Immigration and the Neighborhood

Albert Saiz

Susan M. Wachter
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/real-estate_papers

Part of the Real Estate Commons

Recommended Citation
http://dx.doi.org/10.1257/pol.3.2.169

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/real-estate_papers/81
For more information, please contact repository@pobox.upenn.edu.
Immigration and the Neighborhood

Abstract
Within metropolitan areas, neighborhoods of growing immigrant settlement are becoming relatively less desirable to natives. We deploy a geographic diffusion model to instrument for the growth of immigrant density in a neighborhood. Our approach deals explicitly with potential unobservable shocks that may be correlated with proximity to immigrant enclaves. The evidence is consistent with a causal interpretation of an impact from growing immigrant density to native flight and relatively slower housing value appreciation. Further evidence indicates that these results are driven more by the demand for residential segregation based on ethnicity and education than by foreignness per se.

Disciplines
Real Estate

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/real-estate_papers/81
As with the African American migration from the South to the North of the United States in the first half of the twentieth century (Leah Platt Boustan 2007), the socioeconomic outcomes of recent waves of immigrants could be shaped by the attitudes of natives toward mixed neighborhoods and the resulting degree of residential segregation (Per-Anders Edin, Peter Fredriksson, and Olof Åslund 2003).

The existing economics literature on the impact of immigration has focused on the labor market. A number of studies (Kenneth F. Scheve and Matthew J. Slaughter 2003; Anna Maria Mayda 2006) find that native workers, who are more likely to be in direct competition with immigrants in the national labor market, tend to have negative views on immigration. However, most of the variance in attitudes towards immigrants remains to be explained (Kevin H. O’Rourke and Richard Sinnott 2001; Christian Dustmann and Ian P. Preston 2007). In fact, many of the citizens’ concerns, and much of the debate about immigration, seem to be focused not on national labor factor proportions but on very localized social interactions: the unrest in the Paris banlieues; the rise of anti-immigrant parties and violent groups in areas of dense immigrant settlement in Britain, Belgium, France, Russia, and other countries; the growth of new immigrant ghettos; the

---

Immigration and the Neighborhood

By Albert Saiz and Susan Wachter

Within metropolitan areas, neighborhoods of growing immigrant settlement are becoming relatively less desirable to natives. We deploy a geographic diffusion model to instrument for the growth of immigrant density in a neighborhood. Our approach deals explicitly with potential unobservable shocks that may be correlated with proximity to immigrant enclaves. The evidence is consistent with a causal interpretation of an impact from growing immigrant density to native flight and relatively slower housing value appreciation. Further evidence indicates that these results are driven more by the demand for residential segregation based on ethnicity and education than by foreignness per se. (JEL I20, J11, J15, R23, Z13)
use of the native language in the local community; and the approval of local ordinances related to immigration, just to name a few examples.

After all, immigration is not so much defined by the consumption of foreign labor, which can also be achieved by international trade, international outsourcing, or telecommunications. Immigration is truly defined by the physical presence of immigrants in the host country. While some residents in the country may conceptually oppose foreign trade, international outsourcing, or immigration, natives who do live in immigrant areas engage in further considerations: Are there preferences against living and socially interacting with people of different cultures, language, and ethnic or socioeconomic backgrounds?

If natives exhibit negative preferences toward interacting with immigrants, we may be able to capture this effect through residential choices and housing market dynamics. A vast literature has demonstrated the capitalization of local attributes on housing values (Wallace E. Oates 1969; Sherwin Rosen 1974) and used it to estimate the market valuation of neighborhood characteristics. A number of papers have used housing value differentials between African and European American neighborhoods to measure the extent of racially-based residential preferences.1

Previous papers (Saiz 2003, 2007; Gianmarco I. P. Ottaviano and Giovanni Peri 2006) have shown that immigration has a positive impact on average house value growth in metropolitan areas. This is a quite simple consequence of a local upward sloping supply of housing and population growth in the destination metropolitan areas. Since US citizens are mobile in the long run, Peri (forthcoming) argues that the positive impact of immigration on housing values may suggest a positive productivity effect of immigrants at the city level.

However, within metropolitan areas, sorting makes in unclear a priori whether values in the neighborhoods where immigrants settle should grow at a relatively faster rate. Even if immigrants have preferences for segregation, this should not necessarily imply higher prices in immigrant neighborhoods as long as there are mobile native price arbitrageurs. However, if natives have preferences for ethnic or socioeconomic segregation (Roland J. Benabou 1993) then immigration may actually be associated with a relative negative impact on neighborhood values.

We find evidence that, within metro areas, the growth of a neighborhood’s immigrant share is associated with relatively lower housing value appreciation. This empirical fact is consistent with the idea that natives are willing to pay a premium for living in predominantly native areas.

It is also consistent with reverse causality: immigrants may be attracted to areas that are becoming relatively less expensive. Therefore, in order to generate instruments, we introduce a spatial diffusion model (akin to a biological contagion framework) that provides predictions about the pattern of new immigrant

---

1 The list is too numerous for a detailed itemization: examples of this literature include Luigi Laurenti (1960), Martin J. Bailey (1966), A. Thomas King and Peter Mieszkowski (1973), Brian J. L. Berry (1976), George C. Galster (1977), John Yinger (1978), James R. Follain, Jr. and Stephen Malpezzi (1981), and Daniel N. Chambers (1992). The main thrust of this literature is to distinguish between discrimination against blacks in the housing markets (which implies higher housing prices in black areas) versus “decentralized racism” where white flight is the product of white preferences for racial segregation (which implies lower housing prices in black areas). A good discussion of these hypotheses (and of the alternative “port of entry” explanation for higher prices in minority areas) can be found in David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor (1999).
settlement. Neighborhoods that are spatially contiguous to immigrant enclaves are more likely to subsequently become more immigrant-dense themselves. We deal with potential omitted variables that could be correlated with the instruments by using heterogeneity in the predictive power of the geographic diffusion model as our effective source of identifying variation. For instance, the proximity of a neighborhood to existing immigrant enclaves is a stronger predictor of subsequent immigrant arrivals in metropolitan areas with larger immigrant inflows. This allows us to use the interaction between our measure of proximity to the enclaves and the level of immigrant inflows at the metro level as the effective instrument for expected new immigrant inflows into a neighborhood; hence, we explicitly control for both metro area fixed effects and proximity to existing immigrant communities in the second stage of our 2SLS specifications. Our approach could be made extensive to future research about localized social and epidemiological trends. Geographic contagion models can be used to explicitly model assignment to the treatment in many contexts. Here, we show an example of how researchers can deal with unobservable shocks that may be spatially correlated by providing plausible exclusion restrictions based on the structure of the spatial diffusion process.

The evidence is consistent with a statistically significant causal impact of immigration on neighborhood values. For instance, in an area where the share of the foreign born changes from 0 to 30 percent, housing values can be expected to be about 6 percent lower. This valuation reflects the tastes of the marginal native, and likely represents a lower bound for the willingness-to-pay for segregation of the average native. Consistent with a negative-amenity explanation, we also find that the impact of immigration is concentrated in high-income areas where most residents were non-Hispanic white prior to immigration shocks.

The results are important for understanding the impact of immigration on destination cities and, unfortunately, seem to identify potential challenges for the integration of immigrants. Indeed, recent research finds that immigrant segregation in the US has been on the rise during the last three decades of the twentieth century (Cutler, Glaeser, and Vigdor 2008). The new immigrant ghetto may be partially due to the tendency of immigrants to spatially cluster, but the paper shows that natives have preferences for avoiding immigrant areas. Why? Our final results shed some light on this issue. In our sample of immigrant-dense cities, the correlation (at the census tract level) between the foreign-born share and the share of adults with less than a high-school diploma is 0.49. The correlation between decennial changes in the share foreign born and decennial changes in the share of high-school dropouts is a notable 0.35. The association between changes in immigrant shares and the growth in the share of minorities at the census tract level is similarly strong. The fact that neighborhoods with growing relative concentrations of immigrants are becoming relatively less educated and less white can explain a good deal of the association between immigration and housing values. Areas with less educated populations and more minorities are being increasingly perceived as relatively less attractive places to live (Glaeser and Saiz 2004; John M. Clap, Anupam Nanda, and Stephen L. Ross 2008). Thus, immigrant neighborhoods may not be becoming relatively less attractive because they are populated by the foreign born per se, but because they are more likely to contain populations with perceived low socioeconomic status.
The rest of the paper is organized as follows. In Section I we discuss the data, and in Section II the core results. In Section III, we present further results relating to where and why immigration matters for the evolution of housing values and rents. Section IV concludes.

I. Data

We use census decennial data for the metropolitan areas of the United States at the tract level. A census tract is a small census-defined geographic level which, on average, encompassed a population of about 4,000 inhabitants in the 1990 and 2000 censuses. The version of the data that we use is provided by Geolytics Inc. Census tract geographic definitions change decennially. However, our data are processed so that we keep the geographic tract definitions constant over the years 1980, 1990, and 2000. Census tract and metropolitan statistical area (MSA) boundaries correspond to their 1999 definitions. Census tracts can be interpreted as a geographical measure of neighborhoods and have been used in this sense by previous research.

Several variables concerning the socioeconomic characteristics of the neighborhood and the quality of the housing stock are available and will be used (see Appendix Table 1). We are also able to match the census tract data to geographic data from the United States Geological Survey (USGS) on land use by tract in 1992. Distance to central business district (CBD) is calculated by the authors using the coordinates of the census tracts defined as CBD by the 1982 Census of Retail Trade.

We focus on metropolitan areas and years for which the decennial change in the number of the foreign born amounted to 5 percent, or more, of the MSA population in the previous census: we want to understand neighborhood dynamics in areas where immigration is an issue. In the 2000 census, for example, this included 67 metropolitan areas, encompassing 76.5 percent of all metropolitan immigration inflows (whereas the other 264 metro areas only accounted for 23.5 percent of new immigrants). However, the results are not sensitive to changing this threshold or to using the whole sample. Since many tracts were not included in the 1970 census, we will focus on the last two decades (1980, 1990, and 2000). Overall, we have 34,835 tract observations in 122 MSA-year groups.

The data gives us the best available coverage of the geographic patterns of settlement of immigrants in the United States and their evolution over time. The data is also the most comprehensive source for changes in housing values by neighborhood over the two decades under consideration. Furthermore, for 1990 and 2000, we are able to create immigration counts by origin country and tract using

---

2 The interested reader should consult the online Appendix. In the spatial epidemiological models that form the core of the IV strategy, the patterns of diffusion of relatively small numbers of new immigrants into areas with almost nonexistent significant immigrant concentrations are highly random, and not worth modeling explicitly.

3 Other housing data sources do not identify geographic location, or are limited to a few recent years and a small number of metropolitan areas. None of them identify foreign-born status of individuals moving into a location. We have used housing transaction records from the San Francisco metropolitan area to check on the robustness of our data. The correlation between average values in recent transactions by census tract and the census self-reported average values is a very high 0.95 across 1,008 tracts in the San Francisco CMSA, both in 1990 and 2000. Using changes in sales values across tracts between 1990 and 2000 yields similar results to using the census price data in our specifications as applied to the San Francisco area.
published census tract crosswalks. We then use census micro data (IPUMS) to

cross-tabulate foreign-born status by country with other characteristics (education,

income, ethnicity, English proficiency). We can, therefore, infer local immigrant

average characteristics by tract. Using that information on immigrant character-

istics we examine the implications of immigrant heterogeneity for neighborhood
dynamics.

Finally, we also use micro data on housing quality attributes and foreign-born

II. Immigration and Neighborhood Values

Following conventional racial segregation models (Bailey 1959; Thomas

C. Schelling 1971; Yinger 1976; Paul N. Courant and Yinger 1975; Yoshitsugu

Kanemoto 1980) we are interested in knowing whether changes in a neighborhood’s

immigrant share are related to local changes in home values. One can extend the
classical model\(^4\) to show that immigration may exert a positive effect on average
housing values at the metropolitan level while, due to sorting, also generating relative
price declines in the neighborhoods where immigrants tend to settle. Across
neighborhoods within a city, an empirical positive association between relative
changes in values and changes in immigrant density is hard to interpret. However,
a negative association (controlling for other location and housing quality attributes)
provides an unequivocal sign of native preferences for segregation. Intuitively, a
non-arbitrage condition ensures that prices cannot be lower in a location unless there
is a perceived negative compensating differential: otherwise opportunististic natives
move in until the price gap is bridged.

Empirically, we follow the evolution of average housing values in the census
tracts of high-immigration metropolitan areas in the 1980s and 1990s. In Table 1,
we start by regressing the inter-census (10-year) change in the log of the average
house values in a neighborhood on the change in the share of the foreign born. Using
changes in housing values, and the share of the foreign born, helps us control for
time-invariant omitted variables related to neighborhood quality, the relative valu-
ation of which stays constant across decades. The first-differences model that we
estimate takes the form:

\[
\Delta \ln P_{i,M,T} = \alpha_{M,T} + \lambda \cdot \Delta(ISH_{i,M,T}) + \Delta Z_{i,M,T} \cdot A + X_{i,M,T-10} \cdot B + \xi_{i,M,T}.
\]

Subscripts \(i, M,\) and \(T\) are for neighborhood (census tract), MSA, and year, respec-
tively. \(P_{i,M,T}\) is the average house value in the neighborhood, \(\alpha_{M,T}\) are a group of
MSA-by-year fixed effects (we concentrate on the impact of immigration within a
metropolitan area and year), \(ISH_{i,M,T}\) is the immigrant share, \(Z\) is a vector of housing
stock traits, and \(X\) is a vector of initial socioeconomic characteristics. The regres-
sions are weighted using the initial number of owner-occupied housing units in the
neighborhood as weights, and standard errors are clustered at the tract level. The

\(^4\)This we do in the online Appendix that accompanies this paper.
A. OLS Results

Column 1 in Table 1 shows OLS regressions that only control for MSA-year fixed effects. In column 2, we control for contemporaneous changes in the observable characteristics of the housing stock, because home values are also a function of these factors.
of the physical attributes of the housing units in a neighborhood. While changes in observable housing characteristics may be endogenous to immigration, we want to focus on the impact of immigration on quality-adjusted housing values. Despite the fact that the census offers a relatively small number of quality indicators, our analysis of the American Housing Survey unveiled that additional quality attributes are not markedly different between immigrant and native owner-occupied properties (they are in fact, very slightly better in immigrant-owned properties). More importantly, transitions toward immigrant ownership are never associated with subsequent declines in structural quality attributes.

Column 2—our baseline specification—also controls for the initial housing characteristics and other lagged socioeconomic neighborhood variables in levels. We do not believe in a model where lagged variables in levels have an infinitely durable impact on growth rates, but the valuation of place-specific characteristics had been changing in the last part of the twentieth century, and some of these initial variables are good predictors of subsequent housing value growth. The initial values of the socioeconomic variables should therefore capture evolving trends in the valuation of preexisting neighborhood traits and partially capture the impact of social trends that are unrelated to subsequent immigration levels. The variables that we use in this baseline specification, a total of 44 which we augment later, are specified in Appendix Table 1. The results suggest that a change of one percentage point in the share of immigrants in a neighborhood is associated with a relative decrease of roughly 0.25 log points in the neighborhood’s average housing value.

It is well known that housing values tend to mean-revert (Karl E. Case and Robert J. Shiller 1989; Stuart S. Rosenthal 2008). Likewise, we know that immigrants tend to locate in areas with initially low housing values. We therefore include, in column 3, the initial log of housing values to allow for mean-reversion. More generally, this variable may capture heterogeneity in the evolution of values in neighborhoods of different initial housing quality (which might, for instance, be affected by widening income inequality). Furthermore, immigrants may find more affordable those areas in which housing values are trending down. To mitigate these concerns, column 3 also includes, on the right-hand side, home value growth in the previous decade, and controls for the change in the log of income in the previous decade. The results of the main variable of interest do not change much after the inclusion of preexisting economic and housing trends.

Classical tipping models (à la Schelling 1978) suggest nonlinearities in the impact of minority concentrations. If relatively minor immigration inflows forecast bigger inflows in the future, most of their impact may be concentrated in the initial stages of the process of immigrant settlement. In our data, higher-order polynomials on the change in foreign-born density are never economically significant. This can

---

5 The interested reader is referred to the online Appendix for details.
6 We obviously do not control for changes in socioeconomic characteristics of the neighborhood, since these are endogenous to immigration. In other words, immigration has an impact on housing values because the attributes of the individuals who move into the neighborhoods (the new immigrants) are different. We will think of this impact as the relevant treatment effect of immigration as per current immigration policies. Later, we will discuss through which channels the treatment effect of immigration on local prices may work.
7 The interested reader can consult the online Appendix for further robustness tests.
be appreciated graphically in Figure 1, a scatter plot where the change in the share of the foreign born appears on the horizontal axis and the change in the log of housing values on the vertical axis. To facilitate interpretation, we present averages of both variables within each of 150 quantiles of the variable in the x-axis (about 200 tract-year observations by group). Both variables are partialled out of the baseline controls in Table 1, and the line displays the prediction from an OLS regression. The graphic evidence is fairly consistent with an approximately linear pattern of changes in values.

B. IV Empirical Approach and Results

There are two reasons why immigration inflows may be endogenous to the contemporaneous evolution of neighborhood values. One is reverse causation. Immigrants may be looking for affordable housing and avoid those areas where home values are growing faster than the MSA’s average. The second reason is omitted variables. Moving costs are sunk for newly arriving immigrants. They are, initially, very mobile. Immigrants may tend to select the best new locations in the city: those locations that are experiencing improvements in public goods or amenities, or nicer, high-quality new housing developments. Or, they might be attracted to neighborhoods with improving job prospects. That would lead to an overestimate of the association between the growth in the foreign-born population and neighborhood valuations. Alternatively, omitted variables, such as the changing valuation of neighborhood characteristics that are correlated with immigration, could bias the relevant coefficient downward. While using first differences in housing values should do away with much of the impact of omitted neighborhood characteristics, diverging trends may still be an issue.

FIGURE 1. GRAPHIC RELATIONSHIP BETWEEN CHANGES IN IMMIGRANT DENSITY AND PRICES
To deal with reverse causation and omitted trends, we would optimally like to randomly assign immigration shocks into a group of neighborhoods and analyze the subsequent evolution of housing values. We devise an instrumental variable strategy that tries to emulate that ideal experiment. Immigrants tend to cluster in proximity to where other immigrants live, which is a very well-documented fact both in sociology and economics (George J. Borjas 1995; Markus M. Mobius 2002). There are many reasons for this, most of them having to do with the advantages of proximity to people in the same national, ethnic, linguistic, or socioeconomic group. We take advantage of immigrant clustering to partially predict the patterns of new immigrant settlement in US metropolitan areas. In our model, neighborhoods that are geographically close to existing immigrant enclaves have a higher probability of becoming immigrant areas in the future. We start by defining a variable that proxies for the appeal of a neighborhood to immigrants using the following gravity equation:

$$\text{Pull}_{i,T} = \sum_{j \neq i, j \in M} \left( ISH_{j,T-10} \cdot Area_j / (d_{ij})^\beta \right).$$

$\text{Pull}_{i,T}$ is our estimate of the immigrant geographic gravity pull of a neighborhood $i$ (which is located in a metropolitan area $M$) at time $T$. $\left( ISH_{j,T-10} \right)$ is the share of immigrants in neighborhood $j$ in the previous census, $\text{Area}_j$ is the area (square miles) of the corresponding $j$th census tract, and $d_{ij}$ is the Euclidean distance between neighborhoods $i$ and $j$. Our measure of gravity is a weighted average of lagged immigrant densities in neighboring communities, where the weights are directly proportional to the area of neighboring tracts and inversely proportional to their distance from the relevant neighborhood. We do not have strong priors on the exact magnitude of the coefficient of spatial decay, $\beta$, and so we let the data convey that information.$^8$

The econometric intuition for the identification assumptions can be formally considered using an example with two types of neighborhoods. Neighborhoods of the type $C$ are close or contiguous to existing immigrant enclaves, whereas type $F$ consists of neighborhoods located far from the immigrant enclaves. Denoting with an upper tilde variables that are partialled out of the rest of the explanatory variables and dropping MSA and time subscripts for simplicity, the main equation of interest is:

$$\Delta \ln P_i = \lambda \cdot \Delta ISH_i + \xi_i.$$

Since we are not certain that $\text{cov}(\Delta ISH_i, \xi_i) = 0$, we can use the empirical knowledge that immigrant enclaves tend to expand to contiguous neighborhoods to add the following immigration diffusion equation:

$$\Delta ISH_i = \delta \cdot D_i^C + u_i,$$

$^8$See the online Appendix for details. None of the results is sensitive to the choice of $\beta$ (for instance, identical results are obtained with the Newtonian assumption $\beta = 2$).
where $D_i^C$ denotes a dummy variable that takes value one if neighborhood $i$ is contiguous to an immigrant enclave, and zero if it is located far away. Given this model, under the assumption $\text{cov}(D_i^C, \xi_i) = 0$, $D_i^C$ can be used as an instrument for $\Delta ISH_i$. Analogously, in Table 1, column 4, we present the results of a regression where we use immigrant gravity, $Pull_{i,T}$, directly as an instrument for the change in the immigrant share in a neighborhood. Indeed, neighborhoods that were located close to previous centers of foreign-born settlement were more likely to attract new immigrants subsequently (the $F$-test for the excluded exogenous variable is 76.01). The results are not significantly different from OLS, and become almost identical when we control for lagged immigrant density (column 5), in order to account for the fact that many neighborhoods with high $Pull_{i,T}$ were immigrant-dense initially themselves.

A potential caveat of this naive instrumental variable approach hinges on the exogeneity assumption of $Pull_{i,T}$. It is possible that previous immigrants were attracted to neighborhoods with characteristics that were becoming relatively less valuable to natives, and which are also spatially correlated (note that amenity levels were differenced-out). This is a problem if none of the 44 neighborhood variables that we use as controls capture the effect of such omitted characteristics. In the two-neighborhood-type world, this can be modeled as a direct effect of proximity to the enclave on subsequent changes in housing prices:

$$\Delta \ln P_i = \lambda \cdot \Delta ISH_i + \pi \cdot D_i^C + \xi_i.$$  

If $\pi \neq 0$, the previous exclusion restriction is no longer valid. However, the impact of proximity to an immigrant enclave is heterogeneous, which can be used to generate plausible new exclusions restrictions. For instance, consider now the existence (ex post) of two types of cities: cities with high immigration shocks, and cities with low-immigration shocks. It is a plausible (testable) proposition that immigrant enclaves in high-immigration cities are likely to expand more. We can now use a two neighborhood, two-city model and express the immigration equation thus:

$$\Delta ISH_i = \delta_1 \cdot D_i^C + \delta_1 \cdot D_i^C \times D_i^H + u_i.$$  

Here $D_i^H$ stands for a dummy variable that takes value one if the neighborhood is in a city with a high immigration shock. Note that variables are already partialled out from MSA fixed effects (i.e., $D_i^H$ is also controlled for). Now $(D_i^C \times D_i^H)$ can be used as an instrument for $\Delta ISH_i$, and we can control explicitly for both MSA fixed effects (price trends in high versus low immigration cities) and proximity to the enclave, as captured by $D_i^C$, under the assumption $\text{cov}(D_i^C \times D_i^H, \xi_i/D_i^C, D_i^H) = 0$.

Empirically, we use the heterogeneity in the predictive power of the geographic diffusion model as our effective source of identifying variation. Pull$_{i,T}$ must be a worse predictor of future immigration in neighborhoods that are already heavily immigrant. For example, if a large percent of the population in a tract is already
composed of immigrants, proximity to other foreign-born areas will not likely be predictive of increases in its immigrant density. We model the fact that geographic diffusion of immigration is more likely to go from more immigrant-dense neighborhoods to less immigrant-dense neighborhoods by interacting $\text{Pull}_{i,T}$ with the lagged share of the foreign born.\(^9\)

We use the general MSA level of immigration similarly. If there is no new immigration into the city, we would not expect the gravity pull of a neighborhood to be a particularly strong predictor of future changes in the immigrant share. Therefore, the interaction between $\text{Pull}_{i,T}$ and the relative magnitude of immigration by metropolitan area is likely to improve the predictive power of the geographic diffusion model.\(^10\)

Graphic intuitions for both research designs, maps, and results from first-stages can be obtained in the online Appendix, but we note that this empirical model of the probability of “spatial diffusion” is very successful in this application. In the first stage, interactions of $\text{Pull}_{i,T}$ with the lagged share of the foreign born obtain the expected negative sign (the direction of “diffusion” is from highly immigrant dense neighborhoods to less immigrant dense neighboring communities), and interactions with MSA immigration levels obtain positive signs (spatial diffusion is stronger in metro areas with higher immigrant shocks). The value of $F$-tests (reported in Table 1) easily situate the instruments above the 10 percent maximal-IV-size James Stock and Motohiro Yogo (2005) weak ID test critical values.\(^11\)

Using the interactions of $\text{Pull}_{i,T}$ with the initial share of the foreign born and immigration per capita in the MSA, we can now control for the “gravity pull” of a neighborhood on the right-hand side in the second stage of our 2SLS specification (Table 1, column 6). Results are similar, if slightly smaller, to those in the OLS specifications. Hansen overidentification tests fail to reject exogeneity. Hausman tests fail to reject that the IV and OLS parameters are equivalent and hence we deploy the latter in the regressions below.

### III. Further Results

#### A. Heterogeneous Treatment Effects

In Table 2, we speculate about the possibility that the treatment effect of immigration is different in different types of neighborhoods. Concretely, we interact the change in immigrant density with the initial values of two variables: the share of non-Hispanic white population and the quartile of tract housing values in the previous census within each MSA (the relevant variable takes value zero for the first quartile, and 1, 2, and 3 for the subsequent quartiles). The regressions (columns 1

---

\(^9\) Similar results are obtained using semi-parametric approaches, where we allow the impact of $\text{Pull}_{i,T}$ to differ progressively in tracts with higher immigrant density. We choose to present results with the linear interaction as the simplest approach to modeling the fact that the impact of $\text{Pull}_{i,T}$ should be smaller in neighborhoods with higher immigration shares.

\(^10\) We divide the number of new immigrants in an MSA by its initial population to obtain the relative size of immigration.

\(^11\) At the same time, the first-stage partial $R^2$ of the identifying interactions of $\text{Pull}_{i,T}$ is 0.045: most of the variation in local changes in immigrant density, as used in OLS, is outside of the gravity model. This is comforting considering the similar results obtained with the IV strategy with strong instruments.
The results show the association between growing immigrant density and relative housing value depreciation to be stronger in neighborhoods where the population was predominantly white initially. Similarly, the impact of immigration was stronger in neighborhoods that were initially perceived as more valuable. Consistent with a sorting story, the negative impact of immigrant arrivals on home values was larger in white areas with wealthy individuals. The data are strongly consistent with the view that in neighborhoods that were already minority-dense and in low-SES tracts the marginal natives who had not left earlier displayed lower willingness-to-pay for segregation. In other words, immigration did not have a negative impact on relative home values in areas in which socioeconomic sorting had already taken place earlier. This fact is consistent with local revitalization stories in relatively poor minority neighborhoods (remember that average housing prices are increasing in immigrant cities in absolute terms). We note that similar results are obtained for rents (available in the online Appendix for the interested reader).

Note: Robust standard errors are in parentheses. The table shows regressions where the change in the log of average housing prices between consecutive decennial censuses by census tract is the left-hand-side variable. The explanatory variable of interest is the change in the share of the foreign born by tract between consecutive census years. All regressions include fixed effects for each Metropolitan Statistical Area (MSA) and year combination. The regressions include observations from all census tracts in major immigrant cities, as defined in the text, for the 1980–1990, and 1990–2000 periods.

12 We do, however, substitute the log of lagged income by the log of lagged housing values when using the interaction between immigration and housing value quartiles. The correlation between the log of incomes and the log of values is 0.9, so the two variables play a similar role as controls, and cannot be used together due to multicollinearity problems.
It is interesting to map changes in immigrant concentration to changes in native population. Trivially, the growth in the share of the foreign born implies a commensurate negative change in the share of natives. In order to learn about this issue, therefore, we consider the change in the number of immigrants, natives, and non-Hispanic whites, divided by the original tract population as the relevant measures of local demographic change. In Table 3, columns 1 and 5, not surprisingly, we find that the native and white populations were also growing in areas with more new immigrants. This is not surprising because we can expect new real estate developments to attract a growing share of the city’s population—native and immigrant alike. In fact, the results in columns 1 and 5 are solely driven by census tracts where the population more than doubled between 1990 and 2000, which we will call new-development areas. In columns 2 and 6 we exclude new-development census tracts, which yields
a sample with 95 percent of the original neighborhoods. Excluding only 5 percent of the original sample yields a remarkably different picture. In most metropolitan neighborhoods, immigrant arrivals are now associated with decreasing native populations, especially of non-Hispanic whites. Similarly negative, if stronger, associations are obtained using median regression on the whole sample of neighborhoods (columns 3 and 7). The picture that arises from these regressions is clear: while new housing developments attracted both native and immigrant residents, in the vast majority of urban neighborhoods in America immigration was associated with native flight.

Of course, these associations probably underestimate the causal impact of immigrant arrivals because, as seen, factors that make a neighborhood develop, such as new housing and job opportunities, tend to be attractive to both natives and immigrants ex ante. A more interesting exercise is to use our most demanding IV strategy (as in Table 1, column 6) to assess the impact of exogenous immigration shocks. As expected (in Table 3, columns 4 and 8), these are associated with stronger absolute decreases in the level of native population. Such native flight can be entirely accounted for by a shrinking non-Hispanic white population in these areas. The difference between columns 4 and 8 is quite consistent with the fact that the average immigrant family has about 0.45 native children per immigrant, and with the fact that in the areas where the instrument has most of its bite (neighborhoods with high immigrant concentrations) immigrants tend to be minorities (mostly Hispanic and Asian).

C. Unbundling Immigration

In Figure 2, we lay out the likely avenues through which immigration may be associated with changes in local neighborhood housing values. We have discussed housing quality, omitted variables, and reverse causation. Here we focus on the remaining causal channels. First, as in our model, natives may have preferences
for living with other natives. Second and third, natives may have preferences for living with individuals of the same racial or ethnic group, or with individuals of higher socioeconomic status. This latter preference is consistent with models based on local human capital externalities (Benabou 1993), and with empirical evidence of segregation by income in the United States (Tara Watson 2002). Under these two scenarios, the earlier conceptual discussions are still applicable, but now, rather than foreignness per se, the salient characteristics that determine residential segregation are ethnicity or socioeconomic status.

Finally, another possibility is that immigration effects the quality of any locally provided public good that may experience crowding or peer-effects. For example crime is sometimes cited in opinion-surveys as a negative potential effect of immigrants. However, there is strong evidence that immigrants are less likely to commit crimes (Kristin F. Butcher and Anne Morrison Piehl 1998). Moreover our own analysis of the American Housing Survey also suggests that reported crime decreased in the housing units where immigrants moved in (the interested reader should see online Appendix). Therefore a more important and salient argument to explain a negative association between immigration and neighborhood values may involve actual or perceived changes in the quality of schools.

In Table 4, we provide evidence against the first hypothesis based on simple nativist preferences. If natives simply want to avoid living with foreigners, the association between immigration and prices should be similar for all immigrant groups in the United States. Using the 1990–2000 census tract crosswalk, we are able to produce estimates of immigrants by national group by tract (as defined in the 2000 census) for 1990 and 2000. We then group the data of nationalities with a relatively small number of migrants into broader regional groups. Column 1 shows coefficients from a regression where we control for the changes in the different immigrant shares by nationality. The association between changes in the share of Europeans, South Asians (from the Indian subcontinent), and Cubans and changes in housing values is not statistically or economically different from zero. There is a fair amount of heterogeneity by national origin. These results do not seem consistent with a model of generalized, untargeted nativism.

Can broad trends in school quality or finances (Clap and Ross 2008; Raquel Fernandez and Richard Rogerson 1996) explain our results? If the quality of education was very important to explain our results, we would expect the association between immigrant density and neighborhood values to be stronger between school districts rather than within school districts. In column 2 (Table 4), we show that the results of a regression that includes school district-by-year fixed effects are similar to the earlier estimates. However, the existence of private school alternatives and the fact that we do not have school attendance

---

13 We assign 1990 immigrants to 2000 tracts using the share of the population in 1990 that was contained within the 2000 tract boundary. This inevitably generates measurement error, because immigrants needn’t be distributed within the tract as the rest of the population. Since we also have the actual number of total immigrants in each 2000-defined tract in 1990, we use only observations where our imputation of the total number of immigrants in 1990 is within 10 percent of the actual number (83 percent of the cases). The correlation between our imputed change in the total share of the foreign-born between 1990 and 2000, and its actual change, is 0.99 in this subsample.
### Table 4—“Unbundling” Immigration

<table>
<thead>
<tr>
<th></th>
<th>Δlog value</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Δshare Europe</td>
<td>0.055</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare South Asia</td>
<td>0.017</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Cuban</td>
<td>-0.045</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Middle East</td>
<td>-0.17</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.161)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Mexico</td>
<td>-0.245</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.027)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Africa</td>
<td>-0.289</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.133)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare East Asia</td>
<td>-0.322</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.054)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare China</td>
<td>-0.36</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.073)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare South America</td>
<td>-0.418</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.075)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Caribbean</td>
<td>-0.55</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.056)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Central America</td>
<td>-0.554</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.075)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Dominican</td>
<td>-0.843</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.173)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δshare Philippines</td>
<td>-0.857</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.096)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ(Foreign born/population)</td>
<td>—</td>
<td>-0.233</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.020)***</td>
<td></td>
</tr>
<tr>
<td>Dropout immigrant shock</td>
<td>—</td>
<td>—</td>
<td>-0.377</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.163)***</td>
</tr>
<tr>
<td>Non-Hispanic white immigrant shock</td>
<td>—</td>
<td>—</td>
<td>0.267</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.061)***</td>
</tr>
<tr>
<td>Black immigrant shock</td>
<td>—</td>
<td>—</td>
<td>-0.494</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.066)***</td>
</tr>
<tr>
<td>Asian immigrant shock</td>
<td>—</td>
<td>—</td>
<td>-0.183</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.048)***</td>
</tr>
<tr>
<td>Hispanic immigrant shock</td>
<td>—</td>
<td>—</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.102)</td>
</tr>
<tr>
<td>School district—year fixed effects</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>MSA—year fixed effects</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Δ in housing quality</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing quality at T-10</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other variables in Table 1</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>18,178</td>
<td>18,167</td>
<td>18,167</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.76</td>
<td>0.82</td>
<td>0.83</td>
</tr>
</tbody>
</table>

**Notes:** Robust standard errors in parentheses. The table shows regressions where the change in the log of average housing prices between consecutive decennial censuses by census tract is the left-hand side variable. The explanatory variables of interest are the change in the share of the foreign born by nationality and tract between consecutive census years (column 1). In column 2 we reproduce the results in Table 1 controlling for school district fixed effects. In column 3 we use differences in average education and ethnicity by national group and State to proxy for the “shocks” on these variables by census tract that are associated with immigration. The regressions include observations from all census tracts in major immigrant cities, as defined in the text, for the 1990–2000 period.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.
boundaries do not allow us to completely rule out a school-based explanation, which we leave for future work.

In column 3, we explore the other avenues through which we hypothesize immigration to impact neighborhood dynamics: ethnicity and socioeconomic status (SES). We do know that immigrant neighborhoods contain a higher share of less-educated and minority individuals. For instance, a simple regression with the share of high school dropouts on the left-hand side and the share of the foreign born on the right-hand side yields an estimated elasticity of 0.65 (the t-statistic is 126). We can make use of the data on immigration by nationality to infer the immigration-driven shock to local SES characteristics. Using 1990 and 2000 microdata from the census (IPUMS) we calculate the average share of high school dropouts and racial characteristics by immigrant national group and state of residence. We focus on these variables because other interesting immigrant SES attributes (income, or the ability to speak English well, for example) were found to be extremely collinear to, and therefore well explained by, these two main factors across national groups.

We then proxy the immigrant-driven shock to these characteristics at the tract level using immigrant shares by nationality. This is summarized in the equation: \( S(X)_{ir} = \sum_{c} \Delta f_{ic} \cdot X_{cr} \); the supply shock \( S \) to attribute \( X \), in census tract \( i \), located in State \( R \), is proxied by the sum of the changes in the shares of the foreign born \( f \) by country group \( C \) in the tract, multiplied by the average attributes by country-group and state.

Introducing variables that capture the immigrant-driven supply shock to the local share of individuals who are high school dropouts, and four racial/ethnic group shares (non-Hispanic white, black, Asian, Hispanic), we find evidence that both education and ethnicity seem to matter (we also control for school district fixed effects, in order to unbundle the three channels outlined in Figure 2). The negative association between immigration and changes in prices was focused in neighborhoods where new immigrants were less educated and tended to be minorities. The coefficient on the Hispanic immigrant-driven shock is not significant, but this and the education variables are highly multicollinear (correlation of 0.91), and we cannot reject an impact statistically equivalent to that of the Asian group. These results suggest that the local interplay between immigrants and the cultural, racial, and SES preferences of natives should become a central topic for the economics research on the local impact of international migration.

IV. Conclusions

Is the prospect of having immigrant neighbors attractive to natives? While previous research (Saiz 2003, 2007; Ottaviano and Peri 2006) shows that metropolitan areas with major immigration inflows have tended to experience faster housing valuation growth on average, we do not know much about the impact of immigration on local residential dynamics. In a theoretical model with perfect mobility, immigration need not have any impact on the relative housing values of the neighborhoods where immigrants concentrate. However, if immigrant enclaves are perceived as less desirable places to live by natives, then we should expect a relative negative association between immigration density and housing values.
Using US data, we find that, controlling for metropolitan–area-by-year fixed effects, housing values grow relatively more slowly in neighborhoods with increasing immigrant density. By a simple arbitrage argument, this empirical fact is consistent with the idea that natives are willing to pay a premium for living in predominantly native areas. However, while the valuation of existing amenities is taken out by first-differencing, there is still the possibility that amenity shocks are noncausally associated with growing immigrant density.

We therefore use a geographic diffusion model (akin to an epidemiological contagion model) to generate predictions about the pattern of new immigrant settlement.

### Table A1—Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (SD)</th>
<th>Variable</th>
<th>Mean (SD)</th>
<th>Variable</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in log value</td>
<td>0.478 (0.421)</td>
<td>Change share single attached units</td>
<td>0.020 (0.051)</td>
<td>Share single detached units at T-10</td>
<td>0.623 (0.268)</td>
</tr>
<tr>
<td>Change in (foreign born/population)</td>
<td>0.052 (0.072)</td>
<td>Change share housing units in 2 unit buildings</td>
<td>–0.006 (0.027)</td>
<td>Share single attached units at T-10</td>
<td>0.055 (0.088)</td>
</tr>
<tr>
<td>Change share units with no bedrooms</td>
<td>0.008 (0.023)</td>
<td>Change share housing units in 3–4 unit buildings</td>
<td>0.002 (0.031)</td>
<td>Share housing units in 2 unit buildings at T-10</td>
<td>0.048 (0.092)</td>
</tr>
<tr>
<td>Change share units with 1 bedroom</td>
<td>0.003 (0.057)</td>
<td>Share units with no bedrooms at T-10</td>
<td>0.020 (0.037)</td>
<td>Share housing units in 3–4 unit buildings at T-10</td>
<td>0.045 (0.061)</td>
</tr>
<tr>
<td>Change share units with 2 bedrooms</td>
<td>–0.015 (0.072)</td>
<td>Share units with 1 bedroom at T-10</td>
<td>0.148 (0.124)</td>
<td>Share with bachelor’s degree at T-10</td>
<td>0.208 (0.143)</td>
</tr>
<tr>
<td>Change share units with 3 bedrooms</td>
<td>–0.006 (0.074)</td>
<td>Share units with 2 bedrooms at T-10</td>
<td>0.299 (0.146)</td>
<td>Share high school drop outs at T-10</td>
<td>0.269 (0.161)</td>
</tr>
<tr>
<td>Change share units with 4 bedrooms</td>
<td>0.004 (0.052)</td>
<td>Share units with 3 bedrooms at T-10</td>
<td>0.369 (0.157)</td>
<td>Log family income at T-10</td>
<td>10.156 (0.348)</td>
</tr>
<tr>
<td>Change share units with electric heating</td>
<td>0.057 (0.098)</td>
<td>Share units with 4 bedrooms at T-10</td>
<td>0.138 (0.123)</td>
<td>Share white at T-10</td>
<td>0.815 (0.228)</td>
</tr>
<tr>
<td>Change share units with oil heating</td>
<td>–0.029 (0.070)</td>
<td>Share units with electric heating at T-10</td>
<td>0.227 (0.246)</td>
<td>Share 25 or younger at T-10</td>
<td>0.388 (0.093)</td>
</tr>
<tr>
<td>Change share units with gas heating</td>
<td>–0.025 (0.105)</td>
<td>Share units with gas heating at T-10</td>
<td>0.089 (0.194)</td>
<td>Share 65 or older at T-10</td>
<td>0.112 (0.099)</td>
</tr>
<tr>
<td>Change share units with complete plumbing</td>
<td>0.005 (0.015)</td>
<td>Share units with gas heating at T-10</td>
<td>0.630 (0.312)</td>
<td>Share households family + kids at T-10</td>
<td>0.364 (0.150)</td>
</tr>
<tr>
<td>Change share units with complete kitchen facilities</td>
<td>0.006 (0.017)</td>
<td>Share units with complete plumbing at T-10</td>
<td>0.990 (0.020)</td>
<td>Ownership rate at T-10 (households)</td>
<td>0.672 (0.208)</td>
</tr>
<tr>
<td>Change share units built 10 years ago or less</td>
<td>–0.116 (0.208)</td>
<td>Share units with complete kitchen facilities at T-10</td>
<td>0.987 (0.020)</td>
<td>Vacancy rate at T-10</td>
<td>0.063 (0.064)</td>
</tr>
<tr>
<td>Change share units built 20 years ago or less</td>
<td>–0.021 (0.233)</td>
<td>Share units built 10 years ago or less at T-10</td>
<td>0.308 (0.276)</td>
<td>Log density at T-10</td>
<td>7.092 (1.570)</td>
</tr>
<tr>
<td>Change share units built 30 years ago or less</td>
<td>–0.011 (0.220)</td>
<td>Share units built 20 years ago or less at T-10</td>
<td>0.244 (0.166)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change share single detached units</td>
<td>–0.025 (0.100)</td>
<td>Share units built 30 years ago or less at T-10</td>
<td>0.210 (0.172)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We use interactions of these predictions with other variables as instruments for the actual changes in immigrant density in a neighborhood. Our approach could be made extensive to future research about localized social and epidemiological trends. Geographic diffusion models can be used to explicitly model assignment to the treatment in numerous contexts. Researchers can, in principle, deal with unobservable shocks that may be spatially correlated by providing plausible exclusion restrictions based on the structural characteristics of the spatial diffusion process. Robustness tests can be conducted to rule out alternative interpretations.

The evidence is consistent with a causal interpretation from growing immigrant density to native flight and relatively slower housing value appreciation. Further results indicate that the negative association between immigration and local price growth may be driven more by the fact that immigrants tend to be of low socioeconomic status and to belong to minority groups, than by foreignness per se.

We do not try to model structurally the heterogeneity in native preferences for segregation. In fact, the majority of the native population is inframarginal to substantial changes in residential exposure to immigrant density in the data. Nevertheless, our estimates provide both a proof of the existence of segregating preferences with respect to contemporaneous immigration, and a lower bound for their valuation by the marginal individuals who were actually exposed to changes in immigration inflows and policies.

Given the growing demographic importance of low-skilled and ethnically-diverse immigration flows in many developed nations, the results of the paper suggest that the disappearance of the new immigrant ghetto will be difficult to achieve.

REFERENCES


This article has been cited by: