mexico	15538.44
Araucania	Chile	4527.089	Colon	Panama	14206.61
Tacna	Peru	4491.076	Tamaulipa	Mexico	13529.93
Chiriqui	Panama	4485.421	Tabasco	Mexico	13272.18
Atlantico	Colombia	4463.1	Baja cali	Mexico	13183.52
Caldas	Colombia	4457.753	Tarapaca	Chile	13141.12
Guayas	Ecuador	4377.935	Baja cali	Mexico	12931.24
Corriente	Argentina	4271.22	Atacama	Chile	12640.15
Huila	Colombia	4207.118	Queretaro	Mexico	12111.4
Tungurahua	Ecuador	4151.876	Aguascal	Mexico	12041.04
Quindio	Colombia	4146.981	Chihuahua	Mexico	11961.19
Bolivar	Colombia	4128.348	Jalisco	Mexico	11626.67
Risaralda	Colombia	4126.507	Region me	Chile	11547.72
Misiones	Argentina	4087.74	Sonora	Mexico	11422.69
Jujuy	Argentina	4053.412	Colima	Mexico	11368.88
Tolima	Colombia	4048.406	Aisen del	Chile	10450.27
Canar	Ecuador	4046.895	Durango	Mexico	10239.92
Bocas del Toro	Panama	4037.679	Panama	Panama	9838.929
Chaco	Argentina	4004.932	Guanajuat	Mexico	9585.483
Los Santos	Panama	3910.067	San luis	Argentina	9396.421
Cotopaxi	Ecuador	3889.771	Morelos	Mexico	9142.687
El oro	Ecuador	3884.909	Chubut	Argentina	9090.517
Pasco	Peru	3879.04	Sinaloa	Mexico	8839.887
Arequipa	Peru	3772.823	Distrito	Colombia	8834.176
Carchi	Ecuador	3676.799	Valparais	Chile	8802.368
Cesar	Colombia	3597.037	Moquegua	Peru	8739.193
Formosa	Argentina	3560.928	Yucatan	Mexico	8584.496
Esmeralda	Ecuador	3526.708	San luis	Mexico	8469.441
Herrera	Panama	3470.285	La pampa	Argentina	8245.575
Zamora-ch	Ecuador	3419.025	Santa fe	Argentina	8069.77
Madre de	Peru	3406.235	Bio-bio	Chile	8001.209
Imbabura	Ecuador	3367.316	Cordoba	Argentina	7977.862
La guajir	Colombia	3354.643	Arauca	Colombia	7866.877
Los rios	Ecuador	3352.762	Libertado	Chile	7797.634
Cordoba	Colombia	3293.139	Buenos Aires	Argentina	7450.283
Salta	Argentina	3227.938	Mexico	Mexico	7401.308
Pastaza	Ecuador	3218.985	Rio negro	Argentina	7305.858
Norde de	Colombia	3209.381	Puebla	Mexico	7270.737
Ancash	Peru	3175.512	Los lagos	Chile	7013.758
Cocle	Panama	3133.973	Michoacan	Mexico	6980.882
Loja	Ecuador	3043.689	Veracruz	Mexico	6899.355
Darien	Panama	3020.319	Santander	Colombia	6897.605
Amazonas	Colombia	3008.243	Hidalgo	Mexico	6873.152
Tucuman	Argentina	2974.215	Nayarit	Mexico	6777.711
Manabi	Ecuador	2957.994	Tlaxcala	Mexico	6766.112
Guainia	Colombia	2905.528	Maule	Chile	6513.725
Cauca	Colombia	2826.27	La rioja	Argentina	6380.576
Veraguas	Panama	2774.959	Valle del	Colombia	6368.038
Ica	Peru	2767.031	Coquimbo	Chile	6332.168
Magdalena	Colombia	2740.459	Mendoza	Argentina	6245.443
Napo	Ecuador	2730.418	Meta	Colombia	6180.759
Caqueta	Colombia	2660.213	Zacatecas	Mexico	5959.308
Bolivar	Ecuador	2571.542	Guerrero	Mexico	5956.673
Chimboraz	Ecuador	2559.072	Cundinama	Colombia	5918.706
Guaviare	Colombia	2527.035	Antioquia	Colombia	5899.217
Vaupes	Colombia	2442.433			
Sucre	Colombia	2442.325			
Junin	Peru	2440.957			
Narino	Colombia	2440.74			
Putumayo	Colombia	2433.034			
La libert	Peru	2336.601			
Sucumbios	Ecuador	2299.415			
Morona-sa	Ecuador	2265.742			
Ucayali	Peru	2201.317			
Lambayequ	Peru	2176.718			
Pando	Bolivia	2144.45			
Cajamarca	Peru	2143.893			
Loreto	Peru	2024.41			
Tumbes	Peru	1943.955			
Piura	Peru	1874.891			
Huancavel	Peru	1844.268			
Oruro	Bolivia	1820.605			
Santa cru	Bolivia	1777.119			
Orellana	Ecuador	1733.602			
Tarija	Bolivia	1715.386			
Choco	Colombia	1694.354			
Puno	Peru	1492.099			
Cochabamb	Bolivia	1422.624			
San marti	Peru	1411.274			
Cusco	Peru	1328.072			
Huanuco	Peru	1297.192			
Amazonas	Peru	1277.133			
Ayacucho	Peru	1228.001			
La paz	Bolivia	1148.692			
Chuquisac	Bolivia	1144.664			

D Correlation Matrix

Table 16 shows the correlation coefficients between the variables included in the extended model presented in section 3.4.5.

Table 16: Correlation between variables of the extended model

	urbanization	initial GDPpc	inv_GDP	tertiary_educ	IER
urbanization	1				
initial GDPpc	0.71 (0.000)	1			
inv_GDP	-0.39 (0.000)	-0.43 (0.000)	1		
tertiary_educ	0.63 (0.000)	0.56 (0.000)	-0.25 (0.0001)	1	
IER	-0.23 (0.0002)	0.001 (0.975)	0.03 (0.6395)	-0.2153 (0.0006)	1

In parenthesis, the p-value of the correlation coefficient.

References

- Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer, Dordrecht.
- Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis*, 27(2):93–115.
- Arbia, G. and Piras, G. (2005). Convergence in per-capita GDP across european regions using panel data models extended to spatial autocorrelation effects. Technical report, Istituti di Studi e Analisi Economica, Working paper 51.
- Atienza, M. and Aroca, P. (2012). Es la concentración un problema para el crecimiento de América Latina? (In english: Is the concentration an issue for Latin American growth?). Technical report, Universidad Católica del Norte Chile.
- Baldwin, R. E. and Forslid, R. (2000). The core-periphery model and endogenous growth: Stabilizing and destabilizing integration. *Economica*, 67(267):307–324.
- Baltagi, B. (1995). *Econometric Analysis of Panel Data*. John Wiley Sons, 1st edition.
- Blanco, L. (2011). The impact of spatial interdependence on FDI in latin america. Technical report, Pepperdine University, School of Public Policy Working Papers, Paper 27.
- Bosker, M. (2007). Growth, agglomeration and convergence: A space-time analysis for european regions. *Spatial Economic Analysis*, 2(1):91–110.
- Brühlhart, M. and Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics*, 65(1):48–63.
- Chauvin, J.-P., Glaeser, E., and Tobio, K. (2013). Urban economics in the US and India. Work in progress, Harvard University.
- Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review*, 46(2):213–227.
- Ciccone, A. and Hall, R. E. (1995). Productivity and the density of economic activity. Economics Working Paper 120, Department of Economics and Business, Universitat Pompeu Fabra.

- Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2008). Estimating agglomeration economies with history, geology and worker effects. Technical report, Série des Documents de Travail du CREST, Institut National de la Statistique et des Etudes Economiques.
- Combes, P.-P. and Lafourcade, M. (2012). Revue de la littérature académique quantifiant les effets d'agglomération sur la productivité et l'emploi. (In english: Literature review about agglomeration effects in productivity and employment.). Report, Societe de Grand Paris.
- Crozet, M. and Koenig, P. (2008). Arbitrage cohesion-croissance: Une analyse des régions européennes (1980-2000) (In english: The cohesion vs growth tradeoff: Evidence from EU regions). *Revue économique*, 59(2):331–340.
- Davis, J. C. and Henderson, J. V. (2003). Evidence on the political economy of the urbanization process. *Journal of Urban Economics*, 53(1):98–125.
- Duranton, G. (2007). From cities to productivity and growth in developing countries. Working Paper tecipa-306, University of Toronto, Department of Economics.
- Duranton, G. and Puga, D. (2003). Micro-foundations of urban agglomeration economies. Working Paper 9931, National Bureau of Economic Research.
- ECLAC (2015). Historial background about latin american integration. <http://www.cepal.org/publicaciones/xml/8/4228/cap1.htm>. Accessed: 2015-02-20.
- Estevadeordal, A. and Volpe-Martincus, C. (2009). Trade policy and specialization. Technical report, Inter-American Development Bank IDB, Working Paper series IDB-WP-134.
- Fatás, A. (2002). The effects of business cycles on growth. In Loayza, N. and Soto, R., editors, *Economic Growth: Sources, Trends and Cycles*.
- Fujita, M. and Thisse, J. (2003). Globalization and the evolution of the supply chain: who gains and who loses? Technical report, KIER Working Papers 571, Kyoto University, Institute of Economic Research.
- Gardiner, B., Martin, R., and Tyler, P. (2010). Does spatial agglomeration increase national growth? some evidence from europe. *Journal of Economic Geography*, pages 1–28.

- Glaeser, E. L., Kallal, H., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *The Journal of Political Economy*, 100(6):1126–1152.
- Grossman, G. M. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press. Cambridge MA.
- Guevara, C. and Autant-Bernard, C. (2015). Technological interdependence between south american countries: A spatial panel data growth model. *Revue d’Economie Régionale et Urbaine*, 2(1):181–209.
- Guevara, C. and Jarrin, S. (2011). *Modelo Gravitacional de Comercio para America del Sur. (In english: Gravity Model for South America)*. Editorial Academica Espanola.
- Henderson, J. V. (2000). The effects of urban concentration on economic growth. Technical report, National Bureau of Economic Research Working Paper 7503.
- Henderson, V. (2003). Urbanization and economic development. *Annals of Economics and Finance*, 4(2):275–341.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42.
- Marshall, A. (1890). *Principles of Economics*. London. MacMillan.
- Martin, P. (1999). Public policies, regional inequalities and growth. *Journal of Public Economics*, 73(1):85–105.
- Martin, P. and Ottaviano, G. (1999). Growing locations: Industry location in a model of endogenous growth. *European Economic Review*, 43:281–302.
- Martin, P. and Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review*, 42(4):947–968.
- Massiris-Cabeza, A., Espinoza-Rico, M. A., Ramirez-Castaneda, T., Rincón-Avellaneda, P., and Sanabria-Artunduaga, T. (2012). *Procesos de ordenamiento en América Latina y Colombia (In english: Planning processes in Latin America and Colombia)*. Universidad Nacional de Colombia.
- McCallum, J. (1995). National borders matter: Canada-U.S. regional trade patterns. *American Economic Association*, 85(3):615–623.

- Moran, P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society, series B*, 37(1):243–251.
- Pesaran, H. M. (2004). General diagnostic tests for cross section dependence in panels. Technical report, University of Cambridge, Working Papers 0435.
- Ramírez, M. T. and Loboguerrero, A. M. (2002). Spatial dependence and economic growth: Evidence from a panel of countries. Technical report, Banco de la Republica Colombia.
- Romer, P. (1987). Growth based on increasing returns due to specialization. *American Economic Review*, 77(2):56–62.
- Serra, M.-I., Pazmino, M.-F., Lindow, G., Sutton, B., and Ramirez, G. (2006). Regional convergence in Latin America. Technical report, International Monetary Fund.
- Venables, A. J. (2005). Spatial disparities in developing countries: cities, regions and international trade. *Journal of Economy Geography*, 5(1):3–21.
- Williamson, J. G. (1965). Regional inequality and the process of national development: A description of the patterns. *Economic Development & Cultural Change*, page 3.