ERASMUS UNIVERSITEIT ROTTERDAM

Erasmus School of Economics

Masterthesis Economics & Business

Master Specialisation Urban, Port and Transport Economics

The effect of accessibility to the labor market on the housing market

A study to the province of North-Holland

Name student: Romy Vreeswijk

Student ID number: 444471

Supervisor: M.J.A. Gerritse

Second assessor: F.G. van Oort

Date final version: 06-08-2021

Abstract

In this paper, research is done to the effect of accessibility and diversity on the labor market on the residential property prices. The literature review shows that there is a positive effect of labor market accessibility on the residential property prices, that different sectors of labor do have different effects on the residential property prices and that there is a positive effect of diversity in employment on the residential property prices. In addition, the literature review shows that these effects vary significantly for different types of dwellings. Subsequently, the theories and findings of the literature review are tested, for the province of North-Holland during the period 2004-2017, in the empirical part of this research. In this empirical research, a Repeat Sales model with fixed effects and instrumental variables is used to test the effect of the independent variables on the residential property prices. The data used for this is derived from the NVM and the LISA-bestand. Based on the results, it can be concluded that the sectorial mix in employment has a positive and significant effect on the residential property prices and that the magnitude of this effect varies for different types of dwellings.

Table of contents

1. Introduction

The housing market in the Netherlands has been growing fast for two decades (with the exception of crisis years), but in recent years there has been a massive excessive demand which has resulted in an overheated housing market. Despite the housing stock of about 8 million dwellings and an annual growth in the number of dwellings of one percent, there is a current housing shortage of about 331,000 which will increase to a peak shortage of 418,000 dwellings in 2025, according to ABF Research (CBS, 2021a; Groenemijer, Gopal & van Leeuwen, 2020). To reduce this shortage, plans are being made for large-scale new construction projects.

The accessibility to the labor market is often underexposed when determining the position of these new construction projects, but nevertheless very important and relevant. Not only is general access to the labor market important, but rather the access to different sectors of employment and the diversity within and between those different sectors is crucial. For example, it will be a huge boost for certain households if there is close access to a bank with jobs in the financial sector, while others prefer to live close to a supermarket with job opportunities and third parties prefer to live in a less densely populated area where there are employment opportunities in the agricultural sector. Besides the employment opportunities between different sectors, also the number of jobs within different sectors is important. For example, an employee of a restaurant will have more need for employment opportunities in a hotel than in the agriculture sector. A research by the Dutch statistical institute (CBS, n.d.) shows that in 2019, most people in the Netherlands were employed in the sector of business services, followed by the sectors of healthcare, trade and industry. These, and all other sectors each have their own effects on residential property prices in the area of residence. For example, international studies in the United States show that houses close to (and with people working in) the sectors of manufacturing and heavy industry are generally priced lower than in the services and FIRE (finance, insurance and real estate) sectors (Case & Mayer, 1996; Song & Knaap, 2004). This shows that different jobs are valued in different ways by households. The importance of the accessibility between residential property and employment opportunities is also reflected in facts: at least 80 percent of the working Dutch people consider the travel distance to work as an important factor for work-life balance and long travel times to work is the most mentioned reason for the search for a new job (28 percent), excluding redundancies (Kennisinstituut voor Mobiliteitsbeleid, 2018; SD Worx, 2018).

Besides this explanation, the labor and housing markets are two major topics when talking about urban economics and development that are always up for discussion. It is also therefore that this paper is socially relevant, since the combination of both these important markets is interesting for policy makers in the real estate market, the labor market and in urban development. Since the importance of these subjects, an appropriate policy is of great importance. Besides that, the results of this research

can be used in concrete terms for inner-city development and the question of where to find the value of cities for real estate buyers.

Furthermore, this paper is scientifically relevant, since many papers have been written about the effect of labor market accessibility in general, but only a few papers have been written about the effects between different labor sectors on residential property prices. Moreover, this study takes a more detailed look at the composition of employment opportunities. In addition, many of these studies took place in major countries such as the United States, China and Australia. Since the cities in the Netherlands are relatively close to each other compared to these countries (polycentric areas instead of monocentric areas), these results cannot be applied one-to-one for the Netherlands. Furthermore, most researchers of papers regarding this topic use a (standard form of) the Hedonic Price Method to draw their conclusions. In this paper, the Repeat Sales Method is used. Since this method is rarely used on this topic, this research has a contributing value to the existing literature on the effect of the labor market on the housing market.

The effect of the labor market on the housing market is therefore still very relevant, especially in a country like the Netherlands, where the spatial density is very different from that in larger countries like the United States. Therefore, the main research question of this paper is:

What is the effect of the sectorial mix in employment on the residential property prices?

To provide support in answering the research question, a number of sub-questions have been formed, these are as follows:

- Does job accessibility have an effect on the housing market?
- Are there any differences in the effects on the housing market of the accessibility to different sectors of employment?
- Do the employment sectors have different effects on the prices of different types of dwellings?
- Does more diversity in employment opportunities lead to higher residential property prices?

In order to answer these questions, the research in this paper has been divided into two parts. The first part of the research consists of a literature review in which various theories are discussed that are supported by previous studies. Based on those findings, a number of hypotheses were formed. These hypotheses were then tested in the second part of this study: an empirical research. By performing multiple regression analyses, the hypotheses may or may not have been rejected.

Therefore, the structure of this research is as follows: after the introduction follows the theoretical framework in which various concepts, the literature research and hypotheses are discussed. This is followed by a description of the data and the method used in the empirical research. Subsequently, the results of the performed analyses are presented and interpreted. The results are followed by a synthesis in which both parts of the research come together, the results of the literature research and the results of the empirical research are compared. The last part contains the conclusion in which the research and sub questions are answered and a discussion in which limitations, policy implications and suggestions for further research are mentioned.

2. Theoretical Framework

2.1 The housing market

The housing market is a market that cannot be seen as a 'normal' market. Due to the fact that no house is the same and because of the many characteristics that are taken into account, it is difficult to determine the value of a single house. An important and one of the most used methods to explain the variation in house values is the Hedonic Price Method (HPM). With this method, the value of goods without a fixed market price (but with properties and characteristics that give utility) can be determined. The HPM is based on the idea that consumer goods, in this case dwellings, "are not homogeneous goods and differ with respect to a variety of characteristics" (Rosen, 1974; Visser, van Dam & Hooimeijer, 2008). The HPM consists of a regression equation whose function depends on the determinants of a unit's value. The most used determinants for the value of a dwelling are the structural characteristics of the dwelling, the neighborhood characteristics, the location within the market, the contract conditions or characteristics and the time value is observed (Malpezzi, 2003). Since in this paper the effect of the sectorial mix in employment is assessed, the focus is on the determinant location within the market, and then especially the location within the labor market.

In this research, the Repeat Sales Method will be used. This model is another pricing model that is closely related to the HPM. Repeat Sales models use data in which all units are sold at least twice within a period, with the advantage that certain determinants are (assumed to be) fixed over time (Case and Shiller, 1987; Malpezzi, 2003; Haan & Diewert, 2013). More about this repeat sales method will be discussed in the methodology.

2.2 The labor market and its accessibility

As discussed in the previous section, the interesting determinant of housing prices for this paper is the location of a dwelling within the labor market. This location within the labor market is often seen as the distance to job opportunities as the "accessibility potential to employment opportunities" (Osland

& Thorsen, 2008). For this research, the location within the labor market (or labor market accessibility) is defined as the number of accessible jobs within the maximum desired travel time of households. How this is practically determined for this research, is described in the data and methodology sections.

The first theorem of this research is that there is a significant and positive effect of the number of accessible jobs on the residential property prices. The mechanism behind this theory is as follows: when there are more jobs accessible from a certain area, this area is more attractive for households assuming that the heads of the households belong to the (potential) labor force. When an area is more attractive for households, there will be an increase in demand for housing, resulting in higher residential property prices. In addition, research has shown that within a commuter flow there is a travel time decay and that this does affect land prices. In other words, not only the presence of employment opportunities matters, but also the travel time to work is important for homeowners. Ahlfeldt (2011) has found in his research that commuting flows have a travel time decay function: the spatial weight of employment opportunities decreases as travel time increases. Furthermore, Ahlfeldt and Wendland (2016) found in their research that the "estimated spatial decay in employment potential models is reflective of the cost of commuting in polycentric regions" and that the "decay in commuting probabilities can be inferred from the spatial distribution of land prices and employment".

The traditional and most simplistic models in terms of the labor market assume that cities are monocentric. In other words, cities consist of a centrally located spatial reference point (the centre) in which all activities such as employment and business take place, surrounded by a circular residential area (Giuliano, Gordon, Pan & Park, 2010). In this case, accessibility to the labor market is equated to the accessibility to the central business district (CBD), since everything takes place in the CBD. The more traditional papers therefore start from this theory. In reality, however, the larger and moderner cities and the smaller countries consist of many centres, which all act as "local focus points for business and commercial activity" (McCann, 2013). Most jobs are not at the centre (as assumed in the monocentric city), but are dispersed throughout the whole metropolitan area. A schematic representation of the two dimensions of urban spatial structures is shown in figure 1. Especially in relatively smaller and densely populated countries (like the Netherlands), there are areas with many local focus points in which job opportunities are located. In this case the labor market accessibility is no longer equal to the accessibility to the CBD. This is also reflected in the more recent literature. For example, Agarwal, Giuliano and Redfearn (2012) found, in their research to the usefulness of exploring polycentricities, that the traditional CBD still is the largest employment centre in the region, but that there are many other highly relevant concentrations of employment beyond the CBD. To come to this conclusion, they acknowledge that the measure of labor market accessibility is important since it accounts for the fact that jobs are not generally located in a single node of a region.

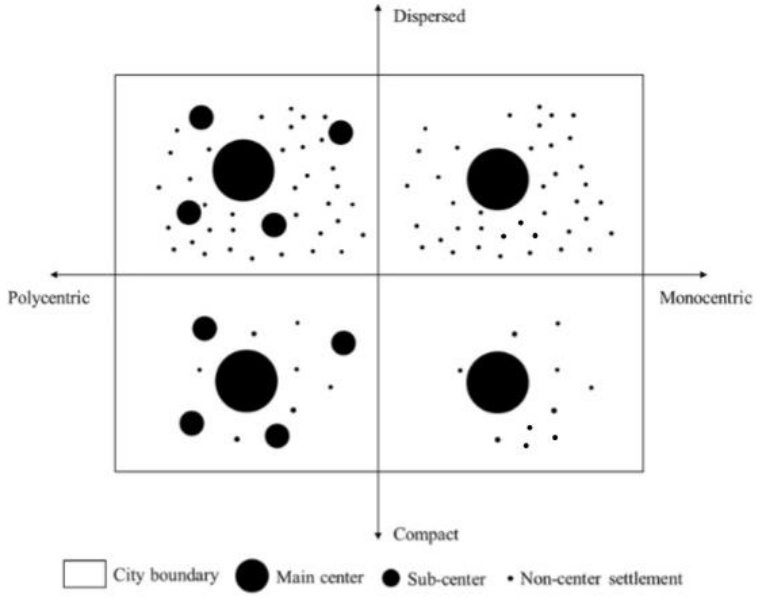

Figure 1: Two dimensions of urban spatial structures, monocentric versus polycentric city. (Li, Xiong & Wang, 2019).

The theory that jobs are not centralized in the city centre but that they are spread all over the area is also important for determining residential property prices. As discussed, spatial variation in dwelling prices is among others explained by urban attraction and labor market accessibility effects. Households determine which regions are relevant in their housing decision, since "households are, ceteris paribus, assumed to prefer a location with favorable job opportunities within a reasonable distance from their residential site" (Osland & Thorsen, 2013). Recent years, several studies have been done worldwide which have taken into account the theory that the spatial structure and local variation in the labor market affect the willingness to pay for dwellings. For China, especially for the region of Beijing, results are found that the number of accessible jobs (that are spread over the city) have a positive and significant effect on the property values (Li, Chen & Zhao, 2019). This implies that areas with more employment opportunities have relatively higher property prices compared to other areas. In other words, the more job opportunities employees can reach, the more they are willing to pay for a dwelling. This is also found for Europe, especially for Norway. Results show that labor-market accessibility has a positive effect and contributes significantly to the variation in housing prices, when a gravity-based labor market accessibility measure is used (Osland & Thorsen, 2008). This measure is used to explain variation in housing prices while accounting for the fact that jobs are not entirely concentrated within the CBD.

This theory does not only seem to hold for China and Norway, but there is evidence to support this theory for the United States as well. Iacono and Levinson (2017) for example have researched accessibility dynamics and in particular the question how changes in employment over time affects

variation in house prices. Just as has been found by the papers discussed, they found a positive and significant effect of labor market accessibility on housing prices. Their results show that, when all else is constant, each additional 100,000 jobs available within a certain range implies a 2,000 US Dollar increase in the sales price of a house. Furthermore, Srour, Kockelman and Dunn (2002) found the evidence to adopt this theory. In their research on the relationship between connection, land prices and location choices they found that the workplace of the household's head plays an important role in location predictions. Their results show a positive and significant effect of the sum of the total available jobs within 30 minutes of travel time on the property values, where the standardized coefficient on labor market accessibility is equal to 0.291. This implies that one standard deviation increase in labor market accessibility contributes to a rise of 29 percent of its standard deviation in property values. Besides that, the signs and magnitudes of the accessibility related variables suggest that labor market accessibility contributes most to the price of a parcel of land. So, it can be concluded that the accessibility to the labor market plays a significant role in the variation in dwelling prices.

In addition, Giuliano et al. (2010) also have tested the effect of aggregate job accessibility on the residential land values. They found a significant and positive coefficient of 0.000111, which implies that an increase of one unit in overall job accessibility will cause an average increase of 0.011 percent in house prices. Last, like Giuliano et al. also Kim and Jin (2019) shows a significant positive effect of job accessibility on housing prices. In their paper, Kim and Jin analyses how job accessibility satisfies housing consumers' needs in Chicago through the use of multiple instrumental variables. The results of their research implies that one increase in job accessibility leads to a 1.5 percentage increase in housing prices.

So the literature generally shows a positive, significant effect of labor market accessibility on prices of dwellings. Positive effects are found for both relative measures (accessibility ratios) and absolute measures (total number of jobs) to explain the variation in house prices. That is why for this research also a significant and positive effect of the number of accessible jobs on the residential property prices is expected. However, the size of the effects cannot be replicated one-to-one since countries like the United States and China are very different from the Netherlands in spatial terms. The extent to which the magnitudes of the effects correspond to each other will therefore be further researched in this study.

2.3 Different sectors of labor

The second theorem of this research is that different sectors of employment do have different effects on the residential property prices. As discussed, the overall effect of labor market accessibility on residential property prices is significant and positive according to the literature. This raises the question whether all sectors can be measured by one single measure. The first mechanism behind this

theory is that different sectors pay different wages. Some sectors pay higher wages than other sectors, resulting in the fact that employees of those sectors can afford more expensive houses than other employees can. The second mechanism behind this theory is that some sectors of employment make living in a certain area more attractive than other sectors. For example, when there are relatively many jobs in sectors like hospitality, culture, sports and recreation in a certain area, it implies that there are amenities and services in that area. In general, these amenities and services have a positive attractiveness to households, resulting in an increase in demand for housing in that area and thus an increase in residential property prices. On the other hand, there are also sectors of employment where households prefer not to live nearby. Given that different sectors of employment represent different types of activities, it can be expected that some sectors will have positive effects (like the example of amenities and services) while other sectors are more known due to their negative externalities (Giuliano et al., 2010). Employment sectors such as heavy industry are not desired near a dwelling and will be therefore avoided by households as the areas that belong to those sectors cause negative externalities such as pollution and (noise) nuisance. In addition to the different wages that belong to the different sectors, the different environmental factors and externalities (both positive and negative) also have an impact on a households' willingness to pay for a dwelling surrounded by these different sectors. Another explanatory mechanism between different sectors of employment and the residential property prices is the establishment of the different sectors. This mechanism is preferred not to be measured in a research, because when measuring this there will occur a bias. Some (more specialistic) sectors, like for example business services, will mainly occur in the relatively larger cities, while sectors such as agriculture will mainly occur in smaller, less densely populated areas (Brueckner, 2011). A bias will occur since residential property prices are in general higher in relatively bigger cities than in the less densely populated areas (Shiller, 2007). Altogether, when taking all mechanisms between different sectors of employment and residential property prices into consideration, it is actually inappropriate to show the (diverse) effect of the labor market by just one single measure. Therefore it would be better to look at the different sectors separately to determine the effect of (different sectors of) employment on residential property prices.

As mentioned before, also for the different sectors of employment already a lot of research has been done on the effects of different employment sectors on the residential property prices all over the world, especially in the United States. For example, it seems like the retail sector has a positive effect on the residential land prices within this country (Srour et al., 2002). Besides the overall effect of job accessibility, the effect of the number of jobs in the retail sector by itself is assessed in this paper. This effect appears to be significant and positive, but relatively small compared to the overall effect of employment on the residential land prices (about 6.5 times smaller). And also the accessibility to the service employment sector seems to have a positive effect on the residential property prices (Song & Knaap, 2004). While the effect of the ratio of the total number of jobs is negative (but insignificant),

the coefficient for the ratio of service jobs is positive and significant, which implies that the sales prices increase when there is an increasing amount of service jobs in the neighbourhood. Service jobs are here defined as population-serving jobs in sectors such as retail, personal services, entertainment, health and other professional and related services. Another paper confirming this is that from Case and Mayer (1996). They found a negative effect of the fraction of residents working in manufacturing compared to the fraction of residents working in services and FIRE (finance, insurance and real estate) sectors, on the nominal house prices.

And also for China the theory that different sectors of employment do have different effects on residential property values is confirmed (Hu, Geertman & Hooimeijer, 2014). In this research the relative housing price effects of, among other things, accessibility in China is assessed. To determine the effect of job accessibility on house prices, job accessibility measures of four sectors of employment are included in the research: education, governmental institutions, financial and business services and the heavy industry sector. A main finding of the paper is that the effect of job accessibility on house price differs depending upon the specific sector of employment. Accessibility to jobs in the three service sectors has strong significant and positive effects, while the sector heavy industry represents both a positive and negative effect. For the closer distance, less than 20 minutes of travel time, the heavy industry sector faces a negative effect on the house prices, while for the distance of 20 till 40 minutes of travel time there is a positive effect of heavy industry. The standardized coefficients of -0.027 and 0.028 indicate that mainly negative externalities are experienced at a shorter distance, while these become positive for the longer distances, which then results in the residential property prices. In comparison, the standardized coefficients of the sectors education, governmental institutions and financial and business services are respectively 0.130, 0.096 and 0.075, so the effects of these sectors are relatively larger compared to both the positive and negative effects of heavy industry.

A more comprehensive paper that covers the effects of different sectors of labor is the paper of Giuliano et al. (2010). Besides the overall effect of labor market accessibility, they have developed in their paper different accessibility measures based on industry sectors to allow for separating out possible different effects on the sales value. The coefficients of all 8 different job accessibility factors were found to be significant, of which five coefficients indicate positive effects and 3 coefficients indicate negative effects. After standardizing the variables (to make it possible to compare the magnitudes of coefficients), the results show that the biggest positive effect on the sales value is found for the wholesale, trade and warehousing industries (a standardized coefficient of 5.024). After this sector, the biggest positive effect is coming from the sector broadcasting and information services, closely followed by the retail and services industry (with standardized coefficients of 3.808 and 3.479). Last, there are positive effects from both the utilities and construction industry (2.416 and

1.487). The biggest negative effect on the sales value that Giuliano et al. found belongs to the jobs in heavy industry. With a standardized coefficient of -6.998, this industry has the biggest effect (positive or negative) on the sales value of all sectors. This is not surprising since this industry is known for its negative externalities such as noise and pollution. Besides the heavy industry, also the sectors public administration and professional services represent negative effects on the residential land values. With standardized coefficients of respectively -4.543 and -4.449, the coefficients with a negative sign are relatively higher than the positive coefficients. Finally, they conclude that the selected job accessibilities indeed have significantly different impacts on the residential land values, but they also make the footnote that the addition of the various sectors explain only 15 percent of the variation in sales values.

The United States and China are not the only countries in which this research has been conducted, a corresponding research has been done for the Netherlands for the city of Rotterdam. Koster and Rouwendal (2012) have researched the value of mixing in employment and residential land use. Besides the overall effect of employment they split up employment in seven different sectors: business services, education and healthcare, government, manufacturing, leisure, retail and wholesale. For the overall effect of employment they found that a standard deviation increase in employment leads to an increase in house prices of about two percent. When the results are presented by sector, both positive and negative effects are found. Sectors that affect house prices in a positive way are leisure, business services, education and healthcare and retail, while negative effects are found for the sectors wholesale, manufacturing and governmental institutions (both descending in size of effects). Although Koster and Rouwendal and the other researchers discussed all used different sectors of employment, there are a number of similarities. The most commonly found result is that the sector heavy industry (manufacturing) has a relatively large negative effect on the residential property prices and the sectors education and healthcare and financial and business services are valued positively. The sectors where there is some discussion between the papers are governmental institutions and wholesale (warehouse), for both sectors are found positive and negative effects.

So the literature generally supports the theory that it is inappropriate to define labor market accessibility by means of just one single overall measure and indicates that the accessibility of different sectors of employment have different effects on the residential property prices. This result is therefore also expected for this study.

2.4 Diversity

The next theorem of this research is that diversity in the sectorial mix in employment does affect residential property prices in a positive way. So far, theories about the effects of the amount of jobs

and the accessibility to different sectors on the residential property prices are discussed. Literature shows that there is a positive effect of labor market accessibility on residential property prices and that this effect is different in sign and size for different sectors. However, those researches did not take into account the distribution of employment sectors across the region. Diversity in employment is a frequently recurring link in the labor-economics literature. For example, diversity of employment is an important element within the localization versus urbanization debate. Localization economies are the "agglomeration economies which accrue to a group of firms within the same industrial sector located at the same area" which are therefore associated with a (more) specialized area (McCann, 2013, p.54; Helsley & Strange, 2014). On the other hand, urbanization economies are "those economies of agglomeration which accrue to firms across different sectors" and are therefore associated with a (more) diversified area (McCann, 2013, p.55). An argument in favor of urbanization is that a more diverse area provides opportunities for unplanned, unpredictable or unusual connections between different sectors of employment which leads to creation of new work, while localization benefits from spillover effects like matching, sharing and learning (Faggio, Silva & Strange, 2017). Some sectors of employment thus benefit in terms of productivity from the presence of other sectors of employment. In addition, it is also desirable for households that there is a diverse supply of labor in the area around their dwelling, as there is an effective overlap between some sectors. For example, if an employee has lost a job in the sector of financial services, this person could probably also work in the sector of business services, but it is unlikely that this person would be suitable for the agriculture sector. In addition, the mix in sectors is also important for, for example, people who work in seasonal sectors of employment. Due to the presence of a mix of sectors, those people have opportunities to work throughout the whole year. Altogether, it is not unlikely that (movements in) the sectorial mix in employment does affect the variation in residential property prices (Case & Mayer, 1996). It has already been shown that housing prices are higher in neighbourhoods where non-residential land uses are evenly distributed and that mixing a variety of amenities and employment opportunities create better urban environments (Wu, Song, Liang, Wang & Lin, 2018; Friedman, 2021). It is therefore that (some) zoning regulators want to move away from "separate facilities philosophy" and move to a "mixed land use philosophy", so the optimal mix of land use and employment activities should be sought, not the regional separation of activities (Cao & Cory, 1981, 1982).

Among others Cervero & Duncan (2004) researched this in their paper about the relationship between indicators of land-use diversity and jobs-housing balance and residential land prices in the United States. One of those indicators is the employment diversity index, which has taken into account the diversity of employees of the sectors retail, services, offices, trade, agriculture and other sectors. The results show that, all else being equal, a single-family parcel in a neighborhood with a maximal mix of employment (diversity index is 1.00) is worth around 8.70 US Dollar more per square meter than a comparable parcel in a neighborhood with minimal mix of employment (diversity index is 0.00). The

same applies, to a lesser extent, to the multi-family parcels. So in general, more diverse use of land (and so in employment) leads to higher residential values.

And also for the Netherlands applies that higher diversity in employment has a positive effect on residential property prices (Koster & Rouwendal, 2012). The results for the region of Rotterdam shows that when the overall effect of employment is measured, a diversity index has a small positive and significant impact on house prices. But when this overall effect of employment is split up in seven different sectors, the positive and significant effect of the diversity index rises and becomes about 4.5 times bigger than with the overall effect of employment. When different sectors of employment are taken into account, it appears one standard deviation increase in employment diversity increases homeowners willingness to pay with about 2.5 percent. This implies that first, the effect of diversity in employment has a significant and positive effect on the house prices and second, diversity in employment has a bigger effect on house prices when the different job opportunities are taken into account.

So, the theory of a positive effect of accessibility to a diversified labor market is thus supported by various papers. It has been shown for the Netherlands but also for other regions worldwide that a more diverse labor market significantly explains the variation in residential property prices. That is why a positive effect of diversity in employment activities on the residential property prices is expected for this research.

2.5 Different types of dwellings

The last theorem of this research is that the effect of accessibility to the labor market (per sector) and the effect of diversity varies for different types of dwellings. Next to different sectors of labor, also the distinction can be made in the effect of labor on different types of houses. It is likely that occupiers of different types of dwellings have different willingness to pay for overall job accessibility, accessibility to different sectors of employment and job diversity. A research that matches with this theory is that of Lin and Cheng (2016). For the region of Taiwan, they researched the relationship between job accessibility and different rent-levels for apartments. Their results show a positive effect of job accessibility on apartment rents (whole sample), but implies that job accessibility affects apartment rents differently depending on the rent level. The coefficients reveal that the effects of job accessibility on apartment rents declines with rent levels. More specifically, job accessibility has a positive effect on median and lower-rent level submarkets, while the effect on higher-rent-level submarkets is negative (but insignificant). It can therefore be concluded that higher valued dwellings are less affected by job accessibility.

A more comprehensive paper that researches the discussed theory is that of Giuliano et al. (2010). They have looked at the effect of 8 different sectors of employment on different sub-markets of dwelling prices. Their sample is divided in three sub-markets: low, middle and high valued dwellings. The results show that all the significant coefficients have the same signs for the low, middle and high valued dwellings within the same sector (the sectors retail and services and professional services do have different signs per sector, but both those coefficients are insignificant). This implies that within a sector the size of the effect differs, but that the positive or negative is constant for dwellings of different price levels. However, the various sectors do have different effects on the three sub-markets. For example, the sector with the largest relative impact on the low values dwellings in broadcasting and information, while for the middle and high valued dwellings the effect is the biggest for the sector wholesale and warehousing.

For this research, there is no use made of the distinction in price-level, but the distinction is made between types of dwellings. A study that does the same is that of Koster and Rouwendal (2012). Besides the effect of different sectors of employment, they considered the effect of different sectors of employment and diversity on the residential property prices for different types of dwellings for the city of Rotterdam. They tested whether the effect of employment on house prices differs for different types of dwellings by looking at the effects of diversity and seven sectors of employment on the house types: apartments, terraced dwellings, semi-detached dwellings and detached dwellings. The results show that all sectors except business services have significant effects on the different types of dwellings. For the sectors manufacturing, leisure and wholesale applies that the sign of the coefficients is the same for all four types of dwellings (respectively negative, positive and negative), where the signs of the coefficients of the other sectors differ per type of dwelling. For example, the education and healthcare sector is valued positively for every type of dwelling except detached dwellings, while the retail sector is valued positively only regarding apartments. The results also show that there is "substantial heterogeneity in willingness to pay for diversity", only apartment occupiers are willing to pay for an increase in diversity, whereas the effects of diversity on the other types of dwellings is negative. These results confirm that some variation in residential property prices is explained by housing attributes and especially by different house types.

So the literature shows that not only the effect on residential property prices varies for different types of sectors, but that also the type of dwelling affects the prices. That is why this result is also expected for this research.

2.6 Hypotheses

So, various studies worldwide have already been conducted into the effect of the sectorial mix in employment on the residential property prices. From the literature discussed it can be concluded that

in general the effect of labor market accessibility on the residential property prices is significant and positive. Besides that, it seems to be inappropriate to define labor market accessibility by means of just one single measure and therefore the results of the literature shows that the accessibility to different sectors of employment have different effects on the residential property prices. In addition, it appears that homeowners appreciate accessibility to a more diverse range of jobs, which also affects the prices of dwellings in a positive way. Last, the results of the literature discussed shows that effects of the employment sectors vary for different types of dwellings and that also the effect of diversity varies for different types of dwellings. So, based on the results of the literature review, the following five hypotheses have been formulated:

Hypothesis I: There is a significant and positive effect of the number of accessible jobs on the residential property prices.

Hypothesis II: The effect of employment on the housing market varies significantly between employment sectors.

Hypothesis III: Diversity in the sectoral mix in employment has a significant and positive effect on the housing market.

Hypothesis IV: The effects of the employment sectors vary significantly for different types of dwellings.

Hypothesis V: The effect of diversity in the sectoral mix in employment varies significantly for different types of dwellings.

These hypotheses will be used to provide a better answer to the research question.

3. Data

To determine the effect of the sectorial mix in employment on the residential property prices, a quantitative research has been conducted. Two different datasets are used to perform the analyses for the period 2004-2017.

3.1 Housing market data

The first dataset is obtained from the NVM, the Dutch cooperative association of real estate agents and valuers that provides various types of information on the Dutch property market (NVM, n.d.). This dataset contains information about among other things the transaction price, the date and year of transaction, the location and various characteristics of real estate properties located in the province of North-Holland for the period 2004-2019. The primary dataset contains 414,561 observations and 65 different variables.

3.1.1 The purification of the data

Before the data could be used, the dataset has been reviewed and purified for irregularities. First, all observations which are categorized for other purposes than housing (residential property) are removed from the dataset. Second, all observations with missing transaction prices, transaction years, location and other important variables have been removed. Then, the descriptive statistics were examined for extreme and incorrect values. Many extremes were found, such as house prices of one euro or a billion euro, houses with zero or 21 toilets and houses with 109 rooms. Furthermore, inaccuracies such as houses with a surface area and volume of zero (or just a few) m2 and m3 were also found in the dataset. Since observations with those extreme values and impossibilities have an unintentionally large effect on the results, these observations have been removed from the dataset. To correct for these outliers, the observations that represent the 1.0 percent lowest and 1.0 percent highest values for the variables transaction price, number of toilets, number of rooms, the surface area and the volume are eliminated. After this correction, all used variables comply with the rule of thumb for outliers that the median and the mean may differ by no more than three standard deviations from each other (Stock & Watson, 2014).

As discussed before, only dwellings that have been sold at least twice are considered in this research (repeat sales method). So the next step taken in the process of data preparation is the removal of dwellings that were only sold once in the period 2014-2017. After the first purification of the NVM dataset, 187,043 observations are left. When running a duplicates report, it turned out that 132,026 unique combinations containing the postal code (zip code), house number and additional numbers or characters were only sold once (table A.1, Appendix A). These single sold observations have been

removed, leaving 55,017 observations of which 25,978 are unique combinations of postal code, house number and additional numbers or characters.

3.2 Labor market data

The second data set of this research is the 'LISA-Bestand'. LISA is the Dutch "national job information system and is a database containing information about all branches in the Netherlands where paid work is performed" (LISA, n.d.). The LISA-Bestand is very useful since the employment data is linked to a spatial component (address data). The primary dataset used for this research contains 8,000,116 observations for the period 2000-2017. The dataset represents for every sector (SBI08 5 digits – classification according to economic activities) the number of available jobs per year, whereby the location of the available jobs is aggregated on numerical zip code level (PC4-level). The LISA-Bestand includes observations for all postal codes that are available for the Netherlands.

3.2.1 The purification and transformation of the data

Just like with the NVM dataset, the first step to be able to use the data was to purify the dataset. All observations with missing or unknown SBI08-codes, postal codes, years and number of available jobs are removed. The next step is the reduction of the number of different sectors or economic activities in the dataset. The primary dataset contains 929 different and specified SBI08-codes, but since only the main sectors are interesting for this research, a number of SBI08-codes are merged. First, all 5 digit codes have been converted to 2 digit codes, according to the standard corporate classification (CBS, 2021b). Then, the 99 remaining SBI08 2 digit codes are divided into 13 different main sectors of employment (table 2). The complete classification of SBI08 into the 13 main sectors can be seen in table A.2 in Appendix A.

Table 2: The main sectors of employment.

Sector	Description
S1	Trade, transport and storage
S2	Production and installation, except industry (S12)
S3	Construction
S4	Agriculture
S5	Hospitality
S6	Information and telecommunication
S7	Financial services; renting of real estate and moveable property
S8	Business services
S9	Government, education and healthcare
S10	Culture, sports and recreation
S11	Personal services
S12	Industry
S13	Other sectors

After the transformation, each observation in the dataset contains the number of jobs per sector (S1 till S13) for a specific zip code and a specific year.

3.3 Combining the housing and labor market

The last step in the preparation of the data is combining both datasets. By combining these datasets, a unique access to the labor market can be calculated for each dwelling. To do this, first the practical definition of labor market accessibility has to be determined. Several studies have shown that the TTR (travel-time ratio) for Dutch employees is between 0.11 and 0.13, which means that the willingness to travel to work is no more than 11 till 13 percent of the hours worked on a regular day (Susilo & Dijst, 2009; 2010). This would mean that for an average working day of 7 to 8 hours, employees would not want to travel longer than about 45-60 minutes. That is why the practical labor market accessibility for this research is determined as the number of total available jobs (per sector) that can be reached within 45 minutes of travel time from a particular dwelling.

The determination of which postal codes are within and which postal codes are beyond the labor market accessibility from the different postal codes in the province of North Holland is done in QGIS. Using the top 250 roads layer from the dataset BGT (Basisregistratie Grootschalige Topografie) from PDOK, the fastest ways between the middle points of all postal codes (PC4) of the Netherlands are calculated by the OD matrix. In doing so, the default speed has been set to 20 kilometers per hour and the maximum speeds of the A-roads (Dutch highways) have been reset from 130 and 120 to 100 kilometers per hour in accordance with the speed change on the Dutch highways (Rijkswaterstaat, n.d.). The output of this OD matrix represents the travel times between all postal codes areas. In this matrix, all observations with origins postal codes other than that from the province of North Holland and the connections with travel times longer than 45 minutes are removed. These results are plotted in figure 3. The map shows the total labor market accessibility for the entire province of North Holland.

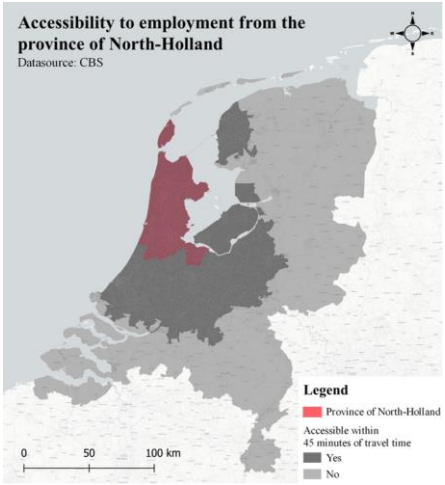

Figure 3: The accessibility to employment from the province of North-Holland within 45 minutes of travel time.

Instead of adding up all the jobs in the relevant range, the next step is the insertion of a travel time decay function. Given the spatial structure of the Netherlands, there will be a (large) partial overlap of the potential labor market between a majority of the dwellings observed. The jobs that explain variation in house prices will therefore be at the edges of the relevant range of the labor market. In addition, it is likely that job opportunities with shorter travel times are valued differently by households than opportunities with relatively longer travel times. For determining the decay function, the paper of Ahlfeldt (2011) has been considered. In this paper, Ahlfeldt uses multiple factors to estimate decay parameters for the city of Berlin. He found that spatial weight of labor market accessibility decreases as travel time increases (figure 4). Since this study does not look at area boundaries, but purely at travel times to job opportunities, the spatial weights associated with the top, solid line in figure 3 are adopted for this research. It is assumed that this decay function also can be used for the data about the Netherlands, since the coefficient from Ahlfeldt's research is approximately equal to the coefficient of the average Dutch decay function: the beta for Berlin is estimated at 0.1042 and the beta for the Netherlands is between 0.11 and 0.13 (Susilo & Dijst, 2009; 2010). To take into account the travel time decay function, the available jobs are all multiplied with the spatial weight associated with the travel time from the particular dwelling to those jobs.

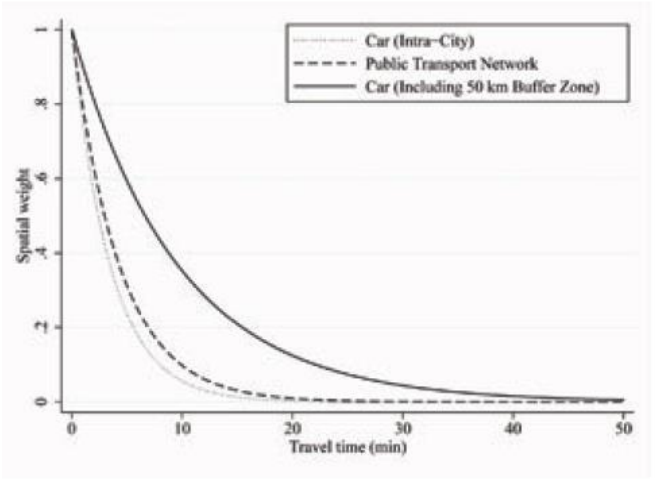

Figure 4: The travel time decay function. Source: Ahlfeldt (2011)

After it is determined which postal codes can be reached within 45 minutes of travel time from other postal codes, the total number of accessible jobs per sector for each postal code from the province of North Holland was added separately for each year. So for each postal code in the province of North Holland the number of available jobs per sector (S1 till S13) and the total number of available jobs that can be reached within 45 minutes of travel time are known, for each year separated. So the total labor market access (per sector) is defined as follows:

$$E_{s,i,t} = \sum_j d_{ij} * A_{s,j,t} \quad (1)$$

Where *d* represents the decay factor belonging to the travel distance between residential area *i* and employment area *j*, *A* represents the number of reachable jobs within the range in sector *s*, in employment region *j* and time *t* and *E* represents the number of employment opportunities (corrected for the travel time decay) in sector *s*, in residential region *i* and at time *t*. Two maps which visualize *E* and *A* for all sectors together are shown in figure A.4 and A.5 in Appendix A (2017 is taken as an example). When comparing those two maps, the difference between those two is clearly visible. When looking at figure A.4, the number of employment opportunities without travel time decay are shown per region. It is striking that the closer an region is to or in the Randstad, the more jobs are available. This is resulting in the lowest number of jobs in the North of the Province while the highest number of jobs is reached in the South of the Province. After the addition of the travel time decay factor (figure A.5), the highest and lowest number of reachable jobs are more spread over the Province. On this map, the areas with the highest number of jobs are more concentrated around the Amsterdamse Zuidas.

Finally, this dataset is merged with the NVM dataset based on year and postal code. Not all observations could be correctly merged together, so a small number of observations were dropped. Therefore, the final dataset consists of 47,572 observations of which 22,538 are unique combinations of postal codes (PC6), house numbers and additional numbers or characters. All those observations contain both variables of the characteristics of the dwelling and the accessibility to the labor market.

3.4 Creating the diversity index

Besides the effect of the overall accessibility and the accessibility to different sectors of employment, this research also takes into account the effect of diversity in employment. The level of diversity is calculated by the Herfindahl-Hirschman-index (HHI). The HHI is a measure of the size of firms or sectors in relation to the industry or market (Rhoades, 1993). The formula of the Herfindahl-Hirschman-index is as follows:

$$HHI = \sum_{i=1}^{n} M_i^2 \tag{2}$$

In which *M* is the market share of sector *i* on the labor market and *n* is the total number of sectors on the labor market. The market share can be calculated with the following formula:

$$M_i = \left(\frac{S_i}{Total\ Available\ Jobs}\right) \tag{3}$$

Where *S* is the total number of jobs in sector *i* and the total available jobs are all jobs of sector 1 to sector 13 added together. Both S_i and the total available jobs are corrected for the spatial weight of employment opportunities, so the travel time decay factor is included in both variables. The HHI has a range from 0 to 1, where a value of 1 indicates that the market is dominated by only one sector. The index will get close to zero when there is a large number of sectors that all have approximately equal

(small) market shares. A lower value of the HHI therefore indicates that there is relatively more diversity. Since it is easier to interpret, for this research the variable diversity (*D*) is created as the reversed value of the HHI. So when the value of the variable diversity moves towards 1, there is relatively much diversity and when this value moves towards 0, there is relatively little diversity.

3.5 Descriptive statistics

3.5.1 Transaction price as dependent variable

The dependent variable in this research is the transaction price of dwellings. After the removal of the outliers, 47,572 observed transaction prices are left. The mean of these observed transactions is 291,857 euro with a minimum observed transaction price of 100,000 euro and a maximum observed transaction price of 1,325,000 euro (table A.3, Appendix A). When plotting these transaction prices in a histogram, it is clearly visible that the figure is right-skewed (figure 5). To correct for this and to account for diminishing marginal returns, the logarithm of the transaction price is taken. After this transformation, the distribution of the dependent variable can be considered as normal (figure 6). The distribution of these transactions over the years can be seen in figure A.6 in Appendix A. In this figure, it is striking that the trend in the figure can be explained mostly by the economic crisis of 2008.

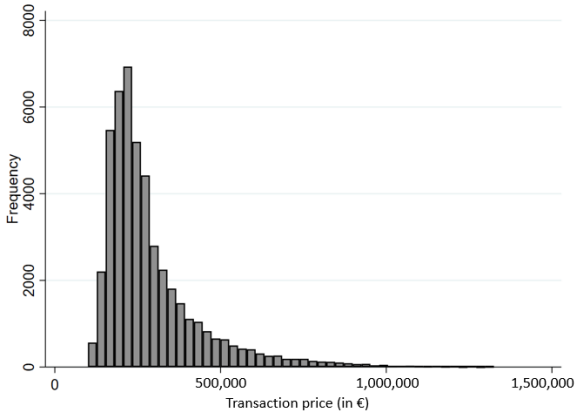

Figure 5: The distribution of the transaction price.

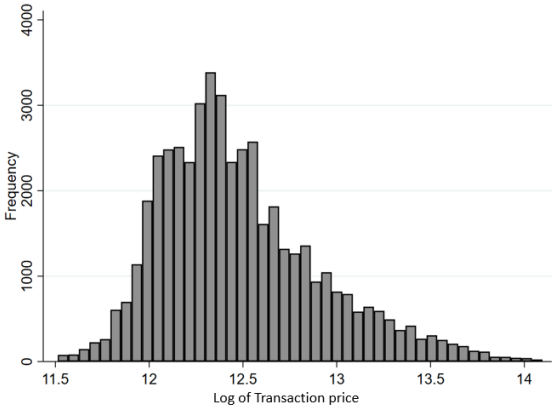

Figure 6: The distribution of the log of the transaction price.

For the fourth and fifth hypothesis, the dependent variable transaction price is divided into four categories of types of dwellings: terraced houses, corner houses, semi-detached houses and detached houses. Of the 22,538 unique dwellings in the dataset, the vast majority of the dwellings are categorized as terraced houses. With an absolute number of 12,590 dwellings, the terraced houses represent about 56 percent of the total number of dwellings in the dataset. After the terraced houses, the corner houses are the most common in the dataset with about 20 percent of the dwellings. This is followed by the semi-detached and detached houses, with 14 and 10 percent respectively (table A.7, Appendix A). The distribution of the four types of dwellings can be seen in figure 8. It is striking that

the terraced and corner houses are located mostly at the same locations, while the semi-detached and detached houses are more widely spread over the province of North-Holland. When looking at the spatial distribution of all houses, it is striking that the spatial distribution of the terraced and corner houses closely corresponds with the distribution shown in figure 7. Those concentrations are in areas where relatively larger cities are located. Furthermore, the extensive descriptive statistics of the different types of dwellings can be found in table A.3 in Appendix A.

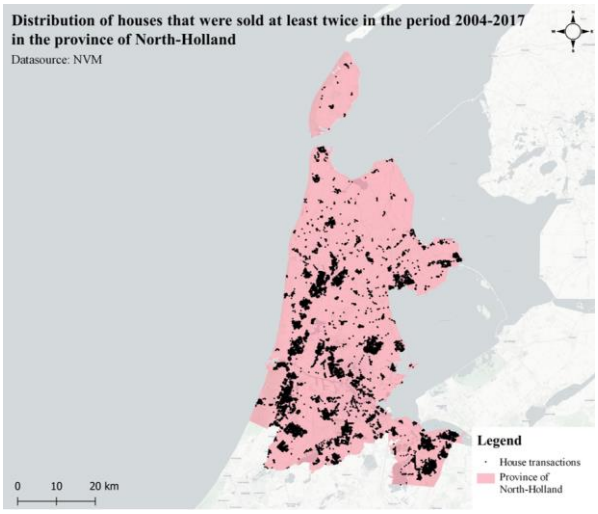

Figure 6: The distribution of the houses that were sold at least twice in the period 2004-2017 in the province of North-Holland.

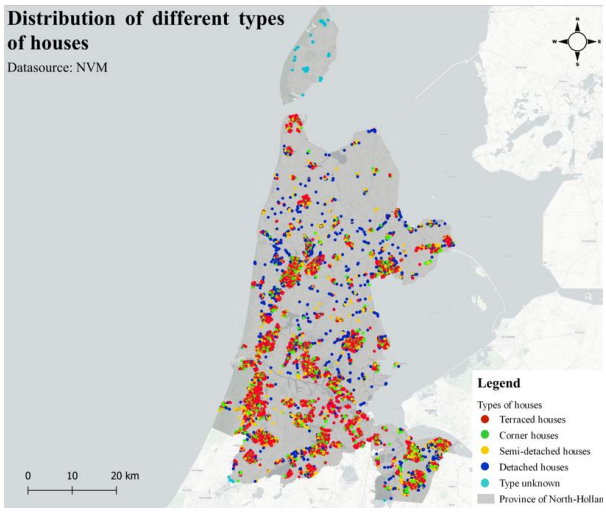

Figure 7: The distribution of the different types of houses that were sold at least twice in the period 2004-2017 in the province of North-Holland.

3.5.2 Total available jobs (per sector) and diversity as independent variable

The first independent variable in this research is the number of total available jobs. As described before, for every dwelling a unique access to the number of total available jobs has been determined. This has resulted in an average access to about 2.5 million jobs, with a minimum of 237 thousand and a maximum of 3.8 million reachable jobs (within 45 minutes of travel time). Second, the available jobs per sector are used as independent variables. A total of thirteen variables are used to determine the effect of thirteen different sectors of employment. The average of those variables vary from just 8 accessible jobs to more than 500 thousand accessible jobs per sector. After multiplying the jobs with the travel time decay function, the average total jobs available within the relevant range is equal to 450, with a minimum of zero jobs and a maximum of 8,432 jobs. Thereby, one unit of the variable total number of jobs available is equal to one distance-weighted job. So one unit is equal to a single job on 0 to 1 minute of travel time, but also equal to for example five jobs at 16 minutes of travel time, ten jobs at 24 minutes of travel time and twenty jobs at 31 minutes of travel time. The extensive descriptive statistics of the independent variables can be seen in table A.3 in Appendix A.

For both the variable of the total number of available jobs and the variables of the different sectors of employment, the level value is used an not the logarithmic value. This was chosen because, especially

for the different sectors of employment, a number of observations contain the value zero for these variables, which would mean that these observations would be lost by implementing the logarithm since the logarithm of zero is not defined.

As described, the third independent variable is the rate of diversity. The accessibility to a diverse labor market is relatively high for the province of North-Holland, since the average diversity is equal to 0.7398 (table A.3, Appendix A). With a minimum diversity of 0.0578 and a maximum diversity of 0.8939 it looks like there is a wide range. When taking a look at the scatterplot, it can be seen that the scatterplot of diversity is left-skewed, with many values between the 60 and 85 percent diversity (figure A.8, Appendix A). To correct for this distribution, the logarithm of the rate of diversity is taken. The distribution of diversity in employment in the province of North-Holland can be seen in figure 9.

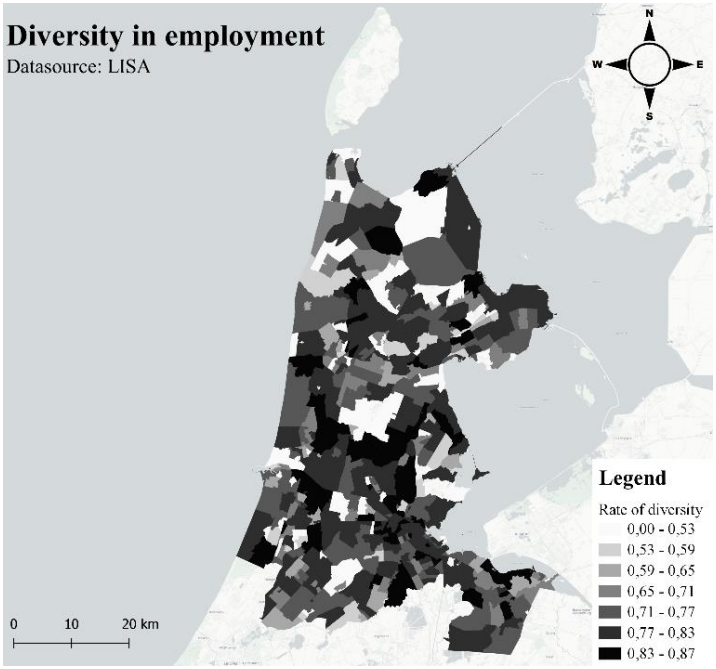

Figure 9: The distribution of diversity in employment (2017).

3.5.3 Control variables

The two control variables that are added are the state of maintenance for indoor and outdoor. These are added, even though all dwelling characteristics are assumed to be fixed, because it is likely that the state of maintenance may change in the period between two transactions. Since this is not corrected by the Repeat Sales Method itself, the indoor and outdoor state of maintenance are added as control variables. Both variables can take values from 1 to 9, where 1 stands for a poor and 9 for an excellent state of maintenance. The data of these variables is retrieved by the dataset of the NVM.

4. Methodology

4.1 Repeat Sales Method

As discussed before, in this research the Repeat Sales Method (RSM) is used. This method uses a model that is closely related to the Hedonic Price Method, with the main characteristic that the data contains only dwellings that have been sold at least twice within a specific timeframe (Case & Shiller, 1987). The biggest advantage of this method is that certain determinants are (assumed) to be fixed over time. As indicated in the HPM, a dwelling's value is determined by multiple determinants like among others the structural characteristics of the dwelling and the neighborhood characteristics (Malpezzi, 2003). It is important to take into account those characteristics, otherwise there may be an omitted variable bias (OVB). For example, an OVB will arise when there are no neighborhood characteristics included in the regression. It then seems like employment has a major effect on residential property prices, while a part of this effect is actually explained by the (not included) neighborhood characteristics (Stock & Watson, 2014). These characteristics are not always measurable and if they are, they are not always readily available. So accounting for price variations in dwellings with varying characteristics is necessary, but not easy in general. This problem is avoided by using the Repeat Sales Method, since this method assumes that structural and neighborhood characteristics of dwellings remain equal over time (Grimes, Sorensen & Young, 2021). Because the differences of dwellings are taken over time, and since the characteristics are assumed to remain the same under the RSM, these variables are dropped in the regression equation, making it therefore redundant to know the exact values of the characteristics (Haan & Diewert, 2013).

Nevertheless, there are also some drawbacks when using the Repeat Sales Method. An important disadvantage is that when the structural characteristics are not taken into account, renovations and transformations to the dwelling are ignored (Malpezzi, 2003). To account for this, two control variables for both the indoor and outdoor state of maintenance have been added. Another disadvantage is that only dwellings that are sold at least twice are taken into account. All dwellings that were only sold once are removed from the dataset, resulting in the use of only a fraction of the potential information on the housing market. A possible consequence of this is that the sample is not representative of the stock of housing since dwellings that are sold at least twice may be systematically different (Haan & Diewert, 2013). Other disadvantages are that there is no information on price levels taken into account or can be derived from the model and that underlying coefficients of characteristics can differ over time (they are assumed to be fixed over time). The disadvantage of the price level is controlled for by adding the price index of sold houses for the period 2004-2017 from the CBS to the model. Despite the above drawbacks, the advantages are believed to outweigh the disadvantages, which is why the Repeat Sales Method is used for this research.

4.2 Model

To understand and determine the model of the Repeat Sales Method, first the model of the Hedonic Price Method is reconsidered. The fundamental Hedonic regression consists of a regression equation of the house value against characteristics of the dwelling that determine the price (Malpezzi, 2003). The base HPM model in log-level form looks as follows:

$$ln(Y_{i,t}) = \alpha_t + \beta * X_{i,t} + \mu_{i,t} \quad (4)$$

Where *ln* represents the logarithm, *Y* is the transaction price of house *i* at time *t*, α is the constant term and μ is the error term, *X* represents a vector of structural housing characteristics and neighborhood characteristics and β represents the coefficient of *X*.

The Repeat Sales Method uses the same method design, only instead of using observations in time *t*, the method uses the difference of observations between time *t* and time *s* (where t > s). As discussed before, one important assumption within the Repeat Sales Method is that structural and neighborhood characteristics of a dwelling are constant over time. When translating this assumption to the regression equation it means that $\beta_{i,t} * X_{i,t}$ is equal to $\beta_{i,s} * X_{i,s}$. Therefore, the term representing the vector of structural housing characteristics and neighborhood characteristics is dropped. For the first hypothesis, the only characteristic that does change over time is the total number of available jobs within the relevant range, so this variable replaces the vector with all characteristics in it. In addition, some control variables are added, as discussed in the data section of this paper. This results in the following regression equation:

$$ln(Y_{i,t}) = \alpha_t + \beta * E_{i,t} + \gamma * C_{i,t} + \mu_{i,t} \quad (5)$$

Where *E* represents the total number of reachable jobs within the relevant range for dwelling *i* at time *t* and *s* and where *C* represents the vector of control variables for dwelling *i* at time *t*. In this regression equation, β and γ represent the coefficients of *E* and *C*.

For the second and fourth hypothesis, the same equation is used where only the variable for the total number of accessible jobs *E* is replaced by the vector *S*. The vector *S* represents the job accessibility variables for the thirteen sectors of employment. For the fourth hypothesis, the transaction price *Y* is taken for the four different types of dwellings. The regression equation using different sectors of employment look as follows:

$$ln(Y_{i,t}) = \alpha_t + \beta * S_{i,t} + \gamma * C_{i,t} + \mu_{i,t} \quad (6)$$

Last, a term for diversity in employment is introduced for the third and fifth hypothesis. Again, the same form of equations is used as in equations 2 and 3, but now with the logarithm of the rate of

diversity in employment as independent variables. This rate is represented by the term *D*. This regression equation looks as follows:

$$ln(Y_{i,t}) = \alpha_t + \beta * ln(D_{i,t}) + \gamma * C_{i,t} + \mu_{i,t} \quad (7)$$

4.2.1 Fixed Effects

Furthermore, fixed effects are added to the regression. Fixed effects are able to solve endogeneity problems by using a transformation to remove unobserved effects of variables that are constant over time across individuals (Woolridge, 2015). Fixed effects can be used to eliminate comparisons across groups of data within the dataset. In this research, year fixed effects are used to control for unobserved effects. By taking year fixed effects, the effects that can be attributed to differences between years are taken out. Examples of these year bounded effects are the annual price index (of residential property), changes of certain rules or a specific event in the given year.

To control for the Repeat Sales Method, house fixed effects are added to the regression. These house fixed effects take out the constant factor of the independent variables across individuals (the single dwellings in this case). Since the constant factors are taken out, only the things that vary over the different measured years are preserved. This leaves the differences between the measured years to analyze, this is exactly what is needed to perform the Repeat Sales Method. This method is chosen over for example the first-difference method, since house fixed effects also account properly for situations in which a dwelling is sold more than two times.

Last, fixed effects for both the indoor and outdoor state of maintenance are added to the regression. These variables are added as fixed effects, and not as control variables, because the variables are not linear but categorical. To correct for the state of maintenance, dummy's are created to correct for each of the nine categories.

4.2.2 Instrumental Variables

Last, some instrumental variables (IV) are added to the regressions. In general, instrumental variables are used to solve endogeneity problems like an omitted variable bias or a simultaneity bias. In doing so, the instruments correct for the causes of the changes in the independent variables and give an explanation for these changes (Wooldridge, 2015). So an instrumental variable has a direct effect on the independent variable, but on the other hand has no direct effect on the dependent variable nor has a direct effect on anything else that might explain the dependent variable. In other words, an instrumental variable explains exogenous variation in the independent variable to identify the impact of the independent variable on the dependent variable (Stock & Watson, 2014). So for this research,

instrumental variables can be used that do explain the variation in the number of jobs (per sector), but do not explain the variation in residential property prices.

A type of instrumental variables that are used in the field of labor economics is the so-called Bartik instrument (or shift-share instrument). Bartik instruments apply to situations where shocks at a higher aggregated level (for example national, continental or even global) hits variables on a lower aggregated level, for example industries, cities, municipalities or even neighbourhoods (Goldsmith-Pankham, Sorkin & Swift, 2020). All those lower aggregated level areas all react differently on that (same) shock due to the different characteristics of those areas. In this research, the annual national employment growth rate is used to explain the variation in the number of available jobs per region (per sector). The Bartik instrument for the total number of available jobs within the range is constructed as follows. First the data of the LISA-Bestand is used to determine the total available jobs in the Netherlands for every year in the period 2004-2017 (figure B.1, Appendix B). Thereafter, the annual number of total available jobs in the Netherlands is used to determine national employment growth rates for every year of the period 2014-2017 (figure B.2, Appendix B). Last, the annual national employment growth rates are used to forecast the regional number of jobs available for each year and each region in the dataset. So the Bartik instrument which applies to the total number of available jobs within the range is defined as the predicted number of jobs for time *t* and region *i*. The definition of the Bartik instrument is also represented in the following formula:

$$\hat{A}_{s,i,t} = A_{s,i,2004} * \left(\frac{(A_{s,NL,t} - A_{s,NL,2004})}{A_{s,NL,2004}}\right) \quad (8)$$

Where *A* represents the number of employment opportunities (jobs) in region *i* and time *t*. Next to the instrument for the total number of available jobs, a Bartik instrument is used for each sector. These instruments are created in the same way as the one for the total number of jobs available, only then separated by sector. That is where the *S* is for in equation 8. For each instrument of the thirteen sectors, the number of jobs in that sector has been used, and for the instrument of the total number of jobs, the values of all these sectors have been added. For the instrument of diversity, the same method is used as for the total number of jobs available, only now the employment opportunities including the decay factor are taken into account.

Before the instrumental variables can be used, it first has to be tested whether they are good instruments or not. In general, there are two requirements for a good instrument: an instrumental variable has to be both relevant and exogeneous. For the IV to be relevant, the thumb rule applies that the F-statistic of the first stage regression (the number of jobs as dependent variable and the instrument as independent variable) have to be higher or equal to 10 (Wooldridge, 2015). The results of the first

stage regressions show that for the instrument of the total number of jobs the F-statistic is significant and equal to 853.6, for the instruments for the different sectors the F-statistics are all significant and vary between 16.332 and about 48,852 and for diversity the F-statistic is significant and equal to 240.5 (tables B.3, B.6 and B.9, Appendix B). It can therefore be concluded that all Bartik instruments are relevant. The second requirement, exogeneity, means that the instrument cannot be correlated to other explanations. Since only one instrument is used for each independent variable, there is no test for overidentification to test for instrument exogeneity. Nevertheless, it is assumed that there is instrumental exogeneity since it is unlikely that the annual national employment growth rate (per sector) has a direct effect on house prices or on other variables in the error term. So, according to both the requirements of relevance and exogeneity, all Bartik instruments are therefore good instruments.

The last test performed is the Wu Hausman test. This test tests whether the instrumental variable is solving an endogeneity problem (or not). This is done by means of a null-hypothesis where the IV estimator is equal to the OLS estimator. The results of the Wu Hausman tests (table B.4, B.7 and B.10, Appendix B) shows that for all instruments, except the instrument for sector thirteen (other sectors), the null-hypothesis have to be rejected, since the values are significant at a five percent significance level. From these results can be concluded that the estimators of IV and OLS are not equal to each other, and thus these Bartik instruments are good and necessary variables. The exception is the instrument for the thirteenth sector (other sectors). This is probably due to the fact that there is a relatively big decrease in the national number of jobs in this sector (98 percent decrease in the period 2004-2017) while there are a lot of observations with zero jobs in this sector (for all years).

In addition, there is a check for a common pitfall: correlation among independent variables, instruments and controls. For the instrument of the total available jobs, the correlation with the independent variable is 0.1328 (table B.5, Appendix B). For the different sectors of employment, the correlation with the dependent variables varies from only 0.0185 to 0.7118 (table B.8, Appendix B). The correlation between the instrument of diversity and the independent variable is equal to 0.0710 (table B11, Appendix B). The correlation between the instrument and sector 5 (hospitality) of 0.7118 is relatively high, nevertheless this value does not fall into the category (very) strong correlations (values between 0.8 and 1.0). Hence it is assumed that there are no (very) strong associations between the Bartik instruments and the independent variables. The highest value of the correlations between the instruments and the control variables is that between the instrument of the total available jobs and the state of maintenance outdoor. Since this value is 0.0402 (table B.5, Appendix B), there is only a weak association between the instruments and the control variables. Therefore the common pitfall does not apply in this case.

5. Results

5.1 Base models

To test the first hypothesis about the effect of the number of accessible jobs on the residential property prices, four statistical models have been established. The four models in table 10 represent a simple OLS regression (1), a linear regression with house, year and state of maintenance fixed effects (2), a linear regression with the instrumental variable (3) and lastly a linear regression with both the fixed effects and the instrumental variable combined (4). When looking at table 10, it is striking that the coefficient for the number of total jobs available (including decay factor) is positive for all of the four models. In addition, the table shows that the coefficient for the number of total jobs becomes bigger as the Bartik instrument is being added and that there is a lack of significance when only fixed effects are added, but the significance returns when fixed effects are added along with the instrument. Since the coefficient of the fourth model is significant and the addition of fixed effects and the instrumental variable are important to counteract (possible) endogeneity problems, this model is chosen to test the first hypothesis.

Table 10: Regression results with the total available jobs as independent variable.

	Dependent variable:			
	Log price			
	(1)	(2)	(3)	(4)
	OLS	FE	IV	FE + IV
Total jobs available	0.00001***	0.00000	0.002***	0.0003***
	(0.00000)	(0.00000)	(0.0001)	(0.0001)
Constant	12.075***		11.793***	
	(0.054)		(0.026)	
Observations	47,572	47,572	47,564	47,564
Year Fixed Effects	No	Yes	No	Yes
House Fixed Effects	No	Yes	No	Yes
State of maintenance Fixed Effects	No	Yes	No	Yes
Instrumental Variable	Yes	No	Yes	Yes
R^2	0.120	0.551	-6.288	0.381
Adjusted R^2	0.120	0.145	-6.288	-0.178
Residual Std. Error	0.401		1.153	
F Statistic	216.946***	1,021.940***		20,029.210***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs include the decay factor.

Therefore, the effect of the total number of jobs available is estimated to be significant, positive and equal to 0.0003. This indicates that one extra job available leads to an increase of 0.03 percent in the residential property prices. Since the decay factor has been applied to this value, this one unit increase is equal to the addition of a single job on 0 to 1 minute of travel time, but also equal to the addition of for example five jobs at 16 minutes of travel time, ten jobs at 24 minutes travel time and twenty jobs at

31 minutes of travel time. The effect of the number of accessible jobs on the residential property prices is therefore positive and significant, as in line with the hypothesis. Therefore, the first hypothesis cannot be rejected.

The second hypothesis states that the effect of employment on the housing market varies significantly between employment sectors. To test this, the same models (1-4) are used as for the first hypothesis, only now the accessibility variable is split into thirteen different sectors of employment. What is striking about the results (table 11) is that when the fixed effects and the instrumental variables are added to the regression, a lack of significance arises under the coefficients of the different sectors of employment. In the fourth model, only four coefficients are significant where only the coefficients of the sectors S3 (construction) and S11 (personal services) are significant at a five percent significance level. Nevertheless, also for the second hypothesis this is model chosen, since it is assumed that the addition of fixed effects and the instrumental variable are sufficiently important to counteract (possible) endogeneity problems.

Looking at table 11, there are a number of things that are standing out. First, there is a roughly equal distribution between sectors with a positive and sectors with a negative effect. There are seven sectors that have a positive effect, significant or not, on residential property prices, namely: S2 – production and installation, S3 – construction, S5 – hospitality, S6 – information and telecommunication, S9 – government, education and healthcare, S12 – industry and S13 - other sectors. The six sectors that have a negative effect on residential property prices are: S1 – trade, transport and storage, S4 – agriculture, S7 – financial services; renting of real estate and moveable property, S8 – business services, S10 – culture, sports and recreation, S11 – personal services. With this, there are some signs of coefficients that were not expected for that sector, but this will be discussed in the discussion. Another striking thing are the relatively large coefficients of the agriculture and other sectors. This is probably due to the fact that for many observations there are no or only a small number of jobs available in these sectors, which led to probably distorted results.

To test whether the coefficients of the different sectors of employment actually effectively differ significantly from each other, as the hypothesis states, the Wald-test is performed. The null hypothesis of this Wald-test tests whether all coefficients of interest are simultaneously equal to zero (Woolridge, 2015). The results of the Wald-test show a p-value of 0.0000, therefore the null hypothesis should be rejected at a significance level of five percent. This implies that the coefficients of interest differ significantly from each other. In other words, the effect of employment on the residential property prices varies significantly between different sectors of employment. Therefore, also the second hypothesis cannot be rejected.

Table 11: Regression results with the different sectors of employment as independent variables.

	Dependent variable:			
	Log price			
	(1)	(2)	(3)	(4)
	OLS	FE	IV	FE + IV
S1: Trade, transport and storage	-0.0002***	-0.0001***	-0.001	-0.001
	(0.00003)	(0.00002)	(0.001)	(0.001)
S2: Production and installation	0.001***	-0.0003***	0.003	0.005
	(0.0001)	(0.0001)	(0.003)	(0.004)
S3: Construction	-0.001***	0.0001**	-0.001	0.002***
	(0.0001)	(0.00005)	(0.001)	(0.001)
S4: Agriculture	0.417***	0.023	6.671	-0.146
	(0.039)	(0.025)	(8.380)	(1.034)
S5: Hospitality	0.0001	0.0002	0.001	0.003*
	(0.0001)	(0.0001)	(0.001)	(0.002)
S6: Information and telecommunication	0.001***	0.00002	0.002***	0.0004
	(0.00005)	(0.00002)	(0.0004)	(0.0004)
S7: Financial services; renting of real estate and moveable property	-0.00000	0.00003	-0.0004**	-0.0003
	(0.00003)	(0.0001)	(0.0002)	(0.001)
S8: Business services	0.001***	0.0001***	0.0002	-0.001*
	(0.00003)	(0.00002)	(0.001)	(0.0005)
S9: Government, education and healthcare	-0.0001***	-0.00003***	-0.0001	0.0003
	(0.00001)	(0.00001)	(0.0002)	(0.0002)
S10: Culture, sports and recreation	0.001***	0.001***	0.001	-0.007
	(0.0002)	(0.0002)	(0.001)	(0.006)
S11: Personal services	-0.001***	-0.0005***	0.004*	-0.007**
	(0.0002)	(0.0001)	(0.002)	(0.003)
S12: Industry	-0.0005***	-0.00003	-0.001***	0.0002
	(0.00004)	(0.00002)	(0.0001)	(0.0004)
S13: Other sectors	-0.066***	-0.029*	-0.099**	0.699
	(0.014)	(0.015)	(0.050)	(1.549)
Constant	12.124***		12.102***	
	(0.052)		(0.056)	
Observations	47,572	47,572	47,564	47,564
Year Fixed Effects	No	Yes	No	Yes
House Fixed Effects	No	Yes	No	Yes
State of maintenance Fixed Effects	No	Yes	No	Yes
Instrumental Variable	Yes	No	Yes	Yes
R^2	0.161	0.555	-0.211	0.237
Adjusted R^2	0.161	0.152	-0.212	-0.453
Residual Std. Error	0.391		0.470	
F Statistic	217.869***	740.687***		12,354.810***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs per sector includes the decay factor.

Last, four models have been established to test the third hypothesis about the effect of diversity in the sectorial mix in employment on the residential property prices. What stands out in these models (table 12) is that the coefficients of the dependent variable are positive for all of the four models and are significant at a significance level of five percent for three of the four models. This in combination with the benefits of fixed effects and the instrumental variable regarding (possible) endogeneity problems, also for this hypothesis the last model is chosen to test whether the hypothesis should be rejected or not. The coefficient of diversity is equal to 1.211, which implies that an increase in diversity of one percent increases the residential property prices with no less than 121 percent rounded. This value is relatively large, but is explainable since the distribution of diversity is extremely left skewed (figure A.8, Appendix A). Because of this distribution and due to the fact the differences in diversity for the individual observations is very small (average difference of 0.0001049), it is unlikely that an increase of one percent in diversity is realistic for an individual observation. This does not affect the fact that the effect of diversity on the residential property prices is both significant and positive, which has led to the fact that the third hypothesis cannot be rejected either.

Table 12: Regression results with diversity as independent variable.

	Dependent variable:			
	Log price			
	(1)	(2)	(3)	(4)
	OLS	FE	IV	FE + IV
Log diversity	0.059***	0.007	0.811***	1.211***
	(0.011)	(0.006)	(0.158)	(0.140)
Constant	12.500***		12.737***	
	(0.004)		(0.050)	
Observations	47,472	47,472	47,467	47,467
Year Fixed Effects	No	Yes	No	Yes
House Fixed Effects	No	Yes	No	Yes
State of maintenance Fixed Effects	No	Yes	No	Yes
Instrumental Variable	Yes	No	Yes	Yes
R^2	0.001	0.552	-0.114	0.252
Adjusted R^2	0.001	0.147	-0.114	-0.426
Residual Std. Error	0.427		0.450	
F Statistic	33.878***	1,024.389***		12,402.590***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

In the first and third hypothesis, the effects of the total number of available jobs and diversity on the residential property prices are estimated separately. In addition to these hypotheses, these variables of interest are also tested conditionally on each other. The results of this regression are shown in the third column of table 13. The two things that are standing out are that both the coefficient of the total number of available jobs and the coefficient of log diversity are about 7 times larger than the coefficient from models 1 and 2 from table 13 (where the variables are still separated), and that both

coefficients of the variables decrease in significance and are therefore no longer significant at a significance level of five percent. Therefore it seems that these effects strengthen each other, but since the F-statistic is relatively low (compared to models 1 and 2) and the coefficients are not significant at a significance level of five percent, this joint model is not preferred over the models where the effects of employment accessibility and diversity are tested separately. However, a Wald-test has been performed which shows that the effect of employment accessibility differs significantly from the effect of diversity on the residential property prices (p = 0.01). This shows, together with a correlation of 0.0637 between the two variables, that indeed two different effects are measured by the two variables.

Table 13: Regression results with total available jobs and diversity combined as independent variables.

	Dependent variable:		
	Log price		
	(1)	(2)	(3)
Total available jobs	0.0003***		0.002*
	(0.0001)		(0.001)
Log diversity		1.211***	8.814*
		(0.140)	(7.005)
Observations	47,564	47,467	47,457
Year Fixed Effects	Yes	Yes	Yes
House Fixed Effects	Yes	Yes	Yes
State of maintenance Fixed Effects	Yes	Yes	Yes
Instrumental Variables	Yes	Yes	Yes
R^2	0.381	0.252	0.008
Adjusted R^2	-0.178	-0.426	-0.889
F Statistic	20,029.210***	12,402.590***	326.260***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs per sector includes the decay factor.

5.2 Different types of dwellings

After testing the hypothesis based on the whole dataset, also a distinction is made between the different effects on different types of dwellings. The effects of the total number of available jobs, different sectors of employment and diversity are split for terraced houses, corner houses, semi-detached houses and detached houses. When looking at table 14, it can be seen that the effect of the total number of available jobs is equal to a 0.03 percent increase when one extra unit is added (including the decay factor). For the different types of dwelling, this effect varies between a 2.0 percent increase and a 1.3 percent decrease when one extra unit is added. Although it looks like there are some differences in effects between the different types of dwellings, all the four coefficients are not significant at a five percent significance level. Therefore, it cannot be stated that there is a

significant difference in the effect of the number of total available jobs on the different types of dwellings.

Table 14: Regression results with the different house types as dependent variables and total available jobs as independent variable.

	Dependent variable:				
	Log price				
	All houses	Terraced houses	Corner houses	Semi-detached houses	Detached houses
Total jobs available	0.0003***	0.006	-0.013	0.020	0.001*
	(0.0001)	(0.005)	(0.044)	(0.068)	(0.001)
Observations	47,564	26,608	9,897	6,568	4,491
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
House Fixed Effects	Yes	Yes	Yes	Yes	Yes
State of maintenance Fixed Effects	Yes	Yes	Yes	Yes	Yes
Instrumental Variable	Yes	Yes	Yes	Yes	Yes
R^2	0.381	0.003	0.002	0.00005	0.114
Adjusted R^2	-0.178	-0.938	-1.022	-1.083	-0.809
F Statistic	20,029.210***	50.957***	4.368	2.396	526.487***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs include the decay factor.

Besides the effect of the total number of available jobs, also the effects of the different sectors of employment is tested. Therefore, the fourth hypothesis states that the effects of different employment sectors vary significant for different types of dwellings. What immediately stands out when looking at table 15 is that there is almost a complete lack of significance among the coefficients for the different types of dwellings. Only 2 of the 51 coefficients are significant at a five percent significance level. This is probably because not all coefficients were significant when the whole dataset was used (for the second hypothesis) in combination with the lower number of observations that arose when categorizing the dwellings. Despite the shifts in signs, the sectors of employment with positive effects are about evenly distributed with the sectors of employment with negative effects for every type of dwelling. Last, results show that coefficients of multiple sectors of employment switching from sign for the different types of dwellings. For example, the sector government, education and healthcare only has a negative coefficient for the corner houses where the other coefficients are positive, while the sector trade, transport and storage has only a negative coefficient for the semi-detached houses where the other coefficients are positive. So, it seems that the effects of the different sectors do differ for different types of houses (in both sign and magnitude), but since there is a lack of significance, the fourth hypothesis has to be rejected.

Table 15: Regression results with the different house types as dependent variables and the different sectors of employment as independent variables.

	Dependent variable:				
	Log price				
	All houses	Terraced houses	Corner houses	Semi-detached houses	Detached houses
S1: Trade, transport and storage	-0.001	0.0001	0.003	-0.008**	0.004
	(0.001)	(0.006)	(0.006)	(0.004)	(0.007)
S2: Production and installation	0.005	0.001	-0.023	0.039**	-0.008
	(0.004)	(0.013)	(0.040)	(0.019)	(0.015)
S3: Construction	0.002***	0.002	-0.003	0.001	0.004
	(0.001)	(0.004)	(0.009)	(0.003)	(0.006)
S4: Agriculture	-0.146	4.137	0.187	35.420	20.681
	(1.034)	(19.495)	(0.226)	(80.103)	(29.187)
S5: Hospitality	0.003*	0.002	-0.0002	-0.012	-0.007
	(0.002)	(0.007)	(0.011)	(0.019)	(0.019)
S6: Information and telecommunication	0.0004	-0.001	0.001	0.001	-0.0004
	(0.0004)	(0.007)	(0.002)	(0.002)	(0.003)
S7: Financial services; renting of real estate and moveable property	-0.0003	-0.002	0.0001	-0.004	-0.019
	(0.001)	(0.004)	(0.003)	(0.009)	(0.026)
S8: Business services	-0.001*	-0.0003	-0.0004	-0.002	-0.003
	(0.0005)	(0.002)	(0.001)	(0.005)	(0.005)
S9: Government, education and healthcare	0.0003	0.0003	-0.001	0.004	0.002
	(0.0002)	(0.0004)	(0.002)	(0.003)	(0.002)
S10: Culture, sports and recreation	-0.007	0.001	0.001	-0.007	0.012
	(0.006)	(0.030)	(0.009)	(0.025)	(0.027)
S11: Personal services	-0.007**	-0.007	-0.014	-0.028	-0.010
	(0.003)	(0.014)	(0.016)	(0.027)	(0.033)
S12: Industry	0.0002	-0.0002	0.001	0.001	-0.0004
	(0.0004)	(0.003)	(0.002)	(0.002)	(0.002)
S13: Other sectors	0.699	-1.916	-0.097		-1.429
	(1.549)	(10.638)	(0.657)		(2.448)
Observations	47,564	26,614	9,895	6,564	4,491
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
House Fixed Effects	Yes	Yes	Yes	Yes	Yes
State of maintenance Fixed Effects	Yes	Yes	Yes	Yes	Yes
Instrumental Variables	Yes	Yes	Yes	Yes	Yes
R^2	0.237	0.122	0.104	0.014	0.025
Adjusted R^2	-0.453	-0.709	-0.819	-1.061	-1.000
F Statistic	12,354.810***	3,131.042***	836.214***	82.759***	79.318***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs per sector includes the decay factor. For the semi-detached houses there is no effect of sector 13, this is due to the fact that there are no observations with more than zero jobs for this type of dwellings.

In both table 11 and table 15 it was noticed that there is a lack of significance in the results of the regression with the thirteen different sectors of employment on the residential property prices. This is perhaps due to the fact that many sectors of employment are tested simultaneously in the regression. The correlations between the variables for the sectors varies from -0.0021 to 0.8357, indicating that one or more variables for the different sectors of employment are (strongly) correlated with each other. Because the effects are interpreted conditionally on the effects of the other variables, it is possible that this poses some problems with regard to the reliability of the effects. Because of this, an additional regression is performed where the thirteen sectors of employment are combined into four different groups of employment sectors. Which sectors belong to which groups can be deduced from the group names in table 16, but also can be found in table C.1 in Appendix C.

The results for this regression, both for all houses and the different types of dwellings, are shown in table 16. It immediately stands out that within the regression with all observations, in contrast to table 10, there are predominantly significant coefficients. Here, only the coefficient for the group personal services and leisure sectors is not significant at a significance level of five percent. Furthermore, it is striking that only the group G3 – business, financial and information sectors show a positive effect on the residential property prices, while the coefficients of the groups G1 – manufacturing and agriculture, G2 – personal services and leisure sectors (not significant) and G4 – trade, transport, government and other sectors are all negative. The result of the Wald-test shows that the coefficients are significantly different from each other ($p = 0.0000$). This result thus confirms that the second hypothesis, which states that the effect of employment on the housing market significantly varies between employment sectors, cannot be rejected.

In addition, these four groups of employment sectors were also used to test the difference in effects on the four types of dwellings. Looking at the results of these regressions, it is striking that the significance obtained in the regression with all observations, is partly lost again. For the terraced houses, three of the four coefficients are still significant at a significance level of five percent. For the corner houses there are only 2 coefficients significant and for the semi-detached and detached houses there are no significant coefficients. A possible explanation for this decrease in significance is that the number of observations per regression decreases when further is moved to the right in table 16. In addition, it is striking that for the groups G1 – manufacturing and agriculture and G4 – trade, transport, government and other sectors, the coefficients are negative for all the four types of dwellings. For the other two groups, the signs of the coefficients vary for the different types of dwellings. Just like the results of table 15, for these results apply that it seems like that the effects of the different sectors do differ for different types of houses (in both sign and magnitude), but since there is a lack of significance, the results confirm the rejection of the fourth hypothesis.

Table 16: Regression results with the different house types as dependent variables and the four groups of different sectors of employment as independent variables.

	Dependent variable:				
	Log price				
	All houses	Terraced houses	Corner houses	Semi-detached houses	Detached houses
G1: Manufacturing and agriculture	-0.001***	-0.001***	-0.001	-0.0003	-0.003
	(0.0002)	(0.0004)	(0.001)	(0.0004)	(0.004)
G2: Personal services and leisure sectors	-0.002	-0.0001	0.004	-0.010	0.019
	(0.001)	(0.002)	(0.003)	(0.007)	(0.022)
G3: Business, financial and information sectors	0.001***	0.002***	0.0003**	-0.0002	0.004
	(0.0001)	(0.0004)	(0.0001)	(0.0003)	(0.005)
G4: Trade, transport, government and other sectors	-0.001***	-0.001***	-0.002**	-0.0005	-0.008
	(0.0002)	(0.0003)	(0.001)	(0.001)	(0.010)
Observations	47,564	26,614	9,895	6,564	4,491
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
House Fixed Effects	Yes	Yes	Yes	Yes	Yes
State of maintenance Fixed Effects	Yes	Yes	Yes	Yes	Yes
Instrumental Variable	Yes	Yes	Yes	Yes	Yes
R^2	0.102	0.055	0.118	0.071	0.009
Adjusted R^2	-0.709	-0.836	-0.787	-0.937	-1.026
F Statistic	5,683.164***	1,869.798***	1,037.717***	452.790***	65.277***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01
Number of available jobs include the decay factor. The exact distribution of the thirteen sectors over the four groups can be seen in table C.4 in Appendix C.

Last, the fifth hypothesis about the effect of diversity on different types of dwellings is tested. Table 17 shows that all four different types of dwelling experience a positive effect from diversity in the sectorial mix in employment, giving that three of the four coefficients are significant at a five percent significance level. The magnitude of the effects on the residential property prices varies between 44 and 162 percent with an increase of one percent in diversity. However, as before, it is unlikely that an increase of one percent in reality for an individual observation applies. Relatively speaking, the biggest effect of diversity can be found at terraced houses, which implies that the owners of these houses have a higher willingness to pay for access to a diverse range of the labor market. All in all, the results show that the effect of diversity in the sectorial mix in employment varies significantly for different types of dwellings, therefore also the last hypothesis is not rejected.

Table 17: Regression results with the different house types as dependent variables and diversity as independent variable.

	Dependent variable:				
	Log price				
	All houses	Terraced houses	Corner houses	Semi-detached houses	Detached houses
Log diversity	1.211***	1.617***	1.041***	0.438	0.888**
	(0.140)	(0.251)	(0.237)	(0.287)	(0.368)
Observations	47,467	26,580	9,879	6,552	4,456
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
House Fixed Effects	Yes	Yes	Yes	Yes	Yes
State of maintenance Fixed Effects	Yes	Yes	Yes	Yes	Yes
Instrumental Variable	Yes	Yes	Yes	Yes	Yes
R^2	0.300	0.155	0.278	0.371	0.221
Adjusted R^2	-0.334	-0.644	-0.463	-0.311	-0.592
F Statistic	15,014.330***	4,259.445***	2,696.287***	2,255.772***	943.047***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

6. Synthesis

In this chapter, the results of the literature review and the empirical research of this research are compared. The first result examined is the effect of the total accessibility to the labor market on the residential property prices. The results of the empirical research of this paper state that there is a positive and significant effect of the total number of accessible jobs on the housing market. This indicates that when more jobs (distance-weighted) can be achieved, house prices will rise. This result is in line with those of the literature review, since the literature review also shows that there is predominantly a positive and significant effect of labor market accessibility on the residential property prices. The most logical explanation for the positive effect of general labor market accessibility on residential property prices is that more employment opportunities make a neighborhood more attractive for households to live in, so that a higher demand for houses in this neighborhood leads to higher residential property prices.

Then, the effect of the total accessibility to the labor market is tested also for different types of houses. The literature does not distinguish between different house types in the same way this research does, but between price levels of dwellings (low, medium and high). However, this can be compared with the results of this study, as it is assumed that the house types used in this study (terraced houses, corner houses, semi-detached houses and detached houses) also have an average selling price from low to high. The literature shows positive effects on median and lower-rent level submarkets and a

negative effect on higher-rent-level submarkets (but insignificant), with the note that these are only from the paper from Lin and Cheng (2016). The results of the empirical research of this study shows that there is only a negative effect of the total number of jobs available on the residential property prices of corner houses, on all the other types of dwellings there is a positive effect. Nevertheless, due to the lack of significance in the results of the empirical research (no significant coefficients), it cannot be stated with certainty that the effect of the total number of jobs differs per type of house, which also makes it difficult to compare with the literature study.

Furthermore, research is done to the effect of different sectors of employment on the residential property prices. The main finding from the empirical part of this study for this topic is that the effect of employment on the housing market does significantly vary between employment sectors. This finding is in line with the conclusions of the literature review. So, the main conclusions of the literature review and the empirical research are similar, yet some differences have been found. For example, there are a number of sectors that have different signs than the conclusions from the literature suggest. For example, the results show a negative effect of the sector business services for both all dwellings and for all four different types of homes, while the literature review shows that there is a predominantly positive effect of business services on the residential property prices. The same applies to the financial services sector, where a negative effect was found for all dwellings as well as three out of four different types of dwellings. The other way around applies for the industry sector. Where the literature draws a clear conclusion that (heavy) industry has a negative effect, this study found a positive effect for both all dwellings and two of the four different types of houses. However, no conclusions can be drawn from these results, since there is a lack of significance in both the situation with all observations and the situation where the distinction is made between 4 types of dwellings.

Because of the lack of significance, these two analyses were also performed in which the sectors were merged into four groups of sectors. Thereby, three of the four coefficients are significant, which means that the results can be compared better with the results from the literature review. Contrary to the first analysis, the group of sectors manufacturing and agriculture (the group that includes the industry sector) now has a negative and significant effect on the residential property prices. In other words, the willingness to pay of households will decrease when more jobs are created in these sectors. This is in line with the conclusions of the literature review and there is a reasonable explanation for this: as more jobs in the manufacturing and agriculture sectors are created, it probably means that there will be more or larger industrial areas in or nearby the residential area. Since these are known for its negative externalities, people do not want to live near them and therefore the willingness to pay for dwellings in these residential areas becomes lower. Another result in line with the conclusions of the literature review is the positive and significant effect of the group of business, financial and information sectors.

This positive effect can be explained by the fact that the jobs in these sectors are likely to be paid relatively well, which allows the employees of these sectors to buy relatively more expensive dwellings. The last significant coefficient is that of the group trade, transport, government and other sectors. This group of sectors have a negative effect on the residential property prices. This result is somewhat contradictory to the findings of the literature, as the conclusions from the literature review state that the education and healthcare sectors both have a positive effect on the residential property prices. In addition, governmental institutions and the wholesale (warehousing) sector have varying positive and negative effects, so in theory a negative effect was not expected. On the other hand, the negative effect of this group of sectors can (partly) be explained. For example, the trade, transport and storage sectors are part of this group. These sectors involve many (large) distribution centers and means of transportation such as trucks. These factors are generally less valued in a residential area. Education is also part of this group of sectors. Initially, this sector will be valued positively, since the children of households can go to a school nearby. But this sector also has to deal with negative externalities, such as nuisance from loitering high school students. Therefore, this negative effect is in contrast to the literature, but can be explained. The last coefficient indicates a negative effect of personal services and leisure sectors on the residential property prices, but since this coefficient is not significant, there is no evidence that this effect is significantly different from zero.

Merging the sectors into four groups has also been applied to the different types of houses. The literature review shows that all sectors except business services have significant effects on the different types of dwellings. For among the sectors manufacturing and wholesale, the sign of the coefficients is the same for all four types of dwellings (both all negative), where the signs of the coefficients of the other sectors differ per type of dwelling. The empirical research of this study also shows that the sectors manufacturing and agriculture and trade, transport, government and other sectors have the same (negative) sign for all types of dwellings. For the other two groups, the sign changes for the different types of houses. However, no real conclusions can be drawn from this, since the majority of the coefficients from the empirical research are not significant.

Finally, research is done to the effect of the sectorial mix in employment on the residential property prices. The literature has shown that there is a predominantly positive and significant effect of diversity in employment on both housing and land values. The results of the empirical research of this study show that this conclusion can be confirmed and that there is indeed a positive and significant effect of diversity on the residential property prices. A possible explanation for the fact that households are willing to pay for a higher rate of diversity is that due to the effective overlap between some sectors of employment, homeowners can switch to another sector of employment that is around when they possibly lose their job when there is a relatively higher rate of diversity. In addition, the sectorial mix in employment is also important for people who, for example, work in a seasonal sector.

Last, also for this variable of interest the distinction has been made between different types of dwellings. The conclusions drawn from the literature review about the effect of diversity on different types of dwellings is mainly based on the paper by Koster and Rouwendal (2012). They found that there is "substantial heterogeneity in willingness to pay for diversity", the paper indicates that only apartment occupiers are willing to pay for an increase in diversity, whereas the effects of diversity on the other types of dwellings is negative. The conclusion that there is a substantial difference in willingness to pay for diversity is completely in line with the conclusion of this study, namely that the effect of diversity in the sectorial mix in employment varies significantly for different types of dwellings. The conclusion of which types of dwellings are positive affected and which types are negative affected is somewhat contradictory to the conclusion of this empirical study, but since only one of the four coefficients of the study by Koster and Rouwendal is significant, no conclusions can be drawn based on these results. For the empirical research of this study, all coefficients are positive and three out of four of them are significant. For all types of houses, the effect of diversity is clearly the biggest for terraced houses and corner houses, while the effects are lowest for semi-detached and detached houses. A possible explanation for this is as follows. Semi-detached and detached houses generally have a higher sales price than terraced and corner houses. To afford these 'more expensive' dwellings, the homeowners generally have higher paying jobs that are mainly concentrated in certain sectors and relatively lower paying jobs are therefore less interesting for these homeowners, which means that some sectors (spoken about employment opportunities) are not interesting for these people. This subsequently results in a lower willingness to pay for diversity and therefore a lower valuation of the sectorial mix in employment.

7. Conclusion

In this research, the following question is chosen as starting point:

What is the effect of the sectorial mix in employment on the residential property prices?

To answer this question, first answer is given to the four sub-questions. The first question was if job accessibility does have an effect on the housing market. Since the results show a positive and significant effect from the total number of available jobs on the residential property prices and therefore the first hypothesis could not be rejected, it can be concluded that accessibility indeed does have a significant effect on the housing market and that this effect is positive. The second question was whether there are any differences in the effect on the housing market of the accessibility to different sectors of employment. Both the analysis with the thirteen different sectors and the analysis with the four groups of combined sectors confirm that the effects of the accessibility to different sectors of employment are indeed statistically significant from each other. Then, the third question was if the employment sectors have different effects on the prices of different types of dwellings. The results show that it seems like that the effects of the different sectors do differ for different types of houses (in both sign and magnitude). Yet it cannot be said with certainty that these differences exist, since there is a lack of significance in the analysis. Last, the fourth question was if more diversity in employment opportunities lead to higher residential property prices. The answer to this question is yes. The results of the third hypothesis show that there is a significant and positive effect of the rate of diversity on the residential property prices. In addition, the last hypothesis shows that even if the dwellings are divided by house type, the effect of diversity remains positive for all types of houses. Therefore, it can be concluded that for both the total housing stock and for the four different types of houses, more diversity in employment opportunities leads to higher residential property prices.

Coming back to the main question about the effect of the sectorial mix in employment on the residential property prices, the answer is as follows: there is a positive effect of the sectorial mix in employment on the residential property prices. In other words, when there is an increase in the sectoral mix in employment (more diversity in employment sectors), an increase in residential property prices is expected as well. This conclusion is grounded on the facts that there is generally a positive effect of labor market accessibility and that the effect on the residential property prices differ significantly for different sectors of employment.

8. Reflection

In this section, the research is discussed. First a number of limitations of this study are mentioned, then a number of policy implications and suggestions for further research are given.

8.1 Limitations

The first limitation of this paper is that no clustered standard errors were used in the empirical research. With spatial data (in this research the dwellings) it is likely that the variable of interest correlates (very) for observations in the same environment. In other words, in this study it is plausible that dwellings in the same street or neighborhood can reach the same number of jobs within 45 minutes of travelling and have approximately the same rate of diversity. Due to this possible correlation in the variable of interest, it is possible that a bias arises. This is certainly not desirable, so it is better to control for this possible correlation due to similarities in a certain environment.

The similarities between houses in the same environment immediately lead to the next limitation for this study, namely the aggregation level of the employment data from the LISA-bestand. The data in the LISA file consists of data about the number of jobs per zip code (PC4-level). This makes the zip code-level the smallest level that was used in the employment data. This means that, for example, all jobs in a zip code area that is less than 45 minutes travelling are included in the study, while there may be jobs on the edges of the zip code area that are reachable in about 46 minutes of travel time or more. Because these jobs are within an 'accessible zip code' no correction has been made for this. This also applies the other way around, for jobs that can be reached within 45 minutes of travelling but are located within a zip code that is not accessible within 45 minutes. This applies not only to the destination (the jobs) but also to the origin (the dwellings). The travel time is calculated between two zip codes, therefore the differences between houses in the same zip code disappear. For example, two houses within the same zip code may be three minutes away from each other. In reality, this means that both houses have different access to the labor market, because a different number of jobs are likely to be accessible within those three minutes. Because they are in the same zip code, it has been assumed, in contrast to the reality, that they have exactly the same access to the labor market within 45 minutes of travelling. In order to make this research more in line with reality, it would be better to do the analysis at a lower aggregation level.

The last limitation of this study has already been mentioned a little above, but is the occurrence of possible endogeneity problems. To counter these possible endogeneity problems, fixed effects and instrumental variables have been added to the regressions of the empirical research of this study. Normally, a Bartik instrument has a coefficient of about one with the variable it predicts, but in this study the coefficients are slightly lower. So despite the fact that all instruments are being labeled as

relevant and exogenous and that the IV regression is significantly different from an OLS regression, there is room for improvement regarding the Bartik instruments. In addition to the national employment growth rates, there are other factors that could explain the number of jobs (per sector) and the sectorial mix in employment, such as concentrated pools of employment opportunities and the quality of the infrastructure (partly taken into account by the use of the roadmap) in certain areas. So, to improve the empirical research concerns possible endogeneity problems, new, additional or improved instrumental variables can be introduced.

8.2 Implications

Since the results from the empirical research of this study and the results from the literature research largely correspond, policy plans can be made and adjusted on these conclusions. In any case, it is important from a policy point of view to consider overall access to the labor market in addition to the presence or construction of amenities in the case of potential new-build locations. In addition, it is important to look at the location of new construction projects in relation to existing employment locations, but also at the location of new employment locations in relation to residential areas. For example, the results show that it is better to keep industrial areas and residential aeras separated. Finally, one of the most important policy recommendations is to maintain and create enough diversity in sectors of employment around residential areas. Looking at the sectorial mix in employment of the province of North Holland (figure 9) and taking into account the nature reserves and the spatial construction of regions, concrete examples for new construction projects would be the areas at the South and West of Purmerend (the areas around Wormer and Wijdewormer). In addition, the results have shown that the effect of the rate of diversity is the biggest for terraced houses. In response to this, the area between Limmen, Molendijk and Akersloot (on the west side of the province) would be a good location for new building projects. In this area there are already many neighborhoods with terraced houses, so new neighborhoods with this type of dwellings would not be out of place. In addition to the positive sectorial mix in employment, there are also the usual factors (such as amenities) that generate a higher willingness to pay for dwellings. Based on the results, the policy advice is to avoid areas around the Noordzeekanaal and areas around the North of IJmuiden. Besides the low diversity in employment, also many industrial areas are located there, two factors that have a negative impact on residential property prices. This is confirmed by case studies, with perhaps the best example being the commotion around the steel manufacturer Tata Steel.

From a scientific point of view, this research mainly contributes to the confirmation and extension of already existing theories. The expansion of the existing literature is mainly achieved when considering the effect of the sectorial mix in employment on all houses as well as on the different types of houses, since the existing literature on this is sparse and often contains insignificant results.

Since this topic is far from being closed after this research and there is still some room for improvement, a number of suggestions for further research have been drawn up. Some of these suggestions arise from the limitations of this paper. First, clustered standard errors can be used instead of robust standard errors to correct for spatial correlation. Second, to get more observation specific data, the aggregation level can be lowered from PC4-level to PC6-level or even single observation level (micro-level). Last, to take into account possible endogeneity problems, new, additional or improved instrumental variables should be introduced.

In addition to the suggestions due to the limitations of this research, there are a number of alternative suggestions for further research. For example, it is interesting to repeat this study for other regions to see whether the results for these areas correspond or differ from the results for the province of Nord of Holland. The province of North-Holland is a fairly closed area because it is surrounded on three sides by seas and lakes. It is therefore interesting to look at similar areas such as the provinces of Friesland and Zeeland, but also areas that have access to larger parts of the country within a certain travel time, such as the province of Utrecht. Besides this, it is also possible to look at the implementation of the various variables of interest. For example, the Herfindahl-Hirschman index is used for the variable diversity in this study, but use could also be made of, for example, the Krugman Specialization Index or location quotients measures. In addition, a decay function was used to give spatial weight to the total number of accessible jobs. This could also involve looking at other or multiple methods to take spatial weights into account. The final suggestion for further research is to ask follow-up questions appropriate to this study. For example, the effect of a combination of the labor market and other markets on residential property prices can be examined.

Bibliography

Agarwal, A., Giuliano, G., & Redfearn, C. L. (2012). Strangers in our midst: the usefulness of exploring polycentricity. *The Annals of Regional Science*, *48*(2), 433-450.

Ahlfeldt, G. (2011). If Alonso was right: modeling accessibility and explaining the residential land gradient. *Journal of Regional Science*, *51*(2), 318-338.

Ahlfeldt, G. M., & Wendland, N. (2016). The spatial decay in commuting probabilities: Employment potential vs. commuting gravity. Economics Letters, 143, 125-129.

Brueckner, J. K. (2011). Lectures on urban economics. London: MIT press.

Cao, T. V., & Cory, D. C. (1981). Mixed land uses and residential property values in the Tucson Metropolitan Region: implications for public policy.

Case, K. E., & Mayer, C. J. (1996). Housing price dynamics within a metropolitan area. *Regional Science and Urban Economics*, *26*(3-4), 387-407.

Case, K. E., & Shiller, R. J. (1987). Prices of single family homes since 1970: New indexes for four cities.

CBS. (2021a). *Ruim 69 duizend nieuwbouwwoningen in 2020*. Retrieved from: https://www.cbs.nl/nl-nl/nieuws/2021/04/ruim-69-duizend-nieuwbouwwoningen-in-2020

CBS. (2021b). *Standaard Bedrijfsindeling 2008; versie 2018 update 2021; Structuur: tweede digit en vijfde digit.* Retrieved from: https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/activiteiten/sbi-2008-standaard-bedrijfsindeling-2008/de-structuur-van-de-sbi-2008-versie-2019-update-2021

CBS. (n.d.a) *Kerncijfers per postcode.* Retrieved from: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/gegevens-per-postcode

CBS. (n.d.b). *Werkgelegenheidsstructuur*. Retrieved from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/banen-werkgelegenheid/toelichtingen/werkgelegenheidsstructuur

Cervero, R., & Duncan, M. (2004). Neighbourhood composition and residential land prices: does exclusion raise or lower values?. *Urban Studies*, *41*(2), 299-315.

Faggio, G., Silva, O., & Strange, W. C. (2017). Heterogeneous agglomeration. *Review of Economics and Statistics*, *99*(1), 80-94.

Friedman, A. (2021). Mixing Land Uses and Dwelling Types. In *Fundamentals of Sustainable Urban Design* (pp. 85-93). Springer, Cham.

Giuliano, G., Gordon, P., Pan, Q., & Park, J. (2010). Accessibility and residential land values: some tests with new measures. *Urban studies*, *47*(14), 3103-3130.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. American Economic Review, 110(8), 2586-2624.

Grimes, A., Sorensen, K., & Young, C. (2021). Repeat sales house price indices: comparative properties under alternative data generation processes. New Zealand Economic Papers, 55(1), 7-18.

Groenemeijer, L., Gopal, K. & van Leeuwen, G. (2020). *Vooruitzichten bevolking, huishoudens en woningmarkt*. Retrieved from: https://www.abfresearch.nl/publicaties/vooruitzichten-bevolking-huishoudens-en-woningmarkt-2020-2035/

Haan, J. de. & Diewert, W.E. (2013). Repeat Sales Methods. In OECD et al. (red.), *Handbook on residential property price indices* (p. 65-72). Luxembourg: Eurostat. DOI: https://doi.org/10.1787/9789264197183-8-en

Helsley, R. W., & Strange, W. C. (2014). Coagglomeration, clusters, and the scale and composition of cities. *Journal of Political Economy*, *122*(5), 1064-1093.

Hu, H., Geertman, S., & Hooimeijer, P. (2014). Amenity value in post-industrial Chinese cities: The case of Nanjing. *Urban Geography*, *35*(3), 420-439.

Iacono, M., & Levinson, D. (2017). Accessibility dynamics and location premia: Do land values follow accessibility changes?. *Urban studies*, *54*(2), 364-381.

Kennisinstituut voor Mobiliteitsbeleid. (2018). *Kriebels tussen woning en werk: wat verklaart het verschil in reistijd?* Retrieved from: https://www.kimnet.nl/over-het-kim/publicaties/papers/2018/11/22/kriebels-tussen-woning-en-werk-wat-verklaart-het-verschil-in-reistijd

Kim, D., & Jin, J. (2019). The effect of land use on housing price and rent: Empirical evidence of job accessibility and mixed land use. *Sustainability*, *11*(3), 938.

Koster, H. R., & Rouwendal, J. (2012). The impact of mixed land use on residential property values. *Journal of Regional Science*, *52*(5), 733-761.

Li, S., Chen, L., & Zhao, P. (2019). The impact of metro services on housing prices: a case study from Beijing. *Transportation*, *46*(4), 1291-1317.

Li, Y., Xiong, W., & Wang, X. (2019). Does polycentric and compact development alleviate urban traffic congestion? A case study of 98 Chinese cities. *Cities*, *88*, 100-111.

Lin, J. J., & Cheng, Y. C. (2016). Access to jobs and apartment rents. *Journal of Transport Geography*, *55*, 121-128.

LISA. (n.d.). *Welkom bij LISA.* Retrieved from: https://www.lisa.nl/algemeen/wat-is-lisa#:~:text=LISA%20staat%20voor%20Landelijk%20Informatiesysteem,waar%20betaald%20werk%20wordt%20verricht.&text=LISA%2Dgegevens%20zijn%20eenvoudig%20te,beleidsmakers%20een%20waardevol%20instrument%20is.

Malpezzi, S. (2003). Hedoninc Pricing Models: a Selective and Applied Review. In A. O'Sullivan & K. Gibb (red.), *Housing economics and public policy* (pp. 67-89). Oxford: Blackwell Science.

McCann, P. (2013). *Modern urban and regional economics*. Oxford University Press.

NVM. (n.d.) *Property market in the Netherlands.* Retrieved from: https://www.nvm.nl/expat/property-market-netherlands/

Osland, L., & Thorsen, I. (2008). Effects on housing prices of urban attraction and labor-market accessibility. *Environment and Planning A*, *40*(10), 2490-2509.

Osland, L., & Thorsen, I. (2013). Spatial impacts, local labour market characteristics and housing prices. *Urban Studies*, *50*(10), 2063-2083.

Rhoades, S. A. (1993). The herfindahl-hirschman index. *Fed. Res. Bull.*, *79*, 188.

Rijkswaterstaat. (n.d.). *De maximumsnelheid in Nederland.* Retrieved from: https://www.rijkswaterstaat.nl/wegen/wetten-regels-en-vergunningen/verkeerswetten/maximumsnelheid#:~:text=Op%20onze%20snelwegen%20geldt%20tussen,van%20100%20km%2Fh%20in.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), 34-55.

SD Workx. (2018). *Nederlandse werknemers besteden de meeste tijd aan woon-werkverkeer in Europa*. Retrieved from: https://www.sdworx.nl/nl-nl/press/2018/20180920-mobiliteit

Shiller, R. J. (2007). *Understanding recent trends in house prices and home ownership* (No. w13553). National Bureau of Economic Research.

Song, Y., & Knaap, G. J. (2004). Measuring the effects of mixed land uses on housing values. *Regional Science and Urban Economics*, *34*(6), 663-680.

Srour, I. M., Kockelman, K. M., & Dunn, T. P. (2002). Accessibility indices: Connection to residential land prices and location choices. *Transportation Research Record*, *1805*(1), 25-34.

Stock, J.H. & Watson, M.W. (2014). *Introduction to Econometrics, Global Edition.* London: Pearson Education Limited.

Susilo, Y. O., & Dijst, M. (2009). How far is too far? travel time ratios for activity participation in the Netherlands. *Transportation Research Record*, *2134*(1), 89-98.

Susilo, Y. O., & Dijst, M. (2010). Behavioural decisions of travel-time ratios for work, maintenance and leisure activities in the Netherlands. *Transportation planning and technology*, *33*(1), 19-34.

Visser, P., Van Dam, F., & Hooimeijer, P. (2008). Residential environment and spatial variation in house prices in the Netherlands. *Tijdschrift voor economische en sociale geografie*, *99*(3), 348-360.

Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. South-Western: Cengage learning.

Wu, J., Song, Y., Liang, J., Wang, Q., & Lin, J. (2018). Impact of mixed land use on housing values in high-density areas: Evidence from Beijing. *Journal of Urban Planning and Development*, *144*(1), 05017019.

Appendix

Appendix A – Data

Table A.1: Duplicates report.

Number of transactions	Observations	Surplus
1	132,026	0
2	46,348	23,174
3	7,722	5,148
4	828	621
5	105	84
6	6	5
8	8	7

Source: NVM

Table A.2: Classification of the main sectors of employment.

Sector	Description	Includes the following SBI sections:
S1	Trade, transport and storage	G – Wholesale and retail (except G45) H – Transport and storage
S2	Production and installation, except industry (S12)	C33 – Repair and installation of machines and equipment D – Production, distribution of and trade in electricity, natural gas, steam and cooled air E – Extraction and distribution of water; waste and waste water management and remediation G45 – Car repairs
S3	Construction	F – construction industry
S4	Agriculture	A – Agriculture, forestry and fisheries
S5	Hospitality	I – Accommodation, meal- and drink supply
S6	Information and telecommunication	J – Information and communication
S7	Financial services; renting of real estate and moveable property	K – Financial institutes L – Renting and trade of real estate N77 – Renting of moveable property
S8	Business services	M – Consultancy, research and other specialized business services N – Other business services (except N77)
S9	Government, education and healthcare	O – Public administration, governmental services and compulsory social insurances P – Education Q – Healthcare and welfare systems
S10	Culture, sports and recreation	R – Culture, sports and recreation
S11	Personal services	S – Other services T – Households as employers; not differentiated production of goods and services by households for own consumption
S12	Industry	B – Extractive industries C – Industry (except C33)
S13	Other sectors	U – Extraterritorial organizations and bodies

Source: CBS, 2021a.

Table A.3: Descriptive Statistics.

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum	Source
Transaction price	47,572	291,857	154,673.1	100,000	1,325,000	NVM
Terraced houses	26,616	254,174	108,898	100,000	1,325,000	NVM
Corner houses	9,897	272,236	127,206	100,000	1,300,000	NVM
Semi-detached houses	6,568	372,310	189,762	100,000	1,325,000	NVM
Detached houses	4,491	440,767	232,228	100,000	1,325,000	NVM
Number of jobs						
Total Jobs Available*	47,572	2,464.247	753.234	236.978	3,843.382	LISA
S1: Trade, transport and storage**	47,572	53.7879	15.6365	5.1232	82.0742	LISA
S2: Production and installation**	47,572	5.9577	1.9964	0.6502	9.6534	LISA
S3: Construction**	47,572	12.9510	4.1640	1.9778	21.3793	LISA
S4: Agriculture**	47,572	0.0008	0.0009	0.0000	0.0025	LISA
S5: Hospitality**	47,572	11.8675	3.3068	1.1051	19.5116	LISA
S6: Information and telecommunication**	47,572	12.6958	4.2985	0.4260	20.0448	LISA
S7: Financial services; renting of real estate and moveable property**	47,572	14.3371	3.9177	1.0463	21.2870	LISA
S8: Business services**	47,572	38.9816	12.1094	2.1419	63.8170	LISA
S9: Government, education and healthcare**	47,572	64.6041	21.4130	7.1738	101.6474	LISA
S10: Culture, sports and recreation**	47,572	6.9658	2.1197	0.5877	11.5858	LISA
S11: Personal services**	47,572	5.7649	1.8937	0.5717	9.2319	LISA
S12: Industry**	47,572	18.2829	5.6592	2.3695	29.9227	LISA
S13: Other sectors**	47,572	0.2276	0.2341	0.0000	0.5488	LISA
Number of jobs corrected for decay function						
Total Jobs Available	47,572	450.309	703.087	0	8,432	LISA
S1: Trade, transport and storage	47,572	98.148	156.157	0	2,643	LISA
S2: Production and installation	47,572	9.751	22.706	0	736	LISA
S3: Construction	47,572	26.086	43.219	0	636	LISA
S4: Agriculture	47,572	0.002	0.039	0	2	LISA
S5: Hospitality	47,572	18.853	48.382	0	1,776	LISA
S6: Information and telecommunication	47,572	18.704	62.006	0	2,372	LISA
S7: Financial services; renting of real estate and moveable property	47,572	23.481	85.434	0	3,055	LISA

S8: Business services	47,572	69.208	136.599	0	3,043	LISA
S9: Government, education and healthcare	47,572	132.963	259.755	0	3,554	LISA
S10: Culture, sports and recreation	47,572	12.639	24.940	0	721	LISA
S11: Personal services	47,572	10.846	16.838	0	346	LISA
S12: Industry	47,572	29.620	69.283	0	2,626	LISA
S13: Other sectors	47,572	0.010	0.299	0	14	LISA
Diversity	47,540	0.7398	0.1164	0.0578	0.8939	LISA
State of maintenance indoor	47,572	7.0587	1.0685	1	9	NVM
State of maintenance outdoor	47,572	7.0597	0.9425	1	9	NVM

* x 100,000 jobs
** x 10,000 jobs

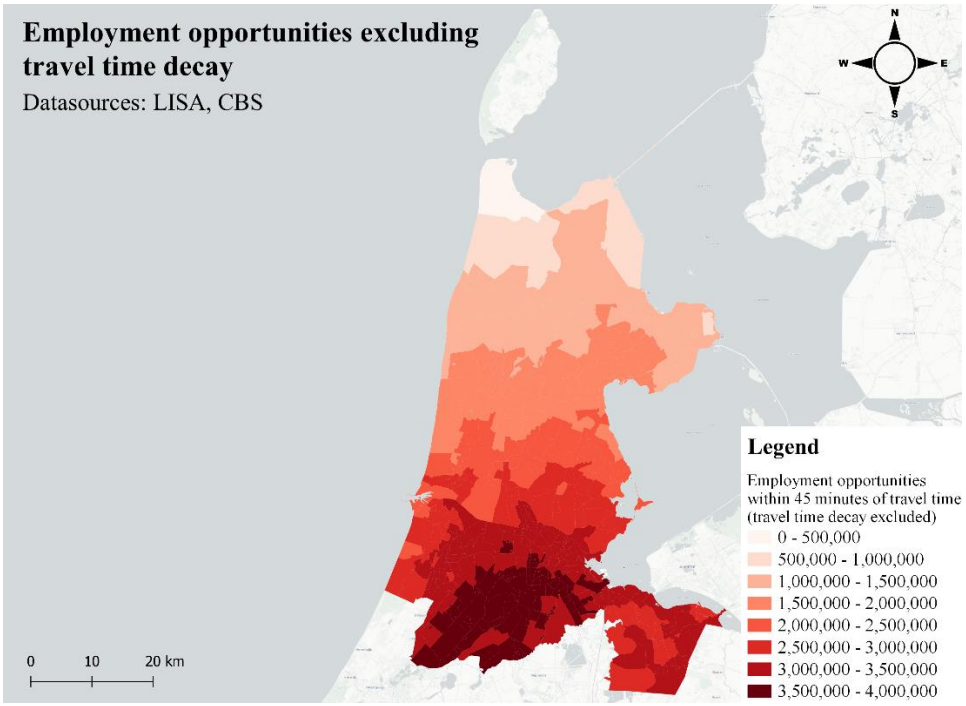

Figure A.4: The number of employment opportunities excluding the travel time decay factor (2017).

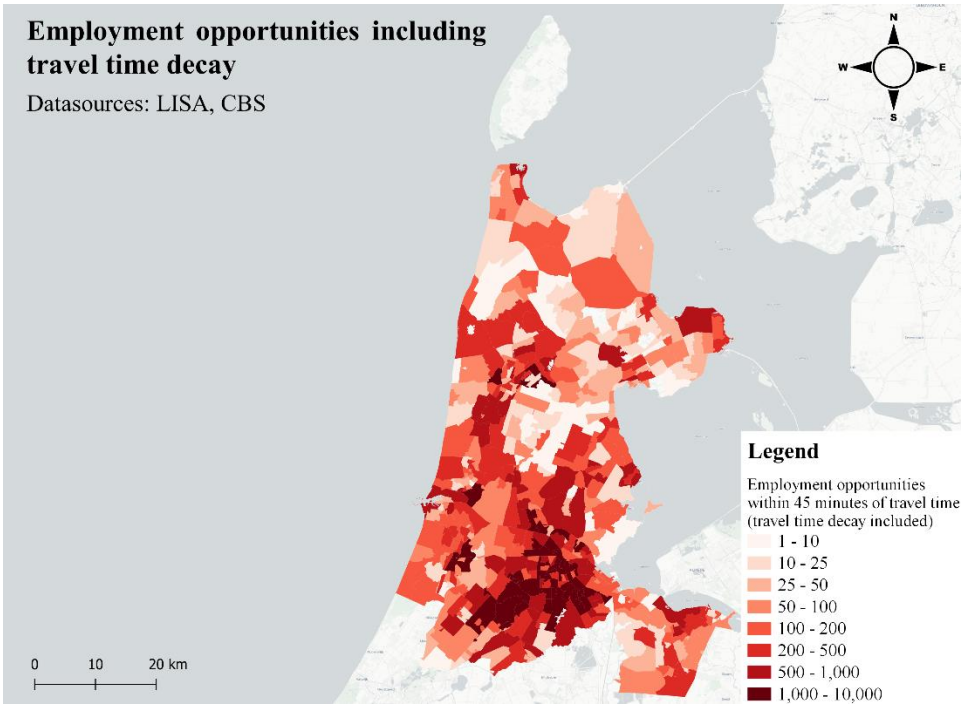

Figure A.5: The number of employment opportunities including the travel time decay factor (2017).

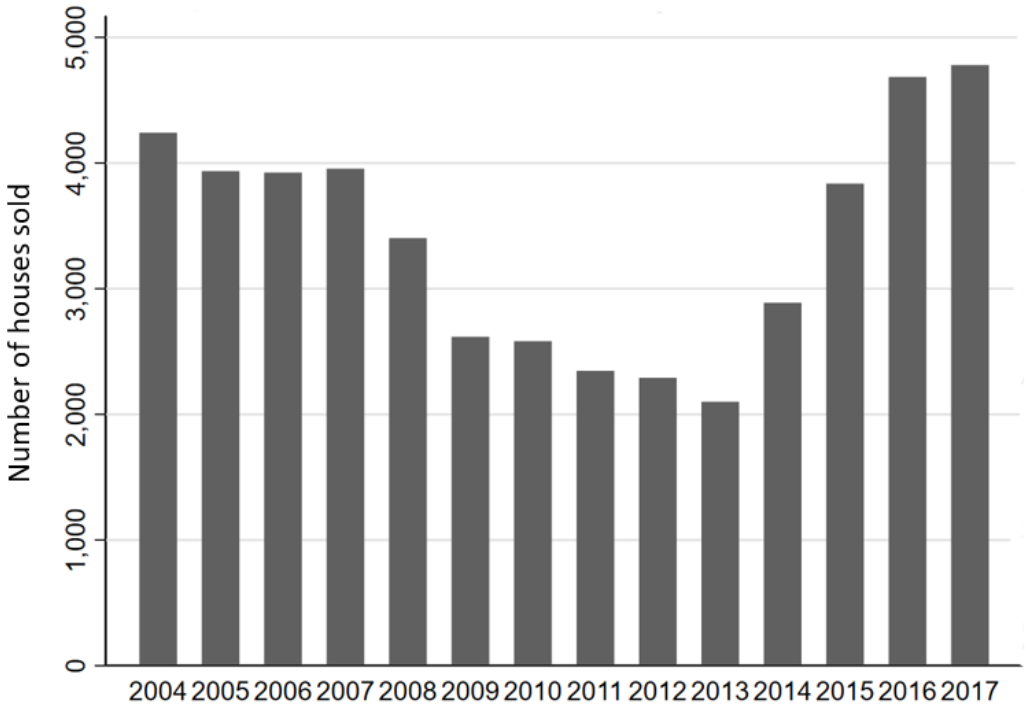

Figure A.6: Distribution of houses sold over the period 2004-2017.

Source: NVM

Table A.7: Distribution of housing types.

Housing type	**Frequency**	**Percentage**
Terraced houses	12,590	55.86
Corner houses	4,642	20.60
Semi-detached	3,138	13.92
Detached	2,168	9.62
Total	22,538	100.00

Source: NVM

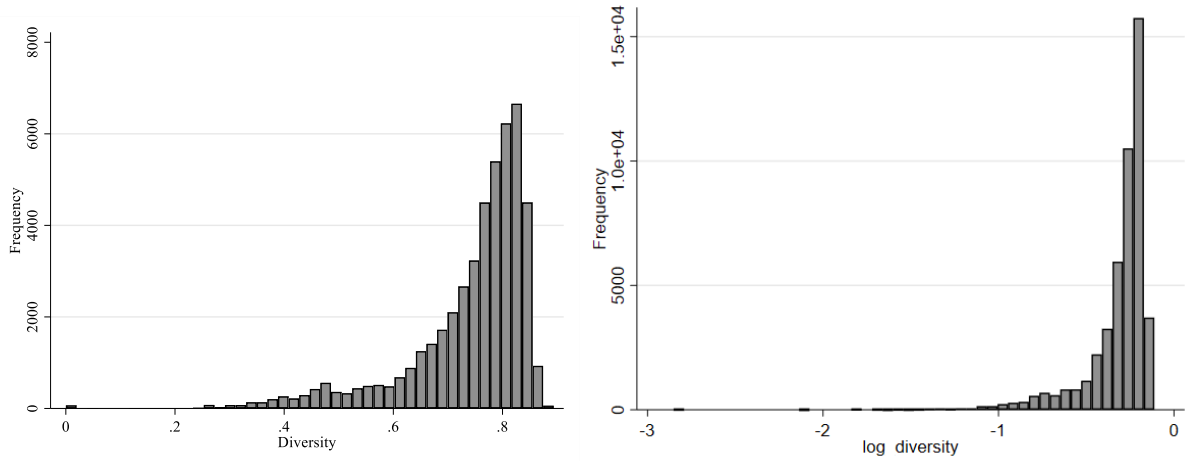

Figure A.8: Distribution of diversity and the logarithm of diversity.

Source: LISA

Appendix B – Instrumental Variables

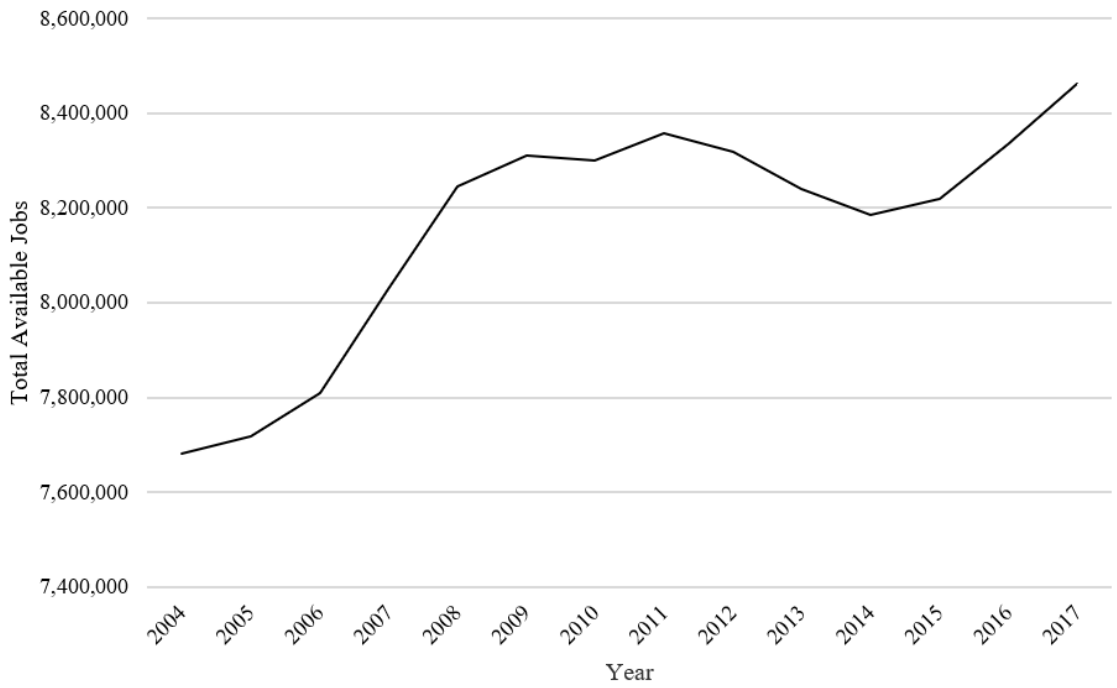

Figure B.1: The trend of the total available jobs in the Netherlands for the period 2004-2017.

Source: NVM.

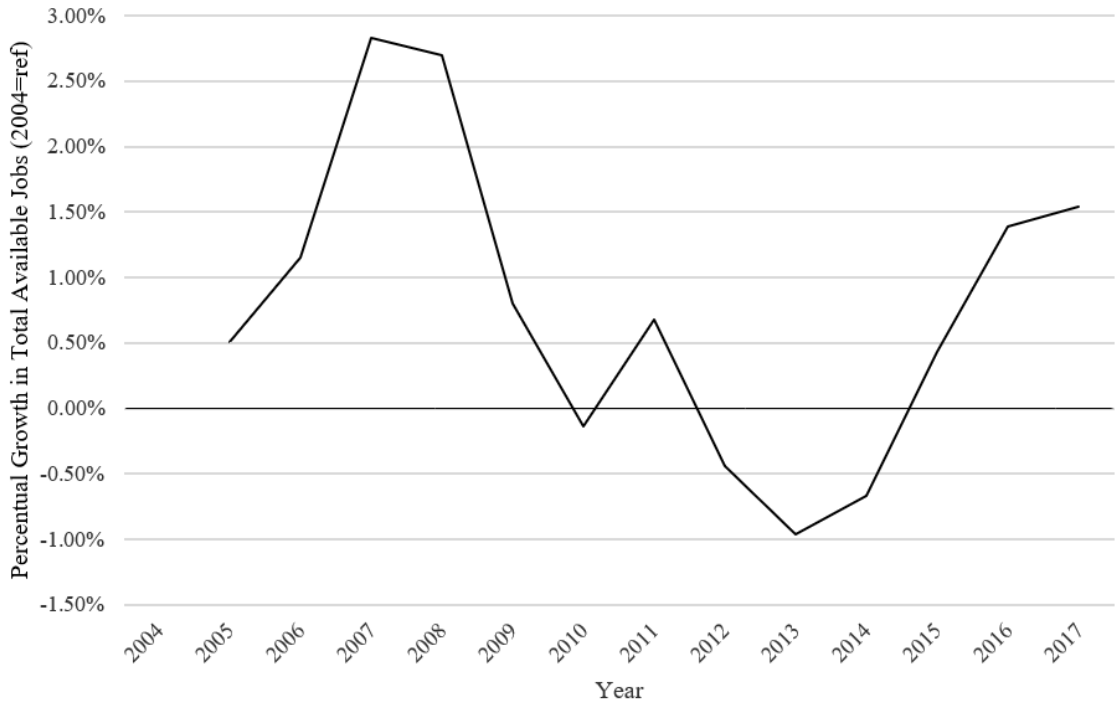

Figure B.2: The trend of the percentual growth in total available jobs in the Netherlands for the period 2004-2017.

Source: NVM.

Table B.3: The result of the first stage regression for testing the instrumental variable for the total available jobs.

	Dependent variable:
	Total available jobs
Predicted number of jobs	0.0001***
	(0.00000)
Constant	145.992***
	(10.895)
Observations	47,572
R^2	0.018
Adjusted R^2	0.018
Residual Std. Error	696.869
First stage F Statistic	853.629***
Note: Robust standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table B.4: The result of the Wu-Hausman test for the instrumental variable for the total available jobs.

	Dependent variable:
	Log price
Total available jobs	0.002***
	(0.0001)
Constant	11.793***
	(0.026)
Observations	47,572
R^2	-6.288
Adjusted R^2	-6.288
Residual Std. Error	1.153
Diagnostic Test:	
Wu Hausman	0.000***
Note: Robust standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table B.5: Correlation table for instrumental, independent and control variables for the total available jobs.

	Predicted number of jobs	State of maintenance indoor	State of maintenance outdoor	Number of jobs available
Predicted number of jobs	1.000			
State of maintenance indoor	0.0293	1.000		
State of maintenance outdoor	0.0402	0.8144	1.000	
Number of jobs available	0.1328	-0.0219	-0.0182	1.000

Table B.6: The results of the first stage regressions for testing the instrumental variables for the different sectors of employment.

	Dependent variable:					
	S1	S2	S3	S4	S5	S6
BI S1: Trade, transport and storage	0.099***					
	(0.001)					
BI S2: Production and installation		0.109***				
		(0.001)				
BI S3: Construction			0.107***			
			(0.001)			
BI S4: Agriculture				0.013***		
				(0.003)		
BI S5: Hospitality					0.110***	
					(0.0005)	
BI S6: Information and telecommunication						0.057***
						(0.0004)
Constant	18.333***	0.903***	4.002***	0.001***	2.398***	9.072***
	(0.736)	(0.102)	(0.196)	(0.0002)	(0.173)	(0.255)
Observations	47,564	47,564	47,564	47,564	47,564	47,564
R^2	0.376	0.340	0.407	0.0003	0.507	0.262
Adjusted R^2	0.376	0.340	0.407	0.0003	0.507	0.262
Residual Std. Error	123.363	18.447	33.285	0.039	33.984	53.264
F Statistic	28,657.590***	24,511.680***	32,637.960***	16.322***	48,852.350***	16,903.490***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table B.6: … continuation

	Dependent variable:						
	S7	S8	S9	S10	S11	S12	S13
BI S7: Financial services; renting of real estate and moveable property	0.171***						
	(0.001)						
BI S8: Business services		0.115***					
		(0.001)					
BI S9: Government, education and healthcare			0.107***				
			(0.001)				
BI S10: Culture, sports and recreation				0.086***			
				(0.001)			
BI S11: Personal services					0.099***		
					(0.001)		
BI S12: Industry						0.070***	
						(0.0004)	
BI S13: Other sectors							0.220***
							(0.001)
Constant	-4.408***	10.048***	26.304***	3.930***	2.153***	9.016***	0.006***
	(0.321)	(0.580)	-1.260	(0.111)	(0.095)	(0.258)	(0.001)
Observations	47,564	47,564	47,564	47,564	47,564	47,564	47,564
R^2	0.455	0.411	0.291	0.315	0.255	0.448	0.370
Adjusted R^2	0.455	0.411	0.291	0.315	0.255	0.448	0.370
Residual Std. Error	63.071	104.876	218.710	20.645	14.530	51.456	0.237
F Statistic	39,722.270***	33,137.890***	19,536.950***	21,860.610***	16,312.510***	38,678.170***	27,883.600***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table B.7: The results of the Wu-Hausman tests for the instrumental variable for the different sectors of employment.

	Dependent variable:					
	Log price					
	(1)	(2)	(3)	(4)	(5)	(6)
S1: Trade, transport and storage	0.0002***					
	(0.00002)					
S2: Production and installation		-0.001***				
		(0.0001)				
S3: Construction			-0.002***			
			(0.0001)			
S4: Agriculture				7.248**		
				-3.153		
S5: Hospitality					0.001***	
					(0.0001)	
S6: Information and telecommunication						0.002***
						(0.0001)
Constant	12.463***	12.488***	12.523***	12.471***	12.457***	12.445***
	(0.003)	(0.002)	(0.003)	(0.005)	(0.002)	(0.002)
Observations	47,564	47,564	47,564	47,564	47,564	47,564
R^2	-0.006	0.0002	0.006	-0.382	-0.001	-0.029
Adjusted R^2	-0.006	0.0002	0.006	-0.382	-0.001	-0.030
Residual Std. Error	0.428	0.427	0.426	0.502	0.427	0.433
Diagnostic Test:						
Wu Hausman	0.000***	0.000***	0.000***	0.007***	0.000***	0.000***

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table B.7: … continuation

	Dependent variable:						
	Log price						
	(7)	(8)	(9)	(10)	(11)	(12)	(13)
S7: Financial services; renting of real estate and moveable property	0.001***						
	(0.00003)						
S8: Business services		0.001***					
		(0.00002)					
S9: Government, education and healthcare			0.0001***				
			(0.00001)				

S10: Culture, sports and recreation				0.005***			
				(0.0001)			
S11: Personal services					0.006***		
					(0.0002)		
S12: Industry						-0.001***	
						(0.00004)	
S13: Other sectors							0.070***
							(0.011)
Constant	12.466***	12.440***	12.464***	12.422***	12.412***	12.506***	12.481***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Observations	47,564	47,564	47,564	47,564	47,564	47,564	47,564
R^2	-0.005	-0.014	-0.008	-0.038	-0.058	0.005	0.002
Adjusted R^2	-0.005	-0.014	-0.008	-0.038	-0.058	0.005	0.002
Residual Std. Error	0.428	0.430	0.429	0.435	0.439	0.426	0.427
Diagnostic Test:							
Wu Hausman	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.981

Note: Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table B.11: Correlation table for instrumental, independent and control variables for different sectors of employment.

	BI S1	BI S2	BI S3	BI S4	BI S5	BI S6
S1: Trade, transport and storage	0.6132					
S2: Production and installation		0.5832				
S3: Construction			0.06379			
S4: Agriculture				0.0185		
S5: Hospitality					0.7118	
S6: Information and telecommunication						0.5121
State of maintenance inside	-0.0169	-0.0044	-0.0244	-0.0051	-0.0059	-0.0079
State of maintenance outside	-0.0184	-0.0053	-0.0270	0.0002	-0.0082	-0.0074

	BI S7	BI S8	BI S9	BI S10	BI S11	BI S12	BI S13
S7: Financial services; renting of real estate and moveable property	0.6746						
S8: Business services		0.6408					
S9: Government, education and healthcare			0.5396				

S10: Culture, sports and recreation				0.5612			
S11: Personal services					0.5054		
S12: Industry						0.6697	
S13: Other sectors							0.6079
State of maintenance inside	-0.0234	-0.0253	-0.0226	-0.0058	-0.0019	-0.0261	-0.0017
State of maintenance outside	-0.0178	-0.0222	-0.0216	-0.0073	-0.0036	-0.0274	-0.0026

Table B.9: The result of the first stage regression for testing the instrumental variable for diversity.

	Dependent variable:
	Log diversity
Predicted employment opportunites	0.012***
	(0.001)
Constant	-0.384***
	(0.005)
Observations	47,457
R^2	0.005
Adjusted R^2	0.005
Residual Std. Error	0.192
First stage F Statistic	240.463***
Note: Robust standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table B.10: The result of the Wu-Hausman test for the instrumental variable for diversity.

	Dependent variable:
	Log price
Log diversity	0.811***
	(0.151)
Constant	12.737***
	(0.048)
Observations	47,457
R^2	-0.114
Adjusted R^2	-0.114
Residual Std. Error	0.450
Diagnostic test:	
Wu Hausman	0.000***
Note: Robust standard errors in parentheses	*p<0.1; **p<0.05; ***p<0.01

Table B.11: Correlation table for instrumental, independent and control variables for diversity.

	Diversity	Predicted employment opportunities	State of maintenance inside	State of maintenance outside
Diversity	1.000			
Predicted employment opportunities	0.0710	1.000		
State of maintenance inside	-0.0019	-0.0359	1.000	
State of maintenance outside	0.0008	-0.0346	0.8144	1.000

Appendix C – Results

Table C.1: Classification of the main groups of employment sectors.

Group of sectors	Description	Sector	Description
G1	Manufacturing and agriculture	S2	Production and installation, except industry (S12)
		S3	Construction
		S4	Agriculture
		S12	Industry
G2	Personal services and leisure sectors	S5	Hospitality
		S10	Culture, sports and recreation
		S11	Personal services
G3	Business, financial and information sectors	S6	Information and telecommunication
		S7	Financial services; renting of real estate and moveable property
		S8	Business services
G4	Trade, transport, government and other sectors	S1	Trade, transport and storage
		S9	Government, education and healthcare
		S13	Other sectors