I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11811

Does Juan Carlos or Nelson Obtain a Larger Price Cut in the Spanish Housing Market?

Catia Nicodemo
Josep Maria Raya

SEPTEMBER 2018

DISCUSSION PAPER SERIES

IZA DP No. 11811

Does Juan Carlos or Nelson Obtain a Larger Price Cut in the Spanish Housing Market?

Catia Nicodemo
Oxford University and IZA

Josep Maria Raya
Tecnocampus, UPF

SEPTEMBER 2018

ABSTRACT

Does Juan Carlos or Nelson Obtain a Larger Price Cut in the Spanish Housing Market?*

Using a unique dataset a non-parametric decomposition, we determined whether immigrants with native name, immigrants with foreign name and natives have different outcomes in Spain's housing market. Results suggest there are significant price discounts for immigrants with native names relative to immigrants with non-Spanish names. As a robustness check we prove that this is not due to the country of birth. We observe that most of the difference in price across immigrant groups remains unexplained, which may imply some form of discrimination (pure or statistical) against immigrants with non-native names.

JEL Classification: R1, R3, J7

Keywords: house price, price cut, discrimination, housing, matching method

Corresponding author:
Catia Nicodemo
University of Oxford
Centre for Health Service and Economics Organisation
Nueld Department of Primary Health Care Science
Radclie Observatory Quarter
Woodstock Road, Oxford, OX2 6GG
United Kingdom
E-mail: catia.nicodemo@economics.ox.ac.uk

* This work was supported by the Spanish Ministry of Economy and Competitiveness under Grant ECO2016-78816R, and Nicodemo received part of the funding by the European Union Horizon 2020 research and innovation programme under grant agreement no 727072 (REMINDER Project)

1. Introduction

Although living in racial and ethnic neighborhoods provides benefits to immigrants (foreign-born individuals; e.g., Gottlieb, 1987; Chiswick and Miller, 1996), it is also possible that discrimination may contribute to the segmentation of immigrants (Saiz and Wachter, 2011). Homebuyers, regardless of their status may pay a premium to avoid having immigrants in their neighborhoods. Discrimination could motivate a difference in the offers that sellers are willing to accept from potential buyers of different nationalities, or sellers may practice statistical discrimination[1] if race or ethnicity are correlated with experience in real estate bargaining or the availability of credit.

In this paper, we take advantage of a unique dataset to determine whether immigrant and native homebuyers obtain different price cuts (the difference between the list and the transaction price) in Spain. Over the past 15 years Spain had a significant number of immigrants. For example, the proportion of immigrants has increased from 1.4% of the population in 1995 to 11.4% in 2007, and 14% in 2008 (INE 2010). The dataset differs from standard housing datasets in two ways. First, in addition to a typical set of dwelling characteristics, the dataset also includes select characteristics of the buyer. Second, the dataset has both information about the list price and the transaction price. We investigate whether native (i.e. Spanish born) and immigrant (i.e. foreign born) buyers obtain different price cuts. Furthermore, we use first and last names to categorize homebuyers as natives, immigrants with native names (INNs), and immigrants with foreign names (IFNs). In this paper, we conduct a comparison of price cuts among these groups (INNs versus natives, natives versus IFNs, and IFNs versus INNs).

Finally, our study is also novel, as we use a matching method. King and Mieszkowski, 1973; Harding et al., 2003; Ihlanfeldt and Mayock, 2009, estimated a hedonic price model that included a dummy variable, indicating the country of birth of

[1] In this case, discrimination may exist and persist between demographic groups, even when economic agents (consumers, workers, employers, etc.) are rational and non-prejudiced. This type of preferential treatment is labeled 'statistical' because stereotypes may be based on the discriminated group's average behavior (Phelps, 1972).

the buyer. A major concern with this approach is the assumption that control variables may not adequately account for differences in neighborhood quality across racial and ethnic groups. To avoid this source of bias we used a non-parametric estimation, developed by Ñopo (2008), to deconstruct differences in price cuts received by natives, INNs, and IFNs into the portions that are due to differences in observable characteristics of both the buyers and the dwellings. An advantage of Ñopo's approach is the matching method that focuses exclusively on individual characteristics, and confines the effects of neighborhood quality to the unobserved portion of the differences in the price cut. Therefore, with our dataset, we can assess the difference in price cuts observed between immigrants and natives. [2]

Our results suggest that people with Spanish first and last names receive similar price cuts, regardless of whether they are natives or immigrants; however, we found evidence of significant price cuts for INNs, relative to immigrants with non-Spanish names. A large part of this difference is documented in the unexplained part, which may imply a form of discrimination (or disparate treatment) against immigrants with non-native names.

The paper is structured as follows. We start with an overview of the Spanish housing market, and the migration phenomenon in Spain. Then, we discuss the methods and the dataset. In the following section, we describe the estimation results and, finally, we end with some concluding remarks.

Literature Review

Usually, newcomers do not located uniformly across their host country; they tend to locate close to groups similar to them, increasing the concentration of immigrants in certain areas. This could influence the economic integration of migrants, the probability of finding a job, etc. While the, literature on the effects of the residential concentration of migrants on their economic and social integration in the US is large (Borjas, 1999; Card and Rothstein, 2007), the literature on these impacts in Southern Europe is scarce. This is mainly due to a lack of detailed data on racial and ethnic minorities and

[2] Furthermore, the Ñopo's method allows us also to consider the reverse causality generate by the chosen area of residence by immigrants, that usually tend to live toghether. With a parametric methods usually an instrumental variable is necessary to take into account this endogeneity.

immigrants, especially at the city level (Bisin et al., 2011).[3] In this study, we analyze the difference between natives and immigrants in one aspect of the housing market: the price cut in the sale price that the buyer receives. Furthermore, we take into account whether the first and last name of the immigrant buyer is similar to a native name.

Names are a rich source of information, with the potential to signal race, ethnicity, or class (Kasof, 1993). Several studies have found that names that are easy to pronounce are judged more positively by people (Mehrabian, 2001). In addition, names can differ in pronunciation across countries (Laham et al., 2012). For example, the Spanish pronunciation of the letter ‘j’ differs from both the English and Latin American pronunciation, while words starting with ‘s’ and followed by a consonant do not exist in Spanish, and are difficult to pronounce for native Spanish speakers. Immigrants with Spanish names may be perceived by natives as familiar, and more culturally integrated than other immigrants, and thus, they may be viewed more favorably (Fryer and Levitt, 2004; Biavaschi et al., 2017).

In this paper, we use information from the Spanish Institute of Statistics (INE, 2013) about the most common (first) names in Spain, to identify names that are foreign.. An immigrant named Javier or Juan Carlos, which are typical Spanish names, according to the INE classification, could give a signal to the other party in the transaction, and could provide an advantage over an immigrant named Nelson or Mohammed. Moreover, the name can potentially provide sellers with information about the nationality of the potential buyers’. Although several papers have analyzed discrimination in European rental markets (Ahmed and Hammarstedt, 2008; Bosch et al., 2010; Baldini and Federici, 2011), our study is the first to analyze transaction price differences in a Spain, and is unique in the larger literature on discrimination in its emphasis on the difference between

[3] There is a strand of literature that has analyzed the effects of immigration on natives in particular on public finances (Dustmann and Frattini, 2014; Dustmann et al., 2016), labor markets (Borjas 1999; Card, 2001; Lewis, 2003; Ottaviano and Peri 2006), the housing market (Sa, 2015), crime (Bell et al., 2013), and health (Giuntella et al., 2018). On the other hand, there is also a large literature concerning how immigrants are discriminated against natives in the labor market (Nicodemo and Ramos, 2013), access to health (Marshall et al., 2005), and the housing market (Diaz and Raya, 2014).

list and transaction prices. Therefore, we focus on potential disparate treatment in the price cut obtained after the bargaining process.

The Spanish Housing Market and Immigration

The booming Spanish housing market was one of the main engines of economic growth in the early 2000s. Growth in the construction sector accounts for around 20% of GDP growth (Raya, 2017). Housing construction in Spain was greater than construction in Germany, France, and Italy combined in 2006. From 2001 to 2008, the total number of dwellings in Spain increased by 20%, despite an increase of only 1% in the population. The peak in housing transactions occurred from 2005 to 2007. According to the Department of Public Works (DPW), a total of 349,118 housing transactions took place in 2011, compared to 706,928 transactions in 2006, and house prices for owner-occupied units in Spain tripled in nominal terms from 1998 to 2007.[4]

Migration flow in Spain

The Spanish housing market offers a unique opportunity to study how natives react to the inflow of immigrants. In the late 1990s and early and mid-2000s, many immigrants came in Spain, resulting in a proportion of 11.4% in 2007 and 13.8% in 2010 (OECD, 2008). As a result, the share of immigrants living in Spain in 1990 was 2.15, and by 2010, this share was 13.8%. No other country in the world saw this more than six-fold increase. In addition, in 2010, Spain ranked second among OECD countries, after the United States, in absolute numbers of annual immigration. In terms of the percentage of the total population that is foreign born (10.3%), in 2010 Spain follows Luxembourg (41.6%) and Switzerland (20.3%); however, the American, Luxembourgish and Swiss inflows were much more gradual (i.e., not concentrated in the early 2000s)[5]. Also, in 2011, 44,6% of Spanish immigrants bought a property[6].

---

[4] The house price index in Spain, as computed by the DPW, is based on appraisals. There is no house price index based on transaction prices.
[5] See Rodríguez-Planas (2012), and references within, to document these figures.
[6] See Censo de población y viviendas 2011 (INE).

Empirical Approach

One important aspect when analyzing and comparing the real estate market of immigrants with that of natives is to adequately consider that immigrants have housing preferences that substantially differ from those of natives. This leaves us with a decomposition problem. The first part of the native-immigrant gap in real estate market outcomes can be attributed to differences in the average socioeconomic background characteristics between the two groups. The second part is due to differences in average returns to these characteristics, which are specifically associated with migration background, and may reflect migrant-specific barriers (e.g., language or bargaining skills, or discrimination). To isolate these two parts, we employed a matching decomposition.

This approach is different from the OLS decomposition proposed by Blinder (1973), Oaxaca (1973), and Juhn et al. (1993). These approaches first place different weights on observations of the groups of interest (see Angrist and Pischke, 2008), and second, the matching decomposition does not specify the regression function as linear. Finally, the matching decomposition allows the resolving of the possible misspecification due to differences in the supports of the empirical distributions of the individuals' characteristics for natives and immigrants. In this regard, the matching decomposition imposes a common support restriction.[7] In contrast, linear decomposition assumes that estimations are also valid in regions of the data where there is no support for individual characteristics. It is important to note that imposing the usual conditional independence assumption is not necessary in this context. Any unobserved variable will contribute to the residual term, i.e., the unexplained part of the gap. For these reasons, we believe that the matching decomposition proposed by Ñopo (2008) better fits our study.

First, the regression function of price cut is no longer specified as linear. Second, the adjusted mean price cut is simulated only for the common support subpopulation. While this latter issue is largely recognized in the program evaluation literature, it has, until recently, not received much attention in decomposition analysis. For example, by not considering the common support restriction, the Blinder-Oaxaca (BO)

[7] Support is essentially the overlap between values of X for the comparison groups (defined by D=1 or 0). Matching only utilizes observations in the region of common support where matched observations can be obtained. That is, unmatched observations are excluded from the analysis. Therefore, we cannot estimate treatment effects (Y) outside the region of common support. As a result, estimated treatment effects may vary substantially, based on the matching method employed.

decomposition is implicitly based on linear extrapolation, and an 'out-of-support' assumption. Put another way, it becomes necessary to assume that the linear estimators of the outcomes are also valid outside of the support region of individual characteristics for which they were estimated.

A further advantage of this matching-based decomposition method is its straightforward interpretation of the explained and unexplained price cut estimates. The unexplained gap is often presented as the differential between prices paid by native and immigrant buyers who have otherwise similar characteristics. This simple interpretation of the unexplained gap is quite accurate in the case of the matching-based decomposition. Ñopo's (2008) approach allows us to decompose the difference in price cuts into the portion due to differences in characteristics, in the region of common support, along with the portion explained by the fact that one group has characteristics that are not shared by any individuals or dwellings in the other group. The remainder––the unexplained portion—is what is typically referred to as the portion of the difference that is caused by pure discrimination, and the portion explained by unobserved characteristics of the buyers (statistical discrimination) in the housing market. The matching approach is more precise in distinguishing between these forms of discrimination because focusing on the region of common support helps to reduce the effects of outliers, while reducing the sensitivity of the results to functional form assumptions.

Our main goal in this study was to determine whether there were any differences between the transaction price and the list price (price cut) provided to natives and immigrants when purchasing a house in Spain, distinguished also on the basis of immigrants with Spanish and non-Spanish names. Our variable of interest was defined as the 'relative price cut' = (list price – transaction price) ÷ list price. We decomposed the difference in the price cuts for natives and immigrants into the part due to differences in the explanatory variables (the explained portion), and the part induced by differences in the coefficients (the unexplained part).

Following the matching procedure, we decomposed the difference in the relative price cut for natives and immigrants. This method recognizes that the regions of support for the distributions of individual characteristics may differ for groups such as natives and immigrants. The approach also controls for differences in support for immigrants and natives, which may occur if immigrants can only afford houses located in specific areas, or if natives have individual characteristics that give them marked advantages in bargaining.

To illustrate the method, consider two groups, natives and immigrants; the differences in the distribution of the price cuts of these two groups, denoted by Δ, are defined as:

$$\Delta = \mathrm{E[Y|N]} - \mathrm{E[Y|M]} = \int_{S^N} g^N(x) dF^N(x) - \int_{S^M} g^M(x) dF^M(x) \quad (1)$$

where Y denotes the price cut and $g^M(x)$ and $g^N(x)$ denote the expected values of the price cut for immigrants (M) and natives (N), conditional on characteristics (x). $F^M(x)$ and $F^N(x)$ are the cumulative distribution functions of the individuals' characteristics, while $dF^M(x)$ and $dF^N(x)$ denote the respective probability measures. Finally, S^M and S^N denote the supports of the immigrant and native distributions of characteristics. As shown by Ñopo (2008), after partitioning the domains of the integrals, the price cut gap can further be expressed as:

$$\begin{aligned}\Delta = &\left[\int_{\overline{S^M} \cap S^N} g^N(x) dF^N(x) + \int_{S^N \cap S^M} g^N(x) dF^N(x)\right] - \\ &\left[\int_{S^N \cap S^M} g^M(x) dF^M(x) + \int_{S^M \cap \overline{S}^N} g^M(x) dF^M(x)\right] \qquad (2)\end{aligned}$$

The expression in (2) can be understood as four additive components of the total price cut differences (as described above):

$$\Delta = \Delta_N + \Delta_M + \Delta_X + \Delta_0 \qquad (3)$$

where Δ_X is the part of the price cut gap, or difference, that can explained by differences in the distribution of the individual characteristics of natives and immigrants over their common support, and Δ_0 is the unexplained part. Thus, the interpretation of these components is analogous to that of the components from the Oaxaca–Blinder decomposition. The traditional explained and unexplained components of the gap are defined only over the common support, i.e., they are calculated only based on those individuals whose combinations of individual characteristics are found among both natives and immigrants. Δ_N is the part of the price cut gap that can be attributed to the existence of buyer profiles for which there are natives, but no immigrants (i.e., out of the common support), Δ_M is the part that is due to the existence of buyer profiles for which there are immigrants, but no natives (i.e., out of the common support). An example of Δ_N is the fact that it is possible to find natives, but no immigrants, in the sample who are single, have higher education, a permanent labor contract, and who live in a new dwelling of 90 m^2 (968.75 ft2), located in a high-income neighborhood. The Δ_N term thus accounts for the part of the price cut gap that exists because natives have a combination of characteristics that is absent among immigrants. Similarly, the Δ_M term accounts for combinations of immigrant characteristics for which there is no comparable group of natives, e.g., an immigrant with only a primary education, working on a temporary contract, and living in an old dwelling of 50 m^2 (538.20 ft2) or less in a low-income neighborhood. In this sense, the values of N and M explain the difference within each group. Different values of N and M, therefore, represent the heterogeneity inside each group. What remains after accounting for these regions that lack common support is either explained by the differences in the distribution of individual characteristics of natives and immigrants over the region of common support (Δ_X), or remains unexplained, (Δ_0). Δ_0 is the unexplained part in the standard Oaxaca-Blinder decomposition, i.e.,$(\beta^N - \beta^M) \cdot X^M$, which is the difference attributable to unobserved characteristics, or to discrimination.

The four components using the matching procedure are obtained following these steps: Step 1 – select one immigrant from the sample (without replacement); Step 2 – select all the natives that have the exact same characteristics, x, as the immigrants previously selected (with replacement); Step 3 – with all the individuals selected in step 2, construct a synthetic individual whose price cut is the average of all of them, and match them to the original immigrant; and Step 4 – put the observations of both individuals (the synthetic native and the immigrant) in their respective new samples of matched individuals. Repeat Steps 1 through 4 until the original immigrant sample is exhausted.

We constructed two different specifications, using the following groups of variables, to create matches for those in our treatment group; obviously, the probability of finding a perfect match decreases as the number of characteristics is increased:

Set 1: Dwelling characteristics (see Table 1 for details): house size, availability of parking, number of rooms, age of construction, year and location dummies (postal codes).[8]

Set 2: Dwelling and individual characteristics: house size, parking, number of rooms, age of construction, education, age, marital status, permanent labor contract, household income, year, and location dummies.

Data

We used a unique dataset, obtained by a housing market intermediary with franchises in most Spanish provinces. The dataset covers the period from 2004 until 2006. This real estate company also has its own mortgage brokerage branch, which belongs to the same group. The company sales are around 4% of the total sales in Spain for the years studied. This market share is the highest of any Spanish real estate company (not including P2P Internet platforms, which do not have information on final prices, appraisals, mortgages, etc.). The Spanish real estate market is very fragmented, and there are regionally few companies that operate throughout the entire country. To assess the representativeness of our sample we compare the appraisal prices in our sample with the official figures of the DPW. The DPW data does not include the whole population of appraisals, but it represents the market quite well (85% of appraisals). As shown in the last two columns of Table 1, our sample has a mean appraisal price of €2,072/m^2, (€22,303/ft^2) very similar to the DPW figure, which is €2,140/m^2 (€23,035/ft^2). The table also provides a comparison of appraisal prices for some large cities, showing a close fit between both sources. Therefore, we believe that smaller places have similar small differences (except for sampling variability). Table 1 also compares our market prices with the prices registered by the INE (which come from the Official Registry of Real Estate Properties). The comparison is just a rough exercise, since INE does not provide prices by city, but only by regions, and we cannot compare the levels because the INE only provides an index based on the initial data of 2007. Notice also that these two prices do not represent the exact same measure. Our sample contains market prices, which are transaction prices. Price registered by the INE may be different from the transaction price,

[8] The exact definition of the variables is in Table 1.

if there are undeclared payments (a very frequent practice in Spain). Despite these definitional differences, the changes in price are remarkably similar.

The average income of the borrowers in our sample is similar to the average compensation in Spain, as reported by the Wage Structure Survey (INE). The dataset is unique as it not only contains the transaction price and the characteristics of the dwelling, but also the list price, time on the market, and characteristics of the buyer (see Table 1). Every observation pertains to a single transaction. That is, the final transaction price can be equal to, or lower than, the list price. The data were collected semi-annually from 2004 to 2006, resulting in 2,896 observations (1,826 from natives and 1,070 from immigrants).

[Insert Table 1 here]

The characteristics of the dwelling include housing size, the age of construction, a dummy variable indicating whether the dwelling has parking (included in the price), and the number of rooms. The characteristics of the buyer include age, level of education, household income, a dummy variable indicating whether the individual is married, a dummy variable indicating whether the individual has a permanent contract, and the country of birth. We defined the buyer as an immigrant if the country of birth was not Spain. We also controlled for the selling time and location. Table 2 shows the variables used in the study, along with their definitions. [9]

[Insert Table 2 here]

We constructed three groups, based on country of birth and names. The first group included natives only, the second comprised IFNs, and the third group is INNs. We used information from the Spanish Institute of Statistics (INE, 2013) about the most common names in Spain, to identify names that are foreign. The Spanish Institute of Statistics annually publishes, a classification of the most common (first) names (in Spain) of people from the 45 most relevant nationalities of origin who live in Spain (19 European, 15 North and South American, seven African and four Asian nations). The names are counted based on the information contained in the Statistical File of the Continuous Register on January 1 of each year. This classification suggests, for example, that Mohamed is the most

[9] Unfortunatly we do not have any information to compute the housing market area using different approaches as the travel-to-work areas, migration data, etc.

common male name for people from Morocco (90 out of every 1,000 people from Morocco in Spain), and Maria Fernanda (6.6 for every 1,000) is the most common name for a woman from Ecuador. Neither of these names is common among native Spaniards. On the other hand, the most common male name for an Ecuadorian, Juan Carlos (4.7 for every 1,000), is also a common Spanish name.

A similar procedure has been followed in the literature on discrimination by race and ethnicity (see, e.g., Bertrand and Mullainathan, 2003; Algan et al., 2013). In our dataset, the most common names are: for natives, Juan, José, and Maria; for INN, José, Luis, and Juan Carlos; and, for IFNs, Mohamed, Nelson, and Edwin. Notice that Nelson or Edwin could have just as easily been born in South America as José or Juan Carlos; Nelson and Edwin are IFNs. In the case of José or Juan Carlos, we assigned Juan Carlos to native or to immigrant, with native name depending on whether the individual was Spanish or not.

Table 2 presents descriptive statistics for the full sample, and for the three groups of buyers (natives, IFNs, and INNs). The dataset includes 1,826 natives and 1,070 foreign-born individuals, of which 492 are immigrants with a foreign name, and the remaining 578 are immigrants with a native name. Table 3 shows that the relative price cut (price cut relative to list price) is slightly higher for INN (5.33%) than for natives (5.23%), while the relative price cut is smallest for INN (4.99%). Also, from the descriptions in Table 3, we show differences in the characteristics of the homes bought by natives and immigrants. We show differences in house size, age, and parking: immigrants are used to living in smaller and older homes with no parking. Finally, we Table 3 shows that immigrant buyers are less educated, on average. Comparing the two groups of immigrants, education levels are higher for those with Spanish names. A key variable determining the potential bargaining power of a buyer is the type of labor contract, which can be either permanent or temporary. Table 3 also shows that the share of buyers with a permanent contract is higher for natives than for non-native buyers. Finally, the percentage of individuals that is married is higher for immigrants than for native buyers.

Table 3 also shows that dwelling characteristics are similar, but differences in individual characteristics are observed. Although IFNs have higher incomes, they have lower percentages of graduates, people with permanent contracts, and people being married.

[Insert Table 3 here]

Results

In this section we discuss the results of the empirical models for the conditional expectation of the price cut, presenting the decompositions of the differences in price cuts among natives, INN, and IFNs.

Determinants of the Price Cut

Table 4 shows the estimation of the relative price cut (greater than zero) received by buyers using OLS[10]. As expected, a longer time on the market is associated with larger price cuts. The significant coefficient (at a 5% significance level) in dwelling characteristics means different valuations for buyer and seller. That is, the buyer and seller give a different valueto this housing characteristic. In our case, larger price cuts were associated with older homes, i.e., the buyer and seller value the age of the dwelling differently. The rest of the dwelling characteristics are insignificant in explaining the price cut, i.e., they are equally valued among buyers and sellers. With respect to individual characteristics, higher levels of education and income for the buyer increase the price cut[11]. Being native or IFN does not have a significant effect on the degree of price cut, with respect to INN (which is the reference category).
[Insert Table 4 here]

Decomposition of the differences in price cut between natives, INNs and IFNs

The OLS estimations, which focus on conditional expectations, suggest that there are no significant differences among relative price cuts for natives, INNs, and IFNs, after controlling for individual and dwelling characteristics. As Ñopo (2008)[12] pointed out, we

---

[10] Results hold if we use a Heckit model to account for selection of buyers with zero price cut because the selection variable is not statistically significant in the Heckit model, so there does not appear to be selection bias in the OLS estimates.
[11] Quantile estimates (McMillen, 2008; Nicodemo and Raya, 2012) across the full distribution of price cuts are available on request. Our estimates show higher price cuts in the middle of the distribution for natives with respect to immigrants.
[12] The software used for our analysis is Stata, and we have programmed part of the code following Ñopo (2008).

must compare individuals with the same characteristics, since there are some combinations of characteristics that are typical of natives, but not for immigrants (either with native or foreign names), and vice versa.

In Table 5, we report the estimations of the relative price cut gap, and the percentage of matched and unmatched cases, for every pair of cases analyzed, and each set of control variables, using the matched method. The first column of Table 5 presents the difference in the average price cut received by each pair of groups (INN versus native, native versus IFN and IFN versus INN). As shown in Table 5, including an additional variable in the matching algorithm reduces the probability to find pairs. This tendency is one of the limitations of the matching method. Table 6 shows the results of decomposing the relative average price cut gap into the four components, Δ_N, Δ_M, Δ_X, and Δ_0. We present estimates for each of the three groups of individuals, and for the two sets of variables defined in the previous section. The confidence intervals are calculated from standard errors of the unexplained relative price cut gap estimates (calculated using the methodology of Ñopo, 2008), which are based on the assumption of normally distributed estimates.

[Insert Table 5 here]

The first column of Table 5 shows that the relative average price cut is 1.85% (Δ), which implies that immigrants with a Spanish name receive a slightly higher price cut than natives.[13] The component due to the unexplained portion (Δ_0) is small. This portion is slightly negative in Set 1, and becomes positive (but is still small) when we add more controls (Set 2). This positive discrimination can be explained by omitted or unobservable factors, such as immigrants' economic needs[14], motivation, expectations, etc.

The Δ_M term represents the portion of the average price cut gap due to the relative price paid by immigrant buyers outside the common support. Table 5 shows that this term is negative when housing and individual characteristics are included. This result suggests that the profile of immigrants with characteristics outside of the region of common support reduced the price cut gap by 4 percentage points.

The Δ_N term represents the part of the relative average price cut gap that can be explained by the fact that there are some combinations of characteristics of natives for

[13] The calculation for this figure is 0.0184=(5,326-5,229)/(5,326*277+5,229*878)/1155)

[14] Akin et al. (2014) suggested that lax lending standards and excessive risk taking during the boom. For example, some unemployed borrowers obtained loans for 100% of the transaction price..

whom there are no comparable immigrants. The Δ_N term switches from negative to positive after controlling for individual characteristics. This result implies that the profile of natives for whom there are no comparable immigrants increases the difference in price cuts between the two groups by 5 percentage points.

The final term in the decomposition, Δ_X, represents the portion of the relative average price cut that can be explained by differences in individuals' characteristics in the region of common support. Δ_X accounts for a significant portion of the relative price cut gap, particularly for Set 3, which combines characteristics of the individual and the dwelling. The negative value for Δ_X implies that natives have a set of characteristics that are associated with relatively larger price cuts than immigrants get.

In column two of Table 6, we estimated the difference in the relative average price cut gap received by natives, with respect to immigrants with a foreign name. The gap (Δ) is positive, and considerably higher—around 4.63%—meaning that natives receive a larger price cut than immigrants with a foreign name.

As discussed above, we decomposed this gap into four components. Focusing on Set 2, all components are positive except for the characteristics of immigrants (Δ_M), which remains negative after controlling for several sets of explanatory variables. The existence of certain profiles of buyers among immigrants that are not present among natives with foreign names increases the price cut gap by approximately 4 percentage points. Δ_N has the opposite effect; the existence of profiles of natives that are not present among IFNs reduces the price cut gap by approximately 4 percentage points.

The groups are similar in their observable characteristics. Indeed, less than 1 percentage point of the price cut gap is explained by differences in characteristics of natives and immigrants in the region of common support. Finally, more than 4 percentage points can be attributed to the effect of unobservable characteristics. To the extent that this result reflects the effects of discrimination (either statistical or pure), IFNs received smaller price cuts than natives, whereas immigrant buyers with native names receive *larger* price cuts. Furthermore, this portion is positive in Set 1 (1.9 percentage points), and becomes still more positive when we add more controls related to buyer characteristics (Set 2). Therefore, a higher price discrimination can be due to the correlation between the fact of being native or IFN, and other buyer attributes that affect the home sales process. Variables such as a permanent labor contract, income, and educational level capture a buyer's financial position, which might be correlated with transaction price. In this case, a higher price cut for the natives is may due to the fact that natives (with respect to IFNs) have better ability to bargaining and maybe a better the finance situation This may simply reflect differences in the buyer's income, wealth, or ability to secure mortgage financing. When we add buyer characteristics, however, this unexplained portion of the price cut gap increases, indicating that, the correlation between the fact of being native or IFN leads to an underestimation of the discrimination part estimated in Set 1. Thus, conditioning for income and wealth, and keeping the price cut constant, would lead to a higher discrimination part. This result is similar when we compare INNs with natives (in this case discrimination favors INNs), but not when we compare IFNs with INNs, this underestimation affects only the native-immigrant comparison[15].

Finally, the last column of Table 6 presents a more interesting result. This column shows the differences in price cuts across the two sets of immigrants. The results are as expected, when taking into account the information from the previous columns. Immigrants with a native name received price cuts that are approximately 6.5 percentage points larger than the price cuts received by immigrants with a foreign name. In this case, the unexplained portion (Δ_0) is positive, meaning that discrimination, or unobserved variables, explains a large (more than 6 percentage points) portion of the difference in the price cuts received by the two groups of immigrants. The difference in individuals' characteristics in the region of common support (Δ_X) is negative; the estimated value of Δ_X decreases the price cut gap by approximately 6 percentage points. This result implies that immigrants with a native name have characteristics that are associated with *lower* price cuts, compared to immigrants with a foreign name. Both factors, Δ_N and Δ_M are positive, which implies that the sets characteristics that are present among one group of immigrants/natives, and are not present in the other group between immigrants/natives, lead to an increase in the price cut gap between the two groups.

[Insert Table 6 here]

[15] Our data also permit us to refine this exercise with a Set 3, in which we can include lender fixed effects (commercial banks, savings banks, and non-bank financial institutions) and down payments. During the boom years (2000–2007), some lenders specialized in high-priced, or subprime, loans. Thus, lenders fixed effects proxies for the buyer's credit worthiness. In this case, the discrimination component increased in all three comparisons, reinforcing this underestimation effect; however, especially in the case where we compare IFN with INN, the number of Ñopo's matches are reduced. The results are available upon request.

Decomposition of the differences in price cut by country of birth

As a robustness check, and to reject the hypothesis that our results are driven by country of birth more than the names, we estimate the same model by immigrants groups in this section. Previous analyses lack some discussion concerning whether there is variation in price cuts across immigrants. In this sense, we need some evidence of whether the impact is driven by foreign name or by the country of origin. One may think that a possible explanation of previous results is that, immigrants with a native name are most likely to come from South America, with characteristics similar to natives and different to the other groups of immigrants. However, the results do not provide this evidence. We estimated a Ñopo decomposition on differences in price cuts between natives and immigrants from Latin America (LA) and between natives and immigrants from Africa (see Table 7).

Natives received price cuts that are approximately 1.0% larger than the relative price cuts received by immigrants from LA. Therefore, LA buyers received lower relative price cuts than INN buyers (by almost 3 percentage points); however, as in the case when we compare natives and INNs, the unexplained (Δ_0) and the explained parts (Δ_X) are small. The most important components are Δ_N and Δ_M (but with an opposite sign); i.e., set characteristics that are present among LA buyers are not present in native buyers.

Differences are more evident when we compare Ñopo's decomposition between natives and Africans, with respect to natives versus IFNs. Natives received price cuts that were 2.9% larger than the relative price cuts received by African buyers. In this case, the unexplained part (Δ_0) is almost negligible (-0.005) and favors African buyers. Compared to natives, two differences between Africans and IFNs are observed. First, IFN buyers received a lower relative price cut than African buyers (by 1.73 percentage points). Second, while in the case of IFN buyers, a larger part of this difference is attributed to the unexplained part, in the case of African buyers, this difference cannot be attributed to this part. The most important components are Δ_N, Δ_M (again with an opposite sign), and the explained part (Δ_X). In particular, the explained part is 5.7 percentage points, is the relative average price cut gap that can be explained by differences in individuals' characteristics in the region of common support. In summary, differences observed between INNs and IFNs cannot be explained by country of birth.

[Insert Table 7 here]

Conclusions and Discussion

In this paper, we analyzed the differences in the price cuts for owner-occupied units received by natives and immigrants in Spain. Our dataset is unique, as we gathered information on the buyer, the transaction, and the list price of the home, in addition to the standard characteristics of the home. The analysis provides new information on the disparate treatment of immigrants in a housing market in which non-natives have come to account for a larger portion of the overall market.

Our study takes advantage of the information on both the list and transaction price of the home to analyze the price cut obtained after the negotiation process between the buyer and seller. Controlling for the characteristics of the home, the neighborhood, and the buyer, we calculated whether immigrants received disparate treatment, relative to natives, when bargaining for a price cut. We analyzed three groups: natives, immigrants with a Spanish name, and immigrants with a foreign name.

Our analysis took advantage of a procedure developed by Ñopo (2008) to decompose the difference in the distribution of price cuts that can be explained by observed characteristics of the buyer. An advantage of this procedure is that it explicitly accounts for differences in the distributions of the characteristics of the buyers across groups. The procedure allows us to decompose the difference in the price cuts that is due to: 1) differences in the characteristics of the groups in the region of common support; 2) the difference, since one group has characteristics that are not shared by individuals of another group; and 3) a portion that cannot be explained by the observed characteristics.

We used this procedure to present evidence for a significant difference in price cuts for INNs relative to immigrants with non-Spanish names. We found that these differences were driven by the names, rather than the country of birth. A large part of this difference is due to unobservable characteristics, which is potentially an indication of a form of discrimination (either statistical, due to unobservables, or pure discrimination). In terms of statistical discrimination, these differences can be attributed to many unobserved factors, such as differences in buyer bargaining power, financial considerations, or differences in institutional knowledge (market experience, such as previous ownership); however, there are certain clues that suggest that some of the previous factors are, to some extent, controlled (unless in the comparison between INNs and IFNs). First, the sample of immigrants belonged to the same immigration wave (very close to the sample period). The average age and the lower standard deviation of immigrants suggest that there are no differences in institutional knowledge (in fact, we know that none of them had previous housing in Spain). In any case, the average age is similar between the two groups of immigrants. It is also difficult to find arguments for differences in the bargaining ability of both types of immigrants, once we rule out the hypothesis of a different area of origin, and after controlling for age, education, or income. Also, it is difficult to find arguments for differences in delinquencies among IFNs and INNs, even more during boom years in which too soft lending standards and excessive risk taking in Spain were documented (Akin et al., 2014). In any case, previous factors are not the only source of statistical discrimination. For example, search costs can be different between both groups of immigrants. Yet, the propensity of migrants to live together which insert special pressure on an area's prices.

No matter what the ultimate reason for receiving a lower price cut in the case of IFNs, the implication is that we have identified a source of disparate treatment in the Spanish housing market. In any case, these results indicate that INN receive favorable treatment in the Spanish housing market. The magnitude of the gap that we have observed in our data is substantial. In a highly competitive housing market in which sellers have few incentives to collect further information about potential buyers, sellers would tend to rank immigrants with non-Spanish names as worse buyers than immigrants with similar characteristics and native names, and therefore be less prone to provide them with price cuts.

Although this is the first paper that documents this type of disparate treatment, future research needs to better understand the mechanisms behind this fact. Discrimination has implications for a range of other social and economic outcomes, including the evolution of differences in wealth, home-ownership, and segregation among groups (which has additional consequences for educational and labor market outcomes).

In summary, Juan Carlos, an Ecuadorian buyer with similar characteristics to Javier, a native, obtains a similar price cut for a similar dwelling. However, both obtain a higher price cut for the same dwelling than Nelson, an Ecuadorian with similar individual characteristics.

References

Ahmed, A. M. and Hammarstedt, M. (2008). Discrimination in the rental housing market: A field experiment on the Internet. *Journal of Urban Economics*, 64(2), 362-372.

Algan, Y. Mayer T., and Thoenig M. (2013). The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France. IZA WP 107.

Angrist, J. D. and Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Baldini, M. and Federici, M. (2011). Ethnic discrimination in the Italian rental housing market. *Journal of Housing Economics*, 20(1), 1-14.

Bell, B., Fasani, F., and Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, 21(3), 1278-1290.

Bertrand M. and Mullainathan S. (2004). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. The *American Economic Review*, 94(4), 991-1013.

Biavaschi, C., Giulietti, C., and Siddique, Z. (2017). The economic payoff of name americanization. *Journal of Labor Economics*, 35(4), 1089-1116.

Black, D. A., Haviland, A. M., Sanders, S. G., and Taylor, L. J. (2008). Gender wage disparities among the highly educated. *Journal of Human Resources*, 43(3), 630-659.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 436-455.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2011). Formation and persistence of oppositional identities. *European Economic Review*, 55(8), 1046-1071.

Bosch, M., Carnero, M. A., and Farre, L. (2010). Information and discrimination in the rental housing market: Evidence from a field experiment. *Regional Science and Urban Economics*, 40(1), 11-19.

Borjas, G.J., 1999. The economic analysis of immigration. In *Handbook of labor economics* (Vol. 3, pp. 1697-1760). Elsevier. Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. Journal of Labor Economics, 19(1), 22-64.

Card, D. (2009). How Immigration Affects U.S. Cities. In: R. P. Inman (ed.), *Urban Enigma: City Problems, City Prospects*. Princeton, NJ: Princeton University Press.

Card, D. and Rothstein, J. (2007). Racial segregation and the black–white test score gap. *Journal of Public Economics*, 91(11-12), 2158-2184.

Chiswick, B. R. and Miller, P. M. (1996). Ethnic Networks and Language Proficiency among Immigrants. J*ournal of Population Economics,* 9(1), 19-35.

Diaz, L. and Raya, J. (2014). Mortgages, Immigrants and Discrimination: An Analysis of the Interest Rates in Spain. *Regional Science and Urban Economics*, 45, 22-32.

Dustmann, C. and Frattini, T., 2014. The fiscal effects of immigration to the UK. *The economic journal*, *124*(580), pp.F593-F643. Dustmann, C. and Görlach, J. S. (2016). The Economics of Temporary Migrations. *Journal of Economic Literature*, 54(1), 98-136.

Fryer, R. and Levitt, S. (2004). The causes and consequences of distinctly black names. *The Quarterly Journal of Economics*, 119(3), 767-805.

Giuntella, O., Nicodemo, C., and Vargas-Silva, C. (2015). The effects of immigration on NHS waiting times. *Journal of Health Economics*, forthcoming

Gottlieb, P. (1987). *Making Their Own Way: Shorthorn Blacks Migration to Pittsburgh*, 1916-30. Urbana: University of Illinois Press.

Harding, J. P., Rosenthal, S., Sirmans, D. F. (2003). Establishing bargaining power in the market for existing homes. *The Review of Economics and Statistics*, 85(1), 178-188.

Ihlanfeldt, K. and Mayock, T. (2009). Price discrimination in the housing market. *Journal of Urban Economics*, 66, 125-140.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3), 410-442.

Kasof, J. (1993). Sex bias in the naming of stimulus persons. *Psychological Bulletin*, 113, 140-163.

King, A. T. and Mieszkowski, P. (1973). Racial discrimination, segregation, and the price of housing. *Journal of Political Economy*, 81, 590-606.

Laham, S. M., Koval, P., and Alter, A. L. (2012). The name pronunciation effect: Why people like Mr. Smith more than Mr. Colquhoun. *Journal of Experimental Social Psychology*, 48(3), 752-756.

Marshall, K. J., Urrutia-Rojas, X., Mas, F.S., and Coggin, C. (2005). Health status and access to health care of documented and undocumented immigrant Latino women. *Health Care Women International*, 26, 916-936.

McMillen, D. (2008). Changes in the distribution of housing prices over time: Structural characteristics, neighborhood, or coefficients? *Journal of Urban Economics*, 64, 573-589.

Mehrabian, A. (2001). *Characteristics attributed to individuals on the basis of their first names*. Genetic, Social, and General Psychology Monographs, 127(1), 59-88.

Nicodemo, C. and Ramos, R., 2012. Wage differentials between native and immigrant women in Spain: Accounting for differences in support. *International Journal of Manpower*, 33(1), 118-136.

Nicodemo, C. and Raya, J. M. (2012). Change in the distribution of house prices across Spanish cities. *Regional Science and Urban Economics*, 42(4), 739-748.

Ñopo, H. (2008). Matching as a Tool to Decompose Wage Gaps. *Review of Economics and Statistics*, 90(2), 290-299.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14, 693-709.

OCDE (2008). *International Migration Outlook, Annual Report 2008*. Paris: OCDE.

Ottaviano, G. I. and Peri, G. (2006). The economic value of cultural diversity: Evidence from US cities. *Journal of Economic Geography*, 6(1), 9-44.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, 62(4), 659-661.

Raya, J. M. (2018). The determinants of foreclosures: Evidence from the Spanish case. *Papers in Regional Science* (in press). https://doi.org/10.1111/pirs.12316

Rodríguez-Planas, N. (2012). Wage and Occupational Assimilation by Skill Level. Migration Policy Lessons from Spain. *IZA Journal of European Labor Studies*, 31 Dec, 1:8.

Sa, F. (2015). Immigration and house prices in the UK. *The Economic Journal*, 125(587), 1393-1424.

Saiz, A. and Wachter S. (2011). Immigration and the Neighborhood. *American Economic Journal: Economic Policy*, 3(2), 169-188.

Table 1: Variable Definitions

Name of the variable	**Definition**
% Relative price discount	Price discount from list to sale price
Characteristics of the dwelling	
House size	Surface area in square meters of the dwelling
Age of construction	Age of the dwelling in years
Time on Market	Time from listing to sale (in days)
Parking	1 if the dwelling has parking included in price, 0 otherwise
Rooms	Number of rooms in the dwelling
Characteristics of the individual	
Age	Age in years of the buyer
Educational level	
Primary Education	1 if the maximum level of education is primary, 0 otherwise
Secondary Education	1 if the maximum level of education is secondary, 0 otherwise
Tertiary education	1 if the maximum level of education is university, 0 otherwise
Household income	Net monthly individual income in euros
Permanent contract	1 if the type of labor contract is permanent, 0 if temporary contract.
Married	1 if the individual is married, 0 otherwise.

Table 2: Descriptive Statistics

	Native	*IFN*	*INN*	*Whole*
Price Cut (%)	5.23	4.99	5.33	5.21[a]
House size	73.40	67.21	65.33	70.74
Parking	0.11	0.06	0.05	0.09
Rooms	2.75	2.86	2.87	2.79
Age of construction	30.11	37.27	38.55	33.01
Time on market	76.43	71.13	66.42	73.53
Individual characteristics				
Age	31.92	33.02	34.21	32.57
Married	0.29	0.38	0.49	0.35
Primary education	0.53	0.84	0.76	0.63
Secondary education	0.36	0.14	0.16	0.28
Tertiary education	0.11	0.02	0.08	0.09
Permanent labor contract	0.52	0.44	0.48	0.50
Household income	506.05	631.63	594.70	545.08
N	**1826**	**492**	**578**	2896

a. This mean is representative since mean comparison tests of price cut (%) between the three groups can not reject that the means are statistically equal.

Table 3: OLS estimation for the amount of discount

	OLS
Immigrants with native name (reference)	
Native	0.047
	(0.197)
Immigrants, foreign name	0.054
	(0.226)
Log house size	-0.277
	(0.338)
Parking	0.24
	(0.217)
Rooms	-0.013
	(0.126)
Age of construction	0.028***
	(0.005)
Time on Market	0.011***
	(0.001)
Individual Characteristics	
Age	0.012
	(0.009)
Secondary education	-0.268*
	(0.157)
Tertiary education	0.383**
	(0.168)
Household income	0.857***
	(0.269)
Permanent labor contract	0.14
	(0.148)
Married	0.087
	(0.056)

Time and Location dummies	YES
R-squared	0.45
F-statistic	38,68
N	2,896

*p<0.10**p<0.05, ***p<0.01

Robuts standard errors in brackets.

Table 4. Price Discount Difference. % Unmatched across Groups

	Difference in Price Discount (%)	% Unmatched, Group a	% Unmatched, Group b
a: Immigrant with Native Name b: Native			
Variable Set 1	1.85%	28%	50%
Variable Set 2	1.85%	60%	82%
a: Immigrant with Foreign Name b: Native			
Variable Set 1	-4.63%	28%	50%
Variable Set 2	-4.63%	60%	72%
a: Immigrant with Native Name b: Immigrant with Foreign Name			
Variable Set 1	6.50%	46%	48%
Variable Set 2	6.50%	76%	79%

Set 1dwelling characteristics, year and region dummies. Set 2 adds Education, Age, Married, Permanent Labor Contract, and Household Income to set 1.

Table 5. Ñopo decomposition

	INN versus N		**N versus IFN**		**INN versus IFN**	
Δ	**0.0185**		**0.0463**		**0.065**	
	Set 1	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
Δ O	-0.0071	0.0153	0.0189	0.0424	0.073	0.068
Δ M	0.0541	-0.0358	0.0416	0.038	0.030	0.0468
Δ N	-0.0234	0.0544	-0.025	-0.041	-0.065	0.0142
Δ X	-0.0051	-0.0153	0.011	0.007	0.027	-0.064
Std.dev	0.03	0.033	0.027	0.03	0.039	0.04

Set 1 adds to dwelling characteristics, year and region dummies. Set 2 adds Education, Age, Married, Permanent Labor Contract, and Household Income to set 1.

Table 6. Ñopo decomposition by region of origin

	N versus LA	**N versus A**
Δ	**0.0100**	**0.029**
Δ O	0.0087	-0.005
Δ M	-0.0608	-0.095
Δ N	0.0605	0.074
Δ X	-0.003	0.057
Std.dev	0.022	0.046