The effect of immigration on the house prices in the Netherlands

Korbee, Leon

The effect of immigration on the house prices in the Netherlands

Leon Korbee
Student number: s1543512

Thesis committee:
Prof.dr. P.W.C. Koning

Leiden University

Faculty of Governance and Global Affairs

Leiden, The Netherlands

June 2021

Table of Contents

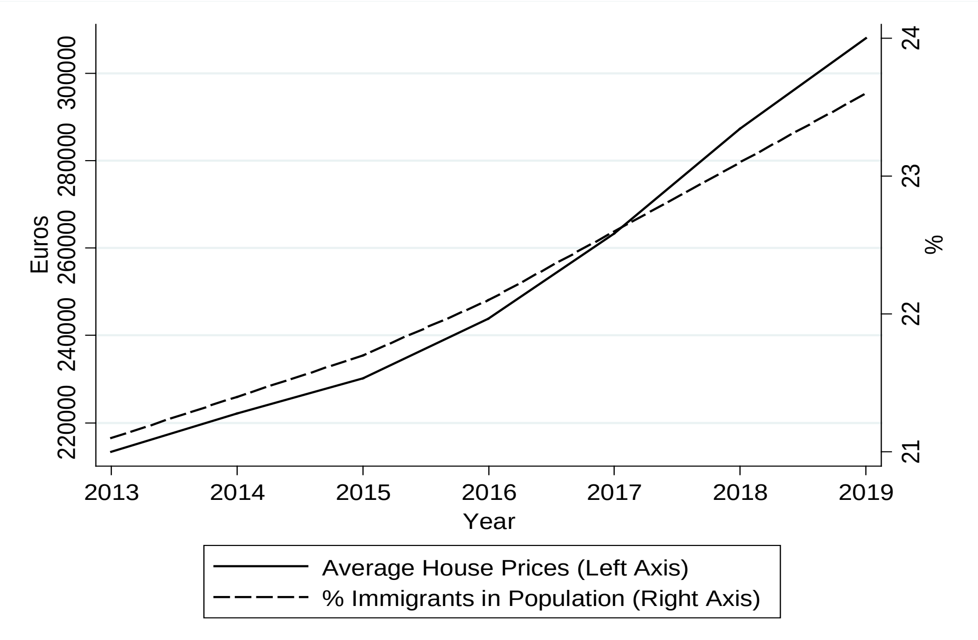
The effect of immigration on the house prices in the Netherlands

Rising house prices are of increasing concern for the Dutch government. Many elements influence the house prices, but it is often unclear how much can be contributed specifically. This research studies the effect of immigration on house prices in the Netherlands from 2013 till 2019. It contributes to understanding the influence recent immigration flows have on the Dutch economy, by estimating its impact on the housing market. Furthermore, it looks for differences of this effect between three geographic areas; municipalities, districts and neighborhoods. The results show a positive effect on the house prices in general because of decreased supply. This effect is negative for house prices in districts and neighborhoods due to the native population moving out (native flight). Natives with the highest disposable income respond to immigration by moving to different districts or neighborhoods. This generates a negative effect on housing demand and decreases house prices is in these areas. There are some differences between provinces in the strength of the effect. No negative relation is found between immigration and crime, or between immigration and housing supply.

1. Introduction

The Netherlands has seen a large rise in both the number of immigrants and the price of houses in the last couple of years. Some researchers think a supply shortage in combination with an increasing population has a major effect on the house prices. According to ABF Research (2021) there are an estimated 330.000 houses needed across country, which translates to around 4%~ of the current housing stock. They think the increasing population, in addition to the decreasing household size, demand for 100.000 new houses build every year. The Central Bureau for Statistics (CBS) calculated the average number of houses build in the last ten years at around 58 thousand, and in the years between 2000 and 2009 at around 76 thousand (CBS, 2001). According to ABF, the difference between houses build and housing demand explains the sharp rise in house prices. This is not undisputed as some think the house prices are more correlated with the financing space of buyers than with the housing shortage (Madsen, 2012; DNB, 2020). The Netherlands has also seen a large rise in the number of immigrants (Zorlu & Hartog, 2001). Figure 1 shows that the stock of immigrants in the Netherlands increased with 13% between 2013 and 2019. The average house price has risen 44% in the same period. This thesis tries to separate the shortage effect from the immigration effect on house prices. By doing so it hopes to give a clearer view of immigration's real influence on house prices.

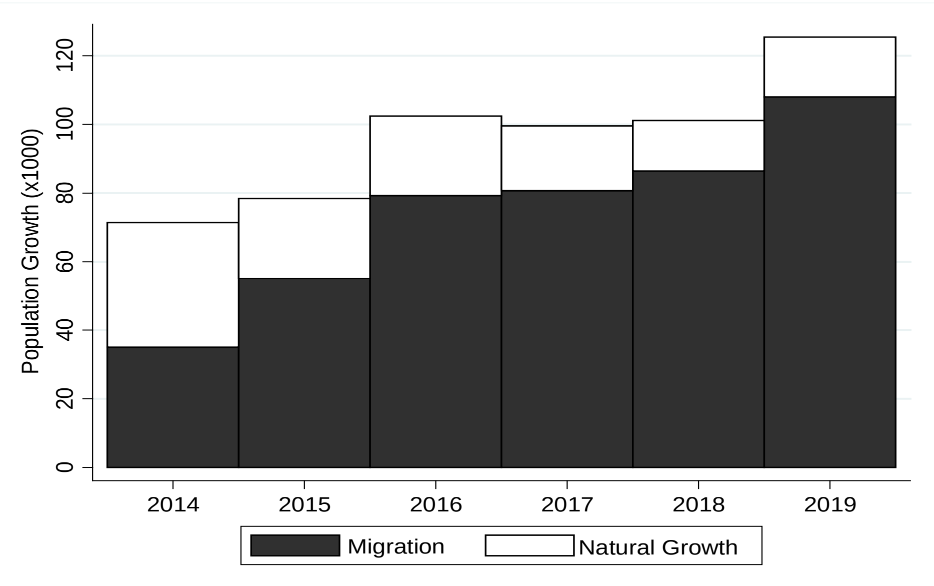
Fig 1. Immigration and House Price. Source CBS; made by author



Even though immigration and house prices are increasing, there are big differences between areas. Large cities have traditionally been the places with the highest increase for both immigration and house prices (CBS, 2021). However, there are major difference between districts and neighborhoods in cities. The interaction between immigration and house prices within versus across local areas, is important (Jeanty et al., 2010). Some neighborhoods see a decrease in house prices in relation to increased immigration (Saiz & Watcher, 2011). They explain this by applying the concept of 'native flight', as seen in the public school systems, onto the housing market. Here it is used to describe natives changing schools often in response to an increase in the immigrant stock. In the housing market, native flight occurs when native-born families respond to inflows of immigrants by moving away from the area. Natives with the largest disposable income are the first to leave and push house prices down (Sá, 2015). Researching the differences between geographical areas can help explain part of the house price movement. This thesis tries to answer the following research question: *"What effect does a change in the immigrant stock have on the house prices on a municipal, district and neighborhood geographic level?"*. The paper is organized as follows. Chapter 2 presents the current literature surrounding immigration, house prices and their relation. Chapter 3 discusses the methodological approach and statistical methods to estimate the effect(s). Chapter 4 describes where the data comes from and how it is used. Chapter 5 reports the results, discussion and additional interesting effects. Chapter 6 concludes.

While natural increase of the population declined in the last decades, the migration balance has a more ambiguous gradient (CBS, 2020). In 2000 the natural increase, birth minus mortality, was around 60 thousand (60%) and the migration balance around 40 thousand (40%). In 2019 natural growth decreased to 17 thousand (13.6%) while the migration balance grew to 108 thousand (86.4%). Migration is since 2014 the primary cause of population growth in the Netherlands as shown in Figure 2. Immigration is of increasing influence on the demographic change. Research towards the effect of immigration on the house prices could prove useful in helping lift some prejudices surrounding it and contribute to knowledge on social geography and planning for local governance. From an academic perspective this study contributes to the existing literature in two ways. Most literature is looking at either the labor market or price impact of immigration (Dustmann et al., 2005-2005-2013). Less research is done towards the effects of immigration on house prices, and the current work is often focused on either the United States or the United Kingdom. Since immigration and the housing market is fundamentally different between states, it is very hard to generalize results across borders. More research towards mainland Europe could add to the broader knowledge and understanding of the effect immigration has on house prices. Secondly, the use of different geographical levels to gain insight in their relative differences contributes to a better understanding of the current ambiguity of the direction of the effect. From a statistical perspective this thesis builds and contributes to the empirical approach used in migration literature, the use of lagged values of immigration flows as instrumental variables to combat potential endogeneity.

Fig 2. Population Growth. Source CBS; made by author.



2. Related work

Kalantaryan (2013) points out that current research on the effects of immigration focusses primarily on labor market outcomes, which is the impact of immigration on employment opportunities and wages. Dustmann and Glitz (2005) find a small disadvantageous effect of immigration on employment and wages below the 20^{th} percentile of the wage distribution, but a slight increase in the upper part of the wage distribution. A slight positive effect on the native wages in total, which they explain through the immigration surplus as well as through immigrants working for less than market price. Another paper by Poot and Cochrane (2005) compares 18 different papers on the subject in a meta-analysis. They find a very small effect overall in international papers. An increase in the share of immigrations into the local labor force by 1% leads to a reduction of 0.1% in wages. The discussed papers propose different reasons for this relatively small effect. It could be that native choose to avoid areas with many immigrants because they are afraid of competition. Immigrants may also chose specific cities because they experienced a positive shock in productivity and/or wage growth. Lastly, it could be that the labor market is more elastic than previously thought.

Existing research on the effect of immigration on house prices is ambiguous. There is no common judgement in the existing economic studies addressing the effect, neither is there consensus on the magnitude or the direction of the effect. Most studies focus on a single country with their analysis, because generalization across countries is difficult. Different laws, economic states and cultures make it hard to generalize. A large part of the work focusses on countries traditionally seen as migration counties, like the USA or Canada. To the best of this author's knowledge there is no economic research done towards immigration and the house prices in the Netherlands. Generally, house prices rise when more people move to an area because the housing supply is likely to be inelastic in the short run. Time lags and legal complexities surrounding the development of new houses increase time costs. The time between increased demand and stationary supply should drive house prices up (Saiz, 2003).

Saiz finds this effects to hold true for the short-run impact of immigration on local housing in his case study in the US. Increased rents for lower to moderate quality housing in Miami between 1979 and 1981 is correlated with the level of immigration. In his research of the immigration effect on the American housing market he sees rent increase 8 to 11% after a specific immigration shock in Miami (Saiz, 2003). This case study was in a specific city at a specific point in time, which makes it difficult to generalize. However, it was one of the first studies on the specific effect of immigration on housing prices. In later research Saiz (2007) looks at a broad region in the US and finds that an immigration inflow equal to 1% of the local population increases the average rents and housing values by 1% and average prices by 2.9 to 3.4%. Rents move first after which house prices follow. Saiz points out that this positive effect of immigration on rents and prices is consistent with the idea that immigrants do not displace natives from gateways cities one-for-one. Gateway cities being cities were two thirds of the immigrants initially settle, and ‘one-for-one’ meaning for every immigrant moving to a gateway city one local resident moves out. He concludes that the housing market response to immigration is of a bigger magnitude than the one in the labor market.

Later research from Saiz and Watcher (2011) find a negative connection between immigration and the changes in house prices and rent. They try to explain their result with three possible explanations. First, natives could have a preference for living with individuals of the same ethnic group and socio-economic status. Second, immigration may attract more crime or decrease the quality of public goods through overcrowding. Third, immigration may decrease the quality of the housing stock. They find the most evidence for the first explanation, which is a preference of natives to live with individuals of the same group. Their focus on a small geographical area shows ‘native flight’ decrease house prices in relation to increased migration. This effect was first seen in public schools around metropolitan areas in the United States, but has gained more attention in recent years in other academic research towards migration (Betts & Fairlie, 2003). Academic research gives various explanations as to why house prices may fall after immigration waves, but see indirect local resident out-migration as an important factor.

Research done by Mussa et al. (2017) find that immigrant inflow into an metropolitan statistical area (MSA) gives a large spill over effect in the surrounding MSA’s. An increase of 1% of a MSA’s population drives house prices up by 0.8% in the MSA itself, but house prices of surrounding MSA’s by 9.6%. This could indicate a large native flight towards surrounding MSA’s and increased levels of home-ownership when moving. Research by Sá (2015) in the UK finds a somewhat similar effect. Sá researches 170 local authorities in England and Wales, using microdata from the UK Labour Force Survey together with Worker Registration Scheme. On a local level, immigration seems to have a negative effect on house prices in the UK. An increase in the stock of immigrants equal to 1% of the local population decreases house prices with 1.7%. The effect is stronger when the immigrant stock has lower education. Sá finds no relation between crime and immigration or housing quality and immigration. She concludes they play no role in the negative effect of immigration on house prices.

Saiz focusses on the American housing market and Sá looks for effects on the UK’s market. A few others search for the immigration effect on house prices in different, but relatively comparable countries. Degen and Fischer (2009) look at districts in Switzerland and find that an immigration inflow equal to 1% of a district's population increases the price of single family homes by about 2.7%. Gonzalez and Ortega (2013) research Spain, and use data at the provincial level. They find that an increase in the foreign-born population equal to 1% of the total population leads to an increase in house prices of 3.2%. Kalantaryan (2013) does research in Italy on a provincial level from 1996 till 2007. She finds that an increase in the concentration of immigrants in the provinces has a positive but declining effect on the average housing prices in provinces. Like the other regional researchers her results indicate that an increase of the immigration population leads to an increase in the average housing price.

In contrasts with research finding increased house prices, some researchers find immigration causes almost no or even decreasing house prices. Akbari and Aydede (2012) find a small and negative effect of immigration on house prices in Canada. They use panel data at census division levels from 1996, 2001 and 2006. Their effect could be a results of the longer period and thus more elastic housing supply in the long run. The effect will be smaller if you use longer periods of time and interval than when using consecutive years. They think the small effect could be caused by the increased supply or by the out migration of the native born. Stillman and Mare (2008) research the effect in New Zealand and use even longer census data from 1986 till 2006 (5 year steps). They find that a 1% increase of the local population, increases house prices with 0.2 till 0.5%. However, they find no evidence that this is caused by the inflow of foreign-born immigrants. They do find a significant correlation between increased house prices and returning Kiwis. In their research they use a relatively large time frame which could cause increased elasticity in the housing supply and hide the immigration effect. On account of this, the time period for this thesis are the consecutive years 2013 till 2019 for a better view of the immigration effect specifically.

The most current all-encompassing work on the effect of immigration on house prices is done by Larkin et al. (2018). They do a meta regression analysis of 474 estimates in 14 different countries and find that the effect of immigration increases house prices on average. However, it is more pronounced at the state or provincial level and moderated by the attitude of natives on immigrants. Similar to the previously discussed studies on different levels of geographical aggregation, they support the notion of decreasing house prices on a local level when hit by increased migration (Saiz & Wachter, 2011; Sá, 2015; Mussa et al., 2017). From the discussed literature the following assumption is made: *"the effect of an increase in the stock of immigrants on housing prices varies with respect to the chosen level of aggregation"*

3. Methods

The regression analyses utilize unbalanced panel data, examining close to a hundred thousand observations over the period 2013 till 2019. Unbalanced panel data could cause attrition issues, problems with non-randomly missing data or selection bias. Attrition, if endogenous, can cause problems if the time frame is (too) long (Winer, 1983). This research does not make use of census data spread 15 years but uses consecutive years between 2013 and 2019. This limits problems with attrition. Additionally, the data does not have attrition in the typical sense. The number of complete observations increase instead of decrease over time, because of better record keeping. Most of the missing data is missing completely at random since the merger of multiple datasets is the major reason for these missing values. However, shifting definitions of areas in local municipalities could insert some attrition or selection bias into the data. To limit selection bias almost none of the observations are removed. By running and comparing multiple different regressions it tries to minimize these problems in general.

To investigate the causal effect of immigration on house prices, two different estimators are used and compared. The first difference estimation (FD) and the instrumental variable estimation (IV). The FD estimator is used to address the problem of omitted variables with panel data. Local house prices are influenced by time-invariant, area specific characteristics, like industrialization. These characteristics are likely to be correlated with immigration in and outflow. The FD approach eliminates unique factors that could potentially affect both immigration and house prices by differencing these factors out (Brüderl & Ludwig, 2015). Furthermore, the estimations should include year fixed effects, to capture trends in house prices, and local area fixed effects. The heterogeneity among local areas is likely correlated with the independent variables and not random. Their inclusion captures the different trends in house prices at the area level (Sá, 2015).

The second method of estimation used, is IV estimation. It makes use of instrumental variables to cut correlations between the error term and the independent variables. Standard linear regression models assume that errors in the dependent variable are uncorrelated with the independent variable(s). If this does not hold true, the estimates could be biased and inconsistent. Since there is no general consensus on the causality of the relation between immigration and house prices, it makes interpreting the correlation unstraightforward. The two-stage least-squares regression uses the instrumental variable to compute estimated values of the problematic predictor(s). Then it uses those values to estimate a linear regression model of the dependent variable (Arellano & Bover, 1998). The literature denotes that the historical settlement pattern of immigrants is highly correlated with the change in the stock of immigrants after controlling for other independent variables (Kalantaryan, 2013). It is only indirectly correlated with the changes in house prices through their relation with the present changes in the immigrant stock. This sort of instrument is called a ‘Bartik’ or ‘shift-share’ instrument and will be further explained in Section 3.1 (Jaeger et al., 2018). The shift-share instrument is picked as the instrumental variable. The validity of the shift-share instrument relies on the assumption that there is no correlation between the settlement pattern and any future change which affects the location choice for immigrants.

3.1 Equations

House prices are calculated somewhat similar to the SPAR-method (Sale Price Appraisal Ratio) (Wal & Tamminga, 2008). This method looks at the change in the relation between the average purchasing price and the average WOZ-worth (appraisal) of sold houses. This makes it unsensitive to differences in the quality of sold houses every year. Equation 1 is used to calculated the average house prices for municipalities as a whole:

$$IP_{mt} = \frac{S_{mt}}{A_{mp}} * 100 \tag{1}$$

Here IP_{mt} is the index price of municipality m for period t, S_{mt} the average purchasing price in municipality m in period t, and A_{mp} the average WOZ-worth for municipality m in valuation moment p. Average purchasing price data on districts and neighborhoods are not available, but WOZ-worth is. Their average purchasing price is calculated by taking the index price for the respective municipality they are located in and multiplying it with their own WOZ-worth.

E.g. $$\frac{103.42 * 128.000}{100} = 132.382$$

The neighborhood ‘Appingedam-Centrum’ with an average WOZ-worth of 128 thousand has an average purchasing price of 132.382 in 2014, based on the price index level of 103.42 for the whole municipality of Appingedam. After calculating the average purchasing price for all areas, the logarithm of this value is taken. Since most areas have some form of spatial dependence and fixed influences, the dependent variable (DV) and independent variable (IV) are likely to be correlated even without the immigration effect. Therefore we estimate the them in first differences, as discussed in the previous section, which makes it the change in the log of the house prices. This eliminates time-invariant and unique factors of areas that affect both the change in the stock of immigrants and house prices.

The economic model follows closely the model by Kalantaryan (2013) and Sá (2015) which in turn is based on the model made by Saiz (2007). Equation 2 is used to estimate the effect of immigration on house prices:

$$\Delta\, ln(HP_{at}) = \beta \frac{\Delta IMM_{at}}{POP_{at-1}} + \gamma X_{at} + \phi_t + \rho_a + \varepsilon_{at} \qquad (2)$$

The dependent variable is $\Delta\, ln(HP_{at})$, which is the change in the log house price of area *a* between years *t-1* and *t*. The main independent variable is the annual change in the stock of immigrants ΔIMM_{at} divided by the initial population POP_{at-1}. So coefficient β is to be interpreted as the percentage change in house price in relation to an annual increase in the stock of immigrant of 1% of the initial population.

X_{at} is a set of lagged socio-economic variables which may influence house prices shown in Table 1. Saiz and Watcher (2011) write that immigration impacts house prices because the characteristics of the individuals who move into the neighborhoods (the new immigrants) are different. This makes changes in socioeconomic characteristics endogenous to immigration. Therefore, in line with their work, these variables are not controlled for but included as lagged levels. These lagged variables are, the percentage of locals claiming state benefits, the percentage of locals in unemployment and the percentage of locals claiming disability insurance. These capture discrepancies in house prices between neighborhoods due to economic different conditions between these areas. The local nonviolent crime rate and the local violent crime rate account for housing demand. The nonviolent crime variable consists of primarily theft offences, while the violent crime variable contains offences from threats to murder. Lastly, the ratio of the number of dwellings in relation to the local population accounts for the housing supply.

ϕ_t is a set of year dummies to capture economic trends in house prices (Kalantaryan, 2013). P_a is a set of area dummies to capture time-invariant area-specific characteristics. ε_{it} is the idiosyncratic error. This panel data error can be interpreted as the unobserved factors that impact the dependent variables while changing over time and across units. To account for correlation within groups, clustered variance estimators are used.

One problem is figuring out the direction of causality between immigration and house prices (Jaeger et al., 2018). In addition, the changes in immigration might be correlated with other unobserved factors. The locational choices of immigrants are not random, and the economy of areas could change in many ways through local shocks. Immigrants could either move towards areas which are already doing well and were house prices are growing, or they could move towards areas with low house prices since it is cheaper. To treat these potential endogeneity issues this thesis uses a shift-share instrument for immigration in which the distribution of immigrants is based on historical settlement patterns. The existence of prior enclaves of immigrants is of great influence on the future flows. For immigrants it is attractive to live among people with the same language and cultural traditions, they can make use of previously established network by earlier immigrants (Bartel, 1989). The instrument of the predicted stock of immigrants is calculated according to equation 3, based on the model of Kalantaryan (2013):

$$IMMIV_{at} = \frac{\sum c\, \lambda_{a,c,t=0,} * \; \Delta IMM_{ct}}{POP_{at-1}} \quad (3)$$

Where $\lambda_{a,c,t=0}$ is the proportion of immigrants born in foreign region *c* that settled in area *a* in the period $t=0$. This represent the historical networks of immigrants from different regions living in specific areas. ΔIMM_{at} represents the change in the immigrant stock born in foreign region *c* at the national level. This gives the predicted change in the stock of immigrants from each foreign region *c* in area *a* in year *t*. So the instrument is the change in the predicted foreign-born population between years *t* and *t-1* relative to the total population in the initial period. This instrument is only valid if we assume two things. The first assumption is that the historical settlement pattern of immigrants is not correlated with changes in economic performance of areas. The plausibility of this assumption is stronger with a greater length of time. The second assumption is exogeneity of the annual changes in the stock of immigrants at the national level to the economic conditions of the local areas. The overall number of legal immigrants in the Netherlands should depend on political decisions (Kalantaryan, 2013).

4. Data Collection

All of the complied data is from the CBS 'key figures of districts and neighborhoods' (KWB) dataset, the National Police database, and the Dutch Cooperative Association of Estate Agents and Valuers (NVM). KWB is made in cooperation with all local municipalities in the Netherlands and contains small scale data. It divides the Netherlands into municipalities, districts and neighborhoods. Municipalities are the third governmental layer as defined by the municipal law, enclosing one village or city. Districts are defined as part of a municipalities containing one or multiple neighborhoods. Most often the districts match with residence areas of the municipality. Neighborhoods are defined as parts of a municipality which are homogeneously demarcated from a building-wise or socio-economic perspective. This means they are split by their dominant function, like residential area, industrial area, recreational area. There could be a mixture of function in certain neighborhoods. Municipalities decide the layout of these areas themselves.

Reclassification of neighborhoods can change borders and significantly increase or decrease the population living within them. The independent variable is calculated through the initial population and immigration stock. Therefore, unnatural changes could influence the outcome disproportionally. These sudden shifts in population are not caused by natural movement and classified as outliers. To find outliers, multiple regressions are run with a different maximum value for the focal independent variable. A value of 5 for the independent variable indicates a change in the stock of immigrants of 500% in relation to the initial population within an area in one year. Since Stata drops all values over 15 itself, we start with a maximum of 10 and decrease the limit to 5, 3 and 2. The effect a very small number of observations have decreases significantly until the limit of 3. Limiting it lower than 3 drops out too many variables with little to no effect. Large shifts in the population stock above a score of 3 are likely caused by reclassification of the area, and are filtered out for districts and neighborhoods. 300% still seems rather high, but the dynamic nature of districts and neighborhoods cause a higher population flow through.

This cutoff point is not fitting for municipalities as they do not see such large changes. Another method of detecting outliers is by measuring Cook's distance for every observation. It finds data points with large residuals and high leverage, which may distort the outcome and accuracy of a regression. The results show that some of the observation are above the $4/N$ and even $Di > 1$ cutoff point, which could indicate that these are outliers. The Cook's D score next to information on the reclassification of municipalities is used to classify outliers in municipal data. These are removed from the regression.

Most of the KWB data is delivered by the municipalities themselves, which results in some wrongly denoted data. For example, areas with zero to fifty inhabitants having 200 people receive assistance benefits. These potential outliers, often located in small districts and neighborhoods, have a disproportional influence on the results. Analyses including these areas gives smaller errors but a distorted view of the actual results. Removing all area's with less than a hundred inhabitants greatly reduces the number of these kinds of outliers. This decreases the total number of observations with 2.38% and potentially brings some bias, but greatly improves the consistency of the estimates.

The primary reasons for missing values are either that the municipalities did not deliver any data or that the layout of the area changed. Table 2.1 shows the three most common data patterns each geographic level has. A 'full' data pattern means the area has data from all years between 2013 and 2019. Municipalities, districts and neighborhoods have a 64%, 63% and 53% full data pattern respectively. A common pattern is just the 2013 observation, which could indicate a reclassification in 2014. Simultaneously, the three year pattern 2017, 2018 and 2019 is also quite common which could indicate that in 2017 there were also reclassifications. Neighborhoods are more dynamic and difficult to track for seven years which could explain the higher level of missing data.

Table 2.2 also shows the numbers of observations and average population over the years for each geographic level. The number observations for municipalities dropped from 408 in 2013 to 327 in 2019. This can mostly be explained by municipal reorganization. Meanwhile, the average population within municipalities increased with 18% through the merger of different municipalities, and by the rising population. The number of districts increased while the average number of people living within them decreased. The number of neighborhoods sharply decreased after 2013, and slightly increased after that. Their average population increased by a very small margin. Since the dataset does not contain data on every level for all years, part of these differences are caused by missing or dropped out data. Table 2.3 shows how the data is spread between the 12 provinces of the Netherlands. Flevoland and Zeeland are the smallest provinces, while Zuid-Holland is the largest province and has the most observations.

Some of the data cannot be used because it is either incomplete, or dropped as an outlier. Table 2.4 shows the number of observations Stata kept, the number of observations left after dropping outliers and the number of observations left after including the control variables. Column (2) shows that the number of observations for municipalities decreased the most after dropping outliers. Municipal reclassifications changed the total number of municipalities in 2013 to 2019, from 408 to 355. Some reclassifications caused major changes in the total number of citizens, which is part of the calculation for the independent variable. Since the number of observations for municipalities is significantly lower than for the other two geographical levels, these changes have great influence and could result in biased estimates. Therefore, a stricter threshold is chosen for municipalities and more observations are filtered out. Districts and neighbourhoods loss relatively less observations that influence the outcome disproportionally and are most likely caused by the reclassification of areas. Column (3) shows that after the socio-economic control variables are added, a large part is dropped with an exception for the district level. This level has no crime statistics available. Municipalities and neighbourhoods do make use of the crime variables which explains the loss of observations after merging two different datasets.

Table 2.4: Number of observations used in the analyses

Area	Number of observations	Number of observations without outliers	Number of observations with CV
Municipalities	2.264	2.153	1.864
Districts	15.596	15.593	15.593*
Neighbourhoods	54.505	54.444	37.742

Notes: Column (1) shows the number of observations Stata kept. Column (2) shows the number of observations after manually deleting outliers. Column (3) shows the number of observations after adding the control variables. *Districts have no data available for crime which explains the similarity between column (1) and (2)

Table 3 contains summary statistics for the main variables used in the analysis after dropping outliers. The smaller levels, districts and neighbourhoods, have a substantially larger spread between the minimum and maximum for all variables. This is because a municipality is the average of all neighbourhoods within its borders. A closer look at the highest maximum and lowest minimum values of other independent and control variables show that they are always located in relatively small neighbourhoods and/or districts. For example, the highest non-violent crime rate is around the Central Station Area of Tilburg. This area does not have a lot of suitable dwellings but does have a high flow through of people, what could explain the relatively high crime rate.

5. Results and Discussion

Table 4 reports the results of the OLS analyses based on Equation 2. The dependent variable is the change in the log of the house prices and the focal independent variable is the change in the immigrant stock relative to the total population in the previous year. The standard errors are heteroscedasticity-consistent (Huber-White's Robust Standard Errors) to prevent bias and inconsistency. They are also clustered by local authority to account for within group correlation. All regressions include year fixed effects to capture economic trends and area fixed effects. The Hausman test shows a high significance level for all three different geographic levels ($p < .05$). The unique errors (u_i) are correlated with the regressors and the null hypothesis is rejected. The difference in coefficients is systematic, which makes the fixed effects model appropriate. The results are robust to the exclusion of local authority fixed effects.

The first columns for every geographical level report the estimates without control variables. The estimate for municipalities is positive and significant at the 5% level. If the stock of immigrants increases with 1% the house prices in a municipality will increase with 0.48% on average. Take for example a municipality with a stock of immigrants equal to 14% of their total population in 2014 and an average house price of 230 thousand. The immigrant stock grows to 15% in 2015:

E.g.
$$\frac{0.475\ (\beta_{Municipality}) * 230.000}{100} = 1.092{,}50$$

This indicates that in 2015 the house prices grew by 1.092,50 because of the increased stock of immigrants. As discussed in the literature, more people moving towards a city or village gives an impulse to the demand side of the housing supply (Saiz, 2003). The average house prices in a municipality most likely go up as houses are not build instantaneously. However, a causal interpretation of these estimates is not possible since the locational choice of immigrants is not random. For the house prices in districts and neighbourhoods, the estimates show both a negative and significant effect at the 1% level. We apply their estimates to a similar example:

E.g.
$$\frac{-0.264\ (\beta_{District}) * 230.000}{100} = -607{,}20$$
$$\frac{-0.178\ (\beta_{Neighborhood}) * 230.000}{100} = -409{,}40$$

House prices in an average neighbourhood with 1.800 inhabitants and a stock of immigrants increasing from 14 to 15% of the total population, decrease with 409,40 euros when the number of immigrants increases with 18. House prices in an average district decrease with 607,20 euros when the number of immigrants increases with 62. Again, the OLS estimates are without using an instrument for endogeneity.

Columns 2 add the socio-economic control variables from Table 1 and obtain very similar results to the ones without. However, for all geographical levels the respective effect gets stronger. Adding the control variables increases the adjusted r-squared, which indicates they improve the model fit more than expected by chance alone. Even though they improve the model, the number of observations is significantly less when the crime variables are included. Additional regression are run excluding the crime control variables, their coefficients are between .01 and .025 weaker than the ones including all control variables. This difference is not large enough to drop them completely from the regression.

Districts show a stronger effect than neighbourhoods, potentially because adjacent neighbourhoods are (too) similar in look and price class. Citizens chose to move to a different district instead of a neighbourhood because they have a larger difference in house quality. It could also be that districts match with residence areas of the municipality, while neighborhoods are more often demarcated as an industrial area or recreational area. These area's are less likely to be influenced by movement since their predominant function is not residential. Overall the house prices move up after increased immigration, most likely to higher demand and lagging supply.

Sá (2015) reports that the full effect on house prices may not be immediate but could increase or decrease in response to immigration over time. To control for this immediate effect, an extra explanatory variable is added in the columns 3. This is the lag of the change in the immigration stock relative to the total population. The coefficients are not significant for any of the geographical levels. It cannot be said that the effect is significantly stronger or weaker over time. Testing the significance of the sum of the coefficient (both immediate and lagged), we reject the null hypothesis that the sum equals to zero at the 1% level for districts and neighbourhoods. There is joint significance between the immediate and lagged change in the immigrant stock relative to the initial population, but it only shows for the instant coefficient.

Table 4: OLS Immigration and House Prices

	Δ ln HP_t								
	Municipalities			Districts			Neighbourhoods		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$\Delta IMM_{at}/POP_{at-1}$	0.475**	0.500**	0.545*	-0.264***	-0.270***	-0.260***	-0.178***	-0.230***	-0.186***
	(0.229)	(0.254)	(0.297)	(0.062)	(0.060)	(0.061)	(0.041)	(0.036)	(0.041)
$\Delta IMM_{at-1}/POP_{at-2}$			0.095			0.002			0.001
			(0.360)			(0.004)			(0.001)
Socio-economic local area controls	No	Yes	Yes	No	Yes[1]	Yes[1]	No	Yes	Yes
Significance level for test b1 + b2 = 0			0.212			0.000			0.000
R^2 within	0.224	0.231	0.174	0.154	0.162	0.148	0.115	0.149	0.128
R^2 between	0.129	0.056	0.001	0.092	0.045	0.018	0.070	0.030	0.012
Observations	2153	1864	1555	15593	15593	12496	54444	37719	30352
Number of groups	406	339	336	3116	3116	2936	11457	8765	8371

Notes: Robust standard errors clustered by area in parentheses. Δ indicates first difference. Regressions include year and local area fixed effects. The socio-economic controls and their sources are described in Table 1.

***Significant at 1%, **significant at 5%, *significant at 10%.

1: There are is no crime data available for districts

OLS estimates for municipalities are positive, while the estimates for the smaller geographic levels are negative. This could be an indication of native-flight, as discussed in the theoretical framework (Saiz & Wachter, 2011; Sá, 2015; Mussa et al., 2017). Even though the coefficient are statistically significant, we cannot interpret them as causal effects. The locational choice of immigrants is not random, they often attract one another. To overcome this problem, the settlement pattern of immigrants from base year 2013 is used to predict the geographic distribution of immigrants currently. Table 5 shows the results of the IV regression with the change in the predicted foreign-born population as an instrument. The instrument is calculated as described in Equation 3.

The first stage regression shows the effect of the change in the predicted foreign-born population relative to the initial population on the change in the foreign-born population in an area. The effect is significant at the 1% level, which indicates a strong correlation between real immigrant flows and the predicted flows based on historical settlement patterns. Moreover the F-statistic for all regressions is above 10, indicating the instrument is not weak.

Table 5: IV Immigration and House Prices

	Δ ln HP_t					
	Municipalities		Districts		Neighbourhoods	
	(1)	(2)	(1)	(2)	(1)	(2)
$\Delta IMM_{at}/POP_{at-1}$	0.754**	0.725**	-0.446***	-0.465 ***	-0.218***	-0.260***
	(0.215)	(0.250)	(0.162)	(0.167)	(0.048)	(0.056)
Socio-economic local area controls	No	Yes	No	Yes	No	Yes
Coefficient IV in first stage regression	1.030***	0.971***	1.825***	1.825 ***	1.336***	1.416***
	(0.061)	(0.077)	(0.360)	(0.362)	(0.078)	(0.076)
R^2	0.211	0.226	0.119	0.123	0.103	0.125
Observations	2143	1856	14960	14960	52152	36006
Number of groups	402	337	2995	2995	11052	8476

Notes: Robust standard errors clustered by area in parentheses. Δ indicates first difference. Regressions include year and local area fixed effects. The instrument is the change in the predicted foreign-born population between years t and t-1 relative to the total population in the initial period. The settlement pattern of immigrants is used to predict the number of foreign-born in each local authority in subsequent years. The F-stat in the first stage IV for columns 1 is 279, 35 and 310 respectively. For columns 2 the F-stat is 154, 34 and 228 respectively.
***Significant at 1%, **significant at 5%, *significant at 10%.

The effect of immigration on house prices in municipalities suggest a positive trend. The same effect suggest a negative trend for districts and neighbourhoods. The average municipality in 2014 had a population of 44 thousand of which 6.6 thousand were immigrants. If this stock were to increase with 440, while the total population remained the same, the average house price (230 thousand) would increase with about 1.610,-. Simultaneously, the districts welcoming these new immigrants would see similar houses decrease 1.035,- in price, and in neighbourhoods with 530,-. These differences show how the effect can be seen as ambiguous when no specific geographic level is picked in advance.

Different geographic levels show opposite results, similar to how previous research on regions was different from research on local authorities. Native flight could be one of the primary causes of decreasing housing prices in relation to immigration. Both the results in the OLS and IV regressions show that while the average municipal house prices rise, the house prices in smaller geographical areas decline when more immigrants move towards them. With just the OLS results we cannot be certain if immigrants move to areas with lower house prices or house prices are lower because immigrants move towards them (Saiz & Watcher 2013; Sá, 2015). The IV results give more certainty about the direction of this relation, and they suggest a native flight effect. The estimates in Table 5 are positive for municipalities and negative for districts and neighbourhoods. The natives with the largest disposable income move out when more immigrants move in. By this movement the house prices in area's with an increased number of immigrants go down. The areas these more wealthy citizens move to most likely see increased house prices through increase demand.

The IV estimates for districts and neighbourhoods are higher than the OLS estimates, especially for the former. The upwards bias in the OLS results might be explained by immigrants locating towards more wealthy areas where house prices are rising faster. The similarity between neighbourhoods in the same district could strengthen the difference between, but not within, different districts. Moreover districts are more often demarcated as residential areas with more houses, which could explain the larger effect for them. Municipalities show a stronger positive result with IV than with OLS. This suggests that immigrant inflow into a district or neighbourhood gives a large spill-over effect towards the surrounding areas. The positive spill-over effect on house prises is larger than the negative effect of an increased immigrant stock. Similar to the results of Mussa et al. (2017), the initial immigrant shock drives the house prises down in certain area's but increases the overall house prices of the municipality more.

5.1 Additional Results

Long Difference

Another approach proposed by Saiz (2003) and Sá (2015) analyses the time effect. This is the timing effect of immigration on house prices taking a longer difference for both the house prices and the immigrant stock. The effect could be underestimated if house prices take time to fully adjust to a change in the immigrant population. Table 6 shows this analyses with the control variables lagged accordingly. The time effect has as dependent variable the log of the house price between years *t* and *t-3* and as focal independent variable the change in the stock of immigrants between years *t* and *t-3* relative to the population in year *t-3*. The OLS estimates are all less 'strong' than the ones obtained with the shorter time effect, only the smaller two stay significant. The effect seems weaker in the long run, which could indicate that house prices do not need time to adjust in neighbourhoods and districts. An instrument similar to the one in Equation 3 is used to address potential endogeneity of the locational choice of immigrants. Instead of the difference between years, the difference between base year t_0 and t_3 is picked in the numerator of the formula.

Table 6: Long Effect of Immigration and House Price

	$\ln HP_t - \ln HP_{t-3}$					
	Municipalities		Districts		Neighbourhoods	
	OLS (1)	IV (2)	OLS (1)	IV (2)	OLS (1)	IV (2)
$(IMM_{at} - IMM_{at-3})/POP_{at-3}$	0.508	1.064***	-0.046***	-0.391**	-0.145***	-0.185***
	(0.387)	(0.326)	(0.018)	(0.160)	(0.038)	(0.055)
Socio-economic local area controls	Yes	Yes	Yes	Yes	Yes	Yes
Coefficient IV in first stage regression		1.10***		2.577***		1.936***
		(0.052)		(0.659)		(0.198)
R^2	0.500	0.385	0.366	-	0.382	0.186
Observations	1218	1218	9153	9153	22591	22591
Number of groups	327	327	2548	2548	7093	7093

Notes: Robust standard errors clustered by area in parentheses. Δ indicates first difference. Regressions include year and local area fixed effects. In the IV regressions the instrument is the change in the predicted foreign-born population between years t and t-3 relative to the total population in the initial period. Robust standard errors clustered by area in parentheses. F-stat for columns (2) are 12, 11 and 14 respectively.

***Significant at 1%, **significant at 5%, *significant at 10%.

The IV coefficients for districts and neighbourhoods in table 6 are smaller than the ones in table 5. This means that the immediate effect for districts and neighbourhoods is stronger than the long effect, although slightly. It seems that no adjustment time is needed, potentially because there is a more dynamic flow through of people in smaller areas. For municipalities, the long term effect is stronger than the immediate effect. Its coefficient is interpreted as a change in the stock of immigrants equal to 1% of the local population during a three-year period generates an increase in house prices of about 1.06% in the same three-year period.

The long term analyses show that after a three-year period the effect of immigration on houses prices is stronger in municipalities, and weaker in districts and neighbourhoods. Typically immigrants buy housing after a certain period of residency, whereas recent immigrants prefer to rent (Akbari & Aydede, 2012). This could explain the strong long term effect for municipalities. The weaker long term effect for districts and neighbourhoods could be explained by the general increase in house price or through a weaker native flight effect. Possibly the native flight effect weakens over the course of time and is strongest at the initial period. These results are not in line with the literature, therefore causal inference cannot be taken with complete certainty.

Crime

Saiz and Watcher (2011) hypothesized that the negative connection between immigration and the change in house prices could potentially be explained through immigration attracting more crime. Table 7 shows the relation between immigration and crime, where the dependent variable is the total numbers of crimes divided by the area's population. There is no crime data available for districts. The immigration effect on municipalities is non-significant and very small for both the OLS and the IV regression. The effect on neighbourhoods is significant and can be interpreted as a 1% increase in the stock of immigrants decreases the total number of crimes with 0.03% for the OLS and with 0.01% for the IV. The average neighbourhood to which 18 extra immigrants move see their total number of crimes per thousand inhabitants lower with about 0.4. The difference between these could be explained by immigrants moving to less expensive neighbourhoods with a higher number of reported crimes. However, there is not enough information available for a causal explanation. The amount of variability in the crime variable explained by the focal independent variable is very small. There is likely a large difference of the effect between areas. Nevertheless, the results suggest, in line with Saiz and Watcher (2011), that there is no positive relation between crime and decreasing house prices through immigration.

Housing supply

Another approach, proposed by Gonzalez and Ortega (2013), studies the effect of immigration on the housing supply. Since a housing shortage is a problem in the Netherlands, immigration most likely has an effect on the housing supply. Table 8 shows the effect of immigration on the housing supply. Here the dependent variable is the change in the stock of houses between years t_{-1} and t in area a, divided by the total population in the initial year. It includes the same lagged control variables but adds a one-year lag of the log of house prices as an additional control. This control variable is added because housing supply responds to a change in price (Gonzalez & Ortega, 2013).

The differences between the OLS and IV estimates suggest a strong endogenous influence over all geographic areas. All coefficients decrease significantly, with municipalities turning negative after IV estimation. Only neighbourhoods stay significant for both analyses, possibly because of the large number of observations. The OLS results suggest that an increase in the stock of immigrants equal to 1% of the initial population leads to 1.51% more available houses in the concerned neighbourhood. If we control for immigrants attracting one another towards certain neighbourhoods, this decreases to 0.35%.

Table 8: Immigration and Housing Supply

	$\Delta HOUSES_{at}/POP_{at-1}$					
	Municipalities		Districts		Neighbourhoods	
	OLS	IV	OLS	IV	OLS	IV
$\Delta IMM_{at}/POP_{at-1}$	0.505	-0.265	1.731***	0.313	1.514***	0.345***
	(0.323)	(0.209)	(0.416)	(0.254)	(0.188)	(0.092)
Observations	1799	1799	14894	14894	35699	35699
Number of groups	336	336	2986	2986	8387	8387
R^2	0.161	-	0.600	0.197	0.431	0.180

Notes: Robust standard errors clustered by area in parentheses. Δ indicates first difference. Regressions include year and local area fixed effects. In the IV regressions the instrument is the change in the predicted foreign-born population between years t and t-1 relative to the total population in the initial period. The dependent variable is the change in the stock of houses between years t and t-1 in area a, divided by the initial population.
***Significant at 1%, **significant at 5%, *significant at 10%.

For an example the absolute number is calculated according to the following formula: ($\Delta HOUSES_{\mathbf{at}}$ = (coefficient/100) x average population)

E.g.

$$\left(\frac{1.51}{100}\right) * 1.800 = 27.18$$

$$\left(\frac{0.35}{100}\right) * 1.800 = 6.3$$

The number of extra dwellings is 27, or 6 after controlling for endogeneity.
This could imply that house prices go down when the number of immigrants in the area increases, after which the number of available dwellings increases (higher supply). The effect seems a lot weaker when accounting for the historical settlement of immigrant in area's. Neighbourhoods with a historically large immigrant stock see a rise in the supply of dwellings when immigrants move in, potentially because the more wealthy natives move out. However, it is difficult to make causal inferences since it was not the primary focus of the research. Additional data and literature should be added before the results can be interpreted causally.

Differences between Provinces

Additional research towards regional differences could help predict how the house prices respond in every province to an increased immigrant stock. To gain more insight into these differences, Table 9 shows the coefficient of the IV regressions for every province on the neighbourhood level. Moreover it shows the average stock of immigrants in each province. The province Friesland jumps out with a significant negative coefficient of -1.39 on the 5% level. The most positive effect is seen in Flevoland with a coefficient of 0.215 which is non-significant. Limburg also has a non-significant positive coefficient, the rest all have negative coefficients. The province with the lowest stock of immigrants within their borders, Friesland (7.4%) has the largest negative effect. While the province with the second largest stock of immigrants, Flevoland (23%), has the largest positive effect. It could suggest that the native flight effect is weaker when the stock of immigrants living in an area is already higher. Noord-Holland (22.8%) and Zuid-Holland (20.2%) both have a fairly large stock of immigrants but a very average (negative) coefficient. Both the most northern provinces, Friesland and Groningen, have a relatively large negative effect, while both southern provinces, Noord-Brabant and Limburg, have a relatively small negative or positive effect. They are not very significant so no real causal inference can be made. Without more specific research on native preference towards immigrants or education level we cannot be completely certain of the interpretation of this effect.

6. Conclusion

Research towards the effect of immigration on house prices has grown in different countries. Understanding the economic effect it has on the Netherlands, helps to clear fog surrounding this controversial topic. Public debate, and in turn public policy, could improve if there is additional research to base decisions on. This research tried to improve the existing body of literature on the effect, by estimating the relationship between immigration and house prices on three different geographic levels: municipalities, districts and neighborhoods. It does so by using two different approach in accordance with the most current literature. Using a large number of observations spread over municipalities, districts and neighborhoods in a seven year timespan this thesis tried to answer the research question.

The estimates show that an increase in the stock of immigrants equal to 1% of the initial population leads to a 0.78% increase of house prices in municipalities. The same increase in the stock of immigrants leads to a 0.46% reduction of house prices in districts and 0.26% in neighborhoods. Similar, but more positive results are found with the OLS estimation. This upward trend is most likely the results of a bias caused through endogeneity. One explanation for this bias is that immigrants locate towards more wealthy areas where house prices are rising faster. The difference in results for districts and neighbourhoods could be caused by the relative similarity between neighbourhoods within the same district. The difference of house quality between districts could be more pronounced. Districts are also more often a residential area with houses, while neighbourhoods have more diverse functions.

One explanation for the contrary effect between the larger and smaller levels is native flight (Saiz & Wachter, 2011; Sá, 2015; Larking et al. 2017). The more wealthy residents move out of districts and neighbourhoods in which immigrants choose to settle down. It decreases the house prices in certain areas while increasing the house prices in areas these natives move towards. Natives moving towards a more expensive area cause a spill-over effect by increase house prices in other areas (Mussa et al., 2017). The native flight argument explains a part of the differences between these geographic levels.

Looking at the effect of immigration on crime, the results underline earlier findings that crime levels decrease after an increase of the stock of immigrants. Immigration also has a increasing effect on the housing supply, possibly because natives move out. For provinces with a smaller stock of immigrant the native flight effect seems to be stronger. In provinces with a relatively large stock of immigrants the effect is weaker. Research towards these results should be examined more closely for real causal inference, which would make an interesting follow up study.

6.1 Research additions

The final notes are dedicated to how potential future research towards this subject could improve, given access to the right information and data. Extra Census data about citizens movement could improve knowledge on this effect. Immigration could have a native-flight effect on the native residents through its effect on house prices, which in turn weakens immigration's impact on the housing demand. Additional data on the displacement effect of native residents through the ratio of movers who were in a different area a year ago, to non-movers, could improve the quality of causal inference. Data on the native preference for immigrants would make the causal inference about differences between regions stronger. Additionally, data on the education level of the immigrants moving into a certain area is useful. In addition to the control variables used in this research, data concerning the primary group of population buying houses, ages 19-34, could explain some of the differences across locations. The variations in the age composition of the population explains differences in housing prices across locations. Specifically the age group 19-34 as the prime home-buying age group. However, available age data sorts ages from 15-24 and 25-45, which is too large of a group for interesting controls. Income data on a local scale is made unavailable for privacy reasons, but could add to the quality of the analysis.

The instrument used in the IV regressions is based on historical settlement patterns from base year 2013. The assumption that the historical settlement pattern of immigrants is not correlated with changes in economic performance of areas is stronger with a greater length of time. It would be better for the instrument if an earlier base year is picked. However, before 2013 there was a difference in the data denotation of the central bureau for statistics. This makes it harder to find old data for similar neighbourhoods and districts without losing many observations. Furthermore, the instrument does have two related shortcomings. The first problem with the shift-share instrument is the lack of data on immigrant backgrounds. Only the regions of the four largest ethnicities are recorded on a local level (Moroccan, Netherlands Antillean & Aruban, Surinamese and Turkish). Other non-western backgrounds and western backgrounds are grouped together. The second problem is that the instrument is unlikely to identify a well-defined causal effect when the nationwide immigrant inflow is relatively stable. This is because the inflow rates of immigrants across cities are often highly serially correlated (Jaeger et al., 2018). Solving this problem needs a clear decomposition of immigrant inflows by origin group. Unfortunately this demands more intensive data than was available.

8. References

Akbari, A. & Aydede, Y. (2012). Effects of immigration on house prices in Canada. *Applied Economics, 44*. 1645–1658.

Arellano, M. & Bover, O. (1998). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29-51.

Bartel, A. (1989). Where do the new US immigrants live? *Journal of Labor Economics, 7*(4), 371–91.

Betts, J.R. & Fairlie, R.W. (2003). Does immigration induce 'native flight' from public schools into private schools? *Journal of Public Economics, 87*(5-6), 987-1012.

Brüderl, J. & Ludwig, V. (2015). Fixed-effects panel regression. In H. Best, & C. Wolf (Eds.), *The SAGE handbook of regression analysis and causal inference* (pp. 327-355). Los Angeles, US: SAGE.

Dustmann, C. & Glitz, A. (2005). *Immigration, Jobs and Wages: Theory, Evidence and Opinion.* United Kingdom. London: CEPR/CReAM.

Dustmann, C., Fabbri, F. & Preston, I. (2005). The impact of immigration on the British labour market. *Economic journal, 115*(507). 324–41.

Dustmann, C., Frattini, T. & Preston, I. (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies, 80*(1), 145–73.

Jaeger, D. A., Ruist J. & Stuhler J. (2018). *Shift-Share Instruments and the Impact of Immigration*. NBER Working Paper No. 24285.

Larkin, M., Doucouliagos, H., Dubelaar, C., Klona, M., Newton, J., Stanley, T., & Vocino, A. (2018). *Do House Prices Sink or Ride the Wave of Immigration?* IZA Discussion Paper No. 11497.

Madsen, J. B. (2012). A behavioral model of house prices. *Journal of Economic Behavior & Organization, 82*(2012), 21-38.

Mussa, A., Nwaogu, U. G., & Pozo, S. (2017). Immigration and housing: A spatial econometric analysis. Journal of Housing Economics, 35, 13-25.

Poot, J. & W. Cochrane (2005). *Measuring the Economic Impact of Immigration: A Scoping Paper.* Hamilton, New Zealand: University of Waikato, Population Studies Centre, Discussion Paper No. 48.

Sá, F. (2015). Immigration and House Prices in the UK. *The Economic Journal (London), 125*(587). 1393-1424.

Saiz, A. (2003) Room in the Kitchen for the Melting pot: Immigration and Rental Prices. *Review of Economics and Statistic, 85*(3). 502–521.

Saiz, A. (2007). Immigration and housing rents in American cities. *Journal of Urban Economics, 61*(2). 345-371.

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

Stillman, S. & Maré, D. (2008). *Housing markets and migration: evidence from New Zealand.* Motu Economic and Public Policy Research. Working Paper 08–06.

Jeanty, P.W., Partride, M. & Irwin, E. (2010). Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. *Regional Science and Urban Economics, 40*(5), 343-342.

Winer, R. (1983). Attrition Bias in Econometric Models Estimated with Panel Data. *Journal of Marketing Research, 20*(2), 177-186.

Zorlu, A. & Hartog, J. (2001). *Migration and Immigrants: The Case of the Netherlands.* Tinbergen Institute. Discussion paper 2001-042.

Internet sources

Centraal Planbureau. (2020). Kansrijks woonbeleid. Retrieved from https://www.cpb.nl

Centraal Planbureau. (2020). CPB-Woningmarktmodel 2020. Retrieved from https://www.cpb.nl

Centraal Bureau voor Statistiek. (2021). *Ruim 69 duizend nieuwbouwwoningen in 2020.* Retrieved from http://www.cbs.nl

De Nederlandsche Bank. (2020). *Huizenprijs hangt meer samen met financieringsruimte koper dan met woningtekort*. Retrieved from http://www.dnb.nl

Wal, van der. A. & Tamminga, W. (2008). *Waarom de gemiddelde koopsom geen huizenprijs-indicator is.* Retrieved from https://www.cbs.nl

ABF Research. (2021). *Terugblik op het ontstaan van het woningtekort.* Retrieved from https://www.abfresearch.nl

Data sources

Centraal Bureau voor de Statistiek. (2020, 15 januari). *Kerncijfer wijken en buurten 2004-2020* [Data file]. Retrieved from https://www.cbs.nl

NVM. (2020). *Marktcijfers koopwoningen* [Data file]. Retrieved from https://www.nvm.nl

De Nationale Politiedienst. (2020, 15 januari). *Misdrijven per wijk en buurt per jaar* [Data file]. Retrieved from https://data.politie.nl

9. Appendix

Table 1: Socio-Economic Control Variables

Variable	Description	Data Source
Assistance Beneficiaries	The number of claimants receiving state benefits as a proportion of the resident population (age 17-65)	Central Bureau for Statistics (CBS), UVW
Incapacitated Beneficiaries		CBS, Institute for Employee Insurance (UVW)
Unemployment Beneficiaries		CBS, UVW
Non-violent crime rate	Number of non-violent crimes committed per 1000 population	National police database
Violent crime rate	Number of violent crimes committed per 1000 population	National police database
Dwelling stock/population	Number of houses available as a proportion of the total population	CBS, Dutch Association of Estate Agents

Table 2.1: Most Common Data Patterns

	Frequencies	**Cumulative Percent**	**Pattern**
Municipality	266	63,64	1111111
	30	70,81	111111.
	12	73,68	1…...
Other Patterns	110	100	
District	2.019	63,27	1111111
	227	70,39	….111
	161	75,43	111111.
Other Patterns	784	100	
Neighbourhood	6.573	52,76	1111111
	1.001	60,80	1……
	825	67,42	….111
Other Patterns	4.059	100	

Notes: '1' represents an observations while '.' represents no observation. A full pattern means observations for all years starting at 2013 and ending at 2019. All data is taken after dropping outliers.

Table 2.2: Number of Observations and Average Population per Level

	Municipality		District		Neighbourhood	
Year	Observations	Population	Observations	Population	Observations	Population
2013	408	41.125	2.553	6.570	10.079	1.655
2014	372	43.977	2.469	6.679	8.825	1.833
2015	364	44.984	2.527	6.510	8.760	1.828
2016	375	42.160	2.550	6.194	8.943	1.720
2017	363	45.311	2.691	6.017	9.333	1.703
2018	352	46.424	2.700	6.087	9.428	1.707
2019	327	49.737	2.656	6.019	9.155	1.702

Notes: All data is taken after dropping outliers.

Table 2.3: Number of Observations per Province

Province	**Municipality**	**District**	**Neighbourhood**	**Total**
Drenthe	83	1.094	2.747	3.924
Flevoland	40	275	1.145	1.460
Friesland	143	1.043	3.251	4.437
Gelderland	357	1.910	8.809	11.076
Groningen	130	655	2.676	3.461
Limburg	220	1.561	4.754	6.535
Noord-Brabant	429	2.573	8.976	11.978
Noord-Holland	322	2.410	8.663	11.395
Overijssel	171	1.609	5.813	7.593
Utrecht	173	1.105	4.395	5.673
Zeeland	90	922	1.950	2.962
Zuid-Holland	402	2.954	11.295	14.651
Total	2.560	18.111	64.474	85.145

Notes: All data is taken after dropping outliers.

Table 3: Descriptive Statistics (2013-2019)

Variable	Area	Observations	Mean	SD	Min	Max
Δlog house price	Municipality	2.153	0,055	0,048	-0,095	0,190
	District	15.593	0,055	0,069	-1,169	1,201
	Neighbourhood	54.444	0,056	0,072	-1,377	1,521
$\Delta IMM_{it}/POP_{it-1}$	Municipality	2.153	0,004	0,005	-0,018	0,064
	District	15.593	0,005	0,072	-0,703	2,652
	Neighbourhood	54.444	0,005	0,041	-0,792	2,490
$\Delta NAT_{it}/POP_{it-1}$	*	54.444	0,007	0,164	-0,933	11,762
$\Delta POP_{it}/POP_{it-1}$	*	54.444	0,012	0,191	-0,974	12.381
$\Delta Dwellingstock_{it}/POP_{it-1}$	*	54.444	0,001	0,032	-2,331	2,172
Unemployment rate	*	64.517	0,020	0,012	0	0,118
Benefits rate	*	64.064	0,018	0,024	0	0,647
Assistance rate	*	64.517	0,043	0,032	0	0,864
Nonviolent crime rate[1]	*	52.481	4,680	9,348	0,171	476,191
Violent crime rate[2]	*	44.439	2,845	4,045	0,105	87,500
Dwelling stock/POP	*	64.523	0,441	0,099	0,0714	3,015

Notes: The dependent and main independent variable are defined at their respective area level, the others are taken from the neighbourhood geographical level.
1 = Contains: Diefstal/inbraak woning, Diefstal/inbraak box/garage/schuur, Diefstal uit/vanaf motorvoertuigen, Diefstal van motorvoertuigen, Diefstal van brom-, snor-, fietsen, Zakkenrollerij, Diefstal af/uit/van ov. Voertuigen.
2 = Contains: Zedenmisdrijf, Moord, doodslag, Openlijk geweld (persoon), Bedreiging, Mishandeling, Straatroof, Overval.

Table 7: Immigration and Crime

	$\Delta(TCRIME/POP)_{at}$			
	Municipalities		Neighbourhoods	
	OLS (1)	IV (2)	OLS (1)	IV (2)
$\Delta IMM_{at}/POP_{at-1}$	-0.003	-0.004	-0.026***	-0.012***
	(0.005)	(0.004)	(0.004)	(0.003)
Observations	1798	1798	31467	31467
Number of groups	335	335	7523	7523
R^2	0.039	0.033	0.046	0.055

Notes: Robust standard errors clustered by area in parentheses. Δ indicates first difference. Regressions include year and local area fixed effects. The dependent variable is the number of reported crimes divided by the population.
***Significant at 1%, **significant at 5%, *significant at 10%.

Table 9: Immigration Effect and Stock of Immigrants per Province

Province	IV Coefficient	Average Stock of Immigrants
Drenthe	-0,039	7,9%
Flevoland	0,215	23,0%
Friesland	-1,390**	7,4%
Gelderland	-0,349***	12,3%
Groningen	-0,418	9,9%
Limburg	0,035	18,7%
Noord-Brabant	-0,152*	14,1%
Noord-Holland	-0,211*	23,2%
Overijssel	-0,346***	10,9%
Utrecht	-0,381	18,3%
Zeeland	-0,364**	15,8%
Zuid-Holland	-0,204**	22,3%
Average over all provinces	-0.260***	16,2%

Notes: The effect of immigration on house prices and the average stock of immigrants at the neighbourhood geographic level split per province.
***Significant at 1%, **significant at 5%, *significant at 10%.