



**HAL**  
open science

## Agglomeration externalities in Ecuador. Do urbanisation and tertiarisation matter?

Carolina Guevara, Stéphane Riou, Corinne Autant-Bernard

### ► To cite this version:

Carolina Guevara, Stéphane Riou, Corinne Autant-Bernard. Agglomeration externalities in Ecuador. Do urbanisation and tertiarisation matter?. 2018. halshs-01872922v2

**HAL Id: halshs-01872922**

**<https://shs.hal.science/halshs-01872922v2>**

Preprint submitted on 24 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1818 – September 2018

## **Agglomeration externalities in Ecuador. Do urbanisation and tertiarisation matter?**

Carolina Guevara, Stéphane Riou, Corinne Autant-Bernard

### **Abstract:**

The paper investigates whether the tertiarisation and the rapid urbanisation faced by developing countries favour agglomeration economies. Focusing on Ecuadorian cantons, a productivity equation is estimated using the GMM model with instruments controlling for endogeneity. The varying impact of industrial concentration, diversity, competition and density across industries is investigated and for the first time, the implication of the level of urbanisation on agglomeration externalities is studied. Stronger effects are found for services. The threshold of urbanisation at which diversity, density and competition agglomeration externalities all generate positive effects was 33% while they seem challenged by congestion in highly urbanised cantons.

### **Keywords:**

Agglomeration externalities, productivity, urbanisation, tertiarisation, Latin America

### **JEL codes:**

O54, L80, O18, L60

# Agglomeration externalities in Ecuador. Do urbanisation and tertiarisation matter?

Carolina Guevara

Contact author: Email: carolina.guevara@epn.edu.ec  
Departamento de Matemática, Facultad de Ciencias,  
Escuela Politécnica Nacional

Stéphane Riou

University of Lyon, UJM Saint-Etienne  
GATE UMR 5824, F-42023 Saint-Etienne, France.

Corinne Autant-Bernard

University of Lyon, UJM Saint-Etienne  
GATE UMR 5824, F-42023 Saint-Etienne, France.

September 18, 2018

## Abstract

The paper investigates whether the tertiarisation and the rapid urbanisation faced by developing countries favour agglomeration economies. Focusing on Ecuadorian cantons, a productivity equation is estimated using the GMM model with instruments controlling for endogeneity. The varying impact of industrial concentration, diversity, competition and density across industries is investigated and for the first time, the implication of the level of urbanisation on agglomeration externalities is studied. Stronger effects are found for services. The threshold of urbanisation at which diversity, density and competition agglomeration externalities all generate positive effects was 33% while they seem challenged by congestion in highly urbanised cantons.

agglomeration externalities; productivity; urbanisation; tertiarisation; Latin America  
O54; L80; O18; L60

## 1 Introduction

The theoretical and empirical literature on agglomeration is extensive. The channels by which agglomeration economies translate into economic efficiency have been widely debated (see blackRosenthal and Strange, black2008), stressing the potential role played by local diversity, specialisation and competition.

However, the survey by blackDe Groot et al. (black2015) points to the fact that at-

tempts to understand these various types of agglomeration economies mainly focus on developed countries, despite that the nature of the agglomeration externalities in developing countries seems to be a relevant issue. Indeed, one can expect agglomeration to have a differentiated effect on developing countries in comparison with developed ones because of the different conditions under which these agglomeration externalities take place. Actually, at the country level, the literature does not provide clear evidence of the existence of a monotonic relationship between sectoral diversification and income. blackImbs and Wacziarg (black2003) detected a robust U-curved relationship between specialisation and GDP per capita. Interpreting this relationship as a causality, diversification could be a major driving force of economic development for countries at intermediate stages of development.

A better understanding of the driving forces at work behind agglomeration in developing countries is therefore requested in order to identify the sectoral organisation that maximizes its productivity effects. Indeed, identifying whether specialisation, diversity and competition accelerate or impede local growth contains some important public policy issues. To this aim, it is important to identify the impact of the industrial structure on firms' performances at a local level. Indeed, it can be questioned whether industrial diversity, which seems to have positive effects on macroeconomic performance, generates similar effects on a more spatially disaggregated level. Income inequality between regions is a significant concern in developing countries, particularly in Latin America. Their industrial structure is frequently presented as part of the explanation for the economic gap between regions. Poor regions suffer from an excess of specialisation, the firms located therein having too few opportunities to benefit from inter-industry relationships. The immediate policy recommendation would be to encourage the emergence of an industrial diversification dynamics in these regions. Such recommendation must be systematically based on robust empirical evidences that the industrial structure of regions has a significant impact upon local growth.

A rising literature tries to fill this gap but there is still little empirical evidence for developing countries at the regional or individual level (see. blackCombes and Gobillon (black2015), blackChauvin et al. (black2014)). In most cases, evidence is provided on the impact of both specialisation (often measured through industrial concentration) and diversity, and when studied, on the competition effect. As detailed below, the underlying mechanisms remain however poorly explored so far.

At the regional level, blackGhani et al. (black2013) point to a higher productivity in Indian districts that display a positive role played by the initial level of specialisation for employment growth. For the Indonesian case, blackWidodo et al. (black2015), based upon firm-level data, observe that local specialisation increases firm's technical efficiency while diversity would reduce it. blackErcole and O'neill (black2016) stress the role of industrial linkages, pointing to the positive impact of related variety for Indonesia. In the Chinese case, blackNing et al. (black2016) put special emphasize on the role of agglomeration effects on the level and spillovers effects of FDI. Their results reveal positive impact of specialisation, diversity and competition. As for the Latin American case, blackDa Silva Catela et al. (black2010) show that both diversity and specialisation have positive effects on the real average wage per worker at the municipality level in Brazil. In Chile, blackSaito and Gopinath (black2009) found that more productive firms in the food industry locate in regions where other food industry

plants agglomerate and in regions that combine industrial diversification with a large market.

More recently, blackSimões and Freitas (black2014) and blackBarufi et al. (black2015) explore, respectively, the drivers of agglomeration economies on regional and individual wages in Brazil. They both find mixed evidence of specialisation-based and diversity-based agglomeration forces although specialisation prevails in the study conducted by blackBarufi et al. (black2015).

These investigations focus however on the manufacturing industry only (blackBarufi et al. (black2015) being however a noticeable exception). The rising share of service sector in developing economies (reaching more than 50% for instance in Latin America) leads to seek a better understanding about how the change of industrial structure in these countries may shape agglomeration economies. For developed countries, blackCombes (black2000) shows that the impact of the local economic structure differs in industry and services: while employment density, competition, sectoral specialization and diversity have a negative impact on growth in manufacturing industries, diversity and density exhibit positive effects on service industries. If the same picture occurs in developing countries, this means that the tertiarisation of these countries would reinforce their agglomeration economies.

In addition, one of the most striking fact in developing countries is the drastic increase in the level of urbanisation and, compared to developed countries, developing countries face huge disparities in terms of urbanisation rate at the regional level. Such features are also likely to shape agglomeration economies. The intuition is twofold. First, a critical mass of urban population may be required to enjoy the potential benefits of agglomeration. A certain level of urbanisation would guarantee the existence of minimum transport and telecommunication infrastructures and banking, financial services or other specific services that matter for the emergence of externalities (blackHenderson, black2000). Secondly, one may expect the existence of an optimal city size which maximizes the economy's per capita output net of congestion losses (blackBertinelli and Black, black2004). Therefore, in a context of increasing spatial concentration of economic activities, a major challenge for developing countries is to reinforce the role of their urban systems as drivers of economic growth (blackDuranton, black2015). It is thus worth investigating whether agglomeration economies vary according to the local level of urbanisation.

In order to explore these issues, this paper studies the emergence and nature of agglomeration effects in the Latin American context. Total factor productivity is regressed on specialisation, diversity and competition indexes while controlling for the effect of density. In order to explore the underlying features that condition the emergence and nature of agglomeration effects, separate estimates of the agglomeration effects in the service and manufacturing industries are provided. Whether a critical threshold of urbanisation would be required in order to benefit from agglomeration effects is then assessed.

Empirical evidence is driven from Ecuadorian cantons. Indeed, Ecuador constitutes a suitable case study as this country shares numerous common characteristics with most of the Latin American countries, such as its industrial composition and its urbanisation process. In Ecuadorian regions, 25% of the total economic activity take place in the

primary activity sector (22% in average in Latin America), 20% in the manufacturing sector (resp. 22% in Latin America) and 55% in the service industry (resp. 57% in Latin America). Most importantly, the services industry in Ecuador recorded an average growth rate of 1.5% (resp. 0.4% in Latin America) in the 2000 decade while the manufacturing industry recorded a 0.9% (resp. 0.5% in Latin America). This shows that Ecuador might be experiencing a process of tertiarisation. The urbanisation process is also similar to the Latin American one. Due to the fact that emerging countries are far from being fully urbanised, the urbanisation growth is rapid (blackHenderson, black2003). The annual average growth rate of urbanisation is around 2,2% in Latin America and around 3,2% in Ecuador.

In addition, when measuring agglomeration economies, one concern is endogeneity. This problem is not systematically addressed in the studies mentioned above (blackCombes and Gobillon, black2015). In this study, one major source of endogeneity that must be controlled is the existence of circular causality between local productivity and the type of externalities estimated. In this purpose, the instrumental variables method is implemented. Estimates are provided using the GMM model and instrument the endogenous variables with a threat index, long lagged population density and spatial lags.

This empirical work is mainly based on the 2010 Economic Census of Ecuador which contains information declared in 2009. By aggregating company data, a two-digit industry database is built on the cantonal level. Based on 7,803 canton and industry pairs, our results support Porter's externalities which argue that the combination of a diverse industrial environment and strong competition leads to positive productivity effects. In this line, specialisation is then detrimental. The magnitude of the externalities from diversity is positive and significant in the service industry, pointing to higher agglomeration forces when tertiarisation raises. Economies of density, measured by the density of firms and employees at the canton and industry levels, are found for the manufacturing sector. Competition leads to local productivity gains, especially for manufacturing sectors. More importantly, the regressions highlight the fact that diversity-based externalities increase with the level to which cantons are urbanised. In contrast, the effects of density and competition are non-monotone across the levels of urbanisation, producing both positive and negative effects. The latter suggest the presence of congestion diseconomies in some cantons.

The rest of the paper is organised as follows. Section 2 presents the various indicators used to identify the different types of externalities and provides statistical facts highlighting the sectoral and spatial organisation of economic activities in Ecuador. Section 3 focuses on the specifics of the model and the econometric methodology. The results are discussed in Section 4, separating first the service and manufacturing industries and distinguishing then the effects according to the levels of urbanisation. The main conclusions are presented in Section 5.

## 2 Spatial and sectoral distribution of economic activities in Ecuador

### 2.1 Measures of agglomeration externalities

The 2010 Economic Census of Ecuador includes information on 511,130 firms. It covers all localities with more than 2,000 inhabitants, with the exception of rural areas. By aggregating firm data<sup>1</sup>, an industrial database is constructed with two-digit ISIC industries (International Standard Industrial Classification). The resulting sample data covers 78 industries including manufacturing and service activities (Appendix A). Despite the availability of data on a very disaggregated geographical scale, using information on the block and city levels would induce too many missing industries. For this reason, data are aggregated at the cantonal level. Ultimately, the resulting panel data covers 78 industries and 221 cantons. Even after this aggregation, however, some industries are not present in all cantons. More specifically, on average 42 industries are absent in each canton.

The average performance of a given sector in a given canton is measured by the total factor productivity. Alternatively, we could restrict our analysis to the labor productivity<sup>2</sup>. It would result in an incomplete picture of the sectoral efficiency sources as a large number of sectors under consideration are capital intensive. Moreover, it is well known that some externalities directly translate into efficiency gains in the use of capital. For instance, sectoral diversity offers a better productive environment by enlarging the inputs market and improving the matching with the needs of firms. The Levinsohn and Petrin (2003)'s method is used. Following this method, the productivity measure is consistent and the simultaneity problem between production and the unobservable productivity is correctly treated. To estimate productivity, firm level data is first considered for the following production function of firm  $i$ .

$$\ln rev_i = \ln L_i \beta_L + \Phi(\ln K_i, \ln inp_i) + \epsilon_i \quad (1)$$

where  $rev_i$  is the level of revenue,  $L_i$  is the number of employees,  $K_i$  is the assets value and  $inp_i$  is the materials value (including electricity, packaging, raw materials, etc.) of firm  $i$ . The term  $\hat{\Phi}$  is approximated with a polynomial in  $(K_i, inp_i)$  capturing the TFP. In order to be consistent with the level of data availability for the other variables, the resulting TFP at the firm level is averaged by industry and canton.

The Marshall-Arrow-Romer (MAR) view of externalities considers that the productive efficiency of a particular industry in a given region is mainly boosted by the possible interactions within the industry in this region. Put differently, spatial proximity facilitates within industry knowledge transmission and technological appropriation. As it is standard in the literature, a relative specialisation index for each pair of sector-canton is built (see. Glaeser et al. (1992); Beaudry and Schiffauerova (2009)). The index measures the share of the canton's employment devoted to the given industry. Note that this measure is the regular location quotient which describes the industrial concentration. To control for the size of the sectors at the canton level and identify to what extent a canton is concentrated in a given sector compared to the

country, it is normalised to the share of national employment allocated to this sector. The industrial concentration index is thus as follows:

$$spe_{s,c} = \frac{empl_{s,c}/empl_c}{empl_s/empl} \quad (2)$$

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

Unlike MAR theory, blackJacobs (black1969) sees industrial diversity of a particular region as the main source of externalities. To capture the degree of sectoral diversity faced by sector  $s$  in canton  $c$ , the inverse of the Herfindahl index is built. Formally:

$$div_{s,c} = \frac{1}{\sum_{s'=1, s' \neq s}^S (empl_{s',c}/(empl_c - empl_{s,c}))^2} \quad (3)$$

where  $S$  is the number of industries. When industries are numerous and of equal size in a region, the value of this index reaches its maximum. Based on this measure, diversity externalities are understood as the effects on productivity due to the presence of different industries<sup>3</sup>. In addition, it is worth noting that industrial concentration and diversity calculated in this way are not necessarily opposed to one other. A canton may host a very large share of a sector and a broad number of other industries as well. Such a canton is therefore both specialised and diversified (blackPanne and Beers, black2006).

In accordance to MAR theory, blackPorter (black2010) argues that specialisation is the main source of externalities but that the market structure matters. Unlike a local monopoly, competition produces emulation and fosters the adoption of the best organization and/or technologies. To approximate these competition effects, full use of the firm level data is made. This enables characterising the size distribution of firms within sector  $s$  in canton  $c$ . After a normalisation at the country level, we have:

$$comp_{s,c} = \frac{1/\sum_i^N (empl_{i,s,c}/empl_{s,c})^2}{1/\sum_i^N (empl_{i,s}/empl_s)^2} \quad (4)$$

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 the country level. It is worth noting that this measure of competition is close to the industrial concentration measure. However, the latter shows the employment concentration in given industries, whereas competition shows the strength of firms within an industry that allows them obtaining productivity gains through innovation.

Moreover, abundant and recent literature provides evidences on the effect of density on local productivity (see blackCombes and Gobillon, black2015). Most existing studies use employment density as an explanatory variable. Formally, it is computed as follows:

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

Observe that a high employment density value may reflect both the presence of few big firms or the existence of a large number of competitive firms. Both configurations may produce positive externalities as they may emerge from both individual interactions in a given firm or between firms. Consequently, we provide the following alternative index which measures density of operating firms in sector  $s$  and canton  $c$ :



$$den_{s,c} = \frac{f_{s,c}/area_c}{f_s/area} \quad (6)$$

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

## 2.2 Spatial statistics of Ecuador

The dynamism of the Ecuadorian economy lies in its large share of service and manufacturing industries, which in 2010 represented 20% and 55% of national production, respectively. Such an economic structure is common to other developing countries in Latin America (blackHernández et al., black2016; blackECLAC-ILO, black2012). This study thus contributes to a better understanding of agglomeration externalities in these countries.

In Ecuador, the share of population living in urban areas was 67.8% in 2010. However, this share hides strong intra-regional heterogeneity. In terms of cantons, the proportion of the population living in urban areas is, on average, 38.9%. This means that there are many more cantons with rural population than cantons with urban population. Figure 1 presents the spatial concentration of economic activities in cantons using firms density (a) and employment density (b). The darker areas correspond to denser cantons. It can be observed that spatial patterns are fairly similar across density measures. Denser cantons are mainly located in the Sierra region of the country, while the less dense cantons are located in the eastern region which has large territories occupied by the rain forest. On the national average, the mean (employment) density of firms in cantons is 3.9 firms (12 workers) per square kilometre and the standard deviation is 13.35 (40), which indicates high heterogeneity between cantons. In the north, the main centre is the capital, Quito (1 in Figure 1), which records a density of 23 firms (121 workers) per square kilometre. Three centres can be distinguished in the coastal region: 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 square kilometre, respectively. Another centre of economic activity is Ambato (4 in Figure 1), recording 18 firms (59 workers) per square kilometre. In the south, the main centre is Cuenca (5 in Figure 1) which has a density of nine firms (36 workers) per square kilometre. In addition, Figure 1 shows higher firm and employment densities in cantons which are close to the economic centres.

The two following figures present the profile of cantons in terms of industrial diversity and productivity.

Basic comparison between Figures 2a and 2b indicates a similarity: *ceteris paribus*, denser cantons exhibit a high level of industrial diversity. Moreover, the closer the cantons are to urban centres, the higher their level of diversity. As expected, the three major urban centres have higher productivity rates than the national average: the capital, Quito (1 in Figure 1), Guayaquil (2 in Figure 1) and Cuenca (5 in Figure 1). The correlation between diversity and density at the cantonal level is 0.24, which is significant. Similarly, density and productivity are correlated, with a significant coef-

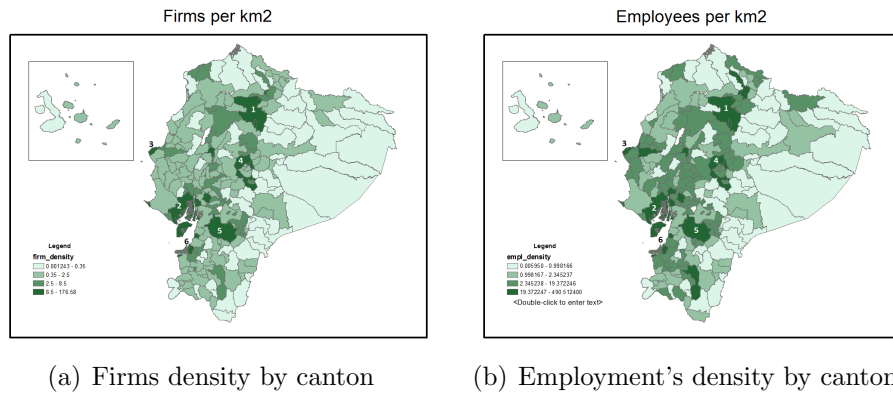


Figure 1: Density of economic activities in cantons  
Elaborated by authors based on the Economic Census 2010, Ecuador.

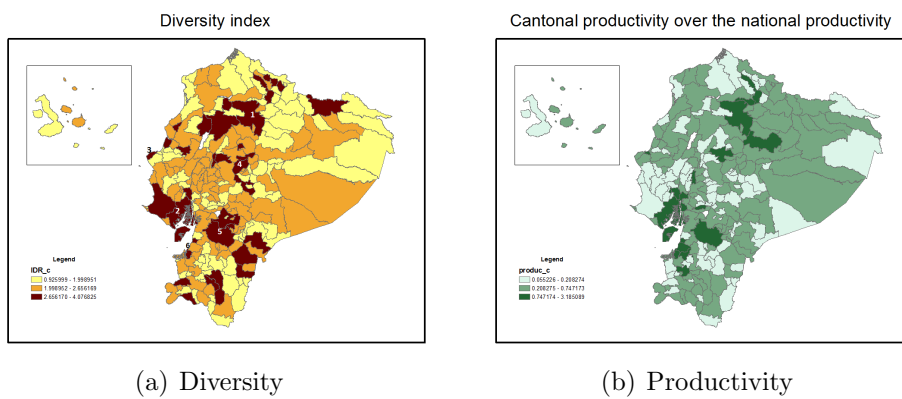


Figure 2: Diversity and productivity of cantons  
Elaborated by authors based on the Economic Census 2010, Ecuador.

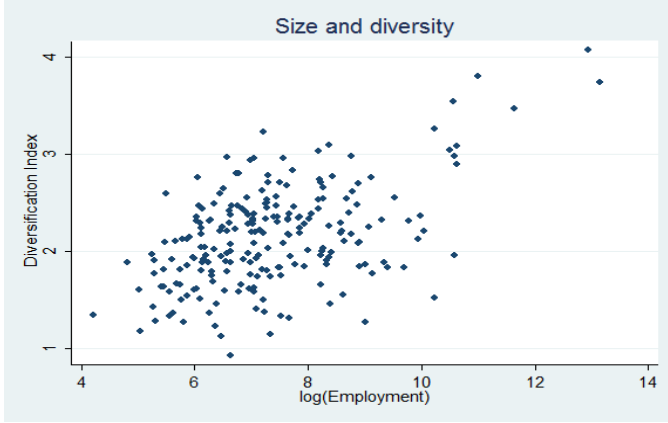


Figure 3: Canton size vs diversification

Elaborated by authors based on the Economic Census 2010, Ecuador.

ficient of 0.15. Overall, highly diversified cantons are larger in terms of economic size (blackHenderson, black1997). As shown in Figure 3, larger cantons present a wider variety of economic activities. The correlation coefficient between size, measured by total employment, and diversity index is 0.5087.

### 3 Econometric methodology

#### 3.1 Econometric issues related to agglomeration economies

In productivity regressions involving agglomeration variables, endogeneity problems may emerge for two reasons (blackCombes and Gobillon, black2015). On the one hand, endogeneity arises when omitted variables affect both agglomeration and productivity. For instance, a high level of productivity in a given industry might be caused by good local conditions, such as transport infrastructures, 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.

In this study, endogeneity concerns can arise from the industrial concentration variable. High sectoral productivity in one canton is likely to attract more firms from the same sector to that canton. This encourages more externalities from specialisation, which will be reflected in higher productivity. As a consequence of the bidirectional nature of the relationship, the variable encompassing the industrial concentration will be endogenous.

In terms of externalities from diversity, the impact of local productivity in a particular industry on the level of industrial diversity of one canton is, a priori, not obvious. Typically, it depends on whether the productivity gains of a given industry which spill-over across sectors are such that they boost their attractiveness. If so, local productivity in a given industry would lead to a more diversified local economy. In the other direction, according to Jacobs, a more diverse industrial environment is beneficial for the exchange of ideas across industries and the emergence of innovations, which, in turn, foster productivity gains. Consequently, the endogeneity of local industrial

diversity by reverse causality remains to be tested.

Regarding local industrial competition, the externalities arising from competition described by Porter may boost innovation within the same industry, but the opposite direction of the relationship is unclear. Local industrial productivity may boost competition through the entry of new firms attracted by the local profitability of existing firms. However, firms would rather locate in places where competition is not strong. Hence, the endogeneity of this variable also needs to be tested.

### 3.2 Specification of the model

In terms of the aforementioned issues of endogeneity, the feasible efficient two-step GMM estimator is used in the regressions. This estimator is efficient (more than 2SLS estimator<sup>4</sup>) in the presence of heteroskedasticity. Thus, it provides consistent estimates by dealing with both potential endogeneity and heteroskedasticity.

The model estimating the effect of each type of externality upon the productivity level of a specific industry  $s$  in a given canton  $c$  is as follows.

$$(TFP_{s,c}) = \alpha + \beta_1 spe_{s,c} + \beta_2 div_{s,c} + \beta_3 comp_{s,c} + \beta_4 den_{s,c} + \gamma_2 p_2 + \dots + \gamma_l p_l + \lambda_2 s_2 + \dots + \lambda_l s_l + selection_{s,c} + u_{s,c} \quad (7)$$

where  $(TFP_{s,c})$  is the Total Factor Productivity in industry  $s$  in canton  $c$ , estimated using the blackLevinsohn and Petrin (black2003)'s method. To control for unobserved heterogeneity, l-1 province fixed effects given by  $p_2 \dots p_l$  are introduced as well as sector fixed effects given by  $s_2 \dots s_l$ . The term  $selection_{s,c}$  is a control variable which allows us to address the issue regarding the presence of many zeros in the 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 equation (see Appendix blackC) proposed by blackHeckman (black1979).

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

First, the 1990 population density (*pop 90*) is used as an instrument. The long-term lagged variable 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 (blackCombes et al., black2010). The endogenous variables are also instrumented with their spatial lags (*spe\_spatial*, *comp\_spatial*, *den\_spatial*, *div\_spatial*). In order to compute the spatial lags, the inverse distance weight matrix is built using the coordinates of the capital of each canton. The spatial lags are likely to be relevant instruments since neighbouring cantons could affect one canton's level of industrial concentration, diversity, density or competition, but the extent to which neighbours' agglomeration externalities could affect the level of productivity of one canton cannot be immediately assumed. It is worth noting that the introduction of

the spatial dimension of these instruments somehow controls for possible spatial correlation in the model. The squares of the endogenous variables could be considered as good instruments because they explain the current patterns of endogenous variables. However, only the density squared ( $den^2$ ) passes the validity test.

Another interesting instrument is the natural threat index of cantons (*threat*) which takes into account causes of natural disasters in each canton. This variable has been compiled on the basis of a joint research study produced by the System of Economic Indicators (SIISE), Oxfam International and Cooperazione Internazionale (COOPI). This index is assumed to be a reasonable instrument because it is correlated with all the endogenous variables. The level of industrial concentration, diversity and density of regions is related to their degree of risk exposure.

In order to have consistent estimates of the productivity equation (black7), the instrumental variables must meet two specific properties. First, they must be highly correlated with instrumented variables (relevance condition), and second, a valid instrument has to be exogenous (orthogonality condition). Although the latter cannot be directly tested, the former is readily testable through the following analysis of correlation coefficients between the endogenous variables and their instruments shown in Table 1.

Table 1. Validity tests of instruments

	Correlation between instruments and endogenous variables			Exogeneity test
	Specialisation	Diversity	Density firms	C-statistic
threat index	-0.0813 (0.000)*	0.1439 (0.000)*	0.1914 (0.000)*	1.25 (0.264)
squared density	0.2685 (0.000)*	0.108 (0.000)*	0.5513 (0.000)*	0.978 (0.323)
Popd 90	-0.1097 (0.000)*	0.2902 (0.000)*	0.8059 (0.000)*	0.042 (0.838)
Spe spatial lag	-0.0045 (0.6895)	-0.0005 (0.9672)	-0.0054 (0.6319)	
Div spatial lag	-0.0101 (0.3673)	-0.0261 (0.0197)*	-0.022 (0.0494)*	0.074 (0.786)
Comp spatial lag	-0.0144 (0.1990)	-0.0253 (0.0242)*	-0.0234 (0.0366)*	0.062 (0.804)
Den firms spat lag	0.2685 (0.000)*	0.108 (0.0001)*	0.5513 (0.000)*	0.923 (0.337)

In parentheses, the p-value of the correlation coefficient (\* p < 0.05) and the p-value of the C-statistic.

Threat index, 1990 population density and squared density are highly correlated with all the endogenous variables, proving to be relevant instruments. The spatial lags of diversity, competition and density are correlated with some of the endogenous variables. In contrast, the spatial lag of industrial concentration is not correlated with any of the endogenous variables. Therefore, it is not considered as instrument.

The orthogonality condition is tested after the estimations through the C statistic or so called 'difference in Sargan' statistic whose null hypothesis is that suspected and additional instruments are valid. As shown in Table 1, the null hypothesis cannot be rejected, so instruments are valid.

## 4 Results

Prior to discussing the results of the estimates, the endogeneity test (Sargan-Hansen statistic) of the variables of agglomeration externalities is presented in Table 2. According to the results, the null hypothesis that industrial concentration and firms' density are exogenous is rejected. They are therefore treated as endogenous variables. Diversity is treated according to the exogeneity test and competition is exogenous and is treated as such.

Table 2. Endogeneity test for agglomeration variables

Model	Sargan-Hansen Statistic	Specialisation	Diversity	Competition	Firms density
TFP whole sample	Chi_sq	251.2	0.749	0.051	97.873
	p-value	0.000	0.3866	0.8215	0.000
TFP Services	Chi_sq	81.447	5.678	0.180	19.769
	p-value	0.000	0.0172	0.6718	0.000
TFP Manufacture	Chi_sq	77.415	0.838	0.051	35.671
	p-value	0.000	0.3599	0.8215	0.000

The robustness checks (shown at the bottom of Table 3) of homoscedasticity<sup>5</sup>, over-identification, weakness of instruments, spatial correlation and colinearity (see Appendix blackB) verify that estimations are reliable. Potential selection bias in the sample is also controlled for by introducing the inverse Mill's ratio, computed using the estimates from a Probit model (see Appendix blackC). It is worth noting that since there are more instrumental variables than endogenous variables ( $Z > X$ ), the GMM estimator is appropriate and more efficient than the 2SLS estimator.

### 4.1 Sectoral composition, density and local productivity

Table 3 shows the estimation results. Column (1) records the estimates of the model using the whole sample. These regression results do not account for industrial heterogeneity. Sectors might be very different in terms of productivity dynamics and backward and forward linkages. A distinction is then made between manufacturing (column 2) and service<sup>6</sup> activities (column 3)<sup>7</sup>.

At global level, productivity gains arising from sectoral diversity are important and significant on the whole sample. This result is in accordance with other studies focusing on Brazil and Chile which show that diversity is the most important driver of productivity (blackDa Silva Catela et al., black2010; blackSaito and Gopinath, black2009). In particular, the positive effect is present in the the service industry. Indeed, the existence of a large variety of activities is often beneficial for the emergence and development of many service industries both for supply and demand linkages and the collaborative environment it offers.

The tertiary sector seems to be ongoing in this economy, encouraged by agglomeration economies (blackMeliciani and Savona, black2014). Service industries have become intermediate inputs for other industries. Thus, the geographical proximity and concentration of customer industries are crucial. Besides, positive effects of diversity at the canton level might arise from knowledge spillovers across industries. As firms provide services to many industries, they can obtain information about what other

Table 3. Estimation of the effect of agglomeration and industry structure on productivity

Dependent variable	(1)	(2)	(3)
Total Factor Productivity	GMM total	Manufacturing	Services
specialisation	-0.126 (-6.02)***	-0.228 (-4.034)***	-0.16 (-5.334)***
diversity	0.353 (8.53)***	-1.141 (-2.159)**	0.378 (7.495)***
firms density	-0.009 (-0.83)	0.0515 (1.766)*	-0.00182 (-0.136)
competition	0.073 (7.15)***	0.254 (5.746)***	0.0689 (5.599)***
selection	-0.270 (-6.17)***	-0.255 (-3.950)***	-0.052 (-1.556)
Constant	8.551 (71.92)***	12.11 (10.511)***	8.422 (62.523)***
N	7803	2198	5605
F	32.01	9.444	11.1
Province specific effects	Yes	Yes	Yes
Industry specific effects	Yes	Yes	Yes
Instrumented	spe,den firms	den firms, spe, div	den firms, spe
Included instruments	comp, div, prov_i, ind_s	comp, prov_i, ind_s	comp, div, prov_i, ind_s
Excluded instruments	menace,den <sup>2</sup> ,pop90	menace, den <sup>2</sup> , pop90	menace, den <sup>2</sup> , pop90
	div_spatial compt_spatial	div_spatial, den_firms_spatial,	div_spatial, den_firms_spatial,
	den_firms_spatial	comp_spatial	comp_spatial
Endogeneitytest	265.044 (0.000)	91.79(0.000)	77.41 (0.000)
K.P(Under-id.test)	1036.6 (0.000)	29.71 (0.000)	563.48 (0.000)
Hansen J statistic	3.532	0.993	3.05
J p value	0.473	0.803	0.55

*t* statistics in parentheses

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

All variables are in logarithm except for TFP.

firms in their own field or even in other fields are doing and the process of imitation and cross-fertilisation takes place.

By contrast, negative effects of both specialisation and diversity-based externalities arise in the manufacturing industry. As recent works show, the effect of agglomeration externalities would depend on the life cycle of the industry (blackBoschma et al., black2005; blackGreunz, black2004; blackForni and Paba, black2002). In this case, it seems that the manufacturing firms in Ecuador are producing standard and mature products. In that sense, neither specialisation nor diversity externalities favour their level of productivity. This could be due to the low quality of transport infrastructure and low levels of innovation impeding to generate significant MAR and Jacobs' externalities.

Regarding the effect of competition, results confirm the existence of Porter externalities. A high level of sectoral competition at the canton level in any sector leads to productivity gains. The positive effect is particularly strong in the manufacturing sector suggesting that local competition is more strongly transformed into emulation and productivity gains in this sector.

As in the great majority of papers devoted to the estimation of agglomeration economies, the local density of firms boosts productivity (blackMelo et al., black2009). For Ecuador, the elasticity of productivity with respect to density is significant in the manufacturing sector only, where it reaches around 5.2%. While international comparisons must be interpreted very cautiously because of different methodological approaches, blackCiccone and Hall (black1996) and blackCiccone (black2002) found that economies of density in Europe and the US measured by the effects of employment density are 4.5% and 5%, respectively. The results obtained for developing countries are usually higher, ranging between 5% and 11%. In particular, according to blackMelo et al. (black2009)'s meta analysis, South American countries tend to experience around 5.6% higher density effects than Europe. Although, Ecuador achieved a 5.2% elasticity in 2010 according to our results, further investigation for recent years is needed. Other South American countries confirm higher estimated productivity elasticities to density such as Brazil with 11% (blackHenderson, black1988; blackDa-Mata et al., black2007), Chile with 0.09 (blackSaito and Gopinath, black2009) and in Colombia (blackDuranton, black2016) with 0.05 (elasticity of wages to respect to population). There may be specific mechanisms operating in these countries which induce higher levels of labour productivity, such as the demographic bonus and internal migration whose complementary nature between unskilled and skilled workers produces positive effects (blackCombes et al., black2013).

## 4.2 Does urbanisation matter?

The preceding sections allow accurately observing the existence of agglomeration benefits for industries in Ecuador's cantonal economies. However, one of the specific feature of developing countries is their strong regional heterogeneity in terms of urbanisation. Ecuador is no exception: in terms of urbanisation, the distribution of cantons is skewed (see figure 4). The second quartile is 0.33, i.e., there are 110 cantons of a total of 221 cantons with an urban population which is lower than 33%.

Given the heterogeneity of Ecuadorian cantons, it is worth investigating whether the externalities identified in the preceding sections are shaped by the level of urban-



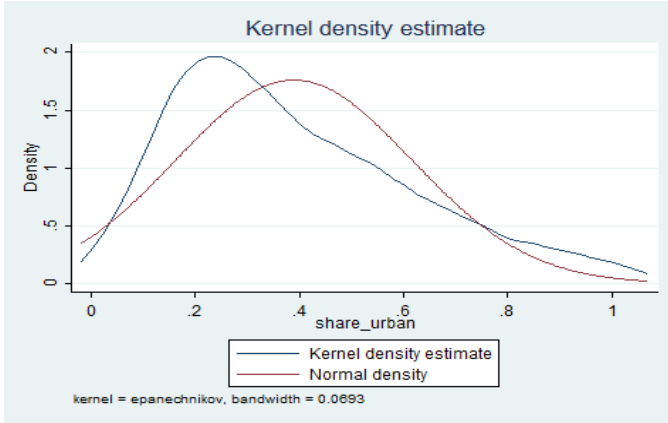


Figure 4: Histogram of urbanisation of cantons in Ecuador  
 Elaborated by authors based on Population and Housing Census of Ecuador, 2010.

isation (*urban\_rate*) of the cantons. The intuition is twofold. First, a critical mass of urban population would be required to enjoy the potential benefits of agglomeration. Secondly, one may expect the existence of an optimal city size which maximizes the economy's per capita output net of congestion losses (blackBertinelli and Black, black2004).

Through an analysis which disaggregates the entire sample into different levels of urbanisation, measured by the share of urban settlements<sup>8</sup>, light is shed on two aspects: i. the relationship between the magnitude of externalities and the level of urbanisation and ii. the threshold of urbanisation at which agglomeration externalities decline.

In order to determine the threshold values, the level of urbanisation is first set at 33%, which is the median of the sample. Some simulations have then been run changing the level of urbanisation and identify two threshold values: 46% and 61%, at which the significance of agglomeration externalities change. In Table 4, the estimates for different levels of urbanisation are presented: lower than 33% (column1); larger than 33% (column 2); larger than 46% (column 3) and larger than 61% (column 4).

Firstly, urbanisation 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 throughout the urbanisation process. In contrast, the effects of specialisation externalities measured through our industrial concentration index slightly change with levels of urbanisation.

The results in column (1) of Table 4 indicate that when a region faces low levels of urbanisation, externalities from diversity and density economies do not occur. This could be explained by the fact that regions with low levels of urbanisation do not provide an appropriate environment for the interaction of different industries since communication and transport infrastructures are still embryonic and the provision of services is scarce. By contrast, the competition is a significant source of positive externalities at low urbanised cantons. As these cantons are small, monopolist firms may not be interested in locating there. Then, knowledge and information could be diffused more easily in small cantons so that high competition can promote positive

Table 4. Agglomeration effects vs. the level of urbanisation (TFP)

Dependent variable	(1)	(2)	(3)	(4)
Total Factor Productivity	urb <33%	urb >33%	urb >46%	urb >61%
specialisation	-0.158 (-4.968)***	-0.210 (-7.660)***	-0.188 (-5.359)***	0.377 (7.322)***
diversity	0.106 (1.487)	0.607 (10.686)***	1.520 (5.119)***	6.585 (8.459)***
firms density	-0.0301 (-1.282)	0.0315 (2.141)**	0.0415 (1.938)*	-0.456 (-5.116)***
competition	0.115 (6.399)***	0.0346 (2.359)**	-0.0347 (-1.454)	-0.239 (-4.727)***
selection (Mill's ratio)	0.0207 (0.507)	0.155 (3.317)***	0.273 (4.443)***	0.658 (5.267)***
Constant	10.82 (10.908)***	8.011 (23.766)***	5.153 (5.585)***	-9.342 (-4.202)***
N	2983	4820	3524	2028
F	15.88 (0.000)	31.53 (0.000)	21.25 (0.000)	7.705 (0.000)
Province specific effects	Yes	Yes	Yes	Yes
Industry ag. specific effects	Yes	Yes	Yes	Yes
K_P Underidentification test	464.75 (0.000)	632.38 (0.000)	81.59 (0.000)	122.37 (0.000)
Hansen J statistic	3.134	4.26	1.456	1.156
J p-value	0.536	0.372	0.692	0.764

*t* statistics in parentheses

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

externalities.

On the other hand, the results suggest that externalities from industrial concentration negatively affect cantons at low and medium levels of urbanisation. The negative effect might arise from the primary sector which represent a large share of the economic activity at those levels of urbanisation.

The results in column (2) suggest that as a canton urbanises, the conditions for diversity externalities improve. A more urbanised environment can benefit from the concentration of different industries for many reasons. First, industrial dynamism is improved through the provision of services and in turn productivity is enhanced. Second, urban areas benefit from better communication and transport infrastructures, which facilitate cooperation, cross-fertilisation of knowledge between firms of different industries. In addition, an urbanised environment offers a larger variety of schools and universities which enhances the recruitment of more skilled workers and facilitates job turnover between industries.

Simulations show that at higher levels than 33%, the density effect becomes positive and significant. However, we can assume that there is a certain level of urbanisation leading to negative effects of firm density on productivity because of congestion diseconomies in big cities. In order to test this hypothesis, the model was run with several levels of urbanisation above 33% and the positive effect of density ceases to be significant at 61% urbanisation rate (fourth column). It turns to be negative and significant at high levels of urbanisation. This result shows clear evidence of congestion effects in larger agglomerations.

The externalities derived from competition are no longer positive for levels of urbanisation higher than 46%. They become significantly negative in cantons with levels of urbanisation above 61%. This negative effect means that competition may become detrimental for firm productivity. As competition heightened, land costs and input prices increase, while firms' market power decreases, inducing lower productivity for firms located in highly urbanised environment.

## 5 Conclusions

Estimating a productivity equation using the GMM model with instruments to control for endogeneity, we investigate whether the tertiarisation and the rapid but uneven urbanisation faced by developing countries would favour agglomeration economies.

The results show that diversity externalities matter more than specialisation externalities, supporting Jacobs' externalities. The strong presence of diversity externalities coming from the concentration of different industries is explained by the fact that most industries in the country interact with one another, leading to high backward and forward linkages. Therefore, the exchange and mixture of ideas from different industries is facilitated in such a diverse industrial environment.

In addition, the innovation-inducing effect of competition promulgated by Porter's theory appears to be relevant not only at the general level but also at the sectoral level. Conversely, cantons and industries do not benefit from localisation economies. This means that there is, in fact, rivalry between diversity and specialisation externalities. In terms of the economies of density, measured by firm spatial concentration, the es-

estimated elasticity of productivity with respect to density is 5% on the cantonal and industrial scale.

More importantly, these agglomeration economies are sensitive to the rising share of service industries in developing countries and to the increase in their urbanisation level. The particularities of the industries determine the specific effects of agglomeration externalities on their productivity. The service industry benefits the most from the agglomeration of different industries. There might be two reasons for that: 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 to be due to a wide sales network. The nature and therefore the magnitude of agglomeration forces would thus evolve as service industries rise. The manufacturing industry benefits more from the geographical proximity of firms and workers within the same industry. The increasing share of service industry in developing countries, boosted by agglomeration economies, enlarges in return agglomeration economies.

For the first time, this study presents the impact of the level of urbanisation upon agglomeration externalities. The results show that no significant diversity effect is recorded in cantons whose urbanisation level is under 33%. As a canton urbanises, several changes occur. The threshold at which diversity, competition and density agglomeration externalities generate positive effect was identified at an urbanisation rate of 33%. At higher rates of urbanisation than 46%, the positive effect of competition disappears, which means that in highly urbanised cantons, congestion effects would prevail.

Our results may have interesting implications for policy makers. Urbanisation should be supported as long as it increases agglomeration economies, and public infrastructures should be developed to reduce the risk of congestion effects. Moreover, as service industries develop, diversity would become a more prevailing agglomeration force. Public actors could thus accompany the tertiarisation of the economy with support for the diversification of local activities in order to strengthen this prevailing source of agglomeration economies.

It is worth noting however that additional research investigations would be requested to assess congestion effects, to observe more recent years and to take into account the dynamic nature of the urbanisation and tertiarisation processes. In addition, due to a large amount of missing values at the firm level, very few variables were available to explain the variability at the firm level, preventing to run reliable estimations on individual data. As a consequence, this study relies on an empirical strategy based on aggregated data. Even if blackMelo et al. (black2009)'s meta analysis shows that there is no difference between aggregated and individual data results, future researches should try to get individual data to grasp more precisely agglomeration economies.

## References

- Barufi, A. M. B., Haddad, E. A., and Nijkamp, P. (2015). Industrial scope of agglomeration economies in brazil. Technical Report 2015 13, University of São Paulo (FEA-USP).

- 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.
- Boschma, R., Neffke, F., and van Oort, F. (2005). Externalities and the industry life cycle. Technical report, CEFR 5th Spring School in Economic Geography.
- Chauvin, J.-P., Glaeser, E., and Tobio, K. (2014). 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. (1996). Productivity and the density of economic activity. Technical Report 1.
- 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.
- Da-Mata, D., Deichman, U., Henderson, V., Lall, S., and Guanghai, H. (2007). Determinants of city growth in brazil. *Journal of Urban Economics*, 62(2):252–272.
- Da Silva Catela, E. Y., Goncalves, F., and Porcile, G. (2010). Brazilian municipalities: Agglomeration economies and development levels in 1997 and 2007. *CEPAL Review*.
- De Groot, H. L., Poot, J., and Smit, M. J. (2015). Which agglomeration externalities matter most and why? *Journal of Economic Surveys*, 30(4):756–782.doi: 10.1111/joes.12112.
- Duranton, G. (2015). Growing through cities in developing countries. *The World Bank Research Observer*, 30(1):39–73.
- Duranton, G. (2016). Agglomeration effects in colombia. *Journal of Regional Science*, 56(2):210–238.
- ECLAC-ILO (2012). The employment situation in latin america and the caribbean. Technical Report 6, Economic Comission for Latin America and the Caribbean and International Labour Organization.

- Ercole, R. and O’neill, R. (2016). The influence of agglomeration externalities on manufacturing growth within Indonesian locations. *Growth and Change*, doi: 10.1111/grow.12145.
- Forni, M. and Paba, S. (2002). Spillovers and the growth of local industries. *The Journal of Industrial Economics*, 50:151–171.
- Frenken, K., van Oort, F. G., and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5):685–697.
- Ghani, S. E., Kerr, W. R., and Tewari, I. (2013). Specialization, diversity, and Indian manufacturing growth. Technical report, World Bank Policy Research Working Paper 6648.
- Glaeser, E. L., Kallal, H., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *The Journal of Political Economy*, 100(6):1126–1152.
- Greunz, L. (2004). Industrial structure and innovation: evidence from European regions. *Journal of Evolutionary Economics*, 14:563–592.
- Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.
- Henderson, J. (1997). Medium size cities. *Regional Science and Urban Economics*, 27(1):583–612.
- Henderson, J. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth*, 8(1):47–71.
- Henderson, J. V. (2000). The effects of urban concentration on economic growth. Technical report, National Bureau of Economic Research Working Paper 7503.
- Henderson, V. J. (1988). Urban development: Theory, fact, and illusion. *London: Oxford University Press*.
- Hernández, R., Hualde, A., Mulder, N., and Sauv e, P. (2016). Innovation and internationalization of Latin American services. Technical report, Economic Commission for Latin America and the Caribbean, United Nations, Colegio de la Frontera Norte.
- 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.
- Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–342.
- Meliciani, V. and Savona, M. (2014). The determinants of regional specialisation in business services: agglomeration economies, vertical linkages and innovation. *Journal of Economic Geography*, 15(2):387–416.
- Melo, P. S., Graham, D. J., and Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration externalities. *Regional Science and Urban Economics*, 39(3):332–342.

- Ning, L., Wang, F., and Li, J. (2016). Urban innovation, regional externalities of foreign direct investment and industrial agglomeration: Evidence from chinese cities. *Research Policy*, 45(1):830–843.
- Panne, G. v. d. and Beers, C. v. (2006). On the marshall-jacobs controversy: It takes two to tango. *Industrial and Corporate Change*, 5(5):877–890.
- Porter, M. E. (2010). *The Competitive Advantage of Nations*. The Free Press.
- Rosenthal, S. S. and Strange, W. C. (2008). *Handbook of Regional and Urban Economics*, volume 4, chapter Evidence on the nature and sources of agglomeration economies, pages 2119–2171. Henderson, J. V. and Thisse, J. F. Elsevier, 1 edition.
- Saito, H. and Gopinath, M. (2009). Plants’ self selection, agglomeration economies and regional productivity in Chile. *Journal of Economic Geography*, 9(1):539–558.
- Simões, R. F. and Freitas, E. E. (2014). Urban attributes and regional differences in productivity: Evidence from the external economics of brazilian micro-regions from 2000-2010. *Journal of Economic Financial Studies*, 2(1):27–39.
- Widodo, W., Salim, R., and Bloch, H. (2015). The effects of agglomeration economies on technical efficiency of manufacturing firms: evidence from indonesia. *Applied Economics*, 47(31):3258–3275.

## Notes

<sup>1</sup>Another strategy consists in using firm-level data. However, some variables at the firm level, needed to explain the individual variability, are generally missing. Thus, only one variable (lifetime) is available to explain the variability at the firm level. Consequently, such a strategy is not adequate for econometric reasons.

<sup>2</sup>Previous versions of the article focuses on the labor productivity. They are available upon request

<sup>3</sup>Note that diversity externalities can also be derived from the so called Related Variety, which is the interaction between sub-industries within one industry (blackFrenken et al., black2007). In the present study, evidence on the diversity externalities originated in the interchange between different industries is provided. However, estimations using Related variety measures were also conducted. Two measures were tested: i) RV at the cantonal and sectoral level based on blackFrenken et al. (black2007) and ii) RV at the cantonal and sectoral level based on the standard measure. Due to the sensitivity of the results to the measure used (Frenken vs Standard), they are not presented here but are available upon request.

<sup>4</sup>Ordinary Least Squares (OLS) and Two Stages Least Squares (2SLS) estimation methods were also used. By comparing, the 2SLS and GMM estimations show quite similar results in terms of coefficients signs and magnitudes while the OLS estimation differs regarding the sign of the competition effect.

<sup>5</sup>Enough evidence is found to reject the null hypothesis of the Pagan-Hall general test. The estimates are then presented with standard errors which are robust to arbitrary heteroskedasticity.

<sup>6</sup>The classification of industries is shown in Appendix blackA

<sup>7</sup>The differences in the coefficient estimates between both industries are statistically significant according to the Chow test. The F statistic of the test is 59.54 and the p-value is 0.000.

<sup>8</sup>According to the Population and Housing Census of 2010, urban settlements are defined by provincial and cantonal capitals under the political and administrative division of the country. Rural areas include parish capitals, other towns, peripheries and sparsely populated settlements

## A List of industries

Industries ISIC 2	classified according to	ISIC	at 2 digits	Industry name
Manufacturing industries				
C10				Manufacture of food products
C11				Manufacture of beverages
C12				Manufacture of tobacco products
C13				Manufacture of textiles
C14				Manufacture of wearing apparel
C15				Manufacture of leather and related products
C16				Manufacture of wood and cork, except furniture; articles of straw and p
C17				Manufacture of paper and paper products
C18				Printing and reproduction of recorded media
C19				Manufacture of coke and refined petroleum products
C20				Manufacture of chemicals and chemical products
C21				Manufacture of basic pharmaceutical products
C22				Manufacture of rubber and plastics products
C23				Manufacture of other non-metallic mineral products
C24				Manufacture of basic metals
C25				Manufacture of fabricated metal products
C26				Manufacture of computer, electronic and optical products
C27				Manufacture of electrical equipment
C28				Manufacture of machinery and equipment n.e.c.
C29				Manufacture of motor vehicles, trailers
C30				Manufacture of other transport equipment
C31				Manufacture of furniture
C32				Other manufacturing
F41				Construction of buildings
F42				Civil engineering
F43				Specialized construction activities
Service industries				
C33				Repair and installation of machinery and equipment
D35				Electricity, gas, steam and air conditioning suppl
E36				Water collection, treatment and supply
E37				Sewerage
E38				Waste collection, treatment and disposal activities
E39				Remediation activities
G45				Wholesale and retail trade and repair of motor vehicles
G46				Wholesale trade, except of motor vehicles and motorcycles
G47				Retail trade, except of motor vehicles and motorcycles
H49				Land transport and transport via pipelines



## B Robustness checks

Regarding the test of under-identification presented at the bottom of Table 3, the Kleibergen-Paap rk Wald statistic shows that both equations are identified. According to the Hansen's J test, the instruments are exogenous. The test of Kleibergen-Paap rk Wald F statistic equals to 17.08 is larger than Stock-Yogo weak ID test critical values for K1=endogenous variables and L1=excluded instruments (12.20). It shows that the instruments are not weak at 5% of maximal bias of the IV estimator relative to OLS.

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

In order to detect whether the estimated models present spatial correlation, a test of the presence of spatial correlation in the residuals for each industry (78 industries) is implemented. The results are presented in Tables black?? and black??. According to the Moran's I test, the null hypothesis of zero spatial correlation cannot be rejected. Then, the residuals of the estimations do not present spatial autocorrelation.

Industries ISIC 2	classified	according	to	ISIC	at	2	digits
		Industry name					
Services industries							
H50				Water transport			
H51				Air transport			
H52				Warehousing and support activities for transportation			
H53				Postal and courier activities			
I55				Accommodation			
I56				Food and beverage service activities			
J58				Publishing activities			
J59				Video and television production, sound recording and music			
J60				Programming and broadcasting activities			
J61				Telecommunications			
J62				Computer programming, consultancy			
J63				Information service activities			
K64				Financial service activities, except insurance and pension			
K65				Insurance, pension funding, except compulsory social security			
K66				Activities auxiliary to financial service and insurance activities			
L68				Real estate activities			
M69				Legal and accounting activities			
M70				Activities of head offices; management consultancy activities			
M71				Architectural and engineering activities			
M72				Scientific research and development			
M73				Advertising and market research			
M74				Other professional, scientific and technical activities			
M75				Veterinary activities			
N77				Rental and leasing activities			
N78				Employment activities			
N79				Travel agency, tour operator, reservation service			
N80				Security and investigation activities			
N81				Services to buildings and landscape activities			
N82				Office administrative, office support			
O84				Public administration and defence; social security			
P85				Education			
Q86				Human health activities			
Q87				Residential care activities			
Q88				Social work activities without accommodation			
R90				Creative, arts and entertainment activities			
R91				Libraries, archives, museums and other cultural activities			
R92				Gambling and betting activities			
R93				Sports activities and amusement and recreation			
S94				Activities of membership organizations			
S95				Repair of computers and personal and household goods			
S96				Other personal service activities			
U99				Activities of extraterritorial organizations and bodies			

		Model	
		Variable	VIF
Collinearity diagnostics		competition	1.82
		specialisation	1.07
		diversity	1.28
		firms density	1.76

Measure of global spatial correlation					
	ISIC two digit Sectors	I	E(I)	sd(I)	z
p-value					
C10	-0.006	-0.005	0.983	-0.002	0.499
C11	0.251	-0.016	0.152	1.751	0.04
C12	-0.775	-0.5	0.099	-0.725	0.234
C13	0.168	-0.011	0.067	2.643	0.004
C14	-0.471	-0.005	0.973	-0.479	0.316
C15	0.104	-0.012	0.072	1.611	0.054
C16	-0.17	-0.005	0.985	-0.167	0.433
C17	-0.084	-0.037	0.074	-0.628	0.265
C18	0.051	-0.01	0.065	0.931	0.176
C19	-0.131	-0.143	0.364	0.032	0.487
C20	0.01	-0.026	0.286	0.126	0.45
C21	0.627	-0.167	0.405	1.959	0.025
C22	0.096	-0.02	0.195	0.596	0.275
C23	-0.018	-0.007	0.982	-0.011	0.496
C24	0.204	-0.029	0.966	0.241	0.405
C25	-0.076	-0.005	0.988	-0.072	0.471
C26	0.108	-0.071	0.112	1.605	0.054
C27	-0.019	-0.042	0.068	0.334	0.369
C28	-0.076	-0.022	0.201	-0.27	0.394
C29	0.029	-0.016	0.122	0.371	0.355
C30	-0.03	-0.053	0.111	0.203	0.419
C31	0.016	-0.005	0.986	0.022	0.491
C32	0.053	-0.01	0.063	0.99	0.161
C33	0.003	-0.009	0.05	0.229	0.409
D35	0.017	-0.009	0.986	0.022	0.491
E36	-0.146	-0.014	0.102	-1.298	0.097
E37	-0.146	-0.014	0.102	-1.298	0.097
E38	-0.194	-0.036	0.042	-3.722	0.16
E39	-0.237	-0.167	0.112	-0.624	0.266
F41	0.07	-0.018	0.035	2.508	0.006
F42	0.046	-0.023	0.243	0.286	0.387
F43	0.131	-0.016	0.127	1.155	0.124
G45	-2.265	-0.005	0.986	-2.292	0.011
G46	-0.251	-0.006	0.988	-0.249	0.402
G47	0.138	-0.005	0.988	0.144	0.443
H49	0.035	-0.006	0.034	1.099	0.136
H50	-0.172	-0.1	0.099	-0.725	0.234
H51	0.079	-0.037	0.067	1.743	0.041

Measure of global spatial correlation

	ISIC two digit Sectors	I	E(I)	sd(I)	z
p-value					
H52	-0.004	-0.011	0.072	0.103	0.459
H53	-0.024	-0.011	0.076	-0.166	0.434
I55	0.492	-0.005	0.989	0.502	0.308
I56	-0.064	-0.005	0.99	-0.06	0.476
J58	0.068	-0.017	0.034	2.508	0.006
J59	0.042	-0.023	0.061	1.063	0.144
J60	-0.236	-0.007	0.991	-0.231	0.409
J61	0.22	-0.005	0.985	0.228	0.41
J62	0.071	-0.042	0.093	1.207	0.114
J63	0.035	-0.05	0.077	1.099	0.136
K64	0.604	-0.005	0.987	0.618	0.268
K65	-0.11	-0.029	0.068	-1.203	0.114
K66	-0.018	-0.037	0.052	0.356	0.361
L68	-0.021	-0.01	0.063	-0.176	0.43
M69	-1.099	-0.005	0.978	-1.118	0.132
M70	-0.012	-0.02	0.028	0.257	0.399
M71	-0.029	-0.011	0.065	-0.288	0.387
M72	-0.067	-0.045	0.09	-0.235	0.407
M73	-0.015	-0.011	0.065	-0.072	0.471
M74	-0.25	-0.006	0.947	-0.258	0.398
M75	-0.02	-0.013	0.026	-0.278	0.391
N77	-0.001	-0.008	0.044	0.157	0.438
N78	0.082	-0.091	0.254	0.678	0.249
N79	-0.025	-0.018	0.062	-0.107	0.457
N80	-0.013	-0.026	0.291	0.042	0.483
N81	-0.1	-0.031	0.065	-1.065	0.143
N82	0.044	-0.006	0.951	0.053	0.479
O84	0.036	-0.005	0.974	0.041	0.483
P85	-0.205	-0.005	0.984	-0.204	0.419
Q86	0.012	-0.005	0.025	0.702	0.241
Q87	-0.066	-0.01	0.965	-0.058	0.477
Q88	0.89	-0.007	0.941	0.953	0.17
R90	-0.051	-0.018	0.147	-0.227	0.41
R91	0.036	-0.01	0.032	1.436	0.076
R92	0.476	-0.005	0.987	0.487	0.313
R93	-0.06	-0.005	0.987	-0.056	0.478
S94	0.01	-0.005	0.026	0.563	0.287
S95	-0.002	-0.005	0.983	0.003	0.499
S96	0.12	-0.005	0.988	0.126	0.45
U99	-0.327	-0.143	0.104	-1.773	0.038

## C 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 the database. To do so, the Heckman's method (Heckman,1979) is applied. 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, then:

$$(1a) Y_{1i} = X_{1i}\beta_1 + U_{1i}$$

$$(1b) Y_{2i} = X_{2i}\beta_2 + U_{2i}$$

$$E(U_{ji}, U_{j'i''}) = \sigma_{jj'}, i = i''$$

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

$$\begin{aligned} E(U_{1i} | X_{1i}, \text{sample selection rule}) &= E(U_{1i} | X_{1i}, Y_{2i} \geq 0) \\ &= E(U_{1i} | X_{1i}, U_{2i} \geq -X_{2i}\beta_2) \end{aligned}$$

Assuming that  $h(U_{1i}, U_{2i})$  is a bivariate normal density:

$$E(U_{1i} | U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{12}}{(\sigma_{22})^{1/2}} \lambda_i$$

$$E(U_{2i} | U_{2i} \geq -X_{2i}\beta_2) = \frac{\sigma_{22}}{(\sigma_{22})^{1/2}} \lambda_i$$

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

and  $\phi$  and  $\Phi$  are the density and distribution function for a normal standard variable, respectively.  $\lambda_i$  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  $\lambda_i$  using a probit analysis for the whole sample. In the 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. According to location theories, the firms' location would be guided by factors such as transport costs, the availability of labour, raw materials and agglomeration economies and diseconomies (Combes et al. 2010). This is the case of this model, the explanatory variables are the level of urbanisation in a given canton, the inverse distance to main cities (Quito, Guayaquil and Cuenca) and industry dummies. Variables such as the cantonal income or variables related to agglomeration economies cannot be introduced in this selection model since they are already taken into account in the original model (7) and could produce high multicollinearity. In Table black1, the results of the Probit model show that the probability that an industry operates in a given canton positively depends on the level of urbanisation and the industrial characteristics. The proximity to main cities, Quito in particular, also determines the presence of industries in a canton. The correct prediction rate is 82.62%. Using the resulting estimates,  $\lambda_i$  can be computed. Then, such a value is included in the main equation (7) that analyses the effects of agglomeration externalities in productivity.

Table 1: Probit model, presence of zeros

	(1) Probit
urban rate	2.39 (39.46)***
Quito	5.07 (6.02)***
Guayaquil	1.15 (0.88)
Cuenca	-0.61 (-1.51)
industry dummies	significant
Constant	-2.97 (-17.29)***
Observations	17238

*z* statistics in parentheses

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