Agglomeration Externalities and urbanization in Ecuador

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

Stéphane Riou
University of Lyon, Lyon
CNRS, GATE Lyon Saint-Etienne, Ecully

Corinne Autant-Bernard
University of Lyon, Lyon
CNRS, GATE Lyon Saint-Etienne, Ecully

This version: September, 2015

Abstract

The spatial agglomeration of economic activities play a crucial role on productivity but the composition of such an agglomeration is what really matters. There exists an ongoing debate between the predominance of the effects of agglomeration from specialization and diversity. This paper contributes to the debate by looking at the economy of Ecuador, a country of Latin American region for which literature gives little attention. We use the sectoral and cantonal information of the Economic Census of Ecuador of 2010. Our productivity equation is estimated using GMM model and we instrument the endogenous variables with a menace index, long lagged population density and spatial lags. The results suggest that all the agglomeration externalities play an important role in the local productivity of sectors. However, the effects are distinct. For instance, at the national scale urbanization economies measured with the diversity index matter more than localization economies measured with the specialization index. We also provide a sectoral analysis to determine the effects in the industry sector and the services sector. Indeed, the elasticities of productivity with respect to the types of agglomeration externalities differ. Interestingly, in order to enjoy the existence of agglomeration externalities, a rate of urbanization is required.

1 Introduction

This paper provides the first estimation of the impact of urbanization (Marshall-Arrow-Romer externalities) and localization economies (Jacob's externalities) on the local performance of Ecuadorian firms.
Agglomeration is often viewed as playing an important role for local economic growth. However, it is now well documented that the effects of a given level of agglomeration may vary across regions depending on how the local economy is structured. The real nature of agglomeration externalities is a frequently debated question.
On the one hand, following Marshall (1890), Arrow (1962) and Romer (1986), agglomeration combined with a strong industrial specialization (localization economies) could be the best mix in terms of positive externalities. The spatial concentration of a given industry in a specific place would facilitate face-to-face contacts and the opportunities to cooperate. As a consequence, ideas and methods of production of a firm would spill over other firms located in the same region. Moreover, the local network of researchers of a given industry would be more dynamic than if it was more global and based on distant relationships. Marshall (1890) also shows that the industrial specialization of regions generates other benefits related to the local labor market and supply linkages (see Duranton and Puga (2003)). Indeed, agglomerated firms of an industry have more possibilities to find workers and inversely, unemployed workers are very likely to get a job. In addition, industrial concentration gives incentives to suppliers to concentrate as well and transportation costs of inputs are then reduced.
On the other hand, Jacobs (1969) consider industrial diversity as an engine for regional growth through urbanization economies. She emphasizes that innovative activity is the result of interactions of close firms from different industries. Significant innovations are often the result of a process of imitation and adaptation of ideas developed in different industries. Moreover, the additional benefit of a diversified region is the availability of a large panel of services which facilitates production.
More than the levels of specialization and diversity, one may also consider that the intensity of competition is the strongest determinants for regional growth (Porter's externalities). Local competition leads firms to innovate or to adopt the most efficient production process in order to stay in the market.

Numerous empirical studies provide estimations of these three types of externalities. Beaudry and Schiffauerova (2009) and De Groot et al. (2007) survey this literature and conclude on a very contrasted picture. Variables used to measure performance, specialization and diversity and the various econometric methods would largely explain the lack of clear conclusions which will be valid in every circumstance. Their survey also points the fact that the attempts to measure the impact of specialization versus diversity are mainly focused on developed countries. However, the nature of the agglomeration externalities in developing countries seems to be a relevant issue. Most of the literature focuses on specialization and diversification as drivers of economic growth and development. The empirical analysis of this important issue both in terms of policy and theoretical implications has been renewed by Imbs and Wacziarg (2003). The authors detect a robust U-curved relationship between specialization and GDP per capita. Put differently, diversification would be a major driving force of economic development for countries at low stages of development. In addition to such aggregate analysis, it is also important to detect the impact of the industrial structure on the firm performances at a local level. Indeed, one may wonder whether industrial diversity which seems to have positive effects on the macroeconomic performance generates similar effects at a more spatially disaggregated level. Income inequality between regions is a main concern in developing countries, especially in Latin American ones. Their industrial structure is frequently presented as part of the explanation of the economic gap between regions. The poor ones would suffer from an excess of specialization, the firms located therein having too few opportunities to benefit from inter-industry relationships. The immediate policy recommendation would be to favor the emergence of an industrial diversification dynamic in these regions. Our view is that this recommendation must be systematically based on robust empirical evidences that the industrial structure of regions matters for local growth. Such empirical evidences are scarce for developing countries. More generally, new studies on agglomeration economies in developing countries are needed to gain knowledge on additional countries[1]. (Combes and Gobillon, 2015). Da Silva Catela et al. (2010) show that both diversity and specialization have positive effects on the real average wage per worker at the municipality level in Brazil. Nevertheless, they observe that diversity is more significant for low-income regions while specialization is more significant for high-income regions. For Chile, Saito and Gopinath (2009) find that more productive firms in food industry locate in regions where other food industry plants agglomerate and in regions that combine an industrial diversification with a large market size. In other words, externalities arise from both specialization and diversity. Last, Pereira and Soloaga (2012) conclude on positive urbanization economies (Jacobs' externalities) in the long-run in

[1] Combes et al. (2013) for China, Chauvin et al. (2013) for India, among others, provide estimations of the impact of density on individual wages. They are not focusing on the role played by the sectoral diversity and specialization.

Mexican regions.

When measuring agglomeration economies, one concern is the endogeneity issue. This problem is not systematically addressed in the studies mentioned above Combes and Gobillon (2015). On the one hand, endogeneity arises when missing variables affect both agglomeration and productivity. Indeed, a high regional productivity may be promoted by good local conditions such as transport infrastructure services and other facilities which in turn increases the level of agglomeration in the region. On the other hand, endogeneity may emerge from reverse causality when local agglomeration is determined by local productivity and in turn determines it. In our study, one major source of endogeneity is the existence of a circular causality between local productivity and the type of externalities we want to estimate. For example, a high sectoral productivity in a region is likely to attract more firms from the same sector to the region. This favors more externalities from specialization which will translate into higher productivity. As a consequence of the bidirectional nature of the relationship, the variable capturing the specialization level will be endogenous. Similarly, the risk of endogeneity of the diversity or the competition variable must be controlled. In this purpose, the instrumental variables method is implemented. We provide estimations using GMM model and we instrument the endogenous variable with a menace index, long lagged population density and spatial lags.
Last, we investigate the specific role of urbanization. Precisely, we make the assumption that the existence of Jacobs and/or MAR externalities depend on the levels of urbanization of Ecuadorian cantons. The intuition is that a critical level of urbanization is required to produce positive externalities as it guarantees the existence of a minimum level of transport and telecommunication infrastructures, of banking and financial services or other specific services. It is also a manner to capture the great heterogeneity of Ecuadorian cantons in terms of urbanization.
Our empirical work is mainly based on the Economic Census of Ecuador 2010 which accounts for information declared in 2009. By aggregating the firm data, we build a two-digits industry database at the cantonal level. Precisely, our estimations are based on 78 industries and 221 cantons. It is worth noting that some cantons do not have all industries operating. The number of industries varies from one canton to another. It could be due to selection bias. The estimations deal with this issue.
From a first polled estimation on 7803 canton and industry pairs, our results suggest the existence of strong positive externalities from diversity and to a lesser extent from specialization impacting on the local productivity of industries. We also find economies of density, measured by the density of firms and employees at

the canton and industry levels that positively influence the local productivity of industries. Competition of firms has a negative effect. Moreover, we conduct regressions by distinguishing the manufacturing from the service industry. The magnitude of the externalities from diversity is positive and significant in manufacturing but much higher in the service industry. Interestingly, our regressions exhibit the non-monotonous effect of urbanization on the various externalities that impact on the local productivity. Precisely, the positive externalities arising from diversity are growing with the level of urbanization of cantons. Last, economies of density occur until a urbanization rate of 61%. Above this threshold, the economies of density cease to be positive suggesting that they are overcompensated by congestion effects. However, the effects of diversity and specialization do remain highly positive, we therefore do not consider that the threshold of 61% represents excessive agglomeration. Finally, the negative effect of competition seems to vanish in highly urbanized cantons.

2 Geographical statistics of agglomeration externalities in Ecuador

2.1 Measures of agglomeration externalities

The analysis of agglomeration economies requires good measures. Following Combes (2000), our indicators are standardized using the national share of sectors. In addition, logarithms of the variables are used in our linear regression so that we can directly analyze the elasticities.

As it is standard, our dependent variable which must approximate the average performance of a given sector in a given canton is measured by the labor productivity. Precisely, our dependent variable is: $produc_{s,c} = \frac{rev_{s,c}/empl_{s,c}}{rev_s/empl_s}$ where $rev_{s,c}$ and $empl_{s,c}$ are respectively the levels of revenue and employment in sector s and canton c.

Many authors measure specialization as the share of employment of a given industry in a region. In order to avoid the bias concerning the size of sectors, we use a relative measure of specialization by normalizing it with the share of the national employment allocated to this sector as in Combes (2000). In this manner we can compare between industries. The specialization indicator is related to the so-called localization economies that come from the intra-sectoral externalities. The formula is the following:

$$spe_{s,c} = \frac{empl_{s,c}/empl_c}{empl_s/empl} \tag{1}$$

where $empl_{s,c}$ is the employment in sector s in canton c and $empl$ is the national employment.

With respect to diversity, it is common to use the inverse of Herfindahl index. Likewise, it is normalized by the sectors share at national level. The indicator measures the level of sectoral diversity of sector s in canton c. It is related to the so-called urbanization economies that come from inter-sectoral externalities. The formula is the following:

$$div_{s,c} = \frac{1/\sum_{s'=1,s'\neq s}^{S} \left(empl_{s',c}/(empl_c - empl_{s,c})\right)^2}{1/\sum_{s'=1,s'\neq s}^{S} \left(empl_{s'}/(empl - empl_s)\right)^2} \tag{2}$$

where S is the number of industries. When industries have the same size in a region, the value of the numerator reach its maximum. It is worth noting that specialization and diversity calculated as before are not necessarily opposites to each other. A canton may host the main part of a sector and a broad number of other industries as well. Such a canton is then both specialized and diversified.

Regarding the Porter's externalities, competition is measured as the inverse of a Herfindahl index of a canton as in Combes (2000). The unity of calculation is now the firm and the indicator captures the level of competition within sector S in canton c. The formula is the following:

$$comp_{s,c} = \frac{1/\sum_{i}^{N} \left(empl_{i,s,c}/empl_{s,c}\right)^2}{1/\sum_{i}^{N} \left(empl_{i,s}/empl_s\right)^2} \tag{3}$$

where $empl_{i,s,c}$ is the employment of firm i belonging to sector s and located in canton c, and $empl_{i,s}$ is the employment of firm i in sector s at national level.

Another measure that allows controlling for the differences between cantons is the density of firms. This variable takes into account the distribution of firms across space. Moreover there is an abundant and recent literature providing empirical evidences that economies of density may be an important determinant of local productivity. In most of the existent works employment density is used. Such a variable captures the concentration of labor force but does not necessarily capture the concentration of economic activities. For instance, consider the case of few large firms with a large number of employees. Thus, the employment's density index indicates that economic activities are highly concentrated while in reality there are only few firms that operate therein. Conversely, firms' density will capture the actual number of operating activities in the city. The standardized indicator is as follows:

$$den_{s,c} = \frac{f_{s,c}/area_c}{f_s/area} \tag{4}$$

where $f_{s,c}$ is the number of firms of sector s located in canton c and $area_c$ is the area of canton c. In the denominator, f_s is the number of firms of sector s and $area$ is the total surface of the country.

For reasons of comparability with existent literature, the employment's density is also used. It is computed as follows:

$$empl\_den_{s,c} = \frac{empl_{s,c}/area_c}{empl_s/area} \tag{5}$$

These measures of agglomeration have two dimensions: sector and canton. They are used in the estimation of the model. The next analysis show spatial statistics considering the cantonal dimension only.

2.2 Spatial statistics of Ecuador

Ecuador is an economy highly dependent on revenues obtained from mining and quarrying activities and agriculture, forestry and fishing; these sectors alone represented 10% and 9% of the national production in 2010, respectively. But the dynamism of the economy lies on the large share of service and manufacturing industries which in 2010 represented 33.7% and 33.9% of the national production, respectively[2]. Regarding the manufacturing industry, 11 provinces are above the national average. The share of the service industry is higher than the national average in 13 provinces. They are then key industries not only at the national scale but also at the sub-national scale.

It is widely argued that the presence of manufacturing and service industries promotes urbanization (Davis and Henderson, 2003; Fajgelbaum and Redding, 2014). In

[2] The calculations are based on Provincial National Accounts of 2009 published by the Central Bank of Ecuador. We provide statistics of provinces because information at cantonal level is not available

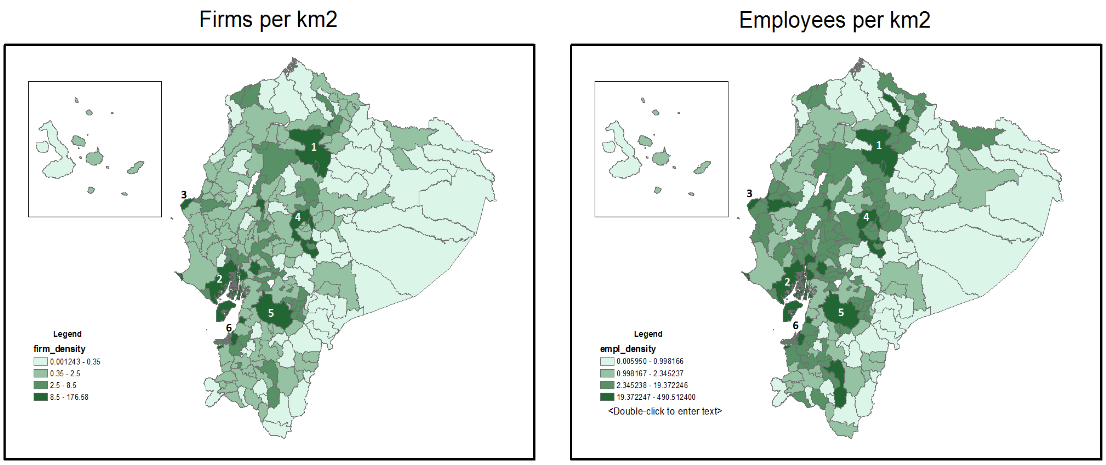
Ecuador, the share of population living in urban areas was 67.8% in 2010. However, if we look at the sub-national geographical level, the scenario is different. Regarding cantons, the proportion of population living in urban areas is 38.9% in average. It means that there are many more cantons with rural population than cantons with urban population. The cantons with more than 50% of urbanization are 63 out of a total of 221 cantons.

In this section, we use global indicators of specialization, diversity, density and productivity of cantons in order to visualize those measures across space. Thus, we use indicators at the cantonal scale instead of indicators at the cantonal and industrial scale presented in Section 2.1.

Figure 1 presents the spatial concentration of economic activities in cantons using firms density (a) and employment's density (b). The darker areas correspond to denser cantons. We observe that spatial patterns are quite similar across density measures[3]. Denser cantons are mainly located in the Sierra region of the country whereas the less dense ones are located in the eastern region with large territories occupied by the rain forest. On national average, the mean of firms (employment) density of cantons is 3.9 firms (12 workers) per squared kilometer and the standard deviation is 13.35 (40) which indicates high heterogeneity between cantons. Indeed, we can distinguish some centers of agglomeration of firms and employment in Figure 1. In the north, the main center is the capital Quito (1 in Figure 1) which records a density of 23 firms (121 workers). In the coastal region, three centers can be distinguished: Guayaquil (2 in Figure 1), Manta (3 in Figure 1) and Machala (6 in Figure 1) which have densities of 20 (100), 28 (121) and 32 (117) firms (workers) per squared kilometer, respectively. Another center of economic activity is Ambato (4 in Figure 1) recording 18 firms (59 workers) per squared kilometer. In the south, the main center is Cuenca (5 in Figure 1) which has a density of 9 firms (36 workers) per squared kilometer. In addition, maps in Figure 1 show higher firm (employment) densities in cantons close to economic centers.

[3]The differences between firms density and employment's density are related to the ranking of some cantons. For instance, Huaquillas is denser than Machala in terms of number of firms. However, it is less dense than Machala in terms of employment. This could indicate that Huaquillas has a large number of small firms that do not generate as much employment as firms in Machala. Thus, when one canton presents high firms density but low employment's density, we can imagine that those firms are small firms that do not have the capacity to employ a large number of workers.

Figure 1: Density of economic activities in cantons



(a) Firms density by canton (b) Employment's density by canton

Source: Economic Census 2010, Ecuador

The specialization index is commonly calculated as the share of an industry s in a canton c. Since there exist some industries that absorb large percentage of local economy, we standardize the indexes with the national average of the given industry s as in Da Silva Catela et al. (2010). The relative specialization index is:

$$IER_c = max_s \frac{x_{sc}/x_c}{x_s/x}$$

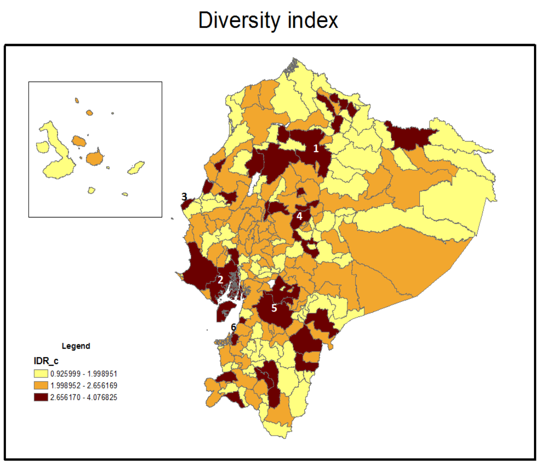
where x_{sc} is the employment in industry s of canton c and x is the national employment.

The diversity index is commonly measured with the inverse of Herfindahl-Hirschman index (Da Silva Catela et al., 2010). This measure is also corrected with the share of a given industry s in the total economy. This formula is not necessarily opposed to that of specialization.

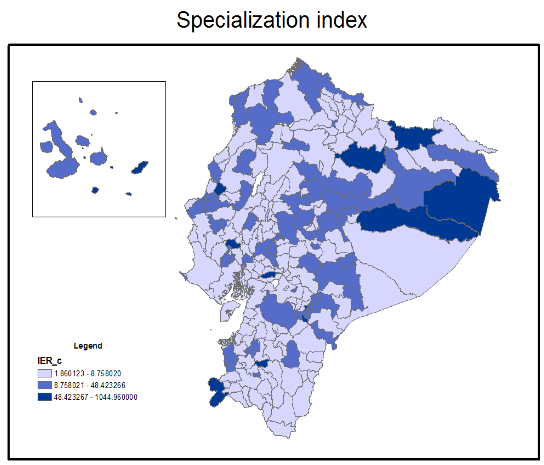
$$IDR_c = \frac{1}{\sum_s |(x_{sc}/x_c) - (x_s/x)|}$$

Figure 2 shows spatial configurations in terms of industrial diversity (a), industrial specialization (b) and productivity (c) of cantons.

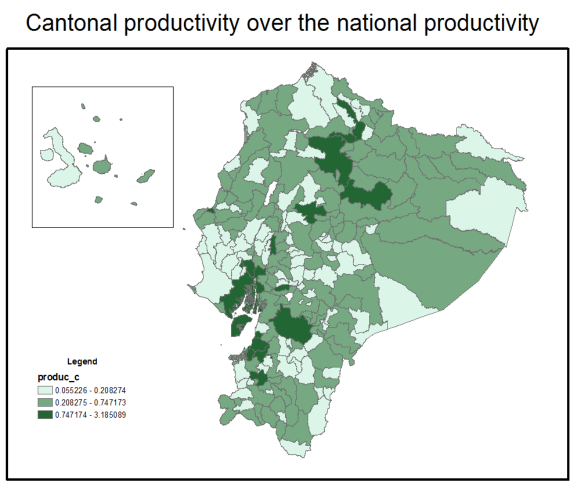
Figure 2: Externalities and productivity of cantons



(a) Diversity of cantons



(b) Specialization of cantons



(c) Productivity of cantons

Source: Economic Census 2010, Ecuador

First, we observe that the identified centers of density have the highest level of diversity. In general, many cantons have diverse economies with interacting industries. Here again, the closer the cantons to centers of agglomeration, the higher their level of diversity. Regarding specialization, only some cantons are identified as highly specialized. They are scattered in the country without a clear spatial pattern.

In terms of productivity, cantons with higher rates than the national productivity are the capital Quito (1 in Figure 1), Guayaquil (2 in Figure 1) and Cuenca (5 in Figure 1) which are economic centers of the country. Putting all this information together, some relationships can be deduced. In the first place, we observe that cantons with high level of diversity tend to have higher productivity. The correlation coefficient is 0.24, which is significant. On the other hand, it seems that there is no link between specialization and productivity provided that the correlation coefficient (-0.029) is not significant. These results show that agglomeration effects at cantonal level coming from diversity of economic activities seem to be positive for productivity in Ecuadorian cantons.

Now, we analyze more closely the characteristics of specialization and diversity of cantons using information at industrial level. Table 1 shows that in aggregate terms, cantons are mainly specialized in Manufacturing, in Arts, entertainment and recreation and in Water supply; sewerage, waste management and remediation activities. The most specialized canton (see Table 2) with a large share of Extraterritorial organizations is Aguarico (Amazon region). This could be explained by high foreign direct investment in complementary activities to oil extraction in this canton. Interestingly, some cantons specialize in activities related to their natural advantage. This means that manufacturing and services have emerged as complementary activities to primary activities. For instance, in Lago Agrio where oil extraction is the main primary activity, remediation services have emerged in such a canton. In canton Coronel where the primary activity is the production of sugarcane, manufacture of paper using the sugarcane waste has emerged.

Table 1: Sectors of specialization of cantons

Aggregate industry	Number of cantons
Manufacturing	43
Arts, entertainment and recreation	30
Water supply, sewerage, waste management and remediation activities	27
Human health and social work activities	22
Transportation and storage	16
Professional, scientific and technical activities	14
Public administration and defence, compulsory social security	12
Information and communication	12
Other service activities	11
Administrative and support service activities	10
Electricity, gas, steam and air conditioning supply	9
Construction	4
Activities of extraterritorial organizations and bodies	4
Accommodation and food service activities	4
Real estate activities	2
Education	1
Total	221

Source: Economic Census 2010, Ecuador
Elaboration: by author

Table 2: Ecuador: Most specialized cantons, 2009

Position	Province	Canton	CIIU2	Sector	IER c
1	Orellana	Aguarico	U99	Extraterritorial organizations	1044.96
2	Manabi	Tosagua	E37	Sewerage	369.5504
3	Sucumbios	Lago Agrio	E39	Remediation activities	169.43
4	Guayas	Santa Lucia	Q87	Residential care activities	123
5	Loja	Zapotillo	E37	Sewerage	120.65
6	Napo	El Chaco	M70	Head offices and management activities	120.30
7	El Oro	Portovelo	R91	Libraries and other cultural activities	94.27
8	Guayas	Coronel	C17	Manufacture of paper and paper products	81.64
9	Pastaza	Arajuno	N79	Travel agency, related activities	69.11
10	Azuay	Guachapala	F42	Civil engineering	61.93

Source: Economic Census 2010, Ecuador
Elaboration: by author

Table 3 lists the most and least diversified cantons in Ecuador in 2009. The most diversified cantons are generally the capitals of provinces. Overall, highly diversified cantons are larger in size. This fact has been underscored in the literature (Duranton and Puga, 2001). As shown in Figure 3 larger cantons present a wider variety of economic activities. The correlation coefficient between size measured by employment and diversity index is 0.5087. In big cantons, different industries benefit from public amenities, transportation services and infrastructure.

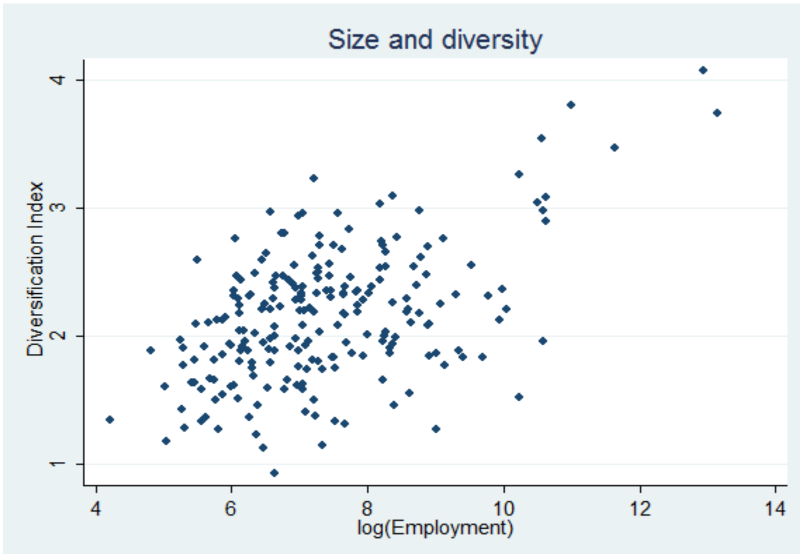
Table 3: Ecuador: Most and least diversified cantons, 2009

Position	Province	Canton	IDR c
1	Guayas	Guayaquil	4.08
2	Tungurahua	Ambato	3.81
3	Pichincha	Quito	3.74
4	El Oro	Machala	3.54
5	Azuay	Cuenca	3.47
6	Imbabura	Ibarra	3.27
7	Cotopaxi	Saquisili	3.24
8	Loja	Catamayo	3.09
9	Santo Domingo	Santo Domingo	3.09
10	Manabi	Manta	3.04
...	...	...	...
212	Morona santiago	Huamboya	1.33
213	Manabi	Pajan	1.32
214	Pastaza	Arajuno	1.29
215	Napo	Quijos	1.28
216	Manabi	Montecristi	1.27
217	Chimborazo	Penipe	1.23
218	Sucumbios	Sucumbios	1.18
219	Zamora Chinchipe	Chinchipe	1.15
220	Sucumbios	Putumayo	1.13
221	Morona santiago	Tiwintza	0.93

Source: Economic Census 2010, Ecuador

Elaboration: by author

Figure 3: Canton size vs diversification



Source: Economic Census 2010, Ecuador

3 Econometric methodology

3.1 Econometrics issues related to agglomeration economies

Throughout the literature on agglomeration economies, it has been stated that the effect of agglomeration is endogenous (Combes and Lafourcade, 2012). The endogeneity problem implies that the assumption that explanatory variables have to be uncorrelated with the error term is violated. In productivity regressions involving agglomeration, this correlation can emerge for two reasons. On the one hand, endogeneity arises when omitted variables affect both agglomeration and productivity. For instance, high levels of productivity in a given industry might be induced by good local conditions, such as transport infrastructure, services and other facilities, which, in turn, increase the level of agglomeration in the region. On the other hand, endogeneity can emerge from a bidirectional causality when local agglomeration is determined by local productivity and, in turn, determines it. The circular causality is theoretically stated by Martin and Ottaviano (2001).

In the case of specialization, a high productivity of a given industry in one canton is likely to attract more firms of the same industry to that canton. Thus, the level of industrial specialization of the canton increases. In the other direction, MAR's externalities coming from knowledge spillovers within the same industry take place which, in turn, increase the workers' productivity in that particular industry of a given canton. As a consequence of this bidirectional relationship between local industrial specialization and local industrial productivity, the variable measuring localization economies is endogenous by reverse causality.

Regarding diversity externalities, the impact of local productivity in a given industry on diversity of one canton is a priori not obvious. It depends on whether gains from productivity of a given industry are reinvested to strengthen and amplify inter-industrial linkages. If so, local productivity in a given industry would induce agglomeration of industries from different fields. In the other direction, according to Jacobs, a more diverse industrial environment is beneficial for exchange of ideas across industries and emergence of innovations, which, in turn, foster growth. In consequence, the endogeneity of local industrial diversity by reverse causality has to be tested.

With respect to local industrial competition, externalities from competition described by Porter may have an important influence on the production of innovation activities in the same industry. But the other direction of the relationship is unclear. The local industrial productivity may have an effect on competition by attracting more firms to the canton through a good indicator of profitability of firms already located in such a canton. However, firms would rather locate in places where competition is not strong. Hence, the endogeneity of this variable has to be tested, as well.

If the endogeneity of these variables is proved, the OLS estimator provides biased and inconsistent parameter estimates. Since our framework consists of industry and canton pairs, we have to analyze the relationships between industrial productivity of cantons and agglomeration externalities in cantons and industries: specialization, diversity and competition.

3.2 Specification of the model

Considering the aforementioned issues of endogeneity, the feasible efficient two-step GMM estimator[4] is implemented in our regressions. Furthermore, this estimator is efficient (more than 2SLS estimator) in the presence of heteroskedasticity. Thus, it provides consistent estimates by dealing with both potential endogeneity and heteroskedasticity.

The specification of the model estimating the effect of each type of externalities on productivity level of a specific industry s in a given canton c is the following.

$$\begin{aligned} produc_{s,c} =& \alpha + \beta_1 spe_{s,c} + \beta_2 div_{s,c} + \beta_3 comp_{s,c} + \beta_4 den_{s,c} + \gamma_2 p_2 + ... + \gamma_l p_l + \lambda_2 s_2 + ... + \lambda_l s_S + \\ & \text{selection}_{s,c} + u_{s,c} \end{aligned} \tag{6}$$

where $produc_{s,c}$ is the labor productivity in industry s in canton c. To control for unobserved heterogeneity, we introduce l-1 province fixed effects that are given by $p_2 ... p_l$[5] and sector fixed effects given by $s_2 ... s_S$. The term $selection_{s,c}$ is a control variable which allows us to address the issue regarding the presence of many zeros in our database (all cantons do not account for the same number of economic activities). This is the inverse of Mill's ratio computed using the estimates of a selection

[4]We use the command of Stata *ivreg2* developed by Baum et al. (2010)

[5]We do not introduce fixed effects for cantons because it reduces the degrees of freedom. Hence, we prefer to control for characteristics of provinces.

equation proposed by Heckman (1979).[6] In order to correct the endogeneity issue derived from circular causality of our endogenous variables, a set of instruments Z_i assumed to be uncorrelated with the error ($cov(Z_i, u_{s,c}) = 0$) is used.

The Z excluded instruments for our endogenous variables ($spe_{s,c}$, $div_{s,c}$ and $den_{s,c}$[7]) are the population density of 1990 ($Popd_{1990}$), the spatial lag of specialization ($\sum_{j=1}^{N} w_{cj} spe_{s,j}$), the spatial lag of diversity ($\sum_{j=1}^{N} w_{cj} div_{s,j}$), the spatial lag of competition ($\sum_{j=1}^{N} w_{cj} comp_{s,j}$), the spatial lag of firms density (the spatial lag of employment's density) ($\sum_{j=1}^{N} w_{cj} den_{s,j}$), squared density ($den^2$), and the menace index (menace_index). They are described in the sequel.

First, we use population density of 1990 (*pop 90*) as an instrument. The long-lag variable[8] is useful to remove the simultaneity that arises from contemporaneous effects. It is unlikely to be correlated with the error term at time t since it was generated at an earlier point in time (Combes et al., 2010).

The endogenous variables are also instrumented with their spatial lags (*div_spatial*, *comp_spatial*, *den_spatial*, *div_spatial*). In order to compute the spatial lags, we construct a distance weight matrix using the coordinates of capitals of cantons. The spatial lags are likely to be relevant instruments since the neighboring cantons could affect one canton's level of specialization, diversity, density or competition. It is worth noting that the introduction of the spatial dimension of these instruments somehow controls for possible spatial correlation in the model.

We also think that the squares of our endogenous variables could be good instruments. The reason is that they explain the current patterns of endogenous variables. However only the density squared (den^2) passes the validity test. The specialization squared, diversity squared and competition squared are dropped because of collinearity.

Another interesting instrument is the natural menace index of cantons (*menace*) that takes into account natural causes of disasters as gliding, flooding, drought, earthquakes, tsunamis and volcanic danger of each canton. This variable has been compiled from a joint work between the System of Economic Indicators (SIISE

[6]The selection equation is estimated by means of a Probit model. It explains the probability to observe the dependent variable using dummies of industries and the level of urbanization of cantons. The regression is presented in Appendix A

[7]In separate models, density is measured with firms density or employment's density.

[8]To recall, the database used for the model corresponds to the census of 2009.

acronym in spanish), Oxfam International and the International Cooperation (COOPI). We presume that such an index is a reasonable instrument because of a twofold character: i. it is positively correlated with all of our endogenous variables and ii. since such an index indicates a probability of occurrence of an incident, the productivity will not be affected if the natural threat does not occur.

In order to have consistent estimates of the productivity equation (6), our instrumental variables must meet two specific properties. First, they have to be highly correlated with instrumented variables (relevance condition). And second, a valid instrument has to be exogenous (orthogonality condition). The latter cannot be directly tested. But the former is readily testable. Through the next analysis of correlation coefficients between the endogenous variables and their instruments shown in Table 4, we are able to test the condition of relevance of our instruments.

Table 4: Correlation between instruments and endogenous variables

	Specialization	Diversity	Density firms	Density workers
menace index	-0.0813	0.1431	0.1914	0.2084
	(0.000)*	(0.000)*	(0.000)*	(0.000)*
squared density	0.2685	0.1264	0.5513	0.6503
	(0.000)*	(0.000)*	(0.000)*	(0.000)*
Popd 90	-0.1097	0.2875	0.8059	0.7556
	(0.000)*	(0.000)*	(0.000)*	(0.000)*
Spe spatial lag	-0.0045	-0.0005	-0.0054	-0.0004
	(0.6895)	(0.9667)	(0.6319)	(0.9736)
Div spatial lag	-0.0101	-0.0275	-0.0220	-0.0302
	(0.3673)	(0.0141)*	(0.0494)*	(0.0071)*
Comp spatial lag	-0.0144	-0.0268	-0.0234	-0.0328
	(0.1990)	(0.0166)*	(0.0366)*	(0.0034)*
Den firms spat lag	0.2685	0.1264	0.5513	0.6503
	(0.000)*	(0.0001)*	(0.000)*	(0.000)*
Den workers spat lag	-0.0035	-0.0430	-0.0627	-0.0663
	(0.7546)	(0.0001)*	(0.000)*	(0.000)*

In parenthesis, the p-value of the correlation coefficient, * $p < 0.05$

Menace index, population density of 1990 and squared density are highly correlated with all the endogenous variables, proving to be relevant instruments. Spatial lags of diversity, competition and density are correlated with some of the endogenous variables. By contrast, the spatial lag of specialization is not correlated with any of the endogenous variables. It could be due to the fact that the level of specialization does not have a defined spatial pattern as shown in Section 2.2. Thus, the specialization of neighboring cantons is not correlated with any of the variables that probably follow a specific pattern over space.

Now, we turn to the expectations of each type of externality. As stated above, according to MAR's model, specialization ($spe_{s,c}$) is expected to have a positive impact on productivity. The concentration of similar industries in a given region implies knowledge spillovers within the industry which, in turn, enhances innovation activity. Conversely, according to Jacob's theory, diversity ($div_{s,c}$) is expected to have a positive sign in the estimation. She argues that the interaction between industries is the main source of spillovers since new ideas from different industries can be recombined and imitated. It induces innovation activity which, in turn, increases productivity. Another source of externalities is competition ($comp_{s,c}$). It is expected to be positive following the Porter's theory. More competition between firms promotes the incentive to innovate. Otherwise, bankrupt will be the destiny of firms. In that sense, the effect of competition depends on the efficiency of firms to overcome strong competition environments. Finally, density ($den_{s,c}$: employment's density or firms density) is considered as the quintessential variable of agglomeration in numerous studies. Density of economic activities takes into account the geography, being the variable that measures agglomeration across space. It is expected to be positive because geographical proximity between firms and employees facilitates the micro-mechanisms (matching, sharing and learning (Duranton and Puga, 2003)) leading to agglomeration economies.

In addition, the aforementioned types of externalities might be related to each other. Then, we explore the statistical relation between them by using the Pearson's correlation coefficient.

In Table 5, we observe that the correlation coefficient between specialization and diversity is low and negative. It means that there is a small probability that Ecuadorian cantons are specialized and diversified at the same time. Furthermore, the correlation coefficient between diversity and competition is higher than that of specialization and competition. The former would be in line with the Jacobs's model meanwhile the latter would be in line with the MAR's model. Regarding firms density and employment's density, the correlation coefficients with diversity

Table 5: Correlation between variables of agglomeration externalities

	Specialization	Diversity	Competition	Firms density	Empl. density
Specialization	1				
Diversity	-0.0223	1			
	(0.0467)				
Competition	0.0286	0.4386	1		
	(0.0107)	(0.000)			
Firms density	0.2080	0.3538	0.6214	1	
	(0.0000)	(0.000)	(0.000)		
Employment's density	0.4706	0.3722	0.4494	0.8839	1
	(0.0000)	(0.000)	(0.000)	(0.0000)	

In parenthesis, the p-value of the correlation coefficient

and specialization are high. This shows that geographical proximity is important for diversified and specialized cantons. Finally, the high coefficient correlation between employment's density and firms density indicates that both measures are almost equal. Then, the results should not differ much across those measures of density.

In Section 4, we proceed to determine the effects of these four variables of agglomeration on the productivity of cantons in Ecuador.

4 Results

Previous to the discussion on the estimation's results, we present the test of endogeneity of our variables of agglomeration externalities. We have stated the reasons of reverse causality that make our variables potentially endogenous. Now, we test whether they have to be treated as endogenous. The endogeneity test is defined as the difference of the Sargan-Hansen statistics between the equation treating the suspect regressor as endogenous and the equation treating the suspect regressor as exogenous. According to the results in Table 6, we reject the null hypothesis that specialization, diversity, firms density and employment's density are exogenous. Then, they are treated as endogenous variables. On the contrary, competition is actually exogenous and then treated as such.

Table 6: Endogeneity test for agglomeration variables

	Specialization	Diversity	Competition	Firms density	Empl. density
Chi_sq	7.46	15.87	2.466	4.383	12.232
p-value	0.0063	0.0001	0.1164	0.0363	0.0005

Robustness Checks

The robustness tests are shown at the bottom of Table 7. First, we run a Pagan-Hall general test statistic to check for homoskedasticity and we find enough evidence to reject the null hypothesis. Besides, the residuals plots show heteoroskedasticity (see appendix B). Then, we present the estimations with standard errors robust to the presence of arbitrary heteroskedasticity.

The Kleibergen-Paap rk Wald statistic[9] testing the underidentification shows that both equations are identified. We also test whether our equation is over-identified. If we reject this hypothesis, then we cast doubt on the suitability of the instruments. The Hansen-Sargan tests for overidentification evaluate the entire set of overidentifying restrictions. For our robust estimators, the Hansen's J statistic is presented. According to such a test, our instruments are exogenous and we can be confident that our specifications are appropriate. In order to determine whether the instruments are weak, the Kleibergen-Paap rk Wald F statistic is also shown. The F statistic is compared to the table of Stock-Yogo weak ID test critical values for K1=endogenous variables and L1=excluded instruments. (See Stock and Yogo (2005) for tables) The statistic shows that the instruments are not weak at 5% of maximal bias of the IV estimator relative to OLS. It is worth noting that since we have more instrumental variables than endogenous variables ($Z > X$), the GMM estimator is appropriate and more efficient than 2SLS estimator. Moreover, according to Moran's plots of residuals presented in Appendix B, there is no spatial correlation in our model.

Since our variables of agglomeration are correlated to one another, we run a test of collinearity which is shown in Appendix B. The common rule of thumb is that Variance Inflation Factors of 10 or higher may be reason for concern. The VIFs for suspected correlated variables are less than 10. We also control for potential selection bias in our sample by introducing the inverse Mill's ratio computed using the estimates from a Probit model (see Appendix A). Then, our estimates are reliable.

[9] This statistic is shown when robust standard errors under heteroskedasticity are calculated.

Table 7 shows the results of our estimations. We present two productivity models: one using firms density (column 1) and other using employment's density (column 2).

Table 7: Estimation of the effect of agglomeration and industry structure on productivity

Dependent variable Productivity	(1) GMM firms	(2) GMM employees
specialization	0.0888 (1.899)*	0.0777 (1.536)
diversity	2.120 (5.023)***	1.996 (4.330)***
competition	-0.0658 (-1.647)*	-0.0492 (-1.120)
firms density	0.0705 (2.976)***	
employment's density		0.0684 (2.975)***
selection (mill's ratio)	-0.293 (-4.904)***	-0.246 (-4.289)***
Constant	-0.243 (-1.996)**	-0.219 (-1.906)*
N	7803	7803
F	74.17 (0.000)	76.25 (0.000)
Province specific effects	Yes	Yes
Industry specific effects	Yes	Yes
Instrumented	spe, div, den firms	spe, div, den workers
Included instruments	comp, prov_i, ind_s	comp, prov_i, ind_s
Excluded instruments	menace, den^2, pop90 div_spatial, comp_spatial, den_firms_spatial	menace, den^2, pop90 div_spatial, comp_spatial, den_workers_spatial
Endogeneity test	26.786 (0.000)	9.193 (0.0268)
K_P (Underidentification test)	97.02 (0.000)	91.61 (0.000)
Weak identification	22.30	22.92
Stock-Yogo critical value 5%	12.20	12.20
Hansen J statistic	3.132	2.204
J p-value	0.372	0.531
P_H (Heteroskedasticity test)	584.69 (0.000)	583.07 (0.000)

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

All variables are in logarithm.

Overall, both estimations are significant according to the F statistic. In the case of Ecuador, all agglomeration externalities have a significant impact on local productivity of industries. According to the aforementioned theories about specialization and diversity, only one of them is conceived to be the source of productivity. However, in Ecuadorian cantons, both specialization and diversity externalities have a positive impact. The specialization in a given industry generates positive externalities as expected by MAR's theory. As we could observe previously in Table 1, most of cantons specialize in activities that allow the development of innovations. This is the case of manufacturing, services of entertainment, professional, scientific and technical activities, among others. But other cantons have a strong share of activities that do not necessarily induce creation of new products and services, for example, public administration, social work activities and activities of international organizations. This could explains the lower effect of localization economies with respect to urbanization economies. Another explanation might be related to higher transport costs caused by low quality of transport infrastructure. This fact could impede full specialization of cantons (Duranton, 2007).

The strong presence of diversity externalities merits detailed explanations. It could be explained by the fact that the interactions between industries in cantons are high. Those interactions are reflected in backward and forward linkages. The backward linkages consist in demand pressure for inputs. Suppliers have to increase their production in order to satisfy demanding firms. The forward linkages consist in supply pressure. The expansion of supply firms would promote the expansion of sourcing companies.

According to a study made by the Central Bank of Ecuador[10], industries could be classified in: *motor* industries which have high backward linkages and low forward linkages; *base* industries which have low backward linkages and high forward linkages; *key* industries which have high backward linkages and high forward linkages; and *island* industries which have low backward linkages and low forward linkages. The results show that 33% of industries are motor industries, 23% of industries are base industries and 10% of industries are key industries. Only 24% of the industries have the island character. Thus, most of industries (76%) are likely to produce positive externalities in this context of strong forward and backward linkages. As a consequence, there would be a high exchange and mixture of ideas coming from different industries. This process so-called cross-fertilization might be the result of diversity externalities.

The effect of competition externalities is significantly negative. This fact contra-

[10]Council of the Production Sector. Input-Output Matrices, 2010

dicts the expectations based on the Porter's theory that high levels of competition promotes innovation activity which, in turn, increases productivity. On the contrary, the negative effect may indicate that a high level of competition outweighs the positive effects on innovation. It seems that competition constitutes a deterrent factor that reduces productivity or even worse expels firms out of the market. This could happen because the level of efficiency of firms is not high enough to stay in the market. Besides, the competitive environment impedes the increase of flows of knowledge as large firms with new ideas protect them in order to gain a larger share of the local market.

Regarding the density of the economic activity, the results slightly change between the estimates using firms density and those using employment's density.

As in the great majority of papers devoted to the estimation of economies of density (Ciccone and Hall, 1995; Combes and Gobillon, 2015), we found that productivity increases with density. In our estimates for Ecuador, the elasticity of productivity with respect to density is stable at 7%, whatever the measure of density. While the international comparison is complicated given the large diversity of articles using different frameworks, we take into account the studies with the closest context to ours (similar dependent variable and measure of density). Considering only the regional dimension and neglecting the industrial one, Ciccone and Hall (1995) and Ciccone (2002) found that economies of density in Europe and US measured by the effects of employment's density are 4.5% and 5%, respectively. The fact that our estimate appears to be higher may indicate that economies of density emerge not only at the local level but also at the industrial level. Moreover, we believe that the level of development of this country might explain the larger effect. There could be specific mechanisms such as the demographic bonus and internal migration (Combes et al., 2013) operating in this country that induce higher level of labor productivity. It is worth noting that this deduction has to be taken very cautiously.

In short, we have observed that agglomeration externalities derived from diversity, specialization and density positively affect the industrial productivity of cantons in Ecuador. Diversity externalities are higher than specialization externalities. It reflects that the interaction between different industries is more dynamic than the interaction within the same industry. On the other hand, competition between firms in the same industry generates detrimental effects. It is worth stressing that these conclusions are drawn from the estimations based on the whole sample but an industrial analysis is required in order to identify the specific effects for each industry.

In Section 5, we deepen such an analysis.

5 The industrial dimension of agglomeration externalities

Based on the whole sample, regressions in the previous section do not account for industrial heterogeneity. Sectors might be very different in terms of productivity dynamics and backward and forward linkages. This section addresses this issue by making a distinction between manufacturing[11] and service[12] industries[13].

Since services are provided to any industry, the effect of diversity externalities is expected to be large in the service industry's productivity. On the other hand, manufacturing industry would benefit more from specialization externalities because said industry would be interested in specific knowledge and techniques developed within its field of expertise. Regarding the effect of competition externalities, it is expected to be specific to each industry. It would depend on the level of competition within them.

The industrial estimations are presented in Table 8 and in Table 9. The former records the results when using firms density and the latter records the results when using employment's density. The columns (i) of both tables present the model for manufacturing and the columns (ii) present the model for services.

---

[11] The manufacturing industry encompasses: Manufacturing and Construction according to the International Standard Industrial Classification of All Economic Activities, Rev.4 (ISIC)

[12] In this analysis, the service industry comprises: Wholesale and retail trade; repair of motor vehicles and motorcycles, Transportation and storage, Accommodation and food service activities, Information and communication, Financial and insurance activities, Real estate activities, Professional, scientific and technical activities, Administrative and support service activities, Arts, entertainment and recreation and Other service activities, Water supply, sewerage, waste management and remediation activities, Electricity, Education and Health services and Public administration and defence, International organizations.

[13] We distinguish between these two large industries because our interest is to identify the intensity of externalities at that level of aggregation. Additionally, we use these industries because they are different between them. Each one has specific characteristics of operation in terms of supply and demand. For instance, the production of manufacturing industry is mainly devoted to final demand whereas the production of service industry is mainly devoted to intermediate consumption. It is worth noting that we take into account the heterogeneity within manufacturing and service industry in the model by including sub-industry specific effects.

Table 8: Agglomeration externalities in Manufacturing industry and Service industry (firms density)

Dependent variable Productivity	(1) GMM Manufacturing	(2) GMM Services
specialization	0.104	0.112
	(1.491)	(1.846)*
diversity	1.884	2.152
	(3.019)***	(4.093)***
competition	-0.173	-0.0294
	(-3.249)***	(-0.895)
firms density	0.134	0.0301
	(3.365)***	(0.785)
selection (Mill's ratio)	-0.560	-0.372
	(-7.113)***	(-6.624)***
Constant	-0.716	-0.432
	(-3.170)***	(-1.877)*
N	2198	5605
F	58.00 (0.000)	49.29 (0.000)
Province effects	yes	yes
Industry effects	yes	yes
Instrumented	spe, den firms	spe, div, den firms
Excluded instruments	menace, den^2, pop90 div_spatial den_firms_spatial	menace, den^2, pop90
K_P (Underidentification)	38.58 (0.000)	70.03 (0.000)
Hansen J statistic	3.250 (0.97)	exactly identified

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Table 9: Agglomeration externalities in Manufacturing industry and Service industry (employment's density)

Dependent variable Productivity	(1) Manufacturing	(2) Services
specialization	0.0786	0.107
	(1.034)	(1.610)
diversity	1.651	2.081
	(2.438)*	(3.400)***
competition	-0.136	-0.0226
	(-2.314)*	(-0.568)
employment's density	0.125	0.0312
	(3.362)***	(0.800)
selection (Mill's ratio)	-0.501	-0.349
	(-6.842)***	(-6.876)***
Constant	-0.663	-0.435
	(-3.133)**	(-1.901)
N	2198	5605
F	61.03 (0.000)	49.94 (0.000)
Province effects	Yes	Yes
Industry effects	Yes	Yes
Instrumented	spe, den workers	spe, div, den workers
Excluded instruments	menace, den^2, pop90 div_spatial den_workers_spatial	menace, den^2, pop90
K_P (Underidentification)	37.14 (0.000)	60.48 (0.000)
Hansen J statistic	2.862 (0.239)	exactly identified

t statistics in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

According to the robustness tests shown at the bottom of Tables 8 and 9, both models are identified and the instruments are valid. We also observe that the differences between the estimations using firms density and employment's density are not large. We rely on the regressions using firms density given the stability of the estimates.

The differences of the coefficient estimates between both industries are statistically significant according to the Chow test[14] The F statistic of the test is 23.90 and the p-value is 0.000. Then, we reject the hypothesis that the coefficient estimates of the manufacturing industry and those of the service industry are equal.

The comments of the results are the following. Firstly, diversity externalities are significant for both manufacturing and services. But the magnitude of these effects is much higher in the service industry. Indeed, firms providing services are most likely to benefit from industrial diversity because of demand linkages. Firms in other industries require multiple services to operate. This fact would boost the productivity of firms that provide those services. The resulting complementary relationship with service industry might be a key element driving positive externalities of diversity. Thus, a great variety of industries in the monopolistic competition in the service industry is beneficial.[15] The services are generally not limited to one specific industry or type of consumer (Combes, 2000). According to the aforementioned study of the Central Bank of Ecuador, services are classified specially as *base* industries whose demand of inputs is low but their production is mainly for the intermediate consumption. These services are telecommunications and information, financial activities, transport and storage, wholesale and retail trade, real state activities and professional, scientific and technical activities.

More importantly, positive effects of diversity might come from knowledge spillovers

[14]We use the Chow test to compare the coefficient estimates of the two subsets: manufacturing industry and service industry. We run one model with the whole data and two models with the subsets, separately. The formula of the Chow test is: $\frac{\frac{ess_{whole}-(ess_{manufacturing}+ess_{services})}{k}}{\frac{(ess_{manufacturing}+ess_{services})}{N_{manufacturing}+N_{services}-2*k}}$ where k is the number of estimated parameters, ess_{whole}, $ess_{services}$ and $ess_{manufacturing}$ are the error sum of squares of the estimations with the whole data, with the service industry and with the manufacturing industry, respectively. $N_{manufacturing}$ and $N_{services}$ are the number of observations in those industries. The resulting test is distributed with $F(k, N_{manufacturing} + N_{services} - 2 * k)$. For this computation, we use the results presented in Tables 7 and 8). We have: $\frac{\frac{8418.97-(2091.961+7691.023)}{45}}{\frac{(2091.961+7691.023)}{2198+5605-2*45}} = -23.90$.

[15]It holds true in the case of a production function à l'Ethier with exchange of intermediate goods and division of labor.

across industries. The wide network of service industry allows firms to obtain information very readily. Since these firms provide services in many industries, they can obtain information of what other firms in their same field or even in other fields are doing and the process of imitation and cross-fertilization takes place. The resulting knowledge spillovers allow imitation, adaptation and improvement of their products or services. For instance, a firm of transport services A learns that other firm B is offering a WiFi service in the bus. So the former has the interest to innovate and contacts a firm of telecommunications C that provides mobile data signal. Firm A contracts the WiFi service provided by C and now firm A can offer said service in the bus. The story does not end at this point. The searching of firm A involved several contacts with other firms in the telecommunications area that did not capture the costumer (firm A). Those firms probably do not ensure a stable connection or high speed that firm C has offered. Then, those firms will search to provide a better service by imitation, adaptation and improvement of what firm C has done.

The manufacturing industry also benefits from diversity externalities because of supply linkages. Manufacturing firms require to outsource inputs from other industries. According to the analysis of industries by the Central Bank of Ecuador, manufacturing firms are mainly classified as *motor* industries, which demand inputs of intermediate consumption and their production is mainly for final consumption.

Regarding the effect of competition externalities, it is negative for both industries. If firms do not achieve a certain market share, the effect of competition on their productivity will be detrimental. We believe that the presence of large firms in the same industry might negatively affect other firms. Large firms might conduct protection strategies that impede knowledge diffusion. The manufacturing industry experiences a higher negative effect than the service industry. It seems that high levels of competition in manufacturing industry outweigh the benefits of competition on innovation. On the other hand, service industry faces lower negative effects but not significant. It could be due to its dynamic sales network. Regarding the effect of specialization externalities, it is slightly higher in the service industry than in the manufacturing industry. In the latter, the effect is not significant. One reason could be that knowledge diffuses more easily in the service industry. This is in line with the lower negative competition effects in this industry.

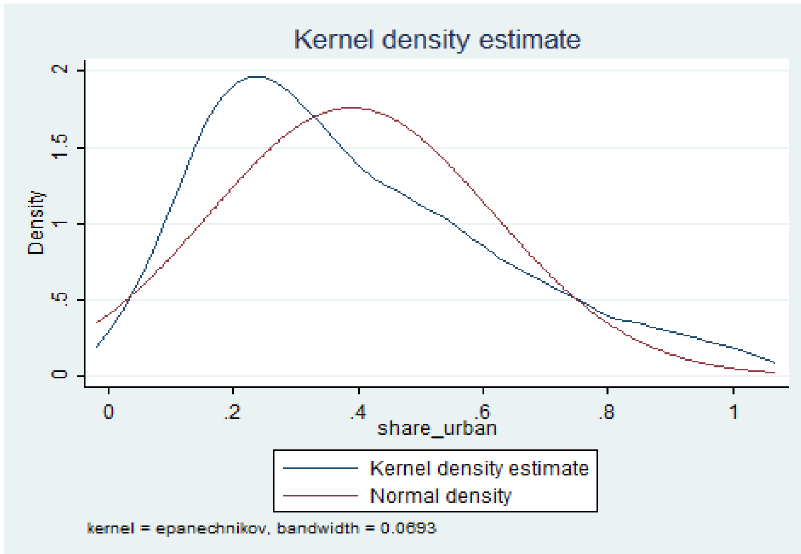
Both firms density and employment's density in the same industry are more relevant in the manufacturing industry than in the service industry. It seems that manufacturing requires geographical proximity to other firms and workers in the same sector in order to benefit from knowledge spillovers, matching and sharing mechanisms. The lower and insignificant effect in the service industry indicates that the interest of this industry is rather the geographic concentration of different industries than the concentration in the same industry.

Summarizing, the specific nature of industries plays an important role in determining the effects of agglomeration externalities on productivity. Four important insights are obtained from the industrial analysis of manufacturing and services. First, diversity externalities produce a noticeable effect in the service industry. Firms in this industry require a wide variety of industries, from which benefits in terms of demand and knowledge spillovers can be exploited. Second, specialization externalities produce significant positive effects for industrial productivity in services. Third, fierce competition might overshadow knowledge diffusion in the manufacturing industry. And fourth, economies of density produce positive externalities in the manufacturing industry.

6 Does urbanization matter?

Throughout the last sections, we could accurately observe the existence of agglomeration benefits for industries in cantonal economies of Ecuador. But it must be considered that cantons are heterogeneous. In terms of urbanization, the distribution of cantons is skewed (see figure 4). The second quartile is 0.33, i.e., there are 110 cantons out of a total of 221 cantons that have a value of urban population lower than 33%.

Figure 4: Histogram of urbanization of cantons in Ecuador



Source: Population and Housing Census of Ecuador, 2010.

Given the heterogenity of Ecuadorian cantons, we investigate whether the detected externalities in previous sections are shaped by the level of urbanization (*urban_rate*) of cantons. The intuition is that a critical mass of urban population is required in order to enjoy potential benefits of agglomeration. Moreover, Bertinelli and Black (2004) state that there exists a city size which maximizes the economy's per capita output net of congestion losses. Put differently, benefits of agglomeration may not be always present because congestion effects also arise from further agglomeration.

Through an analysis decomposing our whole sample in different levels of urbanization, measured by the share of urban settlements[16], we try to shed light on two aspects: i. the change of the effects of externalities depending to the level of urbanization and ii. the threshold of urbanization at which agglomeration externalities decline.

[16]According to the Population and Housing Census of 2010, urban settlements are defined by provincial capitals and cantons capitals under the current political and administrative division in the country. Rural areas include parish headers, other towns, peripheries and sparsely populated settlements

In order to determine the threshold values, we start setting the level of urbanization at 33% which is the median of our sample. Then, we run some simulations changing the level of urbanization and we look at the changes of agglomeration externalities. According to the results, we identify two threshold values. First, a level of urbanization of 46% is the value at which all agglomeration externalities become significant; and second, a level of urbanization of 61% is the value at which one of agglomeration externalities become negative. In Table 10, we present the estimations for different levels of urbanization: lower than 33% (column 1); larger than 33% (column 2); larger than 46% (column 3) and larger than 61% (column 4).

Table 10: Agglomeration effects vs. level of urbanization

Dependent variable Productivity	(1) urb <33%	(2) urb >33%	(3) urb >46%	(4) urb >61%
specialization	0.204	0.297	0.202	0.309
	(2.526)**	(5.401)***	(4.367)***	(4.950)***
diversity	0.196	3.102	2.535	3.011
	(0.238)	(6.398)***	(6.572)***	(6.480)***
competition	0.0365	-0.0665	-0.0298	-0.0280
	(0.937)	(-2.060)**	(-1.033)	(-0.717)
firms density	0.106	-0.0153	0.0719	-0.0371
	(1.709)*	(-0.339)	(1.806)*	(-0.634)
selection (Mill's ratio)	-0.410	-0.305	-0.315	-0.232
	(-5.573)***	(-4.800)***	(-4.185)***	(-2.127)**
Constant	-1.537	-0.740	-1.072	-0.841
	(-1.814)*	(-2.129)**	(-2.894)***	(-2.579)***
N	2983	4820	3524	2028
F	15.66 (0.000)	19.79 (0.000)	19.35 (0.000)	14.79 (0.000)
Province specific effects	Yes	Yes	Yes	Yes
Industry ag. specific effects	Yes	Yes	Yes	Yes
K_P Underidentification test	35.59 (0.000)	122.7 (0.000)	172.1 (0.000)	128.6 (0.000)
Hansen J statistic	1.978	1.580	2.420	4.518
J p-value	0.577	0.664	0.490	0.211

t statistics in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Firstly, we observe that urbanization mainly alters the effects of diversity and density externalities. The sensitivity of these variables could be due to their relationship with urban systems. Cities host a great diversity of industries and density of activities. The effect of competition also changes along the urbanization process. On the contrary, the effects of specialization externalities slightly change across the levels of urbanization.

The results in column (1) of Table 10 indicate that when a region is low-urbanized, diversity externalities do not occur. This could be explained by the fact that low urbanized regions do not provide an appropriate environment for the interaction of different industries since communication and transport infrastructures are still embryonary and provision of services is low. In that context, firms lose connection with a wide variety of suppliers, and workers cannot exchange knowledge. Another element that may impede the benefits of industrial diversity externalities in some low urbanized cantons is their particular geographic location, which curbs the expansion of urban areas. For instance, we mention the extreme cases, Colta and Guamote, cantons with less than 7% of urban population, are located at 3,212 and 3,050 meters above sea level, respectively. Taisha and Huamboya, cantons in the Eastern region with less than 10% of urban population, are located in the heart of the rainforest where road connections are scarce. By contrast, economies of density are significant at low levels of urbanization. This means that the geographical proximity between firms of the same industry is required to increase the industrial productivity. The effect of competition is unexpectedly positive at low levels of urbanization. However, it is not significant.

On the other hand, the results suggest that specialization externalities generate positive effects at any level of urbanization. The significant effect at low levels of urbanization could be due to the fact that some industries of specialization are associated with the primary sector which does not need high levels of urbanization. Overall, the diffusion of knowledge within the same industry is a permanent process across the levels of urbanization.

The results shown in column (2) suggest that as a canton urbanizes, the conditions for diversity externalities improve. A more urbanized environment is appropriate to take advantage from the concentration of different industries for two main reasons. First, the functioning of cities prompts dynamism of industries through a good provision of services and in turn productivity is enhanced. Second, urban areas are favored with communications and transport infrastructure. Hence, population moves to those areas. Thus, the probability of recruiting skilled-workers increases. In addition, industries in large cantons benefit the most from cross-fertilization of ideas of different industries. The exchange of ideas, imitation and adaptation of products and organizational procedures are eased in those cantons.

Surprisingly, the effect of density is insignificant for rates of urbanization higher than 33%. For that reason, we run some simulations to determine the level at which these externalities take place. It was identified at 46%. At that level, the effect of density becomes significantly positive. Thus, all externalities generating positive effects on productivity are significant at this urbanization ratio. However, we presume that there exists a certain level of urbanization leading to negative effects of firms density on productivity because of congestion diseconomies in big cities. In order to test this hypothesis, we run the model with several levels of urbanization higher than 46% and we find that the positive effect of density lasts until levels of urbanization of 61% (see forth column) at which economies of density cease to be significant.

Regarding the effect of competition, its impact becomes significantly negative at higher levels of urbanization than 33%. However, it seems that the negative effect decreases and even loses its importance along the urbanization process. This was not expected since higher levels of urbanization are very likely to lead to tougher competition and lower expected profits which, in turn, would cause greater negative effects on productivity. Nevertheless, the decreasing and non significant negative effect could indicate that competition is no longer a deterrent factor to knowledge diffusion and innovation in highly urbanized cantons. Thus, it seems that industries face a trade-off between negative effects arising from congestion and reduced negative effects of competition in large cantons.
All in all, at higher rates of urbanization than 61%, economies of density produce negative effects in the industrial productivity in cantons but the effects of diversity and specialization remain significantly positive. For this reason, we do not consider such a threshold as excessive urbanization.

7 Conclusions

In this study, the agglomeration effects have been examined considering the industrial structure of cantons. The different types of agglomeration externalities analyzed were: the so-called specialization or localization economies, the so-called diversity or urbanization economies, the density externalities and the competition effect. The empirical results highlight the importance of considering these types of externalities when studying the implications of agglomeration. The conclusions are the following.

The econometric results show that diversity externalities matter more than specialization externalities. The strong presence of diversity externalities is explained by the fact that most of industries in the country interact with each other, leading to high backward and forward linkages. Therefore, exchange and mixture of ideas coming from different industries is eased in such a diverse industrial environment.

Furthermore, cantons also benefit from localization economies. This means that there is no rivalry between diversity and specialization externalities. The effect of the former does not cancel the effect of the latter. Technological externalities (knowledge spillovers) and pecuniary externalities (vertical linkages) within industries of specialization exist and increase productivity. On national average, competition negatively impacts on labor productivity. The innovation inducing effect of competition promulgated by Porter's theory seems to be overshadowed.

Regarding the economies of density measured by firms density and employment's density, we found that the estimated elasticity of productivity with respect to density is 7% at the cantonal and industrial scale. Using a similar framework (dependent variable and measure of density) but disregarding the industrial dimension, estimations for developed countries show that the elasticity of productivity with respect to density is around 5% in Europe and United States. By comparing both results, we state that externalities are higher at the industrial-local level than at the local level only. The larger effect in this country may also be related to its level of development. Some specific mechanisms such as the demographic structure and internal migration (Combes et al., 2013) might be working behind the agglomeration externalities in the Ecuadorian economy.

Important insights are obtained from the industrial analysis of manufacturing and services. The particularities of industries determine specific effects of agglomeration externalities in their productivity. For instance, the service industry benefits the most from the agglomeration of different industries because of two reasons: on one hand, their output is not devoted to a specific industry and on the other hand, knowledge spillovers in this industry are very likely due to a wide sales network. Manufacturing benefits more from the geographical proximity of firms and workers within the same industry.

The present study steps forward to know whether the observed effects change depending on the level of urbanization. Our presumption is that agglomeration externalities hardly occur at low levels of urbanization because a critical mass of urban population is required in order to enjoy the potential benefits of agglomeration. The results indicate that this holds for diversity externalities.

Low-urbanized cantons typically suffer from scarce infrastructure. Hence, the low attractiveness of these cantons impedes the functioning of diversity externalities. As a canton urbanizes, several changes are produced. Communication and transport infrastructure improves, the provision of services increases and more workers move to these cantons. The level of urbanization at which diversity, specialization and density agglomeration externalities generate positive effects was identified at 46%. At higher levels of urbanization, economies of density and competition change. The former is no longer positive most likely due to congestion effects. The negative effect of the latter vanishes which means that competition would not be an impediment to technological and pecuniary externalities in highly urbanized cantons.

Overall, the contribution of this study is to enlarge the knowledge base of agglomeration economies in developing countries by conducting a case study for Ecuador. We estimate an accurate econometric model which deals with potential endogeneity of externalities while controlling for sample selection bias and spatial correlation. Apart from the conventional measure of density, this study proposes an alternative measure which is the firms density. It allows measuring the actual number of economic activities rather than the number of employees. Finally, it deepens the understanding of agglomeration mechanisms by looking at the changing effect of agglomeration externalities according to urban conditions of cantons.

Appendices

A Selection model

The fact that all sectors do not operate in each canton might be a selection issue. Firms' decisions might not be taken randomly. For that reason, it is compelling to deal with the presence of many zeros in our database. To do so, we apply the Heckman's method (Heckman, 1979). It considers a random sample on I observations. Some of them are observed and the others are not observed. If this is a selection process, we have:

(1a) $Y_{1i} = X_{1i}\beta_1 + U_{1i}$
(1b) $Y_{2i} = X_{2i}\beta_2 + U_{2i}$
$E(U_{ji}, U_{j'i''}) = \sigma_{jj'}, i = i''$

Suppose that observable data on Y_{1i} exists if $Y_{2i} \geq 0$. Thus, in the general case:

$$E(U_{1i} \mid X_{1i}, \text{sample selection rule}) = E(U_{1i} \mid X_{1i}, Y_{2i} \geq 0)$$
$$= E(U_{1i} \mid X_{1i}, U_{2i} \geq -X_{2i}\beta_2)$$

Assuming that $h(U_{1i}, U_{2i})$ is a bivariate normal density:
$E(U_{1i} \mid U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{12}}{(\sigma_{22})^{1/2}}\lambda_i$
$E(U_{2i} \mid U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{22}}{(\sigma_{22})^{1/2}}\lambda_i$

where $\lambda_i = \frac{\phi(Z_i)}{\Phi(-Z_i)}, Z_i = -\frac{X_{2i}\beta_2}{(\sigma_{22})^{1/2}}$

and ϕ and Φ are the density and distribution function for a normal standard variable, respectively. λ_i is the inverse of Mill's ratio. It corresponds to the decreasing function of the probability that an observation is selected in the sample.

In order to implement this method, Heckman (1979) suggests to estimate λ_i using a probit analysis for the whole sample. In our probit selection model, the dependent variable is a dummy that equals to 1 if a given industry exists in a given canton and 0 if not. The function *heckman* accommodates this data (Stata, 2015). The explanatory variables are the level of urbanization in a given canton and industry dummies. In Table 11, the results of the Probit model show that the probability that an industry operates in a given canton positively depends on the level of urbanization and the industrial characteristics. The correct prediction rate is 75%. Using the resulting estimates, we compute λ_i. Then, such a value is included in our main equation 6 that analyzes the effects of agglomeration externalities in productivity.

Table 11: Probit model, presence of zeros

	(1) Probit
urban rate	2.37
	(39.78)***
industry dummies	significant
Constant	-2.94
	(-17.01)***
Observations	17017

t statistics in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

B Robustness checks

In order to test the homoskedasticity in the residuals of our model estimations, we used the Pagan and Hall test which indicates that there is heteroskedasticity. In order to verify this result, residuals are plotted against fitted values. The uneven distribution of residuals shown in plots 5 and 6 confirms heteroskedasticity.

Figure 5: Residuals plot (GMM Model den 1)

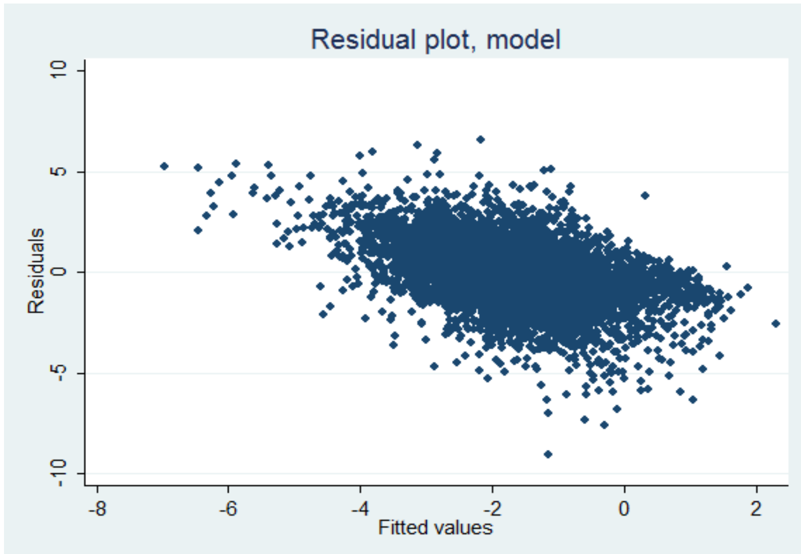
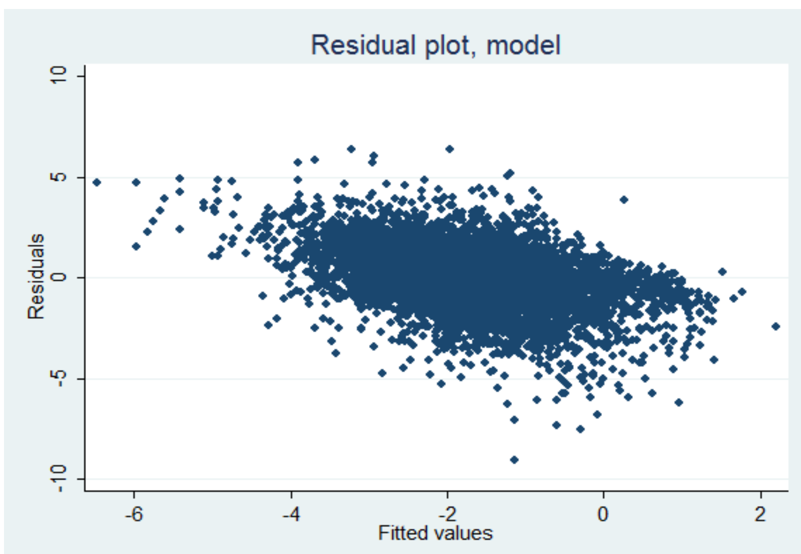


Figure 6: Residuals plot (GMM Model den 2)



According to the Variance Inflation Factor (VIF) presented in Table 12, our explanatory variables do not lead to multicollinearity issues in the model.

Table 12: Collinearity diagnostics

Model (den 1)		Model (den 2)	
Variable	VIF	Variable	VIF
competition	1.81	competition	1.44
specialization	1.07	specialization	1.40
diversity	1.26	diversity	1.34
firms density	1.76	employment's density	1.84

In order to detect whether our estimated models shown in Table 7 present spatial correlation, we construct Moran scatter plots for residuals for each industry (78 industries). They are presented in Tables 7 to 10. Since any linear relationship, either positive or negative, is observed, the residuals of our estimations do not present spatial autocorrelation.

Figure 7: Spatial autocorrelation by industries

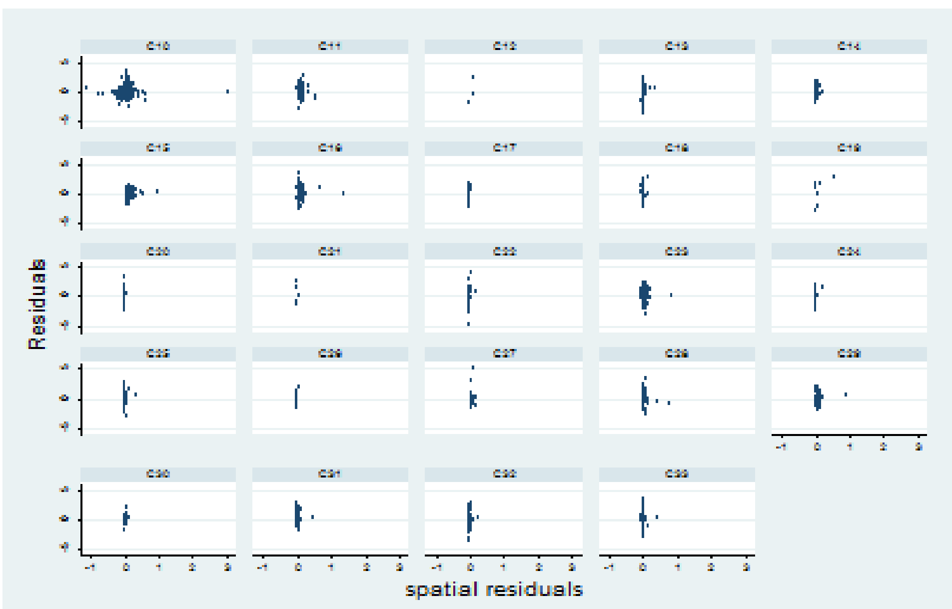


Figure 8: Spatial autocorrelation by industries

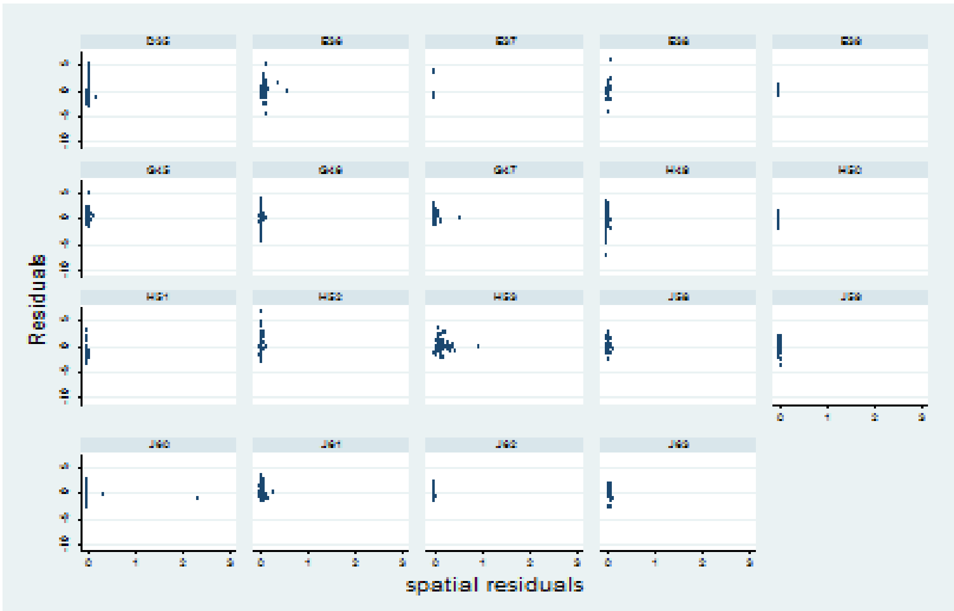


Figure 9: Spatial autocorrelation by industries

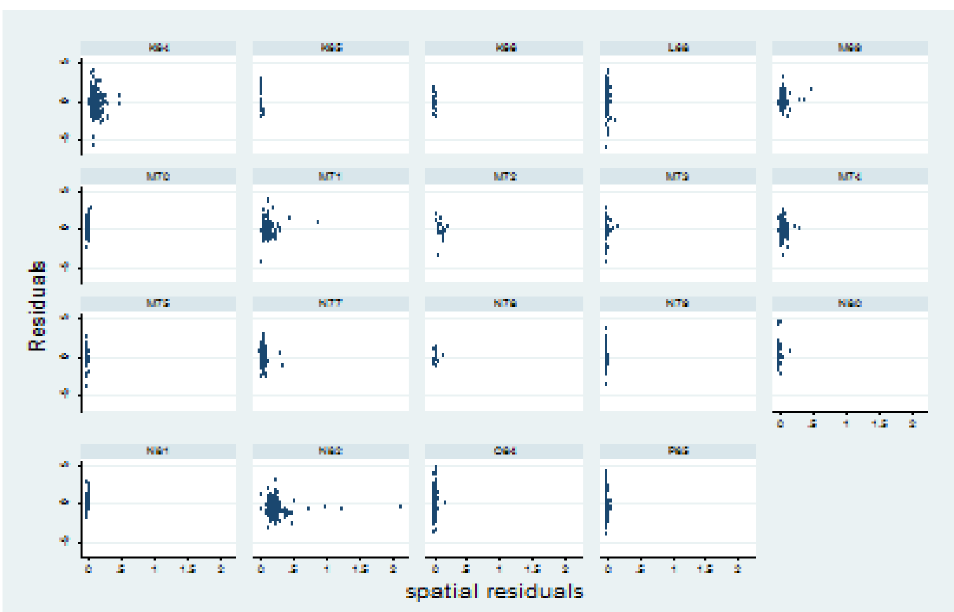
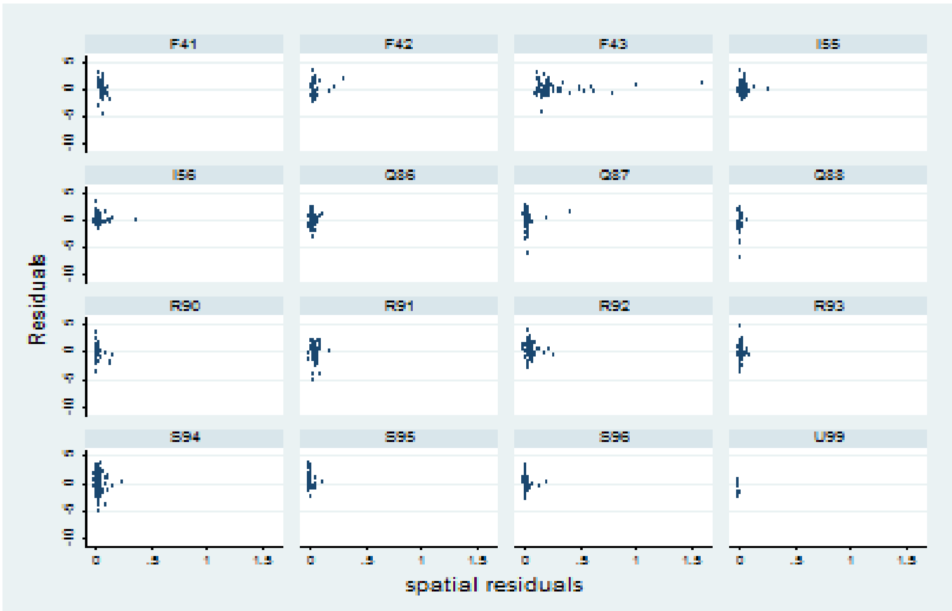


Figure 10: Spatial autocorrelation by industries



References

Arrow, K. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29:155–173.

Baum, C., Schaffer, M., and Stillman, S. (2010). Stata module for extended instrumental variables/2sls, gmm and ac/hac, liml and k-class regression. Technical report, Boston College Department of Economics.

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? the localization versus urbanization debate. *Research Policy*, 38:318–337.

Bertinelli, L. and Black, D. (2004). Urbanization and growth. Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics.

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. (2000). Economic structure and local growth: France, 1984-1993. *Journal of Urban Economics*, 47:329–355.

Combes, P.-P., Demurger, S., and Shi, L. (2013). Urbanisation and migration externalities in China. Technical report, School of Economics Aix Marseille.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In Glaeser, E. L., editor, *Agglomeration Economies*, pages 15–66. University of Chicago Press.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In Duranton, G., Henderson, V., and Strange, W., editors, *Handbook of Urban and Regional Economics*, volume 5.

Combes, P.-P. and Lafourcade, M. (2012). Revue de la littérature académique quantifiant les effets d'agglomeration sur la productivité et l'emploi. (In english: Literature review about agglomeration effects in productivity and employment.). Report, Societe de Grand Paris.

Da Silva Catela, E. Y., Goncalves, F., and Porcile, G. (2010). Brazilian municipalities: Agglomeration economies and development levels in 1997 and 2007. *CEPAL Review*.

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.

De Groot, H. L., Poot, J., and Smit, M. J. (2007). Agglomeration, innovation and regional development. Technical report, Tinbergen Institute Discussion Paper.

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. (2001). Nursery cities: Urban diversity, process innovation, and the life of cycle of products. *American Economic Review*, 91(5):1454–1477.

Duranton, G. and Puga, D. (2003). Micro-foundations of urban agglomeration economies. Working Paper 9931, National Bureau of Economic Research.

Fajgelbaum, P. and Redding, S. J. (2014). External integration, structural transformation and economic development: Evidence from argentina 1870-1914. working paper 20217 2894, National Bureau of Economic Research, working paper 20217.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Imbs, J. and Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1):63–86.

Jacobs, J. (1969). *The economy of cities*. Random House.

Marshall, A. (1890). *Principles of Economics*. London. MacMillan.

Martin, P. and Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review*, 42(4):947–968.

Pereira, M. and Soloaga, I. (2012). Determinantes del crecimiento regional por sector de la industria manufacturera en Mexico 1988-2008 (In english: Determinants of regional growth of manufacturing sector in 1988-2008 Mexico.). Technical report, Centro de Estudios Economicos, El Colegio de Mexico CA.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, 94(5):1002–1037.

Saito, H. and Gopinath, M. (2009). Plants' self selection, agglomeration economies and regional productivity in Chile. *Journal of Economic Geography*, 9(1):539–558.

Stata (2015). *Stata Base Reference Manual, Release 14*.

Stock, J. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In Andrews, D. and Stock, J., editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.*, pages 80–108. Cambridge Univ. Press.