

What drives the spatial wage premium in formal and informal labour markets?

The case of Ecuador

Matano Alessia¹, Obaco A. Moisés², Royuela Vicente³

Abstract:

This article investigates the incidence of agglomeration externalities in Ecuador, a small-sized, middle-income developing country. In particular, we analyze the role of the informal sector within these relations, since informal employment accounts for a significant part of total employment in the developing countries. Using individual level data and instrumental variable techniques, we investigate the impact of spatial externalities, in terms of population density, local specialization and urban size, on the wages of workers in Ecuadorian cities. The results show that spatial externalities matter also for a small developing country. Moreover, analysis of the interaction between spatial externalities and informality shows that, on average, workers employed in the informal sector do not enjoy significant benefits from agglomeration externalities. Finally, by investigating the possible channels behind spatial agglomeration gains we show that the advantages from agglomeration for formal sector workers may well be accounted for by better job-quality matches and, to a lesser extent, by learning externalities. For informal sector workers, our findings also suggest possible gains from job changes, which offset a penalty for remaining employed in the same occupation.

Keywords: Agglomeration Externalities; Developing Economies; Informal Employment; Workers' Wages; Functional Urban Areas; Ecuador.

JEL classification: J31, J46, R23, R12.

Acknowledgments: A. Matano and V. Royuela express gratitude for the support received from Spain's Ministerio de Ciencia e Innovación through action ECO2016-75805-R. M. Obaco acknowledges the support received by the AGAUR. We thank the participants in the ERSA Conference (Groningen, 2017), Geoffrey J.D. Hewings Regional Economics Workshop (Wien, 2017), XLIII International Conference on Regional Science (Seville, 2017), 8th Slovak Winter Seminar of Regional Science (Terchova, 2018), University of Groningen Seminar (2018), ESPE Conference (Antwerp, 2018), EALE Conference (Lyon, 2018) and SAEs conference (Madrid, 2018). Our thanks are also due to Jorge de La Roca and Jordi Jofre-Monseny for very useful suggestions.

¹ Corresponding author. AQR-IREA, University of Barcelona, Barcelona, Spain. Email: amatano@ub.edu.

² GIER (Reserach Group in Regional Economics), University of Cuenca, Ecuador and AQR-IREA, University of Barcelona, Barcelona, Spain. Email: moisesl.obaco@ucuenca.edu.ec.

³ AQR-IREA, University of Barcelona, Barcelona, Spain. Email: vroyuela@ub.edu.

1. Introduction

With the growth of big cities, analysis of the benefits of agglomeration economies has been pursued in a wide range of studies. From the empirical point of view, an extensive literature has analyzed the extent of agglomeration economies, measured by density (or population size) and industrial specialization at the local level. Their results have shown positive impacts of spatial externalities on productivity and wages (see among others Combes, 2000; Combes et al., 2008; Mion and Naticchioni, 2009).

These studies have generally been carried out on developed countries. Less attention has been paid to the role of agglomeration economies in the developing world. However, the topic is relevant since the growth rate of the world's urban development is being driven by urbanization in the developing world. Exploring the role of agglomeration economies in the developing countries is relevant to assessing the importance of urban economies worldwide, since the urbanization process taking place in the developing world is different from the old urbanization, mainly because of high poverty rates and poor-quality institutions (Glaeser and Henderson, 2017).

A few studies have looked into the importance of agglomeration economies in the developing countries, focusing on those characterized by great geographical extension and large populations, such as China, India, Brazil or Colombia (see Chauvin et al., 2017; Combes et al., 2015 and Duranton, 2016). The findings have shown an important role for spatial externalities in fostering productivity and wages, with impacts higher than detected for the developed world. However, these results have been found for emerging market economies, while there is no evidence of studies focusing on other types of developing countries. Moreover, the developing countries are generally characterized by a large proportion of informal economy, which leads to a dual labour market (Fields, 1990; La Porta and Shleifer, 2014). Although the theoretical implications of the presence of informal employment on agglomeration externalities are unclear, it is hardly justified to assume that agglomeration economies are non-existent in the informal sector (Overman and Venables, 2005). Moreover, as Duranton (2009) points out, formal and informal sectors have strong interconnections, thus suggesting that in both sectors agglomeration effects generate benefits. Nonetheless, more formal studies on this topic are lacking (Overman and Venables, 2005).

The aim of this paper is twofold. First of all, we analyze the importance of agglomeration economies for a small-sized, middle-income developing economy, namely that of Ecuador. This country is shaped by a set of characteristics still unexplored in the literature such as limited geographical extension and population size, medium per-capita income, low and increasing urbanization rate, weak industrial activity, and widespread informal employment. Our first goal is to understand whether agglomeration externalities exert impacts similar to those detected for the developed world and/or the emerging economies. Secondly, we analyze the importance of the informal sector within these relations, exploring the heterogeneity of spatial externalities between formal and informal sector workers. Also, we shed light on the channels behind the detected impacts. Our paper contributes to the literature by directly and comprehensively analyzing how spatial externalities affect formal and informal workers in a small developing country.

We use repeated quarterly cross-section data from the Ecuadorean Labour Surveys (ENEMDU) available from 2007 to 2017, together with historical data from the 1950 and 1990 censuses. We take Functional Urban Areas as units of analysis, since they represent a suitable economic definition of cities (OECD, 2013). We assess the extent of spatial externalities by analyzing the impact of two main measures of agglomeration on individual workers' real wages: population density and sector specialization at the local level. We also take into account the area covered by the Functional Urban Areas, which allow for assessment of the impact of urban size for any given level of density.

Two major methodological issues arise in estimating this relationship. First, there could be sorting of workers should the more skilled individuals prove more likely to be located in highly agglomerated areas. The best methodological approach would be to employ fixed effects estimates to control for unobserved individual heterogeneity (Combes et al., 2008; Mion and Naticchioni, 2009). However, in the context of developing countries individual panel data are generally not available (as in the case of Chauvin et al., 2017; Combes et al., 2015; Duranton, 2016). This is also the case of Ecuador. Nonetheless, we will do our best from a methodological point of view to address this issue. The second methodological concern is the endogeneity due to possible simultaneity in individual choices concerning wages and locations. We will take this issue into account by applying an instrumental variable strategy, using deeply lagged values of our agglomeration measures to build our instruments as in Combes et al.

(2008), Combes et al. (2010), Matano and Naticchioni (2012), and Mion and Naticchioni (2009).

Our results lead us to the following findings. First, agglomeration externalities increase productivity and wages in a small developing country like Ecuador. In particular, with regard to urbanization externalities, we find an elasticity estimate for density of 7.1%, whose impact is partly due to a pure size effect. These findings are in line with those found for emerging economies (Chauvin et al., 2017; Combes et al., 2015; Duranton, 2016). As for local industrial specialization, there is also a positive impact, but not precisely estimated and lower than that attributed to urbanization externalities (1.6%).

When we consider the informal sector in our analysis, our findings confirm that it is a key factor since on average workers employed in the informal sector do not enjoy significant benefits from agglomeration externalities. In particular, urbanization externalities exert a positive impact on wages, but lower than for formal workers, and not precisely estimated. As for local sectoral externalities, wages seem to decrease with the increase in the specialization level of the area, even though the effects are not significant.

Finally, we further characterize the results of the analysis, by investigating the channels behind the detected impacts of spatial externalities on wages. We exploit the information derived from a panel subsample of our original data, and look into the matching and learning mechanisms (Combes and Gobillon, 2015; Puga, 2010). For formal sector workers the results show that the wage premium due to spatial externalities is mostly driven by better quality job matches. Nonetheless, in larger cities learning externalities also play an important role. As for informal sector workers, the findings are generally not significant. Nonetheless, they suggest a wage premium through job change in larger cities, while remaining employed in the same occupation entails a wage penalty.

To conclude, the analysis reveals the relevance of taking into account the duality of the labour market in the developing countries when evaluating the impacts of spatial externalities. In the case of a small developing country such as Ecuador, working in the formal sector appears to be crucial to enjoy the high and significant benefits from agglomeration economies. From a policy perspective, these findings suggest that policies

designed to incentivize the formalization of firms might have a direct impact on productivity and workers' salaries as well as represent a key channel to obtain benefits from agglomeration externalities.

The paper is structured as follows. Section 2 introduces the reference literature on spatial agglomeration in developed and developing countries. Section 3 describes the case study, presents the data and defines the main variables used for the empirical analysis. Section 4 illustrates the methodology, describes the empirical analysis and sets out the results, while Section 5 concludes.

2. Related Literature

The idea of agglomeration economies fostering productivity and wages has been widely investigated in both the theoretical and empirical literature. Two of the factors most frequently analyzed are urban agglomeration and local sectoral specialization.

Marshall (1890) was a pioneer in pointing out the productivity gains that may arise in bigger cities or from the concentration of a specific industry in a given location. The channels through which this occurs, formalized by Duranton and Puga (2004), are: the learning mechanism, which reflects the idea of knowledge spillovers and face-to-face interactions in agglomerated areas that enhance human capital; the matching mechanism, showing that agglomerated areas offer conditions for a better match between workers and firms; and the sharing mechanism, lying in the advantages generated by sharing indivisible goods such as facilities and risk in new investments, thereby cutting individual and firm costs. These mechanisms lie behind both urbanization and specialization positive spillovers, the former being associated with highly dense areas, and therefore across industries, while the latter occur within specific industries located in the same area.

From an empirical point of view, many works have analyzed the role of spatial externalities in fostering productivity and wages, using both aggregated data (see Ciccone and Hall, 1996; Ciccone, 2002; Combes, 2000; among others) and individual level data (see Combes et al., 2008; Combes et al., 2010; De la Roca and Puga, 2017; Glaeser and Marè, 2001; Mion and Naticchioni, 2009, Matano and Naticchioni, 2012, 2016). These studies take into account the empirical issues that arise in identification of the role of such externalities. In particular, they address the endogeneity of the relationship by using an instrumental variable strategy, while the role of the sorting of skilled workers

in highly agglomerated areas is examined by means of a fixed effect strategy when using individual level panel data. The results have shown that worker sorting captures a large part of the impact imputed to spatial externalities on wages. Further, research has also revealed the relevance of dynamic gains in the biggest cities (De la Roca and Puga, 2017).

However, these studies have in general been carried out on the developed countries.⁴ Less attention has been paid to the role of spatial externalities in affecting wages in the developing countries. Nevertheless, most of the urban population growth is driven by the growth of cities in the developing countries and, interestingly, not in the largest cities (Royuela and Castells-Quintana, 2015). Thus, in line with Glaeser and Henderson (2017), we deem it important to study whether agglomeration economies in the developing countries exert the same impacts on wages and productivity as in the developed world. Few works have addressed this issue. Chauvin et al. (2017) use cross-sectional individual level data and IV estimates to analyze the difference in urbanization economies between similar-sized cities in the US, China, India and Brazil from 1980 to 2010. They show higher elasticities of wages with respect to urban density for China and India (around 17% and 8% respectively) and lower for Brazil (2.6%), which is close to the US estimate (4.3%). Combes et al. (2015) focus on China and find agglomeration impacts on native wages within a range of 9%-12%.⁵ Duranton (2016), taking the case of Colombia, finds an elasticity of wages with respect to population size of 5%, while Ahrend et al. (2017) find a similar magnitude in the case of Mexico (4.2%).⁶

In contrast to the works analyzing developed countries, these studies are unable to control for unobserved individual heterogeneity due to the lack of a longitudinal data structure. Nonetheless, they make use of a wide range of individual level controls to take into account individual heterogeneity as far as possible (see Combes and Gobillon, 2015, for a discussion). Furthermore, a major concern in the case of the developing countries

⁴ For an exhaustive review see Rosenthal and Strange (2004), Melo et al. (2009) and Combes and Gobillon (2015).

⁵ They also examine the impact of sectoral specialization at the local level and find elasticities of wages with respect to specialization of around 6%.

⁶ Note that these studies are not strictly comparable with each other, nor indeed with ours, due to the use of different spatial variables in each of the cited articles. For instance, Chauvin et al. (2017) investigate the wage impact of both density and population (or density and share of educated workers), while Duranton (2016) investigates the impact of density, market access and share of educated workers, controlling for the city area. Moreover, other related studies are Au and Henderson (2006), who analyze the case of China showing an inverted U relationship between wages and city size, and Lall et al. (2004) who, using firm-level manufacturing data, look into the role of spatial externalities for economic productivity in India, disentangling the sources of agglomeration economies between those arising at the firm level, at the industry level and at the regional level.

lies in the informal economy, which implies the presence of a dual labour market, informal workers representing around half of the labour force (La Porta and Shleifer, 2014; Maloney, 2004; Overman and Venables, 2005). Informal labour is usually represented by low-educated workers (Fields, 1990; Khamis, 2012) and associated with high levels of vulnerability and poverty, low wages and productivity, and lack of legal protection (Bacchetta et al., 2009; Fields, 1990). Consequently, in our view, a thorough analysis of the relevance of spatial externalities in a developing economy should take this dimension into account (Glaeser and Henderson, 2017).

From the theoretical point of view, the literature has yet to reach consensus on the relationship between agglomeration externalities and the informal economy. Overman and Venables (2005) discuss this issue, arguing that the informal economy might be expected to reduce the gains from agglomeration economies, in line with the rationale of the Harris-Todaro (1970) model. In particular, they stress that, according to this model, the urban real wage is higher than real earnings in agriculture because of institutional rigidities or efficiency wage considerations. This urban wage premium in turn attracts to the city people who seek “formal sector” jobs. Migrants who fail to enter the formal job sector remain unemployed or work for a significantly lower wage in the urban “informal sector”. This leads to the burgeoning of a mass of low-wage and low-productivity urban labour,⁷ which represents a cost of urbanization, in terms of congestion and higher rents. As a consequence, Overman and Venables (2005) argue that the presence of a large pool of unemployed/underemployed workers could reduce the benefits from agglomeration. However, they point out that these effects may not be fast enough to offset the advantages deriving from agglomeration.⁸ Overman and Venables (2005) also point out that the informal sector itself could contribute to agglomeration economies because of the vitality of this sector, the existence of networks of small firms benefitting from labour market pooling and the role of the informal economy shown in the literature on clusters in developing countries. This spillover effect is also stressed by Duranton (2009), who argues that the formal and informal sectors are closely interconnected, so positive effects could arise within both sectors.

⁷ A similar argument is raised by Chauduri and Mukhopadhyay (2010), who stress that within dual economy models with an urban informal sector, the latter is viewed as a residual sector that absorbs all workers who did not succeed in entering the formal sector, thus offering a very low competitive wage (less than the rural or formal sector wage rates).

⁸ This is more likely to occur for formal sector workers than for informal sector workers, the latter being characterized by significantly lower wages.

From the empirical point of view, the role of the informal economy within the relation between spatial externalities and wages has generally been neglected. One exception is Duranton (2016), who uses individual level data for Colombia to carry out an indirect test performing estimations of the spatial externalities' impact on wages vis-à-vis, on the one hand, all workers, and on the other hand only workers with a written labour contract. His findings show that spatial externalities have a lower impact when only workers with written contracts are taken into account (around 3.7% compared to 5% for all workers), thus inferring higher benefits due to agglomeration economies for informal workers.⁹

3. Description of the data and definition of variables of interest

In this paper we focus on Ecuador, a small-sized, middle-income developing country. The total population size (around 16 million inhabitants) and GNP per-capita (6,000 US\$ in 2015) are lower than the average of the developing countries.¹⁰ Also, Ecuador is ranked among the countries with the lowest economic complexity indexes¹¹ and is characterized by low industrialization and urbanization rates as compared with other Latin America countries. Further, it is included in the group of countries that have gone through rapid urbanization since 1960.¹²

We use data from the Ecuadorean Labour Surveys (*Encuesta Nacional de Empleo, Desempleo y Subempleo - ENEMDU*) provided by the National Statistics Institute of Ecuador (INEC). These are quarterly surveys (repeated cross-section data) that contain

⁹ Other related papers are by Garcia (2019) who, like Duranton (2016), uses individual level data to study the case of Colombia, finding positive agglomeration benefits (in terms of density) for the informal economy and negative ones for the formal economy. However, he does not analyze the channels behind the detected impact, which is one of the purposes of this paper. Also, Harris (2014), using firm level data for Nairobi's handicraft industry, finds disadvantages in terms of agglomeration externalities in relation to the informal economy. Further, Bernedo Del Caprio and Patrick (2017) use firm level data for Peru to show lower benefits from agglomeration externalities for the informal economy.

¹⁰ The average population size for the developing countries is around 56 million inhabitants, while the GNP per-capita was 7,346US\$ in 2015. The list of developing economies is from the *World Economic Situation and Prospects 2014*, UN, Statistical Annex. Furthermore, Ecuador also has a lower GNP per capita than the average of the other Latin American countries (7,185 US\$ in 2015), while emerging economies such as Colombia and Brazil show values close to or above the average (respectively 7,130 US\$ and 10,100 US\$). Nonetheless, Ecuador is included in the group of upper middle-income countries since its GNP per-capita is higher than 3,895 US\$ (World Bank).

¹¹ Ecuador has the third lowest level of the Economic Complexity Index in Latin America, after Nicaragua and Bolivia (Observatory of Economic Complexity, MIT: https://atlas.media.mit.edu/en/rankings/country/eci/?year_range=2011-2016).

¹² Throughout the period of the analysis (2007-2017) the urbanization process is characterized by faster population growth for small and median cities, with a decreasing tendency over time (Obaco and Díaz, 2018, and Royuela and Ordóñez, 2018).

detailed information on Ecuadorean workers: labour income, hours worked, age, gender, ethnicity, education, occupation, previous migrant status and informal employment status. We also have information on the workers' area/city of residence and sector of employment. To construct our instruments for analysis we combine these data with data from the population censuses of 1950 and 1990.¹³

As spatial units of analysis, we use an international harmonized concept of economic cities, namely Functional Urban Areas (FUAs hereafter)¹⁴ as in Ahrend et al. (2017). We use the 28 FUAs in Ecuador identified in Obaco et al. (2019) (see Figure A1, in the Appendix).¹⁵ The most populated FUAs are the capital city, Quito, and Guayaquil, with more than 1.5 million inhabitants each, located in the Andean and Coastal region respectively.¹⁶

The period of the empirical analysis is from 2007 to 2017.¹⁷ We restrict the sample to workers located in the FUA areas (males and females) aged between 15 and 64 years old who perform one job.¹⁸ We exclude workers employed in the public sector since their wage is set nationally, and focus on employees and self-employed workers. We clean the dataset by dropping observations with missing data in our variables of interest as well as observations below (above) the 1st (99th) percentile of the workers' real wage distribution, hours worked and real wages per hour.¹⁹ We end up with a sample of 377,273 observations for the period of analysis – around 8,800 per quarter.

Informality is defined on the basis of the latest methodology implemented by the INEC (following the 2013 ILO guidelines): the informal sector includes workers

¹³ 1950 saw the first population census. The data are from the National Institute of Statistics and Census of Ecuador (INEC).

¹⁴ Throughout the paper we will use “Functional Urban Area” and “city” as synonymous.

¹⁵ The Functional Urban Areas are defined according to the OECD methodology. First urban cores are identified, using satellite images, based on continuous grid cells of high population density (minimum 500 inhabitants per km²) and total population size of at least 25,000 inhabitants. Second, a travel-time approach is considered to build polycentric FUAs and define their hinterlands' boundaries. The final FUAs represent both the urban cores and the connected municipalities forming the hinterland.

¹⁶ Quito and Guayaquil are the largest metropolitan FUAs. The rest of the Ecuadorian FUAs are medium-sized and small, with a substantial gap in terms of population with respect to the two biggest ones (Cuenca, the third FUA in the country, has around 500 thousand inhabitants).

¹⁷ This time span avoids the years of the previous economic crises (the peak of the great recession was in 1999) and dollarization process (introduced as national currency of Ecuador in January 2000).

¹⁸ We drop workers with more than one job (5% of total sample).

¹⁹ We deflate wages by using the Ecuadorian national CPI. The base year is 2014. CPI information is provided by the Ecuadorian Central Bank. We also drop singleton group observations in terms of city-sector-year-quarter since they cannot be used in the empirical analysis (less than 2% of sample observations).

employed in firms with fewer than 100 employees having no tax identification number (*Registro unico del contribuyente*).²⁰ We also employ an alternative measure for informality associated with the worker status rather than with the firm status, taking informal workers to be those with no access to the social security system. In this way we are able to assess how the results change when taking into account informal workers employed in the formal sector.

We use the following measures of agglomeration effects. First, to proxy urbanization externalities we use the FUA population density ($Dens_{c,t}$). This information is provided by the INEC. Moreover, we also control for the land area of the FUA as suggested in Combes and Gobillon (2015), since it provides additional information on urbanization externalities. In particular, the impact of density, keeping the FUA area constant, reflects the benefits arising from increasing the number of people in the city, while the impact of the FUA area, keeping density constant, reflects the benefits from enlarging the size of the city (i.e. from proportionally increasing city area and population).

The other measure of agglomeration externalities is a specialization index, which proxies sector specialization at the local level. We compute it as in Combes (2000), Matano and Naticchioni (2012) and Mion and Naticchioni (2009):

$$Spec_{c,s,t} = \frac{empl_{c,s,t} / empl_{c,t}}{empl_{s,t} / empl_t}$$

where c stands for the city, s for the sector, and t for time. The specialization index is the ratio between the share of sectoral employment in the total employment in any city c and the corresponding share at the national level.²¹

Table 1 presents the average of the real wages per hour (in US dollars) for the workers' categories considered in the empirical analysis, at the beginning, middle and end of the

²⁰ The INEC has applied different definitions of informality over time. The previous definition was classifying workers employed in the informal sector as those employed in firms with 10 or fewer employees with no full accounting records or tax identification number. The current definition (INEC, 2015) follows the 2013 ILO guidelines and includes in the formal sector firms having tax identification numbers, even though they may not have full accounting records, in order to take into account cases where there is some legal justification. The classification divides workers into formal and informal sector workers, homeworkers and non-classified workers. Formal and informal workers account for 90% of our sample, while 4% are homeworkers and 6% are non-classified workers. In the paper we only consider formal and informal sector workers.

²¹ Sectors are defined according to the ISIC version 3.1 at two-digit level. We use the ENEMDU surveys to compute the specialization index.

period under analysis, as well as for the entire time span of the analysis (2007-2017) combined with their relative presence in the sample.

Turning to the composition of the sample (Table 1, last column), it may be noted that males constitute around 62% of the labour force. The dominant ethnicity is mestizo (89%). Around 75% of the sample is constituted by low/medium-educated workers (educational levels are up to high school at most), while about 47% of the workers are employed in unskilled occupations, 38% in medium-skilled, and 15% in high-skilled ones. In terms of sector distribution, a relatively high percentage of workers (30%) are employed in the wholesale and retail trade sector and, in general, in the service sector (in total around 65% of workers); manufacturing, mining and agriculture account for just 26% of the total sample of workers (16% of which in manufacturing), construction for 9%. As for the dual labour market, workers in the informal sector represent around 37% of the total workforce in the sample. Also, their relative presence has been declining over time (from 40% in 2007 to 36% in 2017). Moreover, 43% of workers are self-employed, while 57% are employees. Considering the regions, 54% of the workers are located in the Andean region, 42% in the Coastal region, and the remaining 4% in the Amazonian region. Finally, 32% of the workers declare they have not always lived in the city where they currently reside.

As for wages (Table 1, columns 1 to 4), a general increase in real wages per worked hour throughout the period analyzed may be noted. In addition, there is evidence of marked heterogeneity according to the dimensions considered: males earn around 9% more than females; white workers have a significant premium compared to the other ethnicities (11% above the average); wages increase with the worker's level of education and job skill intensity, as well as previous migrant status of the worker. In terms of the dual labour market, wages are significantly higher (about 61% more) for formal than for informal sector workers. Further, wages are lower for the self-employed than for the employees. As for the economic sectors, apart from the mining and quarrying sector, wages are higher in the service sector than in manufacturing. In terms of regions, wages are higher in the Andean region followed by the Amazonian region (due to oil extraction).

[Table 1 around here]

4. Empirical Analysis

4.1. The impact of spatial externalities on average wages

In this section, we use data on repeated cross-sections for Ecuador from 2007 to 2017 to estimate a pool cross-section model. We carry out a two-step methodology similar to that of Combes et al. (2008). In the first step, wages are regressed over a set of individual observable characteristics and a set of city-sector-time dummies. In the second step the time-averages of the estimated city-sector-time fixed effects are regressed over the (time-average of the) spatial variables. More specifically, the first step is:

$$\log(W_{i,t}) = \alpha + X_{i,t}' * B + A_{c(i)}' * \Phi + CST_{c(i),s(i),t} + \varepsilon_{i,t} \quad (1)$$

where $\log(W_{i,t})$ is the log of real wage per hour worked by worker i at time t . $X_{i,t}$ is the vector of worker's characteristics including: age, age squared, years of education, female dummy, migrant dummy (indicating whether the worker has previously lived somewhere else), five categories for ethnicities, twenty-seven categories for occupation (ISCO 88 at two-digit level), and job category (self-employed/employee) dummies. $CST_{c(i),s(i),t}$ are the city-sector-time²² dummies, while $\varepsilon_{i,t}$ is the error term. From this first step regression we retrieve the estimate of the city-sector-time fixed effects and we compute weighted time-averages for each city-sector cell, using as weights the number of observations in each cell. The obtained time-averaged fixed effects are then used as dependent variable in the second step estimation.²³ The specification for the second step is:

$$\widehat{CS}_{c,s} = \alpha + \delta * \log(Dens_c) + \theta * \log(Spec_{c,s}) + \gamma * \log(Area_c) + A_c' * \Phi + \nu_{c,s} \quad (2)$$

where $\widehat{CS}_{c,s}$ is the vector of the time-average values of the city-sector-time fixed effects retrieved from equation (1), $\log(Dens_c)$ is the log of the time-average population density of city c , $\log(Spec_{c,s})$ is the log of the time-average specialization index for city c and sector s , while $\log(Area_c)$ is the log of the FUA's land area. δ , θ and γ are the parameters of

²² Year-quarter.

²³ In the second step we use time-averaged fixed effects because instruments are time-invariant. As a robustness check, we have also followed an alternative estimation strategy where in the first step we introduce city-sector fixed effects, as well as time dummies, and retrieve the estimates of the city-sector fixed effects to be used as dependent variable in the second step. Results do not significantly change and are available upon request.

interest (elasticities) that capture the extent of agglomeration effects on wages. Finally, A_c are dummies for the three natural regions of Ecuador (Andean, Coastal and Amazonian) that control for differences in average wages across the regions.²⁴

The OLS estimation of equation (1) may be affected by two main issues. First, there might be sorting of skilled workers into highly agglomerated areas. In our data, we cannot directly address this point since we would need a panel dataset to control for individual unobserved heterogeneity. Nonetheless, we do our best by introducing a wide range of worker control variables in line with Duranton (2016) and Glaeser and Resseger (2010). Second, there might be a matter of endogeneity arising from the possible simultaneity in individual choices concerning wages and locations. We address this point by using an instrumental variables strategy. As instruments, we use the density defined using cities' historical population (in 1950) and the degree of specialization in 1990.²⁵ The idea is that lagged levels of our independent variables are correlated with the current levels of spatial variables, but they are assumed not to influence productivity and wages today (Combes et al., 2008, Mion and Naticchioni, 2009).²⁶

Table 2 shows the results of the estimation of model (1). Column (1) shows the first step estimation, while the results for the second step estimates are presented in columns (2) to (7). In particular, columns (2) to (4) show the OLS estimates where we introduce

²⁴ It might be argued that the introduction of regional dummies in the estimation could capture and blur the impacts of spatial externalities on wages, considering also the limited number of FUAs available in Ecuador. Nonetheless, in the case of Ecuador (as in other countries, such as Italy), the introduction of these dummies proves to be relevant since the country is characterized by structural differences across regions that are reflected in significant differences in terms of average wages, regardless of other factors. In particular, the Coastal region is characterized by significant lower average wages with respect to the other regions due to the massive migration that historically occurred, worsening labour market conditions, while the Amazonian region is characterized by higher wages due to the oil extraction activity that prevails. Besides, Ecuador has two major cities, Quito and Guayaquil, located in the Andean and Coastal regions respectively, which stands in the way of a major effect of regional dummies in capturing the impact of urban primacy. As a check, we carried out the empirical analysis excluding the regional dummies; the results showed lower magnitudes and significance levels, as well as goodness of fit, reassuring us about the importance of including these control dummies in the second step estimation.

²⁵ In order to run estimates with no more than two endogenous variables, which is already challenging and might make results hard to interpret, as pointed out by Angrist and Pischke (2010), we decided to treat land area as exogenous. Nonetheless, if we instrument the land area using the core area of the FUAs in 1975, the results remain consistent with land area generally capturing a larger part of the impact of density. These estimates are available upon request.

²⁶ It might be argued that a lag of 17 years for the specialization variable is not sufficient for the instrument to be valid. Therefore, we suggest taking the results with some caution. Nonetheless, and reassuringly, Chauvin et al. (2017) compared IV estimates of the wage impact of spatial variables using as instruments either historical variables (lagged around 100 years) or more recent variables (lagged 30 years), finding no significant differences across them. Besides, Ecuador went through a structural change in the '90s, introducing dollarization and experiencing massive international emigration.

the spatial variables sequentially, while column (5) to (7) show the corresponding IV estimates.²⁷

The OLS results show that the variables capturing agglomeration effects have positive and significant coefficients. In particular, when including density alone (column (2)), we find an elasticity of 7.7%, which is not affected by introduction of the specialization variable in the estimation (column (3)). Moreover, the elasticity of wages with respect to the specialization index is 3.5%. Introduction of the log of the FUA area (column (4)) causes a reduction of the density elasticity from 7.6% to 6.5%, while the elasticity of the city area stands at 2.7%, indicating that urbanization externalities are generated by both higher concentration of workers given the size of the area and, to a lesser extent, by bigger city size. When applying the IV estimates (columns (5) to (7)) the results remain consistent as far as urbanization externalities are concerned (elasticities of 5.4% for density and 2.8% for the FUA area in column (7) respectively). This is consistent with previous empirical findings indicating that endogeneity does not appear to be a major concern in analysis of the agglomeration impacts on wages (Combes et al., 2010, Matano and Naticchioni, 2012, Melo et al., 2009). As for specialization, the coefficient is lower in magnitude (between 1.6% and 2.2%) and no longer statistically significant.²⁸

These results indicate that urbanization externalities are at work also in a small developing country like Ecuador, with magnitude of impacts in line with those found for the emerging developing countries (see Chauvin et al., 2017; Combes et al. 2015; Duranton, 2016). The magnitude of the specialization index is also consistent with that found in other studies based on a similar definition of sector specialization at the local level (Matano and Naticchioni, 2012; Mion and Naticchioni, 2009).²⁹

[Table 2 around here]

4.2 Agglomeration effects and informality

²⁷ In the second step estimates are weighted by the time-average number of workers in each city-sector-year-quarter cell.

²⁸ Table A1 in the Appendix shows the IV first stage estimates for columns (5) to (7) of Table 2.

²⁹ In this paper we focus on two of the main spatial externality variables – density and specialization. We also introduce a control for the area of the FUAs, which captures the effect on wages of the size of the FUAs keeping density constant. Nonetheless, there might be other spatial variables affecting wages, such as human capital externalities and market access. We do not insert them in the specification because these variables are also endogenous and we lack adequate instruments. Also, if treated as exogenous, they prove not to be significant due to collinearities with the density variable (see Table A2 in the Appendix).

In this section, we directly test the role of informality within the relationship between spatial agglomeration and wages for a developing country. We employ the estimation strategy shown in the previous subsection, with the difference that, in order to analyze the informal sector, in the first step we now introduce the city-sector-time fixed effects interacted with the informal dummy. This allows us to retrieve two sets of city-sector-time fixed effects, one for each category of workers, which we use in the second step to run separate estimations for formal and informal workers. Table 3 shows the results of the second step estimation.³⁰ ³¹ In addition, Table A4 in the Appendix shows the same estimates not controlling for the individual observable characteristics in the first step, to check for the sorting of workers into observable characteristics for formal and informal sector workers. A simple comparison between Table A4 and Table 3 shows a marked reduction in the coefficients of the urbanization externalities variables, evidencing positive sorting into large and high dense cities. The extent of the reduction is particularly striking for formal sector workers. As for specialization, on the contrary, there is evidence of negative sorting particularly pronounced for informal workers. This means that the lowest skilled informal workers work in highly specialized areas in Ecuador.

Having controlled for the sorting of workers, we now go on to discuss the results of Table 3. Columns (1) to (3) present the second step estimates by OLS, where spatial variables are again added incrementally, while columns (4) to (6) show the corresponding IV estimates.

The OLS estimates indicate that the previous results essentially concerned formal sector workers, whose elasticities with respect to the spatial variables proved positive and significant in both the OLS and IV estimates. In particular, in the IV estimates the elasticities for density and city area stand at 4.8% and 4.3% respectively, while the specialization elasticity stands at 3.6% (top panel of Table 3, column (6)). As for the informal sector workers (bottom panel of Table 3), there is evidence of positive and marginally significant urbanization externalities due to density in OLS estimates

³⁰ For the sake of space, we do not show the results of the first step estimation of Table 3, which is qualitatively the same as the one shown in Table 2. The only difference is that some singleton group observations are now dropped from the estimation, because fixed effects are distinguished by formal and informal sector workers' status. Hence, the first step estimation uses 372,439 observations. These estimates are available upon request. The same will also apply to the next estimates, with no significant changes in the first step results. Also, the number of observations in the second step will be accordingly reduced.

³¹ Table A3 in the Appendix show the IV first stage estimates for columns (4) to (6) of Table 3.

(elasticity of 3.6% in column (3)) which, however, turn out not to be precisely estimated in the IV regressions. Besides, the coefficients for the other spatial variables are not significant in both the OLS and IV estimates. These results show that workers employed in the informal sector do not enjoy the same benefits from agglomeration externalities as formal sector workers.

[Table 3 around here]

So far, we have defined informality from the firm point of view, taking informal workers to be those employed in the informal sector, i.e., in firms with fewer than 100 employees and with no tax identification number. However, informality can also be defined from the worker point of view. As an alternative definition for informality, we now take informal workers to be those who do not have access to the social security system (No-SS). Consequently, there is a possibility for informal workers to be employed in the formal sector.³² We modify the previous specification by introducing in the first step city-sector-time fixed effects interacted with both the informal sector dummy and the No-SS dummy. As in the case of the analysis for informality, we retrieve city-sector-time fixed effects for formal sector workers with and without access to the social security system, and we run separate estimations for these categories of workers. In this way, we are able to assess whether there are differences in the benefits of spatial externalities between formal and informal workers who are employed in the formal sector. Table 4 shows the results. The top panel presents the second-step results for workers with access to the social security system, while the bottom panel presents the results for workers with no access to the social security system. The results show small and non-significant differences in the gains from spatial externalities between formal workers and informal workers employed in the formal sector. Interestingly, informal workers show slightly higher point estimates. Hence, the lack of benefits from agglomeration externalities concerns only informal workers employed in the informal sector.³³

These findings show that the relevant dimension of informality to take into account when evaluating the differential impact of agglomeration economies on wages is the

³² Workers with no access to the social security system constitute around 58% of the total sample. Of these, 43% are employed in the formal sector.

³³ The latter results are somewhat in line with the findings of Duranton (2016), who detects higher benefits from urbanization externalities for informal workers defined as workers without a written labour contract.

type of sector where the workers are employed. This outcome could be related to both specific characteristics of the formal sector, which allow workers to better exploit the benefits of agglomeration externalities (bigger firms, higher level of human capital, higher level of productivity and technology, better links across firms, etc.) and to specific characteristics of the informal workers employed in the formal sector, which make them better able to reap the benefits of agglomeration economies (such as higher education and skill levels).³⁴

[Table 4 around here]

Summing up, these results show that informal workers employed in the informal sector do not enjoy the same wage premium from agglomeration externalities as formal and informal workers employed in the formal sector. One explanation for these findings might be an education effect, considering that the more educated workers (both formal and informal) are more likely to be employed in the formal sector.³⁵ Hence, evidence of higher returns to agglomeration externalities for the more educated workers might (at least in part) account for the higher returns to spatial externalities for formal sector workers that we detected. We directly test this hypothesis. Applying our two-step empirical strategy and interacting, now, the city-sector-time fixed effects with a dummy for high education in the first step, we test whether in Ecuador there is evidence of higher returns to agglomeration externalities for more educated workers. We also test this hypothesis considering separately formal and informal workers (Table A5 in the Appendix). When considering all the workers, the second-step results show that high-educated workers enjoy higher benefits from spatial externalities. However, these findings concern only workers employed in the formal sector, while for the informal sector we fail to detect any significant impact of agglomeration economies on wages for

³⁴ In particular, in our sample it is possible to observe positive sorting of informal workers into the formal sector, since the average years of education are 11.5 for informal workers employed in the formal sector and 9.4 for informal workers employed in the informal sector (while the average comes to 12.7 years for the formal workers).

³⁵ High-educated workers are defined as workers with a level of education above high school (i.e. technical, university and post-university education). They account for 25% of the sample. Moreover, in our sample, 86% of the high-educated workers are employed in the formal sector, while the proportion is just 55% for the low-educated workers. Besides, 23% of the informal workers and 42% of the formal workers employed in the formal sector are highly educated.

either high- or low-educated workers.³⁶ Hence, education does not seem to be a driver of our findings.

We will now go on to present a more detailed picture of the mechanisms through which agglomeration externalities impact on formal and informal workers' wages.

4.3 Analysis of the channels of the spatial wage premium: matching, learning and informality

In this section, we aim to shed light on the channels behind the impacts found so far for formal and informal sector workers: we analyze the role of matching and learning mechanisms in the spatial wage dynamics (Combes and Gobillon, 2015; Puga, 2010). To this end, we make use of the data of a panel subsample of our original dataset. The ENEMDU survey is designed to sample families living in specific buildings/dwellings (*viviendas*), resampling of 25% of them being performed quarterly, following a 2-2-2 panel structure: a building/dwelling may be sampled for two consecutive quarters, then left for the two successive quarters, and subsequently again sampled for two more quarters. This means that if a family has not moved house, it could be interviewed at most four times, covering a total time span of six quarters. We take the opportunity of the survey structure to identify working members of households sampled more than once and analyze their wage dynamics.³⁷ As in the previous sections, we focus on workers aged between 15-64 while eliminating outliers³⁸ and extreme observations in terms of real wages, hours worked and real wages per hour worked. With this procedure, we arrive at a panel of 81,533 observations for 39,294 workers residing in the same place.

To the best of our knowledge, this is the first time that the channels behind the spatial wage premium have been analyzed for a developing country by means of an individual panel dataset. Nonetheless, we must point out that the results of this analysis are limited

³⁶ It is worth noting that the number of high-educated informal sector workers is relatively low (14% of high-educated workers), and therefore the analysis is carried out over a lower number of observations compared to the other estimates. Hence, these estimates have to be taken somewhat gingerly.

³⁷ More specifically, in order to identify household persons interviewed more than once, we select the persons living in the same building/dwelling –included in the panel subsample–, belonging to the same family, having the same position inside the family, with same sex, ethnicity, birth city and (should he/she have previously migrated) reporting the same city as the last place he/she has been living in.

³⁸ In particular, we identify outlier workers as those reporting data non-consistent over time regarding age, education, gender and previous migration status.

to the short-run outcomes of the wage dynamics of stayers. Besides, we suggest taking these results with some caution due to the relatively small size of the sample.

We investigate whether the wage premium due to spatial externalities in the short run can be attributable to better job-matches between workers and firms or to learning mechanisms generated while remaining employed in the same job category. To analyze this issue, we use the data on the specific type of occupation of the worker, as we cannot know whether he or she remains in or changes firm. Therefore, our analysis evaluates the impact on wages of a change in type of occupation, which may well proxy a better match between workers and firms in terms of tasks, while not addressing the impact of job change across firms within the same occupation, which might reflect a better match in terms of alternative dimensions, such as wages, working conditions, etc. The focus of our analysis lies, then, in detection of a wage premium derived from a specific matching linked to the type of occupation, which may occur across firms and also within a firm. We build a dummy signalling job-change with a value equal to 1 when a worker changes occupation, considering 2-digit level occupations defined according to the ISCO88 classification. We employ the same estimation strategy used in the first part of the analysis, where now in the first step the city-sector-time fixed effects are interacted with both the informal sector dummy and a job-change dummy. We then retrieve city-sector-time fixed effects estimates for formal and informal sector workers, whether experiencing a job change or not, and use them to run the second step separately for each category of worker.

With this specification we can see whether wages react positively to a change in occupation, thus suggesting that the wage increases due to a better quality match, or if they increase when remaining employed in the same kind of occupation, pointing to the presence of learning effects.³⁹ In addition, we decided to focus on workers not moving between formal and informal sectors when going through job change. In this way we avoid confusing the impact of an occupation change with that of a change in formality status.⁴⁰

Table 5 shows the results. Columns (1), (2) and (3) present the second step IV estimates for workers who change job, while columns (4), (5) and (6) present the second

³⁹ As for descriptive job change statistics, it is worth noting that around half the job changes for both formal and informal sector workers represent an upgrade in the occupational position.

⁴⁰ Observations accounting for workers who move from the formal to the informal sector or the other way around represent 15% of the total number of observations. We carried out a robustness check using all sample observations and the results remain robust. We do not show these estimates for the sake of conciseness. They are available upon request.

step IV estimates for those who do not change job.⁴¹ Moreover, the top panel of Table 5 presents the results for formal sector workers, the bottom panel for informal sector workers.⁴²

As for the formal sector workers, job-quality match appears to be a relevant channel for their wage increase. In fact, the elasticities of wage with respect to both urbanization externalities (density and area) and specialization are positive, significant and high in magnitude. In particular, in column (3) they come to 10% for density, 12.6% for specialization and 6.9% for city area. Considering the wage change when the worker remains employed in the same kind of occupation (top panel of Table 5, columns (4) to (6)), there is evidence of a wage premium due to urbanization externalities. However, it is lower in magnitude with respect to the one detected for job changes (elasticity of 6.4% for density in column (5), absorbed by the city area effect in column (6)). Hence, in more urbanized areas job changes entail a higher wage premium compared to remaining in the same job for formal sector workers, while agglomeration advantages resulting from sectoral specialization are only achieved through job changes. This outcome points to the matching channel as the most important factor behind the urban wage premium for this category of workers. Nonetheless, there is also evidence of learning effects in the bigger cities.

As for the informal sector workers, job-quality match also seems to be relevant, since the estimates are positive and high in magnitude for urbanization externalities. However, they are never precisely estimated, probably due to the limited sample size. On the contrary, remaining employed in the same occupation (bottom panel of Table 5, columns (4) to (6)) entails a wage penalty with respect to both the specialization and density variables. This could be related to the presence of an excessive mass of informal sector workers in larger areas, à la Harris-Todaro (Overman and Venables, 2005) and/or, with regard to specialization, to the negative sorting of informal workers we detected.

[Table 5 around here]

⁴¹ We do not show the first-step estimates for the sake of space. They are qualitatively the same as previous estimates. They are available upon request.

⁴² There was an interruption in the ISCO data classification between 2012 and 2013. We therefore mapped the ISCO08 classification into the ISCO88 classification for the years 2013, 2014 and 2015. Since we estimate the impact of a job change on wages according to ISCO classification, we decided to run a robustness check using the original ISCO classification (ISCO88 up to 2012 and ISCO08 as from 2013) to compute the job change, while at the same time excluding the observations affected during the change period. The results are shown in Table A6 in the Appendix and remain robust to the classification considered.

To sum up, analysis of the channels through which spatial wage premiums arise shows a difference between workers employed in the formal and informal sectors. As for the formal sector workers, urbanization externality impacts are driven by a better quality of job matching and, to a lesser extent, by positive learning externalities. Sectoral specialization at the local level entails a wage premium only through job changes. For the informal sector workers, the overall results suggest that the non-significant impact detected in the first part of the analysis for urbanization externalities may be due to a balancing effect between a wage premium due to job changes (even if not precisely estimated) and a wage penalty for remaining employed in the same job.

5. Conclusions

In this paper we have explored the role of agglomeration economies on the wages of workers for a small-sized, middle-income developing country, Ecuador. We have found positive and significant impacts on wages deriving from spatial agglomeration. In particular, the elasticity of wages with respect to density is as great as 7.1%, partly captured by the size of the area effect. The elasticity of wages with respect to specialization is 1.6%. Moreover, we have also addressed the role of the informal economy in this relationship, analyzing the heterogeneity of the spatial wage premium across formal and informal sector workers in order to take into account the interaction between spatial agglomeration and the presence of a dual labour market – a common characteristic of most developing countries. Our findings show that informal sector workers are penalized, since the benefits accruing from spatial externalities are low and insignificant. The same does not apply when considering informal workers employed in the formal sector, for whom we detect a non-significant difference in the spatial wage premium with respect to formal workers. The latter outcome might be related to the characteristics of the formal sector, which enhances the workers' capacity to reap the benefits of agglomeration economies.

Furthermore, we also sought to identify the channels through which these wage gains occur across worker categories. The results show evidence of higher gains deriving from better job match in denser, larger and more specialized cities for formal sector workers. There is also evidence of positive learning externalities in the bigger cities. As for the

informal sector workers, there seems to be evidence of some gains from job changes in larger cities, which counterbalance a wage penalty when not changing occupation.

From a policy perspective, these findings suggest the need to design policies to facilitate the process of formalization of firms, which might well be a key factor in increasing workers' productivity and wages, and to allow for significant benefits from spatial externalities. Our findings also point to the costs associated with agglomeration based on the inflow of workers into the informal sector.

References

Angrist, J.D., Pischke J-S., 2010. Mostly harmless econometrics, online blog, <http://www.mostlyharmlesseconometrics.com/2010/02/multiple-endogenous-variables-what-now/>.

Ahrend, R., Farchy, E., Kaplanis, I., Lembcke, A., 2017. What Makes Cities More Productive? Evidence from Five OECD Countries on the Role of Urban Governance, *Journal of Regional Science* 57(3), 385-410.

- Au, C.C., Henderson, V., 2006. Are Chinese Cities Too Small?, *Review of Economic Studies* 73, 549-76.
- Bacchetta, M., Ernst, E., Bustamante, J., 2009. *Globalization and Informal Jobs in Developing Countries: a joint study of the ILO and WTO*, WTO ISBN: 978-92-870-3691-9, ILO ISBN: 978-92-2-122719-9.
- Bernedo Del Caprio, M., Patrick, C., 2017. *Agglomeration and Informality: Evidence from Peruvian firms*, Andrew Young School of Policy Studies, Research Paper Series, 17-13.
- Ciccone, A., 2002. Agglomeration effects in Europe, *European Economic Review* 46 (2), 213- 227.
- Ciccone, A., Hall, R., 1996. Productivity and the density of the economic activity, *American Economic Review* 86 (1), 54-70.
- Chauduri, S., Mukhopadhyay U., 2010. The Harris-Todaro Migration Model and Introduction of the Informal Sector. In: *Revisiting the Informal Sector*, ch.3, Springer, New York, NY.
- Chauvin, J. P., Gleaser, E., Ma, Y., Tobio, K., 2017. What is different about urbanization in rich and poor countries? Cities in Brazil, China, India, and United States, *Journal of Urban Economics* 98, 17-49.
- Combes, P.P., 2000. Economic structure and local growth: France, 1984 -1993, *Journal of Urban Economics* 47 (1), 329-55.
- Combes, P.P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters!, *Journal of Urban Economics* 63 (1), 723-742.
- Combes, P.P., Duranton, G., Gobillon, L., Roux, S., 2010. Estimating agglomeration effects with history, geology, and worker fixed-effects, in: Edward L. Glaeser (Eds.), *Agglomeration Economics*, Chicago, Chicago University Press, 15-65.
- Combes, P.P., Démurger, S., Li, S., 2015. Migration externalities in Chinese cities, *European Economic Review* 76 (1), 152-167.
- Combes, P. P., Gobillon, L., 2015. The Empirics of Agglomeration Economies, in: J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, Vol. 5 (Amsterdam: Elsevier-North-Holland), 247-348.
- De la Roca, J., Puga, D., 2017. Learning by working in big cities, *The Review of Economic Studies* 84 (1), 106-142.
- Duranton, G., 2009. Are Cities Engines of Growth and Prosperity for Developing Countries, in: M. Spence, P. Annez, and R. Buckley, (Eds.), *Urbanization and Growth*, the Commission Growth and Development, Washington, D.C.
- Duranton, G., 2016. Agglomeration effects in Colombia, *Journal of Regional Science* 56 (2), 210-238.
- Duranton, G., Puga D., 2004. Micro-foundations of urban agglomeration economies, in: J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, Vol. 4 (Amsterdam: Elsevier-North-Holland), 2063-2117.

Fields, G., 1990. Labour Market Modelling and the Urban Informal Sector: Theory and Evidence, in: D., Thurnham, Salome, B. and Schwarz, A. (Eds.), *The Informal Sector Revisited*, Paris, OECD.

Garcia, G.A., 2019. Agglomeration Economies in the Presence of an Informal Sector: the Colombian Case, *Revue d'Economie Régionale et Urbaine*, 2, 355-388.

Glaeser, E. L., Henderson, J.V., 2017. Urban economics for the developing World: An introduction, *Journal of Urban Economics* 98 (1), 1-5.

Glaeser, E. L., Maré, D. C., 2001. Cities and skills, *Journal of Labor Economics* 19 (2), 316-342.

Glaeser, E. L., Resseger, M. G., 2010. The complementarity between cities and skills, *Journal of Regional Science* 50 (1), 211-244.

Harris, J., Todaro, M., 1970. Migration, Unemployment, and Development: A Two Sector Analysis, *American Economic Review* 107 (1), 126-142.

Harris, J., 2014. The messy reality of agglomeration economies in urban informality: Evidence from Nairobi's handicraft industry, *World Development* 61, 102-113.

ILO, 2013. *Measuring Informality: A statistical manual on the informal sector and informal employment*, Book ILO, ISBN: 9789221273882.

INEC, 2015. *Actualización metodológica: Empleo en el sector informal y la clasificación de los ocupados según sectores*, Nota técnica, Quito, Ecuador.

Khamis, M., 2012. A note on informality in the labour market, *Journal of International Development* 24 (1), 894-908.

La Porta, R., and Shleifer, A., 2014. Informality and Development, *Journal of Economics Perspectives* 28 (3), 109-126.

Lall, S.V., Shalizi, Z., and Deichmann, U., 2004. Agglomeration Economies in the Indian Industry, *Journal of Development Economics* 73, 643-673.

Maloney, W.F., 2004. Informality Revisited, *World Development* 32(7), 1159-1178.

Marshall, A., 1890. *Principles of Economics*, Book London: MacMillan.

Matano, A., Naticchioni, P., 2012. Wage distribution and the spatial sorting of workers, *Journal of Economic Geography* 12 (2), 379-408.

Matano, A., Naticchioni, P., 2016. What drives the urban wage premium? Evidence along the wage distribution, *Journal of Regional Science* 56 (2), 191-209.

Melo, P. C., Graham, D. J., Noland, R. B., 2009. A meta-analysis of estimates of urban agglomeration economies, *Regional Science and Urban Economics* 39 (3), 332-342.

Mion, G., Naticchioni, P., 2009. The spatial sorting and matching of skills and firms, *Canadian Journal of Economics* 42 (1), 28-55.

Obaco, M., Diaz Sanchez, J.P., 2018. An Overview of Urbanization in Ecuador under Functional Urban Area Definition, *Region*, 5(3), 38-48.

Obaco, M., Royuela V., Vítóres, X., 2019. Identifying Functional Urban Areas in Ecuador using a varying travel time approach, *Geographical Analysis*, forthcoming.

OECD, 2013. Definition of functional urban areas (FUA) for the OECD metropolitan database, Paris: OECD Publishing.

OECD, 2017. Employment Outlook, Annex Chapter 3.

Overman, H., and Venables, A., 2005. Cities in the Developing World, CEP Discussion Paper, n.695, London Perry.

Puga, D., 2010. The magnitude and causes of agglomeration economies, *Journal of Regional Science* 50 (1), 203-219.

Rosenthal, S.S., and Strange W.C., 2004. Evidence on the nature and sources of agglomeration Economies, in: J.V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Vol. 4 (Amsterdam: Elsevier-North-Holland): 2119-2171.

Royuela, V., and Castells-Quintana, D., 2015. International migrations and urbanisation: 1960-2010, *International Journal of Global Environmental Issues* 13, 150-169.

Royuela, V., Ordóñez, J., 2018. Internal migration in a developing country: A panel data analysis of Ecuador (1982-2010), *Papers in Regional Science* , 97(2), 345-367.

Tables

Table 1: Descriptive Statistics. Real Wage per Hour and Sample Composition by Worker Categories

Worker data	Mean Real Wage per Hour (US dollars)				Sample Composition
	2007	2012	2017	All	%
Total	2.20	2.29	2.61	2.37	100
Gender					
Male	2.24	2.39	2.68	2.45	61.65
Female	2.14	2.14	2.49	2.24	38.35
Ethnicity					
Indigenous	1.68	1.89	1.89	1.79	4.39
White	2.54	2.93	3.19	2.64	3.24
Mestizo*	2.22	2.31	2.66	2.40	88.61
Black	1.63	1.98	2.24	2.02	2.43
Mulatos**	1.71	1.94	2.11	2.00	1.33
Education level					
At Most Literacy	1.42	1.65	1.74	1.47	0.28
School	1.62	1.84	2.00	1.81	30.55
High School	2.00	2.13	2.37	2.18	43.68
Technical	2.57	2.61	3.02	2.89	1.40
University	3.22	3.07	3.62	3.32	23.17
Post-Univ. Degree	5.67	4.79	6.13	5.72	0.91
Occupation					
Legislators, Professionals and Technicians	3.74	3.60	4.35	3.98	14.41
Agricultural and Fishery Workers	2.03	2.19	2.40	2.20	38.36
Machine Operators, Elementary Occupations	1.84	1.97	2.22	2.01	47.23
Job Category (1)					
Dependent workers	2.12	2.38	2.80	2.49	56.67
Self-employed	2.31	2.19	2.38	2.22	43.33
Job Category (2)					
Formal Sector Workers	2.53	2.62	3.04	2.76	62.95
Informal Sector Workers	1.71	1.72	1.84	1.71	37.05
Economic Sector					
Agriculture, Fishing	1.72	1.91	2.02	1.87	8.62
Mining and Quarrying	4.09	3.85	4.24	4.27	0.47
Manufacturing	2.16	2.27	2.63	2.34	16.37
Electricity, Gas and Water Supply	3.69	2.98	2.68	3.10	0.11
Construction	2.08	2.24	2.43	2.28	9.18
Wholesale and Retail Trade	2.11	2.18	2.48	2.25	29.61
Hotels and Restaurants	2.13	2.25	2.44	2.26	7.39
Transport, Storage and Communications	2.23	2.35	2.52	2.36	9.80
Financial Intermediation	3.74	3.61	4.59	3.98	1.56
Real Estate, Renting and Business Activities	2.63	2.68	3.01	2.80	6.93
Private Education, Health and Social Work	2.83	2.91	3.65	3.20	5.51
Other Service Activities	1.98	1.92	2.47	2.11	4.45
Region					
Andean	2.38	2.35	2.80	2.53	53.90
Coastal	1.99	2.22	2.37	2.15	41.90
Amazonian	2.18	2.44	2.34	2.45	4.21
Migrant					
Yes	2.26	2.36	2.61	2.44	32.53
No	2.17	2.26	2.60	2.34	67.47
Total Obs.	20,213	30,423	41,783	377,273	

Notes: Occupation is defined according to ISCO88, sector according to ISIC 3.1. *Mestizo is the category for mixed white/indigenous race. **Mulato is the category for mixed white/black race.

Table 2: OLS and IV regressions of wages on the spatial variables. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1st step OLS	2nd step OLS	2nd step OLS	2nd step OLS	2nd step IV	2nd step IV	2nd step IV
Log(Density)		0.077*** (0.011)	0.076*** (0.011)	0.065*** (0.011)	0.071*** (0.014)	0.071*** (0.014)	0.054*** (0.015)
Log(Specialization)			0.035*** (0.012)	0.035*** (0.011)		0.022 (0.015)	0.016 (0.015)
Log(Area)				0.027*** (0.008)			0.028*** (0.008)
Female dummy	-0.162*** (0.002)						
Age	0.037*** (0.001)						
Age squared	-3.89e-04*** (6.71e-06)						
Education	0.032*** (0.000)						
Migrant dummy	0.041*** (0.002)						
Ethnicity dummies	yes						
Job category dummies	yes						
Occupation dummies	yes						
City*sector*time dummies	yes						
Regional dummies	no	yes	yes	yes	yes	yes	yes
Observations	377,273	904	904	904	904	904	904
R-squared	0.342	0.170	0.185	0.210	0.170	0.183	0.203
Weak identification test (F-value)					388.5	129.6	138.2

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time-average of the city-sector-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector cell.

Table 3: OLS and IV regressions of wages on the spatial variables and informal sector. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step	2nd step	2nd step	2nd step	2nd step	2nd step
	OLS	OLS	OLS	IV	IV	IV
Formal sector workers						
Log(Density)	0.079*** (0.011)	0.075*** (0.011)	0.056*** (0.011)	0.077*** (0.013)	0.072*** (0.013)	0.048*** (0.015)
Log(Specialization)		0.041*** (0.013)	0.041*** (0.012)		0.046*** (0.017)	0.036** (0.016)
Log(Area)			0.042*** (0.006)			0.043*** (0.006)
Observations	843	843	843	843	843	843
R-squared	0.123	0.140	0.188	0.123	0.140	0.188
Weak identification test (F-value)				513.6	132.1	155.3
Informal sector workers						
Log(Density)	0.038* (0.020)	0.039* (0.020)	0.036* (0.018)	0.031 (0.025)	0.030 (0.025)	0.024 (0.022)
Log(Specialization)		0.004 (0.015)	0.004 (0.015)		-0.012 (0.021)	-0.014 (0.021)
Log(Area)			0.010 (0.015)			0.011 (0.014)
Observations	620	620	620	620	620	620
R-squared	0.017	0.017	0.019	0.016	0.015	0.016
Weak identification test (F-value)				264.4	122.4	187.1

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time average of the city-sector-informality-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-informality cell.

Table 4: OLS and IV regressions of wages on the spatial variables, informal sector and access to the social security system. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step	2nd step	2nd step	2nd step	2nd step	2nd step
	OLS	OLS	OLS	IV	IV	IV
Formal sector workers with access to the social security system						
Log(Density)	0.053*** (0.012)	0.051*** (0.012)	0.039*** (0.013)	0.047*** (0.013)	0.044*** (0.014)	0.031** (0.014)
Log(Specialization)		0.023* (0.012)	0.022* (0.012)		0.028* (0.017)	0.022 (0.016)
Log(Area)			0.024*** (0.007)			0.025*** (0.007)
Observations	733	733	733	733	733	733
Weak identification test (F-value)				574.9	112.9	125.2
Formal sector workers with no access to the social security system						
Log(Density)	0.068*** (0.014)	0.064*** (0.014)	0.058*** (0.015)	0.057*** (0.018)	0.053*** (0.017)	0.042** (0.018)
Log(Specialization)		0.038*** (0.014)	0.039*** (0.014)		0.050** (0.021)	0.043** (0.020)
Log(Area)			0.031*** (0.007)			0.032*** (0.007)
Observations	664	664	664	664	664	664
Weak identification test (F-value)				447.3	109.7	122.0

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time average of the city-sector-informality-no_ss-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-informality-no_ss cell.

Table 5: OLS and IV regressions of wages on the spatial variables, informal sector and job changes. Dependent variable in the first step: log of real worker's wage per hour.

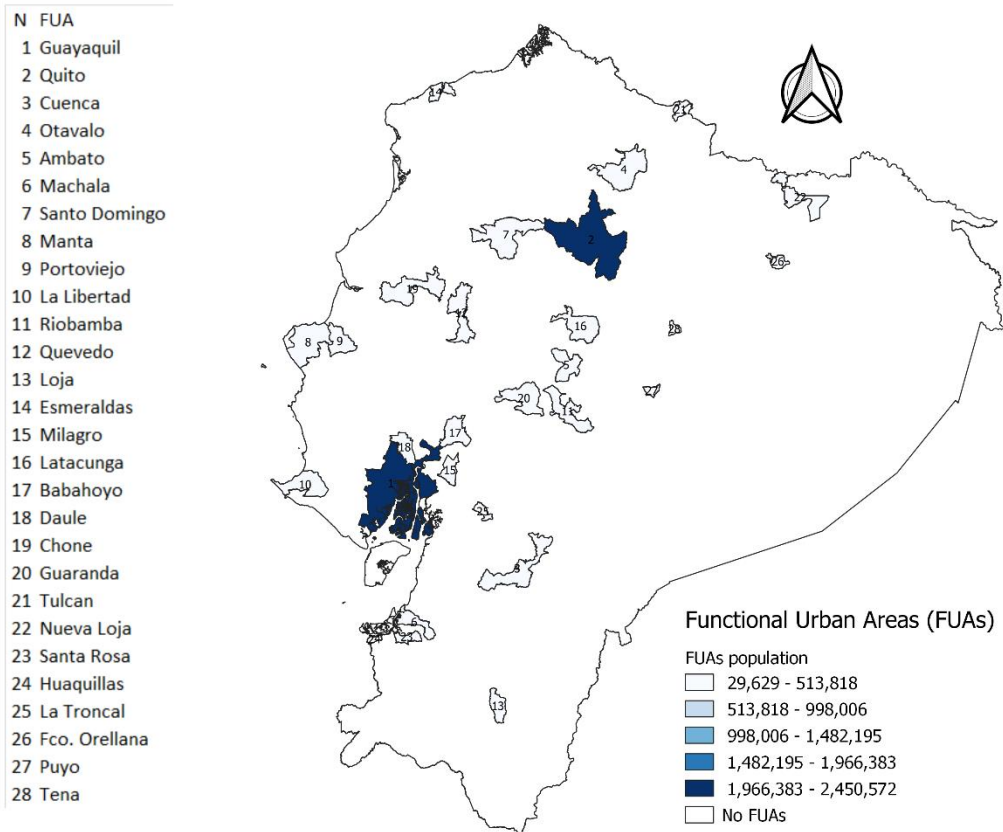
	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step	2nd step	2nd step	2nd step	2nd step	2nd step
	IV	IV	IV	IV	IV	IV
Formal sector workers						
	Job change			No job change		
Log(Density)	0.156*** (0.057)	0.157*** (0.052)	0.102* (0.057)	0.068* (0.037)	0.064* (0.037)	0.030 (0.038)
Log(Specialization)		0.125** (0.052)	0.126*** (0.049)		0.026 (0.030)	0.018 (0.031)
Log(Area)			0.069*** (0.022)			0.063*** (0.015)
Observations	258	258	258	301	301	301
Weak identification test (F-value)	219.3	118.2	114.0	258.9	97.9	118.0
Informal sector workers						
Log(Density)	0.143 (0.102)	0.126 (0.112)	0.074 (0.101)	-0.069 (0.055)	-0.087 (0.055)	-0.110** (0.055)
Log(Specialization)		-0.049 (0.068)	-0.057 (0.069)		-0.154*** (0.051)	-0.160*** (0.053)
Log(Area)			0.065 (0.044)			0.038 (0.024)
Observations	125	125	125	177	177	177
Weak identification test (F-value)	72.0	34.1	35.5	132.8	73.7	71.2

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time average of the city-sector-informality-job change-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-informality-job change cell.

Appendix

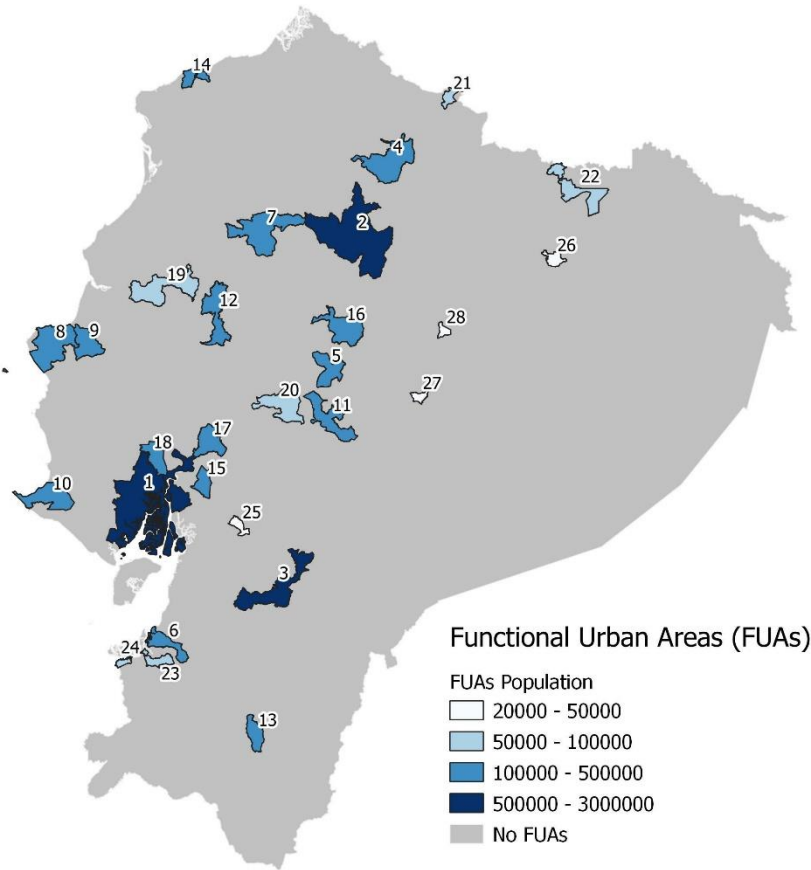
Figures

Figure A1. FUAs in Ecuador



Notes: FUAs ordered by size.

- No FUA
- 1 Guayaquil
- 2 Quito
- 3 Cuenca
- 4 Otavalo
- 5 Ambato
- 6 Machala
- 7 Santo Domingo
- 8 Manta
- 9 Portoviejo
- 10 La Libertad
- 11 Riobamba
- 12 Quevedo
- 13 Loja
- 14 Esmeraldas
- 15 Milagro
- 16 Latacunga
- 17 Babahoyo
- 18 Daule
- 19 Chone
- 20 Guaranda
- 21 Tulcan
- 22 Nueva Loja
- 23 Santa Rosa
- 24 Huaquillas
- 25 La Troncal
- 26 Fco. Orellana
- 27 Puyo
- 28 Tena



Tables

Table A1: First stage estimates of Table 2: Columns (5) to (7)

Dependent Variable	Column (5)	Column (6)		Column (7)	
	log(Dens)	log(Dens)	log(Spec)	log(Dens)	log(Spec)
Log(Density 1950)	0.464*** (0.024)	0.463*** (0.024)	-0.035** (0.017)	0.457*** (0.023)	-0.022 (0.017)
Log(Specialization 1990)		0.018 (0.024)	0.593*** (0.035)	0.015 (0.025)	0.599*** (0.035)
Log(Area)				0.022 (0.023)	-0.048*** (0.015)

Table A2: OLS and IV regressions of wages on the spatial variables. Dependent variable: log of real worker's wage per hour. Additional control variables: market access and share of high-educated workers.

	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step OLS	2nd step OLS	2nd step OLS	2nd step IV	2nd step IV	2nd step IV
Log(Density)	0.074*** (0.013)	0.062*** (0.012)	0.071*** (0.014)	0.058*** (0.022)	0.048*** (0.018)	0.053** (0.026)
Log(Specialization)	0.035*** (0.011)	0.036*** (0.011)	0.035*** (0.011)	0.015 (0.015)	0.016 (0.015)	0.016 (0.015)
Log(Area)	0.029*** (0.009)	0.026*** (0.009)	0.028*** (0.009)	0.029*** (0.009)	0.026*** (0.009)	0.026*** (0.010)
Share of high-educated workers	-0.122 (0.097)		-0.121 (0.097)	-0.047 (0.127)		-0.046 (0.130)
Log(Market access)		-0.010 (0.021)	-0.009 (0.021)		-0.023 (0.025)	-0.021 (0.027)
Region dummies	yes	yes	yes	yes	yes	yes
Observations	904	904	904	904	904	904
R-squared	0.211	0.211	0.211	0.205	0.203	0.204
Weak identification test (F-value)				89.62	124.2	72.54

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time-average of the city-sector-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector cell. Market access is defined as the weighted average of cities' income (average wage of the FUA multiplied by population in the FUA), weighted by the inverse of the distance between FUA's, while the share of high-educated workers is defined as the share of workers with a level of education higher than high school.

Table A3: First stage estimates of Table 3: Columns (4) to (6)

Dependent Variable	Column	Column (5)		Column (6)	
	log(Pop)	log(Pop)	log(Spec)	log(Pop)	log(Spec)
Formal sector workers					
Log(Density 1950)	0.451*** (0.021)	0.446*** (0.020)	-0.025 (0.018)	0.439*** (0.018)	-0.012 (0.018)
Log(Specialization 1990)		0.039** (0.018)	0.594*** (0.036)	0.035* (0.018)	0.603*** (0.035)
Log(Area)				0.029 (0.021)	-0.053*** (0.016)
Informal sector workers					
Log(Density 1950)	0.462*** (0.028)	0.462*** (0.029)	-0.047** (0.020)	0.461*** (0.029)	-0.040* (0.022)
Log(Specialization 1990)		0.003 (0.039)	0.610*** (0.040)	0.003 (0.039)	0.613*** (0.039)
Log(Area)				0.003 (0.024)	-0.032* (0.016)

Table A4: OLS and IV regressions of wages on the spatial variables and informal sector. No individual covariates at the first step. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step	2nd step	2nd step	2nd step	2nd step	2nd step
	OLS	OLS	OLS	IV	IV	IV
Formal sector workers						
Log(Density)	0.132*** (0.016)	0.130*** (0.016)	0.103*** (0.016)	0.153*** (0.019)	0.154*** (0.019)	0.121*** (0.021)
Log(Specialization)		0.020 (0.016)	0.020 (0.015)		-0.009 (0.020)	-0.022 (0.019)
Log(Area)			0.060*** (0.010)			0.058*** (0.010)
Observations	843	843	843	843	843	843
R-squared	0.146	0.148	0.202	0.144	0.141	0.191
Weak identification test (F-value)				513.6	132.1	155.3
Informal sector workers						
Log(Density)	0.061** (0.024)	0.060** (0.024)	0.055** (0.022)	0.052* (0.030)	0.049 (0.030)	0.038 (0.027)
Log(Specialization)		-0.017 (0.020)	-0.017 (0.020)		-0.060** (0.027)	-0.064** (0.027)
Log(Area)			0.019 (0.018)			0.021 (0.017)
Observations	620	620	620	620	620	620
R-squared	0.021	0.022	0.028	0.020	0.012	0.016
Weak identification test (F-value)				264.4	122.4	187.1

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time average of the city-sector-informality-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-informality cell.

Table A5: OLS and IV regressions of wages on the spatial variables and education. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)
	2nd step	2nd step	2nd step	2nd step	2nd step
	IV	IV	IV	IV	IV
High-educated workers					
	All workers			Formal sector workers	Informal sector workers
Log(Density)	0.090*** (0.020)	0.084*** (0.020)	0.062*** (0.022)	0.054** (0.022)	-0.065 (0.051)
Log(Specialization)		0.040 (0.025)	0.025 (0.024)	0.022 (0.025)	-0.016 (0.069)
Log(Area)			0.042*** (0.009)	0.052*** (0.009)	0.027 (0.020)
Observations	643	643	643	613	286
Weak identification test (F-value)	475.5	113.8	118.3	110.3	70.6
Low-educated workers					
Log(Density)	0.048*** (0.016)	0.048*** (0.016)	0.035** (0.017)	0.043*** (0.016)	0.028 (0.023)
Log(Specialization)		0.015 (0.015)	0.010 (0.015)	0.044** (0.018)	-0.020 (0.021)
Log(Area)			0.025*** (0.009)	0.037*** (0.007)	0.011 (0.014)
Observations	837	837	837	753	583
Weak identification test (F-value)	360.6	134.7	140.1	158.3	194.4

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable in columns (1) to (3) is the weighted time average of the city-sector-high education-time fixed effects retrieved from the first step, weighted by the number of observations in each cell, while the dependent variable in columns (4) to (5) is the weighted time average of the city-sector-high education-informality-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variable. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-high education cell for columns (1) to (3), and in each city-sector-high education-informality cell for columns (4) to (5).

Table A6: OLS and IV regressions of wages on the spatial variables, informal sector and job changes. Original ISCO codification and excluding individuals during the change in ISCO versions. Dependent variable in the first step: log of real worker's wage per hour.

	(1)	(2)	(3)	(4)	(5)	(6)
	2nd step	2nd step	2nd step	2nd step	2nd step	2nd step
	IV	IV	IV	IV	IV	IV
Formal sector workers	Job change			No job change		
Log(Density)	0.152*** (0.050)	0.156*** (0.045)	0.103** (0.047)	0.072* (0.038)	0.069* (0.038)	0.036 (0.038)
Log(Specialization)		0.091** (0.043)	0.089** (0.040)		0.017 (0.030)	0.009 (0.031)
Log(Area)			0.067*** (0.019)			0.059*** (0.015)
Observations	260	260	260	294	294	294
Weak identification test (F-value)	201.5	114.9	110.4	247.2	98.7	118.8
Informal sector workers						
Log(Density)	0.112 (0.103)	0.106 (0.111)	0.058 (0.100)	-0.051 (0.058)	-0.069 (0.058)	-0.094* (0.057)
Log(Specialization)		-0.021 (0.076)	-0.030 (0.078)		-0.152** (0.060)	-0.159*** (0.062)
Log(Area)			0.064 (0.046)			0.045* (0.025)
Observations	128	128	128	175	175	175
Weak identification test (F-value)	74.0	35.9	37.2	127.9	67.9	65.7

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications include a constant term. In the first step occupation is codified according to ISCO 88, two-digit level, sector according to ISIC 3.1, two-digit level. In the second step, the dependent variable is the weighted time average of the city-sector-informality-job change-time fixed effects retrieved from the first step, weighted by the number of observations in each cell. Independent variables in the second step are the time-average values of the spatial variables. Three regional indicators (Andean, Coastal and Amazonian) are also included. In the IV estimates instruments are population density in 1950 and the specialization level in the 1990. Estimates are weighted using as weights the average number of individuals in each city-sector-informality-job change cell.