Segregation and urban spatial structure in Barcelona

Miquel-Àngel Garcia–Lopez, [a,b] Rosella Nicolini,[a] José Luis Roig [a1]

[a] Departament Economia Aplicada, Universitat Autònoma de Barcelona
Edifici B - Campus UAB -08193 Bellaterra (Barcelona), Spain.
Correspondent author: Rosella Nicolini (e-mail rosella.nicolini@uab.cat)

[b] Institut d'Economia de Barcelona IEB (Universitat de Barcelona)

This version: December 2018

Abstract

We conduct an empirical analysis to assess the degree of segregation among the different communities in Barcelona taking into account the spatial dependence of the features associated with the neighborhood's status. We build an original database by gathering information for the period 1947–2011. Estimations emphasize that Barcelona increasingly shows a spatial-dependent segregation pattern based mostly on the gentrification of the high-skill workers rather than on the ethnicity of the different communities.

Keywords: *Urban spatial structure, segregation, immigration*

***JEL Classification**: R14, R23, F22*

[1] We are grateful to Mrs. Sara Plaza for the invaluable support in data collection. We also thank K. Lang, B. Margo, A. Rambaldi as well as participants at NARSC (2016, Minneapolis), at seminars at University of Queensland and at Macquerie University for fruitful suggestions. All remaining errors are our own responsibility. Financial support from grants ECO2014-52506-R, ECO2014-52999-R, 2017SGR207, 2017SGR1301, XREA, and XREPP is gratefully acknowledged.

1. Introduction

Since the beginning of the 20th century, Barcelona has been a preferred destination for internal and international migrants. Recent literature on the topic has emphasized the way in which ethnic segregation can pose several challenges for the social cohesion of modern towns (Topa and Zenou, 2015). Although empirical evidence regarding US cities has allowed extensive discussions of the trends (Cutler et al., 2008a, 2008b, de la Roca et al., 2014, Lewis and Peri, 2015), contributions regarding European cities are fewer (Boeri et al., 2015, Glitz, 2014, Musterd, 2005, Musterd et al. 2017).

We propose to detect the patterns of distribution of the different communities within Barcelona referring to the economic and spatial features of its neighborhoods. Under this perspective, we are interested in assessing the potential consolidation of patterns of urban segregation by means of the analysis of features describing the neighborhoods economic status knowing that each spatial unit host three different communities for the overall period of analysis.[2]

Barcelona represents a particularly interesting case in the Mediterranean area, given the availability of data about important foreign communities since the mid-20th century and the city's consolidation as a multiethnic society across time. To our knowledge, this type of issue has not been addressed under the historical and spatial perspective yet. The spatial dimension is relevant because it provides a solid ground to develop the analysis tracing back to the idea of homophily inside each spatial unit and for the closest ones.

Aiming to examine conditions affecting the potential self-reinforcing processes of citizens' location driven by the homophily aptitude, in our empirical work we analyze the determinants of the community-distribution pattern and seek to provide a reasoned interpretation of the outcome.

To that end, we construct an original database by merging historical statistics released by township administrative services across time.

[2] In this study, we intend as *ethnic* the origin of citizens by place (nation) of birth. When merging several ethnic groups for analysis purposes, we label them as *communities*. Also, remark that we are considering Spanish and Catalan born citizens as separated communities. In this way, we are taking into account the social and cultural dimension that leads to a differentiation of the *identity* between these two groups. More information can be found in Conversi (1990), and Moreno and Arriba (1996).

This study aims to deliver evidence about the determinant of neighborhood composition for the case of a representative European city like Barcelona. We develop an econometric exercise in order to track the potential changes of the neighborhood status by the huge immigrant inflows experienced by the town in 2000s. Different from the US case, in Barcelona, the evolution of the urban structure, the local housing policy and the urban public transport system made land organization move towards a gentifricated-type structure hinted by the dichotomy between high versus low skill citizens rather than by the ethnic differences among communities.

In what follows, we provide a brief overview of literature in Section 2, after which we describe our database and some preliminary statistics that justified our empirical strategy in Section 3. In Section 4, we propose a theoretical framework whereas in Section 5 we discuss our econometric results. Section 6 concludes the paper.

2. Literature review

To investigate the features that drive the location decisions of communities in Barcelona, we seek to identify aspects of neighborhood status that makes one place more attractive than another for a specific group of people. The change and persistence of the status of urban neighborhoods from an economic perspective is surveyed by Rosenthal and Ross (2015).[3] It involves analyzing the changes in per-capita income or population or employment, for instance.

Glaeser et al. (2008) and Brueckner and Rosenthal (2009) conclude that -on average- the economic status of neighborhoods in a city decreases with distance from the city center, with the presence of a public transport system, and when the stock of housing is younger.[4] Those features are more likely to affect the local real estate market with dampening prices, rents and attracting lower-income households. At the same time, however, most European cities have protected central historic districts that provide a unique urban amenity which counterbalances the other effects (Brueckner et al. 1999). As a result, those locations attract high-income households that contribute to a different economic status for the neighborhood. Glaeser et al. (2008) provide quantitative

[3] Determining the economic status of a location to a large extent involves the idea of being able to both measure the location's performance in economic terms—for example, resident incomes or employment status—and compare those measures across time or locations, if not both.
[4] Both sets of authors used the average income in the neighborhood to measure its economic status.

evidence about the impact of those three types of mechanisms. Their figures reveal that housing demand is less elastic to income than commuting costs, meaning that the provision of public services (e.g., public transport) tends to be more effective in dense, central areas. That pattern implies that city centers, in being more accessible than the suburbs, attract more low-income households, whereas high-income households are more prone to live in the suburbs.

Such patterns of location nurture effects of segregation. As Rosenthal and Ross (2015) posit, social dynamics are a self-reinforcing influence that favors the endogenous formation of high- and low-income communities.[5]

Those model types stem from Schelling's tradition (1971), which highlights two typical models of spatial configurations: spatial proximity models and isolated neighborhood models. In the former, an individual's utility is high when the share of individuals in her own group in the nearest neighborhood exceeds an established threshold and low otherwise. As such, racial composition preferences always yield to segregation outcomes. In the latter, individual preferences for neighborhood racial composition are based on a step function. That model suggests that the heterogeneity of individuals regarding their preferences for integration is connected with the critical value of their preferences for neighborhoods with other individuals in their own group, which ultimately changes their utility.

Extending Schelling's tradition with an empirical analysis run for a panel of US cities and suburbs for the period 1970–2000, Card et al. (2008) model a setting in which a stable integrated neighborhood emerges as the majority group—in their case, whites—has a strong preference to live in integrated communities. However, if the representation of minorities in the integrated neighborhood becomes too high to be supported by majority group preferences for integration, then the members of the majority group will relocate, and the community will transform into an all-minority segregated location. That type of framework is quite useful for identifying potential changes in neighborhood status in terms of population composition due to perturbations such as migrant influx. Such flows can affect integrated neighborhoods, especially unstable ones, and transform them into segregated communities.

[5] Along similar lines, Bayer, Fang, and McMillan (2014) assessed that black middle-class neighborhoods are more likely to develop when there is a sufficient mass of black high-income households.

An important dimension of segregation in both US and European cities is ethnicity. Among the various theories available to explain ethnic segregation, evidence from Sweden (Edin et al., 2003) and Denmark (Damm, 2014) has shown that ethnic enclaves boost the labor market performance of immigrants in light of ethnic network effects. That theory, which ranks among the most popular to explain ethnic neighborhoods, implies that the degree of segregation is proportional to the cultural distance between the immigrant group and the host society. It also suggests that as immigrants increasingly assimilate into host societies, they live in progressively less segregated neighborhoods (Cutler et al., 2008a).

Yet another theory explaining ethnic segregation holds that as ethnic minorities experience discrimination in the housing market they can reinforce the concentration of immigrants in some parts of a city.

In the case of European cities, the effects of housing policies need to be also taken into account because the accessibility to public housing can partly explain the distribution of immigrant groups throughout the cities (Musterd, 2005). In cities in southern European countries such as Spain and Italy, where public housing is almost nonexistent, the location decisions of migrants depend to a larger extent on the available supply of cheap rental housing. Consequently, less favored immigrant groups tend to concentrate in old parts of the cities, or in peripheral working-class neighborhoods where migrants can find entry into old, poorly maintained housing (Martori and Apparicio, 2011).

Another important dimension of segregation has to be read in connection with schooling. Munster (2009) emphasizes that education is a key policy to promote integration, but evidence is at odds. Rongvd (2007) discusses the low level of ethnic segregation in Copenhagen but a high level of school segregation because of school choice options. In the same line, Burguess et al. (2005) identify that children are more segregated at schools than in the neighborhood and school segregation is associated with family income.

3. Data and preliminary evidence

Original and unique, our database represents a blend of different sources of data. It includes all available data either at districts (*distritos*) or neighborhoods (*barrios*) for Barcelona for 1947, 1965, 1970, 1986, 1995, 2001, 2008, and 2011.

Our chief data sources were the Barcelona Statistical Yearbooks. Data in the yearbooks have been extracted from the city census, when available, or local administrative records (*padrón*). Because of the structure of the sources, we located individuals based on their place of residence. As for the territorial dimension, the identification of districts and neighborhoods has not been constant over time. On the one hand, neighborhood identification was very irregular until 2007, when the current model consisting of 73 neighborhoods was introduced. On the other, the identification of districts has been slightly more constant and consistent across time.

By contrast, to embed a longer historical dimension in our database with a consistent geographical structure across time, we construct a pseudo-panel by relying on the district structure as of 2011. Our strategy to build the pseudo-panel is presented in the Appendix.

An additional value of our original database is information about Barcelona's bus transport system. We tracked the evolution of present-day bus lines using data originally extrapolated from existing evidence,[6] and we created statistics about the number of lines serving a district or neighborhood in each year of our database. The relevance of such information relates strongly to the importance of bus-public transport in the history of Barcelona, where the use of public transport has always been a distinguishing feature of low- and high-income households (Fernández i Valentí, 2006).

Our data cover a quite extended period (1947–2011), during which Barcelona experienced an important evolution in terms of urban structure. First, numerous neighboring municipalities came to accommodate inhabitants who left the city for several reasons, including to search for more affordable rents and to become owners of individual dwellings. Second, Barcelona became a preferred destination for domestic migrants and households looking for jobs in the 1950s–1960s and for international

[6] Historical aspects and technical details, among other things, are available at http://www.autobusesbcn.es/ (by Mr. José Mora).

migrants in the 2000s. We track the evolution of the share of foreign-born citizens in Barcelona across the decades (Figure B.1 in the appendix).

As for the immigrant communities in Barcelona (Figure B.2 in the Appendix), people from the European Union were most densely populous until the 1990s, when, following their massive influx into Spain, Latin Americans achieved the greatest density of the groups. Meanwhile, the density of the Asian community in Barcelona has increased at a relatively stable pace even during the last decades when the group of Latin American people recorded a slight decrease.

We provide preliminary evidence about the degree of ethnic segregation in Barcelona by computing the index of dissimilarity (Index D) proposed by Duncan and Duncan (1955) (Table B.3 in the Appendix).

The degree of segregation among all ethnic groups has lessened across time. Immigrants have usually been the most segregated communities. This dynamics is likely to be associated with the heterogeneous composition of this group and different languages and culture (above all when referring to the growing importance of Asian) make the integration process more complicated.

4. Theoretical setting

We introduce a simple theoretical setting that helps to understand and interpret the results of our empirical exercise. The setting retrieves the building blocks of the von Thünian framework as described in Fujita and Ogawa (1982). Space is represented by a real line X=(-∞, ∞). We model an extended structure for the central business district (CBD). At the origin of the real line we settle the proper CBD, clustering high-value services, while the surrounding area at a distance $a>0$ ($a \in X$) from the CBD hosts lower value and more land consuming services. Each consumer-worker supplies its unit of labor in one of the two activities but she consumes both types of services (combined in the composite good $Z_i(x)$). The population is split into three ethnic groups i, namely, group c, group s and group e. Each group i is represented by a continuum of consumer whose density is defined for each location $x \in X$ with $x>a$ as $n_i(x)\geq 0$.

We assume that citizens commute daily to the CBD and the surrounding area to purchase the composite good $Z_i(x)$, to supply labor and get their income (w_i) exogenously determined. [7]

As usual, we shape the canonical land competition among the three groups to settle in the different urban plots as close as possible to the CBD. We assume the existence of an absentee landlord. The total finite land occupied at $x \in X$ by the three groups is normalized to one in a way that:

$$n_c(x)S_c(x) + n_s(x)S_s(x) + n_e(x)S_e(x) = 1,$$

where $S_i(x)$ – with $i=\{c, e, s\}$- is the size of the land plot of each group-type at location $x \in X$.

In accordance with their preferences and their budget constraints, agents' location choices are defined on the basis of the maximum amount of rent $R_i(x)$ they are willing to pay to settle at $x \in X$.

Consumers share the same type of utility function that involves the consumption of the land $S_i(x)$ at $x \in X$ and the purchase of the composite good $Z_i(x)$ whose price is normalized to one. Then, for each group i with $i=\{c, e, s\}$ the correspondent utility function is defined as:

$$U_i = \frac{1}{\alpha^\alpha \beta^\beta} S_i(x)^\alpha Z_i(x)^\beta \qquad with\ \alpha, \beta \in (0,1).$$

(1)

Commuting to the CBD and location A implies incurring in transport costs t. In this respect, the budget constraint of each type of agent is defined as follows:

$$w_i = Z_i(x) + R_i(x)S_i(x) + tx + t(x-a) \qquad (2)$$

Each group of agents maximizes (1) subject to (2) with respect to $S_i(.)$ and $Z_i(.)$. From the first order conditions, we get:

$$S_i(x) = \frac{\alpha[w - tx - t(x-a)]}{R_i(x)}\ ;\ Z_i(x) = \beta[w - tx - t(x-a)]. \qquad (3)$$

Consider that the reservation utility for all groups is $\bar{U} = 1$. By referring to the indirect utility function and, for the sake of simplicity, assuming that $\beta=(1-\alpha)$, it is possible to define the bid rent function $R_i(x)^*$ for each group that identifies the highest

[7] This extension is in the spirit of the discussion presented in Duranton and Puga (2015): consumers need to commute frequently for job and shopping purposes.

rent they are ready to pay for a unit of land at $x \in X$. This function takes the following form:

$$R_i^*(x) = [(w_i - tx - t(x-a)]^{\frac{1}{\alpha}}, \;\; for\ i = \{c, e, s\}. \qquad (4)$$

Equation (4) assesses that the auction process for land among the three groups assigns the land plot at $x \in X$ to the group with the highest $R_i(x)^*$. In this respect, at equilibrium, there exists a complete specialization of land if the three groups are featured by a different value for $R_i(x)^*$. This condition implies that the location $x \in X$ is exclusively assigned to only one of the three groups and this implies that $n_i(x)S_i(x) = 1\ for\ i = \{c, e, s\}$. If not, two or more optimal group-rents are identical because of identical group-wages. Hence, mixed solutions (the *integrated* district in Fujita and Ogawa (1982)) appear entailing a negative dependence between two (or more) densities. Ceteris paribus, the higher the wage, the higher the bid-rent. Overall, the rent is higher for the plots closest to the CBD and lower for the most remote ones. Finally, in order to learn more about consumer's density, we plug (3) and (4) into the complete-specialization equilibrium condition and we obtain:

$$n_i^*(x) = \frac{1}{\alpha}[w_i - tx - t(x-a)]^{\frac{1-\alpha}{\alpha}}\ for\ i = \{c, e, s\}. \qquad (5)$$

Equation (5) establishes that the population density reduces as we move farther from the CBD and, by construction, the other location A. This effect is shaped by the income of the group, the level of transport costs and the importance individuals assign to the land plot location in their utility function. The way these determinants will impact on the location decision of each single group will be the objective of our econometric analysis.

5. Econometric estimations

Our empirical exercise tackles our research question in two steps. First, we investigate the way in which Barcelona's urban spatial structure and neighborhood status shaped population density distribution within the city and in selected sub-communities. To encompass a sample of different type of citizen-groups in Barcelona, we privilege to focus the empirical analysis on ethnic communities (population – mostly

Catalan- Spanish and Immigrants) and two skill-type groups (illiterate and high-skill).[8] In light of our previous results, the second exercise seeks to assess whether segregation has occurred among communities in Barcelona by exploiting the spatial dependences in a panel data. An important part of our value added is to develop a spatial econometric analysis to a panel data.

We begin with estimating (6) that represent a linear approximation of (6): [9]

$$LnD_{jit}(x) = \alpha_0 + \alpha_1 Ln(x_{j0t}) + \alpha_2 Ln(x_{jat}) + \boldsymbol{\alpha_j X_{jt}} + \mu_j + \delta_t + \varepsilon_{jit}. \tag{6}$$

in which D_{jit} is the density of ethnic community (or group) i in spatial unit j at time t, x_{j0t} is the distance from j to the CBD whereas x_{jat} is the distance from j to location A (here the Port). Vector $\mathbf{X_{jt}}$ corresponds to the set of explanatory variables for spatial unit j at time t including density of bus-line service and income, time fixed effects (δ_t), spatial fixed effects (μ_h), and errors (ε_{jit}).[10] In accordance with our theoretical setting α_1 and α_2 are expected to be negative.

Our first set of regressions estimate Equation 6 for the complete pseudo-panel covering 1947–2011 (Table D.1 in the Appendix).[11] . The positive elasticity of the port and surrounding areas for Spanish-born residents and Plaça Catalunya for immigrants emphasizes that both communities have for decades only weakly valued such proximity in their location decisions because of the high rents of those areas precisely due to value of accessibility for the social and commercial activities already located there. More generally, fixed effects are also likely to embed a reputation effect for a spatial unit that could be easily refer to the *status* effect that kept unchanged in time, and, then, they drive this outcome.

[Table 1 Part I and Part II about here]

[8] About the relationship between ethnic and skill-degree groups refer to Appendix C.

[9] Unfortunately, we are not able to perform the estimations of the changes of the population density against the corresponding changes of the selected estimators because of our limited available numbers (10) of spatial units constant across time.

[10] In Equation (2) the distance-variables x_{tj0} and x_{tj1} are time-dependent. Their values changes over time for two principal reasons: the changes in the spatial-urban structure we are dealing with (and, therefore, the change of the centroids we are selecting for computing those distances) and the expansion of the urban territory across decades.

[11] Estimations were run for all immigrant communities as a whole. Performing the analysis for all of those groups separately could have prompted the problem of statistical significance due to problems of territorial representativeness of some groups.

Both parts of Table 1 highlight a sample of local determinants for the population distribution of different groups in 2001–2011.[12]

The improved availability of data during that period affords information at the neighborhood instead of only the district level. Hence, we are able to better qualify the neighborhood status. Then, we can take into account an index of local average income (*Ind_Income_per_cap*), the interaction term between that index and the distance from the CBD, and a measure of public transport represented by the density of bus lines *(Ln_BusLine_density)*. The index of individual local income, as a measure of the economic status of a specific neighborhood, proxies the expenditure capacity of people who decide to live there. According to the theoretical setting, this estimated coefficient is expected to be negative. Along similar lines, the interaction term is likely to better capture the tradeoff between income and distance, and when statistically significant, that term is expected to be positive due to the combination of the two previous negative effects.

Instead, the bus line service is perceived as a proximity transport infrastructure that turns to be very convenient for daily or highly frequent (short) trips, while the metro or train services are reserved for long daily trips as documented in Fernández i Valentí (2006).

However, between population and bus density endogeneity is possible since the density of bus line services might be associated with the demand for such services, informed by the size of the population living in the vicinity. We investigate that problem by performing a robust Durbin–Wu–Hausman test that confirmed the existence of endogeneity. To instrumentalize that variable, in the spirit of Card et al. (2014) we construct an ad-hoc instrument for any time t and for any neighborhood i by computing at any time t and for any neighborhood i the log of the sum of bus line density of the other neighborhoods (i.e., $j \neq i$) at time t. The rationale for that type of instrument is that it holds amid the effect of spatial proximity. The quality of public transport services in other districts—above all, the nearest ones—is somewhat associated with bus line density in a selected neighborhood, but not likely to be the principal reason to prefer

[12] Another interesting result to investigate will be the degree of persistence of one community in the same urban plot. In order to address this issue, one needs to dispose of more time-series data for each community at neighborhood level. Now, available data only cover two years (in one decade) and they are highly correlated. Then, we leave it for future research.

living in one neighborhood instead of others. To assess the robustness of the instrument, we ran an *F* test, and the instrument passed the test for the communities studied. [13]

Both parts of Table 1 present the results of our estimations and tests performed for Barcelona's overall population, Spanish-born residents, immigrants, plus additional evidence for illiterate residents and high-skill residents.

We performed estimations by considering fixed effects estimations and instrumental variables (IV) with (district) fixed effects. Although we use data at the neighborhood level, according to Barcelona's administrative structure, most services are provided at the level of districts, of which a few grant services based on historical tradition. In Barcelona, districts accumulate the memory of origins, traditions, and the economic and civil status of people living there, whereas neighborhoods have frequently changed over time. Therefore, to take into account that cultural background that affects the status of multiple neighborhoods, we include that type of fixed effect in our estimations.

To discuss the results of our estimations, the distance from the CBD is always statistically significant and the elasticity against population density negative, in line with the theoretical setting for a monocentric structure. Nevertheless, the magnitude of the estimation differs considerably across the skill-intensive subgroups. The elasticity of the high-skill group is the greatest by far. High-skill workers have clearly sought to settle in proximity of the CBD, which could indicate a gentrification process effected by that group of residents and, in turn, a segregation effect because they often record higher incomes than other groups. Estimation results also emphasize that the effect has been exclusive for the CBD (i.e., Plaça Catalunya) since estimations referring to the other point of interest (i.e., the Port) are not statistically significant when considering our two key locations, including district-fixed effects.

Both other variables referring to the economic status of a neighborhood are statistically significant in our preferred estimations. The index of income per capita present a negative coefficient, as the theoretical framework predicts. In line with results for distance to the CBD, the estimated coefficient for income replicates a similar framework, thus again suggesting a gentrification process concerning the high-skill group, largely because of the larger magnitude of the estimated coefficients. Estimations

[13] A discussion about the IV-exclusion restriction issue is provided in Appendix E.

referring to the interaction term also align with the expected results in reference to a monocentric setting.

Bus line density is statistically significant for all groups; the estimated coefficients are positive, and the estimated elasticity is in a range of 0.48–0.60. Therefore, the quality of public transport positively affect the location decisions of all residents.

The second set of estimations seeks to disentangle the potential spatial dependence in location decisions for our subgroups of communities. We aim to identify the extent some characteristics of the status of a unit *j* and its surrounding spatial unit(s) may have an impact on our variable of interest in the selected spatial unit *j*. The rationale of this exercise goes back to the idea of *status* of a neighborhood as discussed in Rosenthal and Ross (2015) as well as the concept of *hybrid* spatial unit in the Schelling's tradition. Taking into account the Barcelona urban spatial structure and the coexistence of different communities in Barcelona, we are interested in exploring their degree of integration intended as spatial co-location. To that end, we used spatial econometric techniques introducing group-interaction effects in the urban spatial model estimated previously. As suggested in Lesage and Pace (2009), the most convenient way to explore the potential existence of spatial dependence among observations is to begin with a spatial Durbin model (SDM).[14] One way to formalize a SDM to be estimated in our setting is

$$LnD_{jit} = \alpha_0 + \rho \boldsymbol{W} LnD_{jit} + \boldsymbol{\alpha} Ln\boldsymbol{X_{jt}} + \boldsymbol{\theta} \boldsymbol{W} Ln\boldsymbol{X_{jt}} + \delta_t + \boldsymbol{\mu_j} + \varepsilon_{jit} \qquad (7)$$

In (7) vector $\boldsymbol{X_{jt}}$ embeds to the set of explanatory variables for spatial unit *j* at time *t* (including the distances from unit j to the CBD and the Port, the income, the bus density and the density of the other community in the same spatial unit). In (7) $\boldsymbol{\mu_j}$ is a vector either of fixed effects to be estimated or, in the random effect case, a $N(0, \sigma_{\mu j}^2)$ distribution. In the set of explanatory variables, we include the distance from the centroid of each spatial unit to the two selected points of interests (i.e., CBD in Plaça Catalunya and Port of Barcelona). In that set of variables, we also include the density of other communities located in the same spatial unit *j*.

We begin by estimating the SDM by testing whether fixed or random effects were more appropriate. By construction, the SDM nests the spatial autoregression model (SAR) and spatial error model (SEM).

[14] Spatial econometrics techniques are widely used when taking into account the spatial interaction of location features (as in van Duijn and Rouwendal, 2013) or when exploiting the spatial features of panel data (as in Ciżkowicz et al., 2017).

On the one hand, the SDM and SAR encompass a spatial lag dependence for selected variables. On the other, the SEM assumes that the spatial lag dependence rested in the structure of the errors—namely, in the omitted variables embedded in the errors.

The last part of analysis is organized as in the previous step: we begin with performing estimation for the pseudo-panel for 1947–2011 (Table 2) and then we mainly focus on the last decade (2001–2011) –Table 3- in which the immigrant inflows is more important. To disentangle the potential spillovers effects arising across communities, our analysis will be performed by splitting the overall sample into three groups: Catalans,[15] Spanish-born residents, and immigrants.

[Table 2 and 3 about here.]

Results in Table 2 qualify the SDM model as the most suitable framework to identify the group-location determinants for the period 1947-2011. They also detect an important spatial-dependence component in the Spanish and immigrant groups. However, important changes appears in the most recent years. In Table 3, for the period 2001-2011, the SAR or SEM with fixed effects have to be preferred to the SDM.[16] Estimation results emphasize that spatial autocorrelation matter for the Spanish community, but not for immigrants or Catalan citizens, for which the spatial dimension becomes relegated to unobservable variables in the errors. In addition, the level of income per capita of the corresponding neighborhood turn to be an important determinant for the groups of Spanish and immigrant citizens, yet without playing any role for the Catalan community. Therefore, in the last decade communities lost their potential complementary dimensions; the presence of one community has acted as a deterrent for the others.

Hence, in line with the results of Rosenthal and Ross (2015) and Card et al. (2008), Barcelona has shifted from integrated neighborhoods to segregated ones. That dynamic has clearly been more important when referring to general profiles of the important mass of foreign citizens arriving in Barcelona to find work. In the 2000s, the composition of international immigrants is heterogeneous: an important share of that group came with higher education degrees.[17] For instance, some were chief executive officers of Barcelona-

[15] The Catalan group replaces the *population* variable to avoid collinearity problems in this exercise.
[16] The structure of the panel and estimation technique involved privileging fixed effects by neighborhood instead of by district.
[17] Fernández–Huertas Moraga (2014) reported that immigrants contributed to increased human capital in Spain, yet at a decreasing rate, during the immigration boom. Also, the average years of schooling of immigrants were always greater than those of natives.

based multinational or professional firms in the service sector. They developed their professional careers independently, and therefore, their potential complementarity - previously existing- with the other groups progressively disappeared.[18] The estimation results referring to the high- skill group presented in Table F.3 in the Appendix endorse that argument. Unfortunately, however, we could not disentangle high- versus low-skilled workers in the group of immigrants. We can proxy this effect by focusing on the density of all high-skill residents. Our findings indicate that (i) income has been a positive driving force in location decisions, whereas (ii) a self-selection tendency for high-skill workers to isolate from the illiterate group can be deduced from the negative marginal effects associated with the average density of illiterate residents in neighboring locations.

6. Conclusions

We develop an exploratory analysis about the degree of segregation in the coastal Mediterranean city of Barcelona, Spain, using information referring to the population composition of the city dating back to the mid-20th century.

The first set of estimations stresses that location determinants in population distribution have been historically associated with the economic and social interest of two important poles: Plaça Catalunya and the city's Port. In line with other results in the literature, features representing the status of a neighborhood matter in location decisions.

Our empirical analysis allows us to identify that though the effects of segregation have been self-reinforcing in recent decades in Barcelona, they more likely stem from skill and income levels, not ethnicity. In particular, one set of estimations using a pseudo-panel of data from 1947 to 2011 identifies the beginning of a gentrification process involving the group of high-skill workers.

For policymakers, our results provide an interesting insight. Recent urban segregation in Barcelona appears to have been principally associated with differences in

[18] According to data at hand, only 25.4% of immigrants in Barcelona before 1970 had a professional or administrative occupation involving a higher education degree. By contrast, in the 2000s more than 52% of the foreign population had a secondary or higher degree. Furthermore, the percentage of foreign migrants with a higher education degree was greater than that of natives: 15.7% against 12.7%.

citizens' skill instead of ethnic features. Public administration could reverse that trend and avoid the radicalization of this segregation process, by triggering the equal accessibility to education. However, implementing that strategy is not easy, and its effectiveness risks being limited, for the assignment of students to a school is not a neutral mechanism. For Barcelona and Catalunya, López-Torres et al (2017) conclude that there exists a strategic effect driving parents' school choices that do not depends only on the proximity of the school center to their residence, but also on quality and reputation effects enhanced by the territorial competition of the different educational centers. This second effect is more relevant when considering as strategic variable the family income. In the case of Barcelona, for instance, Calsamiglia and Güell (2018) discussed the problem in cases in which family background and income matter. Their data revealed that most households apply for schools in their neighborhoods and that only the most advantaged ones choose their preferred institutions. As it stands, the gentrification skill-driven landscape is self-reinforcing.

Nevertheless, our empirical analysis is not exhaustive. We explored some possibilities, but others remain to be investigated. To that end, researchers should develop a device to sort out income-type segregation mechanisms to examine the potential coexistence of different job complementarities among communities. One strategy could be to exploit available information about job occupation in order to approximate job skills and thereby identify the mentioned potential complementary patterns. However, additional data are required to be more conclusive about that issue.

7. References

Bayer, P., Hangming F., and McMillan, R. (2014): "Separate When Equal? Racial Inequality and Residential Segregation", *Journal of Urban Economics*, vol. 82, pp.32–48.

Boeri, T., de Phlippis, M., Patacchini, E. and Pellizzari, M. (2015): "Immigration, housing discrimination and employment", *Economic Journal*, Vol. 125, pp. F82-F114

Brueckner, J. and Rosenthal, S.S. (2009): "Gentrification and Neighborhood Cycles: Will America's Future Downtowns Be Rich?", *Review of Economics and Statistics,* vol. 91(4), pp. 725- 743.

Brueckner, J., Thisse, JF., Zenou, Y. (1999) "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory," *European Economic Review*, vol. 43(1), pp. 91-107, Burguess, S, Wilson, S., Lupton, R. (2005): "Parallelism ? Ethnic segregation in schools and neighbordhoods", *Urban Studies,* vol.42(7), pp. 1027-1056.

Calsamiglia, C. and Güell, M. (2018): "Priorities in school choice: The case of the Boston mechanism", *Journal of Public Economics*, vol. 163, pp. 20-36.

Card, D., Devicienti F. and Maida, A. (2014): "Rent-sharing, holdup, and wages: evidence from matched panel data", *Review of Economic Studies*, vol. 81 (1), pp. 84-111.

Card, D., Mas, A. and Rothstein, J. (2008): "Tipping and the Dynamics of Segregation". *Quarterly Journal of Economics*, vol.123(1), pp. 177-218.

Ciżkowicz, P., Ciżkowicz- Pękała, M, Pękała, P. and Rzońca, A. (2017): "The effects of special economic zones on employement and investment: a spatial panel modeling purposes", *Journal of Economic Geography,* vol. 17(3), pp. 571-605.

Conversi, D. (1990): "Language or race ? The choice of the core values in the development of Catalan and Basque nationalisms", *Ethnic and Racial Studies*, vol. 13(1), pp. 50-70

Cutler, D., Glaeser, E.L. and Vigdor, L. (2008a): "Is the melting pot still hot? Explaining the resurgence of immigrant segregation", *Review of Economics and Statistics*, Vol. 90(3), pp. 478-497

Cutler, D., Glaeser, E.L. and Vigdor, L. (2008b): "Whe are ghettos bad? Lessons from immigrant segregation in the United States", *Journal of Urban Economics*, Vol. 63, pp. 759-774

Damm, A.P. (2014): "neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants", *Journal of Urban Economics*, vol. 79, pp. 139-166.

de la Roca, J., Ellen, I.G. and O'Regan K. (2014): "Race and neighborhoods in the 21st century: What does segregation mean today?", *Regional Science and Urban Economics*, Vol. 47, pp. 138-151

Duncan, O. and Duncan, B. (1955): "A methodological analysis of segregation indexes", *American Sociological Review*, vol. 20, pp. 210–217.

Edin, P-A., Fredriksson, P. and Aslund, O. (2003): "Ethnic enclaves and the economic success of immigrants- evidence from a natural experiment", *Quarterly Journal of Economics*, vol. 118, pp. 329-357

Fernández i Valentí, R. (2006): *"100 años de autobuses en Barcelona (1906 -2006)"*, mimeo.

Fernández-Huertas Moraga, J. (2014): "Immigrant selection over the Business Cycle: the Spanish Boom and the Great Recession", *Cuadernos Económicos del ICE*, vol. 87, pp. 57-84.

Fujita M. and Ogawa H. (1982): "Multiple equilibria and structural transition of non-monocentric urban configurations" *Regional Science and Urban Economics,* vol. 12, pp. 161-196.

Glaeser, E. L., Kahn,M. E. and Rappaport, J. (2008): "Why Do the Poor Live in Cities? The Role of Public Transportation", *Journal of Urban Economics*, vol. 63(1), pp.1-24.

Glitz, A. (2014): "Ethnic segregation in Germany", *Labour Economics*, Vol. 29, pp. 28-40

Lesage, J. and Pace, R. (2009): "*Introduction to Spatial Econometrics*" CRC Press, Taylor and Francis Group, Boca Raton (FL), US.

Lewis, E. and Peri, G. (2015): "Immigration and the Economy of Cities and Regions" in G. Duranton, J.V. Henderson, W. Strange (Eds). *Handbook of Regional and Urban Economics,* Vol. 5A, Elsevier North Holland, pp. 625-685

López-Torres, L., Nicolini, R. and Prior, D. (2017): ""Does strategic interaction affect demand for school places? A conditional efficiency approach", *Regional Science and Urban Economics*, vol. 65, pp. 89-103.

Martori, J.C. and Apparicio, P. (2011): "Changes in spatial patterns of the immigrant population of a Southern European metropolis: The case of the Barcelona Metropolita Area (2001-2008)", *Tijdschrift voor Economisch en Sociale Geografie*, Vol, 102, pp. 562-581.

Moreno, L. and Arriba, A. (1996): "Dual identity in autonomous Catalonia", *Scottish Affairs*, vol.17, pp. 22-41.

Musterd, S. (2005): "Social and ethnic segregation in Europe: levels, causes, and effects", *Journal of Urban Affairs*, Vol. 27, pp. 331-348.

Musterd, S., Ostendorf, W. (2009): "Residential segregation and integration in the Netherlands" *Journal of Ethnic and Migration Studies*, vol. 35(9), pp. 1515 -1532.

Musterd, S., Marcińczak, S., van Ham, M. and Tammaru, T. (2017): "Socioeconomic segregation in European capital cities. Increasing separation between poor and rich", *Urban Geography*, vol.38(7), pp. 1062-1083.

Rangvid, B.S. (2007). " Living and learning separately? Ethnic segregation of school children in Copenhagen", *Urban Studies*, vol. 44(7), pp. 1329-1354.

Rosenthal, S.S. and Ross, S.L. (2015) : *"Change and Persistence in the Economic Status of Neighborhoods and Cities"*, in: G. Duranton, V. Henderson W. Strange (eds.). *Handbook of Regional and Urban Economics, vol. 5*, Amsterdam: Elsevier Publisher, pp. 1047-1120.

Schelling, Th. C. (1971): "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, vol.1, pp. 143-186.

Topa, G. and Zenou, Y. (2015): *"Neighborhood versus network effects"*, in: G. Duranton, V. Henderson W. Strange (eds.). *Handbook of Regional and Urban Economics, vol. 5*, Amsterdam: Elsevier Publisher, pp. 561-624.

van Duijn, M. and Rouwendal, J. (2013): "Cultural heritage and the location choice of Dutch households in a residential sorting model" *Journal of Economic Geography*, vol. 13 (3), pp. 473-500.

Table 1. Estimation results: panel (2001-2011) – Part I

	Ln population density			Ln Spanish-born density			Ln Immigrant density		
	FE	IV	IV(FE)[ab]	FE	IV	IV(FE)[ab]	FE	IV	IV(FE)[ab]
Constant	40.49*** (4.86)	16.96*** (3.61)		38.46*** (5.20)	15.25*** (3.72)		43.71*** (5.41)	21.49*** (3.67)	
Ln_Distance_CBD	-3.79*** (0.69)	-2.03*** (0.67)	-2.99*** (0.602)	-3.81*** (0.74)	-2.06*** (0.71)	-3.03*** (0.66)	-4.17*** (0.77)	-2.37*** (0.71)	-3.30*** (0.68)
Ln_Distance_Port	-0.35 (0.489)	0.78** (0.31)	-0.15 (0.41)	-0.26 (0.52)	0.87*** (0.33)	-0.06 (0.44)	-0.54 (0.55)	0.40 (0.34)	-0.32 (0.46)
Ind_Income_per_cap	-0.14*** (0.04)	-0.07* (0.04)	-0.12*** (0.032)	-0.14*** (0.04)	-0.08* (0.04)	-0.13*** (0.04)	-0.15*** (0.04)	-0.09** (0.04)	-0.14*** (0.04)
Ind_Income_per_cap*Ln_Distance_CBD	0.01*** (0.005)	0.008 (0.005)	0.014*** (0.004)	0.02*** (0.005)	0.008* (0.005)	0.014*** (0.004)	0.02*** (0.005)	0.01** (0.005)	0.02*** (0.004)
Ln_BusLine_density		0.56*** (0.1)	0.48*** (0.112)		0.58*** (0.11)	0.47*** (0.12)		0.61*** (0.12)	0.52 *** (0.13)
- Robust Durbin-Wu-Hausman test		20.09 ***			19.33 ***			12.50***	
<u>Instrument:</u>		*Log_BusLine_density_oth_distr*	*Log_BusLine_density_oth_distr*		*Log_BusLine_density_oth_distr*	*Log_BusLine_density_oth_distr*		*Log_BusLine_density_oth_distr*	*Log_BusLine_density_oth_distr*
First stage F-test		171.23***	252.25***		171.23***	252.25***		171.23***	252.25***
TIME DUMMIES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	District		District	District		District	District		District
F-test FE vs OLS	6.17***			5.28***			4.9***		
Errors	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust	Robust
R-squared	0.34	0.51	0.54	0.34	0.12	0.52	0.44	0.65	0.59
Obs	219								

*[a] Results for instrument-robust inference are available upon request; [b]Centered R-squared. Legend: *** 1%. ** 5%, * 10% degree of significance*

Table 1. Estimation results: panel (2001-2011) – Part II

	Ln Illiterate density			Ln high-skill density		
	FE	IV	IV(FE)[ab]	FE	IV	IV(FE)[ab]
Constant	30.12*** (5.43)	11.53*** (3.58)		51.96*** (6.61)	25.57*** (5.14)	
Ln_Distance_CBD	-2.85*** (0.79)	-1.30* (0.69)	-2.13*** (0.70)	-5.54*** (0.95)	-4.01*** (0.98)	-4.66*** (0.86)
Ln_Distance_Port	-0.24 (0.579)	0.52 (0.33)	-0.09 (0.49)	-0.35 (0.70)	1.35*** (0.48)	-0.158 (0.60)
Ind_Income_per_cap	-0.11** (0.045)	-0.06 (0.04)	-0.10*** (0.04)	-0.20*** (0.053)	-0.14*** (0.054)	-0.19*** (0.045)
Ind_Income_per_cap*Ln_Distance_CBD	0.011** (0.005)	0.005 (0.005)	0.01** (0.004)	0.02*** (0.006)	0.018*** (0.007)	0.02*** (0.005)
Ln_BusLine_density		0.50*** (0.124)	0.41*** (0.14)		0.58*** (0.14)	0.50*** (0.165)
- Robust Durbin-Wu-Hausman test		10.39***			11.51***	
Instrument:		*Log_BusLine_density_oth_distr*	*Log_BusLine_density_oth_distr*		*Log_BusLine_density_oth_distr*	*Log_BusLine_density_oth_distr*
First stage F-test		113.03***	158.73***		113.03***	158.73***
TIME DUMMIES	YES	YES	YES	YES	YES	YES
Fixed Effects	District		District	District		District
F-test FE vs OLS	3.23***			4.10***		
Errors	Robust	Robust	Robust	Robust	Robust	Robust
R-squared	0.65	0.76	0.74	0.43	0.59	0.58
Obs	146					

*[a] Results for instrument-robust inference are available upon request; [b] Centered R-squared. Legend: *** 1%. ** 5%, * 10% degree of significance.*

Table 2: Population location in Barcelona: pseudo -panel (1947-2011)
Estimation method: Quasi-ML;

	Ln Spanish-born pop. density	Ln Immigrant pop. density	Ln Catalan pop. density
Model	*(SDM with RE)*	*(SDM with RE)*	*(SDM with FE)*
Main			
Constant	1.33 (1.466)	2.52 (2.07)	
Ln Spanish-born pop. density		0.68*** (0.114)	0.82*** (0.06)
Ln Immigrant pop. density	0.19*** (0.053)		-0.031 (0.05)
Ln Catalan pop. density	0.84*** (0.07)	0.01 (0.178)	
Ln Distance from CDB	YES	YES	YES
Ln Distance from Port	YES	YES	YES
TIME DUMMIES	YES	YES	YES
WX			
Ln Spanish-born pop. density		-2.52*** (0.53)	-0.82** (0.33)
Ln Immigrant pop. density	-0.81*** (0.303)		0.42* (0.246)
Ln Catalan pop. density	0.76** (0.387)	1.42** (0.63)	
Ln Distance from CDB	YES	YES	YES
Ln Distance from Port	YES	YES	YES
TIME DUMMIES	YES	YES	YES
ρ	0.72*** (0.235)	0.88*** (0.229)	-0.42 (0.314)
RE vs FE (Hausman test)[19] (prob>χ^2)	0.186	0.992	22.39**
Fixed effects			District
SDM vs SAR test (Wald test)	57.42***	32.18***	70.17***
SDM vs SEM test (Wald-type test)[20]	44.63***	27.83***	64.82***
R-sq	0.94	0.96	0.93
Observations	80	80	80

*Legend: *** 1%. ** 5%, * 10% degree of significance.*

[19] H_0: Difference in coefficients not systemic.
[20] No linear form.

Table 3: Population location in Barcelona: panel (2001-2011)
Estimation method: Quasi-ML;

	Ln Spanish-born pop. density	**Ln Immigrant pop. density**	**Ln Catalan pop. density**
Model	*(SAR with FE)*	*(SEM with FE)*	*(SEM with FE)*
Ln Spanish-born pop. density		-0.04 (0.14)	-0.21* (0.11)
Ln Immigrant pop. density	-0.09*** (0.02)		-0.03 (0.034)
Ln Catalan pop. density	-0.07* (0.04)	0.02 (0.06)	
Ln Ind_Income_per_cap	-0.33** (0.144)	-0.83*** (0.271)	-0.18 (0.25)
Ln Distance from CDB	YES	YES	YES
Ln Distance from Port	YES	YES	YES
TIME DUMMIES	YES	YES	YES
ρ	0.56***		
λ		1.49***	-0.17
RE vs FE (Hausman test)[21] (prob>χ^2)	643.58***	23.56***	32.63**
Fixed effects	Neighborhood	Neighborhood	Neighborhood
SDM vs SAR test (Wald test)	5.96	5.06	2.67
SDM vs SEM test (Wald-type test)[22]	10.61**	5.98	2.54
SAR vs SEM test (LM robust)		LM Error: 173.8/LM Lag: 0.697	LM Error: 0.075/LM Lag: 0.007
R-sq	0.81	0.22	0.46
Observations	219	219	219

*Legend: *** 1%. ** 5%, * 10% degree of significance.*

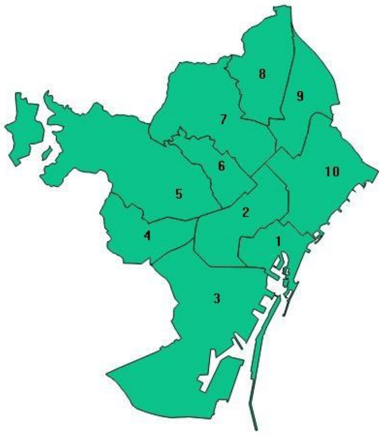
[21] H_0: Difference in coefficients not systemic.
[22] No linear form.

Appendix:

Segregation and urban spatial structure in Barcelona

December 2018

Figura A.1 : Map of Barcelona by district (as of 1984)



Legend:

1: Ciudad Vella
2: Eixample
3: Sants-Montjuic
4: Les Corts
5: Sarrià- San Gervasi
6: Gràcia
7: Horta-Guinardó
8: Nou Barris
9: San Andreu
10: San Martí

Appendix A. Bus line data organization

Available information dates to the beginning of the 1900s. From material available online (at http://www.autobusesbcn.es/), we selected all bus lines active on a regular basis in Barcelona for more than 12 months. That way, we excluded from our sample lines active only during the summer and experimental lines that ran for only a few weeks. We tracked the evolution of each line over time without distinguishing lines already in existence or that were tramways converted into bus lines.

Once all available lines were identified, for each year of our sample we mapped each bus route on the corresponding city plan and identified the districts or neighborhoods that they serve. Then, we totaled the number of lines serving one district or neighborhood.

For bus lines that connect destinations out of Barcelona, we exclusively recorded the urban part of their routes.

Appendix B. Descriptive statistics

Tables B.1 and B.2 present evidence that clarifies demographic and housing trends.

Table B.1: Selected descriptive statistics for variables, 1947–2011

	Variables	Mean	Std deviation	Min	Max
1947	Population density	32435.3	30459.4	3718.3	105036.2
	Catalan density	20158.54	17862.29	2416.125	61053.34
	Spanish density	11444.96	12143.3	1150.77	42248.57
	Immigrant density	625.29	571.79	95.08	1727.7
	# Bus lines per spat. unit	9.3	4.595	4	20
1965	Population density	30860.86	24242.79	6244.761	95114.23
	Catalan density	18289.5	13878.42	3866.109	52940.73
	Spanish density	12194.18	10335.82	2122.405	41232.76
	Immigrant density	377.18	264.54	67.7	940.73
	# Bus lines per spat. unit	15	6.1938	6	29
1970	Population density	29461.22	19413.46	6496.107	78182.12
	Catalan density	16365.44	10277.12	3639.27	40330.82
	Spanish density	10771.99	7963.62	1737.793	32537.72
	Immigrant density	374.44	227.99	80.99	848.06
	EU density	134.362	100.07	12.313	277.185
	Latin American Density	70.296	50.154	9.923	157.832
	Asian density	10.444	9.443	0.828	34.246
	African density	20.92	24.96	2.13	75.43
	# Bus lines per spat. unit	17.25	6.312	7	31
1986	Population density	26039.04	15926	189.66	52523.81
	Catalan density	17096.18	11019.37	129.569	37517.29
	Spanish density	8340.07	5078.27	51.96	16321.62
	Immigrant density	512.75	311.10	165.87	1140.8
	# Bus lines per spat. unit	9.71	5.73	1	29
1991	Population density	25056.7	15281.19	225.55	50151.43
	Catalan density	16629.99	10677.62	163.35	35291.36
	Spanish density	7657.69	4673.95	50.14	14912.89
	Immigrant density	643.29	395.73	212.59	1339.73
	EU density	80.912	73.10	0	296.724
	Latin American Density	58.036	46.417	0.1913	185.4545
	Asian density	59.709	109.71	0	656.36
	African density	50.94	95.35	0	431.83
	# Bus lines per spat. unit	9.5	6.09	1	30
2001	Population density	23469.06	14988.18	33.954	56885.65
	Catalan density	15191.92	9903.66	25.483	37750.7
	Spanish density	6421.87	4623.88	7.77	20451.48
	Immigrant density	1664.88	1353.04	517.72	4726.00
	EU density	166.35	156.14	0	581.95
	Latin American Density	579.19	509.73	0.07	1902.67
	Asian density	95.539	248.768	0	2044.62
	African density	160.27	214.88	0	1341.53
	# Bus lines per spat. unit	7.507	4.952	1	28
2008	Population density	25186.4	15673.6	77.57	59024.3
	Catalan density	14466.7	9225.4	57.97	36253.3
	Spanish density	10719.76	7337.75	19.60	29206.36
	Immigration density	4472.80	3352.25	1125.03	12283.29
	EU density	692.72	696.697	0.35	3148.46
	Latin American Density	1558.67	1357.6	2.940	6618.644
	Asian density	375.678	582.118	0.5207	4554.545
	African density	350.75	473.21	1.04	3230.97
	# Bus lines per spat. unit	8.342	4.969	1	29
2011	Population density	24956.64	15467.52	74.558	59442.82
	Catalan density	14448.05	9116.32	56.007	35713.85
	Spanish density	5163.304	3582.03	10.361	15455.7
	Immigrant density	4521.75	3113.65	1121.05	11223.9
	EU density	654.323	633.645	0.2100	2339.708
	Latin American Density	1296.551	1187.131	1.8202	6118.644
	Asian density	430.34	654.29	0.434	5099.27
	African density	337.28	414.31	1.04	2859.16
	# Bus lines per spat. unit	8.245	4.832	1	29

Table B.2: Density of buildings and real estate properties in number per km²

Variable	Mean	Std. Dev.	Min	Max
Stock dwellings up to 1950	4892.06	4940.48	35.76	1695.31
Dwellings built between 1951-1960	1726.71	1913.61	0.91	8776.37
Dwellings built between 1961-1970	3352.70	2965.72	1.12	11911.39
Dwellings built between 1971-1980	2952.71	2237.46	2.24	8326.15
Dwellings built between 1981-1990	762.17	582.05	1.26	2640.00
Dwellings built between 1991-2000	625.78	694.89	22.00	4039.00
Dwellings built between 2001-2011	497.95	422.32	0	1395.06
Stock vertical buildings up to 1921	65.20	106.24	0	482.73
Vertical buildings built between 1921-1950	46.73	63.50	0	242.85
Vertical buildings built between 1951-1970	22.46	18.70	0	64.71
Vertical buildings between 1970-1995	7.60	10.41	0	47.55
Stock horizontal buildings up to 1921	15.73	27.49	0	127.27
Horizontal buildings built between 1921-1950	14.25	18.13	0	76.19
Horizontal buildings built between 1951-1970	37.62	33.43	0.197	127.62
Horizontal buildings built between 1970-1995	34.82	31.34	0.197	113.99

Although our statistics rely on official data at the most basic territorial level, it was not always possible to access details about the ethnic composition of the population. From our data, we gleaned that, in 1947, Barcelona played host to small foreign communities, undoubtedly related to intense shipping activities associated with its commercial port. It was important Mediterranean hub, as well as one of the final stops on transatlantic routes that have served thousands of migrants traveling to North and South America (Table B.1).

In the when Barcelona started to host job-seeking Spanish migrants from elsewhere in Spain, followed by foreign immigrants after. Spanish-born citizens have tended to cluster in Barcelona's peripheral districts (e.g., Nou Barris, Sant Andreu, Sant Martí),[23] whereas immigrants have gravitated mostly toward more central districts (e.g., L'Eixample, Gràcia).

However, differences among the three groups have nearly disappeared in some other urban districts. A possible explanation relates to the progressive movement of households outside Barcelona's urban area to more metropolitan ones. Another relates to the real estate market. Table B.2 presents the evolution of building construction in Barcelona, the creation flows of which show a stable downward trend. Furthermore, in recent decades, the construction of horizontal real-estate properties, instead of vertical ones, has become a trend. A joint reading of both pieces of evidence suggests that the movement of people outside Barcelona's urban area aligns with arguments presented at the beginning of this section: that people move from municipalities to

[23] Refer to the map of Barcelona in Figure A.1.

improve their real estate properties, typically by privileging individual dwellings. Lastly, the rise in density of Spanish communities in the 2000s could be associated with the mentioned golden age of the Spanish economy and the boom of the construction sector that, once burst, induced rapid decay in the size of those communities.

Differences in location decisions appear in accordance with the different ethnic groups.

From 1991 to 2011, the immigrant share per district almost multiplied by ten, and their distribution pattern deeply changed. Nowadays, their presence is relatively important in town-center districts and in the part of the periphery closed to the Port of Barcelona and the maritime area. It is important to stress that the Port of Barcelona (with the surrounded area) consolidated as a prominent cluster of economic activities over time. Instead, Plaça Catalunya plays the role of center for social, cultural, leisure and service activities for the city (refer to Figure B.1).

Figure B.1: Share of foreign-born citizens by district as of 2010 (%) (source: our database)

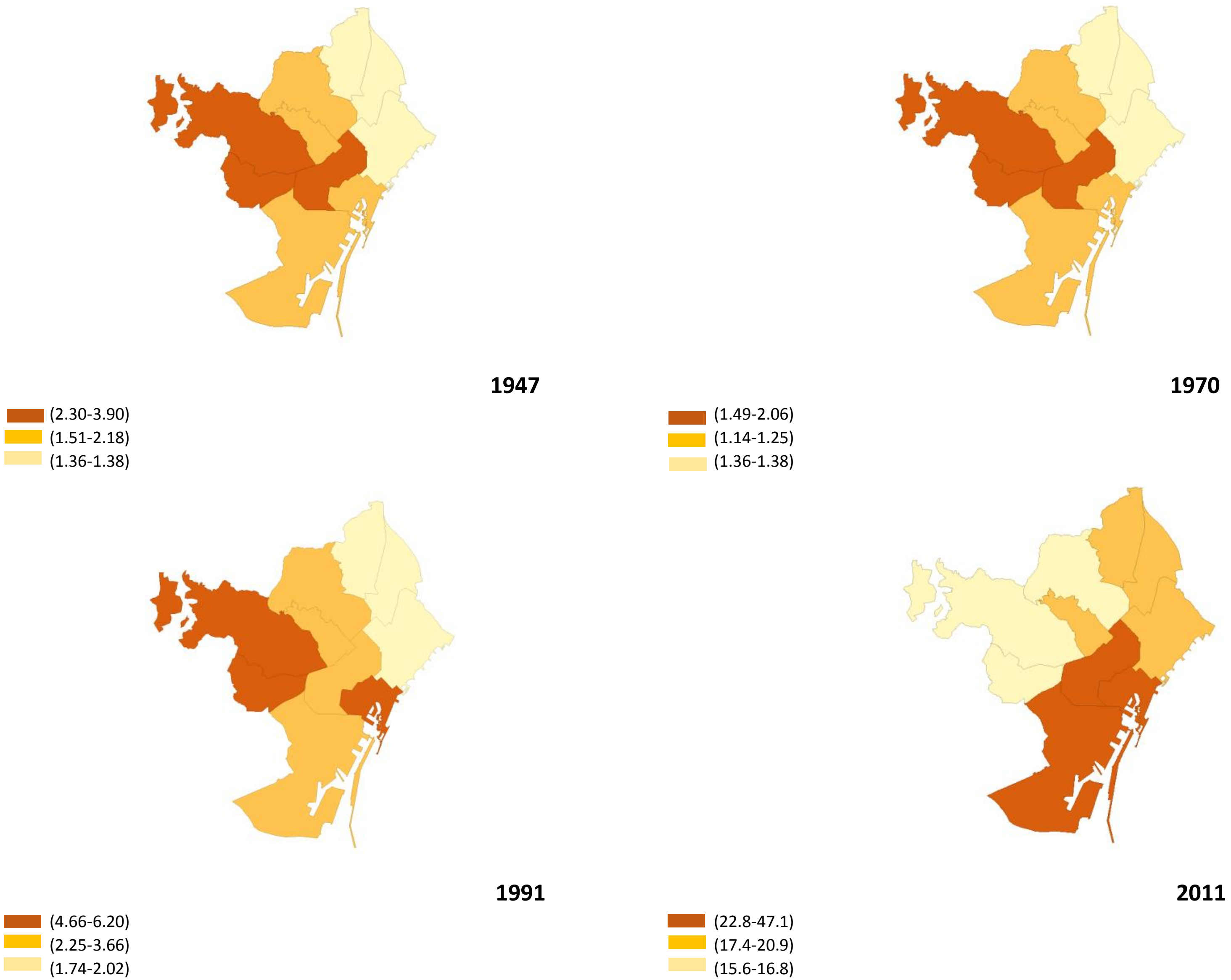
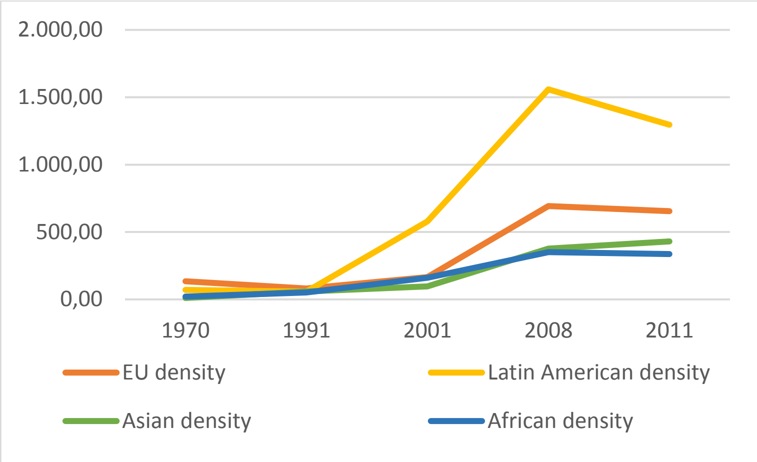


Figure B.2: Density growth of principal immigrant groups in Barcelona (source: our database)



The index of dissimilarity (Duncan and Duncan, 1955) is the most common measure of segregation when referring to an urban environment. Its principal advantage is to be independent from the population composition and it is quite reliable for comparisons over time. For a selected city at time t for any pair of ethnic groups (M, N) in a territorial unit i (for n units), Index D is constructed as follows:

$$D_t = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{M_{it}}{M_t} - \frac{N_{it}}{N_t}\right| \tag{1}$$

As presented in Equation 1, Index D assumes continuous values in (0, 1), with 0 being the most equal situation and 1 the most dissimilar. The interpretation is quite straightforward; the index provides a measure of the proportion of the population of ethnic group N that needs to be displaced in order to negate the degree of dissimilarity between M and N in neighborhood i. An Index D greater than 0.6 usually indicates the presence of a high degree of segregation in a city; with an Index D below 0.3, the degree of segregation in a city is low.

We compute Index D for the three major ethnic groups (Catalan, Spanish and Immigrants). Results appear in Table B.3:

Table B.3: Index of dissimilarity (Duncan, 1955)

	Catalan		**Spanish**
	Spanish	**Immigrants**	**Immigrants**
1947	0.05		
1965	0.10		
1970	0.12	0.34	0.43
1986	0.15		
1991	0.15	0.30	0.40
2001	0.15	0.19	0.26
2008	0.14	0.22	0.13
2011	0.14	0.22	0.22

Appendix C. Population composition

A further interesting assessment is the combination of skilled and unskilled persons in each ethnic group we consider in our analysis. Direct information about this issue is unavailable. However, we can perform a simple correlation exercise between the density of high-skilled group (by district) and the one of three ethnic groups (Catalan, Spanish and Immigrants) for the period 1947-2011. The Table C.1 presents the results including their statistical significance:

Table C.1: Correlations

	Log_dens_Catalan	Log_dens_Spanish	Log_dens_Immigrants
Log_dens_Spanish	0.8507***		
Log_dens_Immigrants	0.2507**	0.1826	
Log_dens_High_skilled	0.3849**	-0.0576	0.8872***

Level of Significance: **** 1%; ** 5%

Correlation results identify that there is a statistical significative positive correlation between the density of high-skilled persons and the density of Catalan and Immigrants by district. Rather the association between high-skilled persons and immigrants is higher than the one with the Catalan group. Therefore, these preliminary statistical results emphasize that the immigrant group is very heterogeneous as well as the high-skilled one. Put differently, it does not seem that the skill-background can be correlated with one specific ethnic group (above all natives, here Catalans)

Appendix D. Spatial-unit conversion

In Section 3, we discuss the strategy we adopted to create the pseudo-panel data. We follow the US Bureau Census TIGER/Line program (footnote 5). Our relationship files were built by following the method of the US Bureau Census that uses the TIGER/Line program, which aims to make, for instance, 2000 census tracts comparable to 2010 census tracts. More information on the method is available at https://www.census.gov/geo/maps-data/data/relationship.html.

In 2011, Barcelona was organized into 10 districts that can be adjusted and made to fit the urban territory for the previous years. However, as relationship files or conversion tools to match old districts with the 2011 districts or vice versa are unavailable, we have to devise a sort of conversion criterion. Using geographical points of reference, we identify an equivalence criterion that combines district boundaries and land surfaces. Then, we use those shares to convert the 1947 variables to the 2011 district boundaries as a weighted sum.

Below we present a visual example of the conversion process: we begin with an original map (here 1947) that we overlap to the most recent urban-district structure (here 1984). The result is the shadowed map that represent the urban structure in 1947 according to 1984-distric criteria.

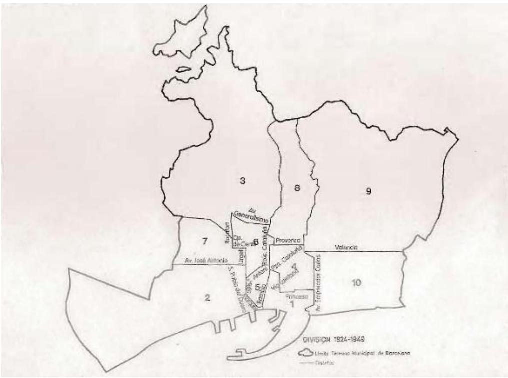


Figure D.1: Original map of Barcelona in 1947 by district (Source: Official Township Yearly Statistics)

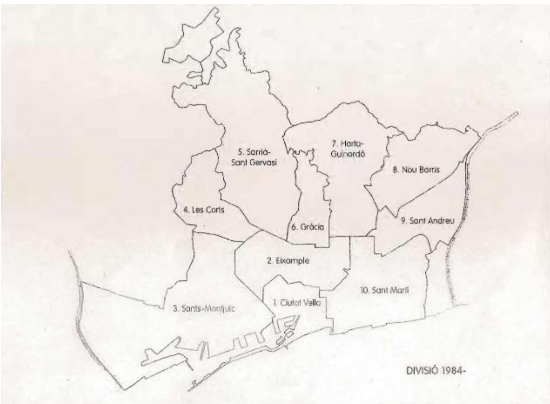


Figure D.2: Current structure of Barcelona by district (from 1984 onward) (Source: Official Township Yearly Statistics)

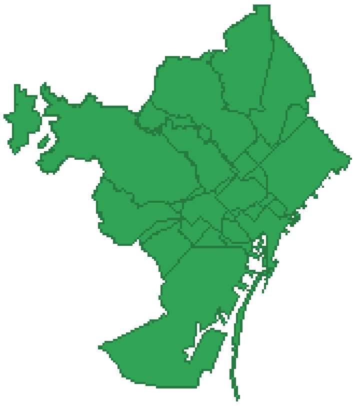


Figure D.3: Administrative structure of Barcelona in 1947 according to 1984-district organization (Source: own elaboration)

Table D.1 presents preliminary estimations results by using the pseudo-panel data

Table D.1: Estimation results: pseudo panel (1947-2011)

	OLS			FE		
	Ln Population density	**Ln Spanish-born density**	**Ln Immigrant density**	**Ln Population density**	**Ln Spanish-born density**	**Ln Immigrant density**
Constant	13.98*** (0.968)	12.26*** (1.29)	11.99*** (0.86)	2.04 (2.516)	3.28 (3.14)	3.34 (3.86)
Ln Distance_CBD	-0.39*** (0.129)	-0.20 (0.146)	-0.33** (0.137)	-0.07 (0.113)	-0.14 (0.141)	0.38** (0.173)
Ln_Distance_Port	-0.17** (0.08)	-0.28*** (0.102)	-0.47*** (0.106)	0.939*** (0.253)	0.728** (0.315)	-0.12 (0.388)
TIME DUMMIES	YES	YES	YES	YES	YES	YES
Fixed Effects				District	District	District
F-test FE vs OLS				41.56***	40.65***	14.40***
Errors	Robust	Robust	Robust	Robust	Robust	Robust
R-squared	0.30	0.30	0.86	0.49	0.65	0.94
Obs	80	80	80	80	80	80

*Legend: *** 1%, ** 5%, *10% degree of significance*

Estimations taking into account district fixed effects present a better level of goodness-of-fit, but most estimations of elasticities lose their statistical significance, or when they are significant, their coefficients turn out to be positive. District fixed effects are expected to capture at least part of the unidentified heterogeneity of our model that refers to the local economic and social environment that has driven people to locate in a particular urban neighborhood.

Appendix E. Exclusion restriction in IV strategy

One potential risk of the implementation of our IV strategy for estimations in Table 1 (in the main text) is the violation of the exclusion restriction. Then, we need to check that our instrument does not suffer from spatial autocorrelation. In order to address this issue, we computed the

Moran Index of our instrument for each of the three years of estimation (2001, 2008 and 2011) separately. Overall, the Moran I coefficient is around 0.15 and the pictures does not depict that our instrument suffers from spatial autocorrelation problem and, hence, our estimation IV strategy holds.

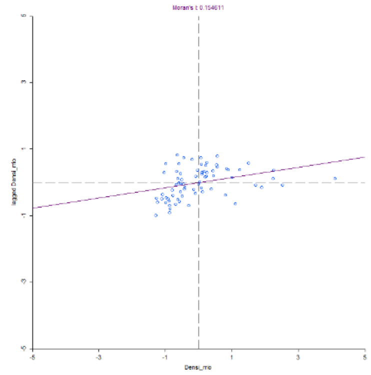


Figure E.1: Moran index for 2001 (0.15461)

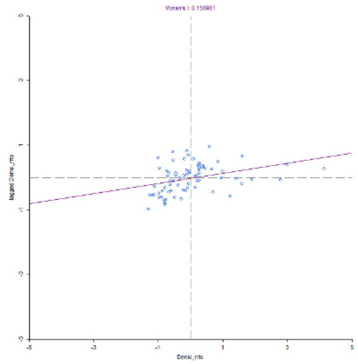


Figure E.2: Moran index for 2008 and 2011 (0.15698)

Appendix F. Spatial econometrics

From a technical viewpoint, we applied an SDM method for a spatial panel. In order to perform the estimation it is required running the *xsmle* command in Stata developed by Belotti, Hughes, and Piano Mortari (2016). Some tests were run by using the commands *spregsarxt* by Elhmessih Shehata and Mickaeil (2013).

Table F.1: Marginal effects: pseudo-panel 1947–2011
Estimation method: Quasi-ML;

	Ln Spanish-born pop. density	**Ln Immigrant pop. density**	**Ln Catalan pop. density**
	(SDM with RE)	*(SDM with RE)*	*(SDM with FE)*
Direct			
Ln Spanish-born pop. density		0.44** (0.172)	0.86*** (0.066)
Ln Immigrant pop. density	0.128** (0.056)		-0.048 (0.055)
Ln Catalan pop. density	0.96*** (0.083)	0.19 (0.213)	
Indirect			
Ln Spanish-born pop. density		-2.30** (1.10)	-0.55*** (0.128)
Ln Immigrant pop. density	-0.63* (0.336)		0.20* (0.113)
Ln Catalan pop. density	1.28*** (0.479)	1.65* (0.946)	
TOTAL			
Ln Spanish-born pop. density		-1.85 (1.244)	0.31*** (0.103)
Ln Immigrant pop. density	-0.50 (0.353)		0.15* (0.08)
Ln Catalan pop. density	2.24*** (0.533)	1.85* (1.09)	

*Legend: *** 1%. ** 5%, * 10% degree of significance.*

Table F.2: Marginal effects : panel 2001–2011
Estimation method: Quasi-ML;

	Ln Spanish-born pop. density
	(SAR with RE)
Direct	
Ln Immigrant pop. density	-0.092*** (0.020)
Ln Catalan pop. density	-0.07* (0.040)
Ln Ind_Income_per_cap	-0.33** (0.14)
Indirect	
Ln Immigrant pop. density	-0.04*** (0.01)
Ln Catalan pop. density	-0.03 (0.02)
Ln Ind_Income_per_cap	-0.14** (0.07)
TOTAL	
Ln Immigrant pop. density	-0.13*** (0.03)
Ln Catalan pop. density	-0.09* (0.06)
Ln Ind_Income_per_cap	-0.46** (0.19)

*Legend: *** 1%. ** 5%, * 10% degree of significance.*

Table F.3: Location decisions of high-skill residents: Panel, 2001–2011
Estimation method: Quasi-ML;

	Ln high-skill density
Model	*(SDM with FE)*
Main	
Ln Illiterate density	-0.103 (0.075)
Ln Ind_Income_per_cap	0.96*** (0.158)
Ln Distance from CDB	YES
Ln Distance from Port	YES
TIME DUMMIES	YES
WX	
Ln Illiterate density	-0.53** (0.235)
Ln Ind_Income_per_cap	-0.90 (0.626)
Ln Distance from CDB	YES
Ln Distance from Port	YES
TIME DUMMIES	YES
ρ	0.85*** (0.178)
RE vs FE (Hausman test)[24] (prob>χ^2)	371.08***
Fixed effects	Neighborhood
SDM vs SAR test (Wald test)	7.68**
SDM vs SEM test (Wald-type test)[25]	8.48**
R-sq	0.64
Observations	146
Marginal effects	
Direct	
Ln Illiterate density	-0.14* (0.075)
Ln Ind_Income_per_cap	0.95*** (0.153)
Indirect	
Ln Illiterate density	-0.64* (0.263)
Ln Ind_Income_per_cap	-0.067 (0.61)
Total	
Ln Illiterate density	-0.78*** (0.27)
Ln Ind_Income_per_cap	0.89 (0.647)

*Legend: *** 1%. ** 5%, * 10% degree of significance.*

References

Belotti, F., Hughes, G., Piano Mortari, A. (2016): "*Spatial Panel Data Models using STATA*", CEIS Tor Vergata Research Paper Series, vol.14(5), n.373, March.

Elmessih Shehata, E. A. and Mickaiel, S.K.A. (2013): "*SPREGSARXT*: STATA module to Estimate Maximum Likelihood Estimation Spatial Lag Panel Regression", available at https://ideas.repec.org/c/boc/bocode/s457762.html

[24] H_0: Difference in coefficients not systemic.
[25] No linear form.

Supplemental material:

Segregation and urban spatial structure in Barcelona

December 2018

A) Urban change

According to Busquets (2004), all southern European towns share the feature of complex urban development. In particular, moving from the traditional city center, each city follows a distinctive pattern of residential development. Barcelona is one such city. In the 1950s and 1960s, a consequence of massive migrant inflows, first from the rest of Spain and then from abroad, fueled the clustering of immigrant groups in peripheral areas of the city. Such migration gave rise to shantyism, or the creation of informal satellite communities that adjoin the established core of the city (L'Eixample), among other forms of peripheral growth. Shantyism is a direct consequence of the arrival of thousands of job seekers that Barcelona's formal real estate system could not accommodate, thereby allowing the amount of substandard housing to skyrocket.[26] Spreading from the hills surrounding the city up to Montjuïc, the seafront, and some spaces in L'Eixample, Barcelona's shanty communities were the first enclaves in which initial groups of immigrants began to cluster and, in a sense, mark a starting point of our analysis.

With data about dwellings properties, we sketched a general picture of the urban change that Barcelona has experienced. We focus on three milestone years:

- 1940, the end of the Spanish Civil War and beginning of the second dictatorial regime;
- 1970, the end of the high internal migration period; and
- 2011, a representative year of the current situation, following both the 1979 introduction of democratic municipal governments to implement urban planning and the real estate bubble in the middle of Spain's profound internationalization.

We referred to the urban structure in 2011, with 73 neighborhoods, and for each selected year, we mapped the percentage distribution of the stock of residences across the various neighborhoods. The three maps appear in Figure A.1.

The changing distribution of the stock of residences indicates that Barcelona has enlarged its urban territory over time in an inland direction. The urban core—the place with the highest concentration of dwellings—has followed that expansion. In 1920, the inner core was El Raval,[27] which now corresponds to the historical core of the city. The construction of new properties progressively displaced the residential barycenter away from the Roman perimeter outward. In 1940, the core residential neighborhood was L'Eixample, whereas in recent decades it has been Gràcia.

Along with that movement, the construction of residential dwellings in peripheral areas belonging to the city's external belt has increased. That trend has clearly aligned with an urban transformation spurred by the need to accommodate more national and international immigrants in those areas.

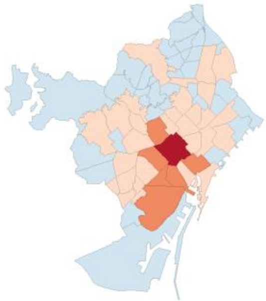
[26] Interesting material referring to this particular historical period is available at http://www.barraques.cat/en/, sponsored by the Museo d'Història de Barcelona.

[27] The map depicting 1900 data is not so different from the current one. For one, the core place of concentration (El Raval) remains unchanged. With a settlement dating back to the Roman origins of the city, El Raval has been part of the active commercial and civil life of Barcelona for centuries (Busquest, 2004).

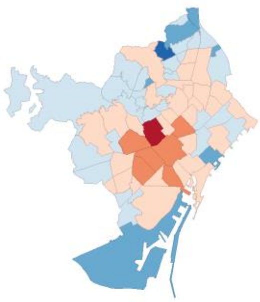
Figure A.1: Stock of dwellings

Legend. Highly concentrated areas: red-shadow areas; Low-concentrated (or empty) areas: light blue areas

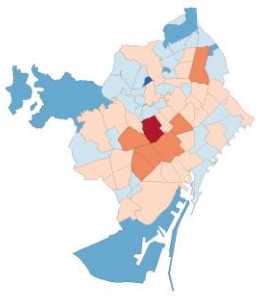
1940



1970



2011



Appendix B. Spatial statistics (G Statistics)

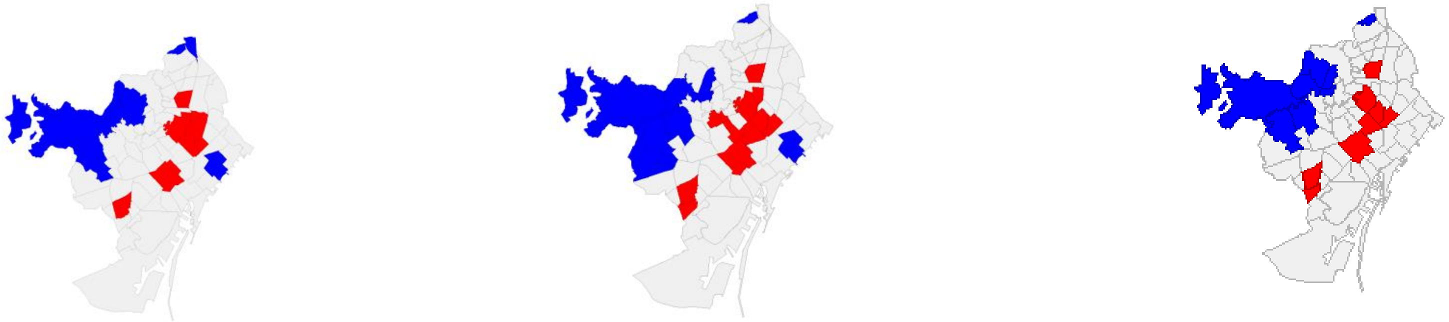
We performed a preliminary statistical analysis of density distribution in Barcelona's population and among selected ethnic communities, as well as of other relevant variables. We used local G statistics (Getis and Ord, 1992), which allowed us to identify correlations between proximate statistical values and to identify local clusters. Such statistics can be used as a measure of local association. We computed G statistics by considering the adjacency matrix as a weight matrix with row-standardized weights. Results appear in Figures B.1–B.8 The most remote neighborhoods have been more likely to be less densely populated, whereas the contrary holds true for central neighborhoods.

Concerning the educational level of Barcelona's inhabitants, we detected a persistently high cluster of graduates in the city center and clusters of illiterate persons in the periphery. Overlapping the maps depicting ethnicity and education revealed that a highly clustered area of graduates has coincided with highly clustered areas of Catalans and a part of the ones of immigrants. By contrast, Spanish have been prone to locate in the peripheral parts of city common to clusters of illiterate people.

Catalans have been highly clustered in the city center or central area in a north–south direction. Spanish-born residents have been more prone to cluster in scattered places, quite often in peripheral areas, but the specific location choices have been not persistent across time. Instead, the distribution patterns of immigrants do not vary a lot across the decade. Statistics determine that there is a strong persistent clustering area in the central and in the south-central urban neighborhoods above all in proximity to the commercial port.

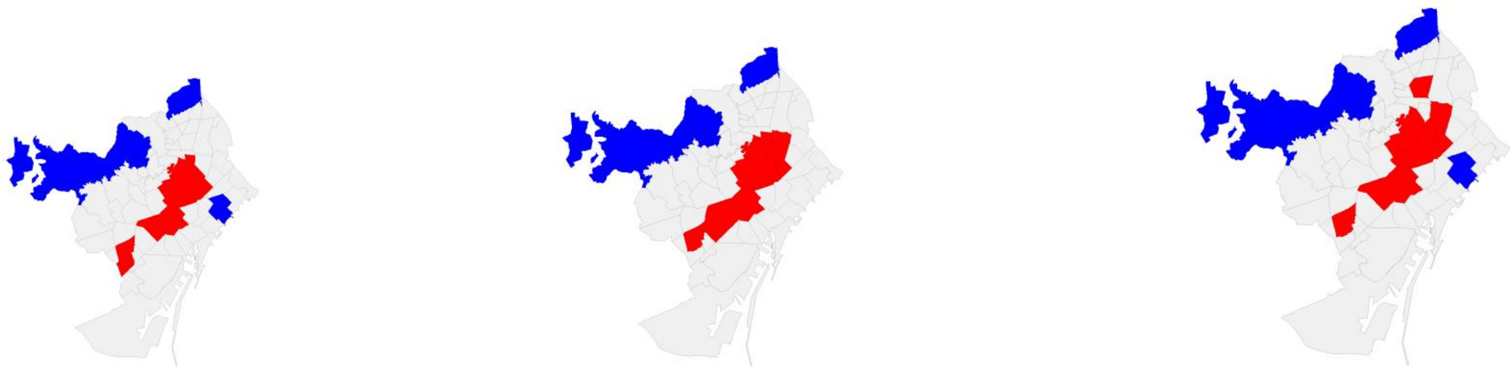
Lastly, we performed exploratory spatial analysis for two other variables: bus line density and income distribution. As concluded in Section 3, the distribution of bus line density has been quite scattered and not seemed to follow any particular clustering process. By contrast, income distribution has been strongly clustered and persistent; the center-west periphery of Barcelona emerged as a clustering area for high-income households and the northwest part as one for low-income households. High-income areas have also had low population density, and low-income areas have mostly coincided with high concentrations of illiterate people, as well as Spanish citizens and Latin American immigrants.

Figure B.1
Population density: Local G-Statistics
Source: Own database. Legend: Red: High; Blue: Low.



2001 **2008** **2011**

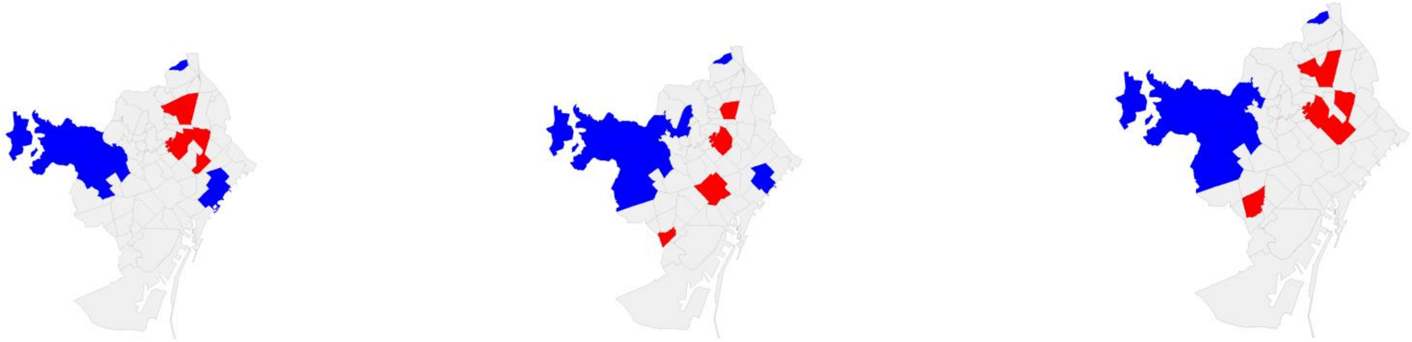
Figure B.2
Population density for people born in Catalunya: Local G statistics
Source: Own database. Legend: Red: High; Blue: Low.



2001 **2008** **2011**

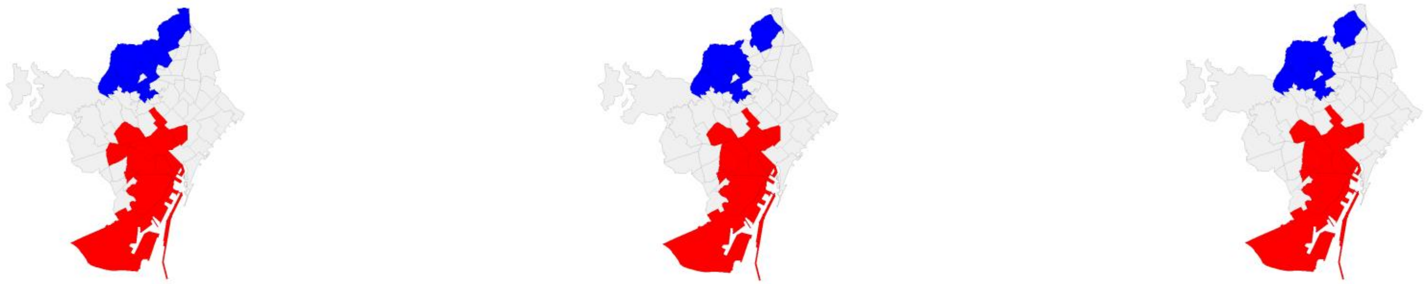
Figure B.3

Population density for people born in Spain: Local G Statistics
Source: Own database. Legend: Red: High; Blue: Low.



2001 **2008** **2011**

Figure B.4
Population density for Immigrants: Local G-statistics
Source: Own database. Legend: Red: High; Blue: Low.



2001 **2008** **2011**

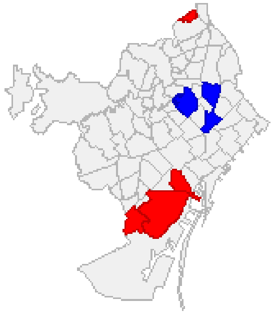
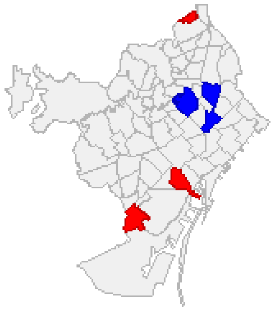
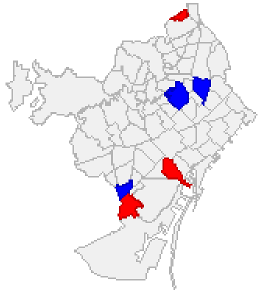
Figure B.5
Graduate density: Local G-Statistics
Source: Own database. Legend: Red: High; Blue: Low.

2001 **2011**

Figure B.6 Illiterate density: Local G statistics
Source: Own database. Legend: Red: High; Blue: Low.

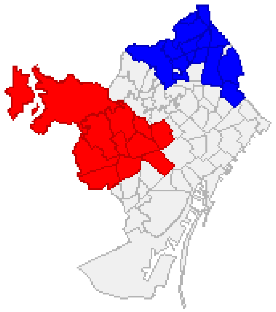
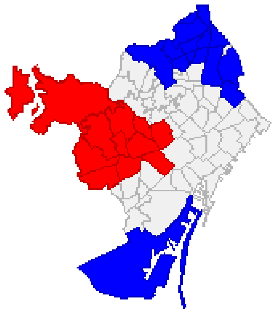
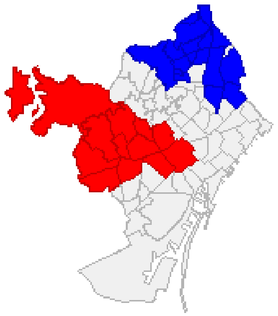
2001 **2011**

Figure B.7
Spatial density for bus-line indicator: Local G statistics
Source: Own database. Legend: Red: High; Blue: Low.

2001 **2008** **2011**

Figure B.8
Indicator for disposable income per capita: Local G statistics
Source: Own database. Legend: Red: High; Blue: Low.

2001 **2008** **2011**