The external effects of football stadiums as consumption and production amenities

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis Urban, Port and Transport Economics

Name: Sven van Kessel
Student ID number: 546733

Supervisor: J. van Haaren
Second assessor: xxx

Date final version: 08-10-2022

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

Amenities have been widely investigated, and it has often been suggested that they have a wide range of effects on the neighbourhoods in which they are located. However, football stadiums as amenities have mostly been disregarded in the Netherlands. As municipalities spend a lot of money on these football stadiums, it is important to acknowledge and understand their impact on the surrounding areas. Therefore, the external effects of football stadiums as consumption and production amenities in Dutch urban regions is investigated in this thesis. Moreover, the research will be expanded by splitting the effects of the stadiums into positive and negative externalities by means of proxy variables. Also, the producer advantages side of football stadiums will be introduced. In terms of consumption amenities, house prices are collected, whereas for the stadium as a production amenity, the total number and composition of jobs are obtained.

In this thesis, the stadium locations for all Eredivisie clubs of the 2020/2021 season have been plotted in QGIS, after which their neighbourhoods, adjacent neighbourhoods and municipalities were determined. Also, the positive and negative externalities are quantified as the average percentage of sporters in a municipality and the number of negative news articles per football club, respectively. The effects of football stadiums on surrounding areas have turned out to be a particularly difficult issue, because the level of variance is very low. Very little new stadiums are being built, making it harder to investigate the effects in the current market. No evidence has been found that football stadiums increase house prices on a local level, but an increase in house prices in the municipality is statistically significant. Additionally, it has been found that football stadiums have no significant effect on local economies. Moreover, the results suggest that it is possible to split the stadium effect into positive and negative externalities, but the attempt in this thesis was not successful. To sum up, the results are inconclusive, and differ in a certain way from the prevailing literature.

However, it should be noted that causal interpretations are not possible, as the models might suffer from several problems like reverse causality. It remains questionable whether stadiums can increase house prices and attract jobs, or whether stadiums are built in places where a lot of people already live, or where there are many jobs and thriving businesses. In addition to that, the modifiable areal unit problem and omitted variable bias are expected to play a crucial role. Additional research is needed to be able to draw clear conclusions. For this future research, a repeated sales method amongst others is likely to deliver accurate results. An alternative methodology for the splitting of the effect is required as well. Different proxy variables may be used for the positive externalities, whilst a different approach is advised for the negative externalities. Instead of only computing the negative articles for the municipalities with football stadiums, a full media analysis that focuses on the sentiment of each article for all municipalities is suggested.

Contents

Introduction

Football plays a major role in the everyday life of billions of fans all around the world. A football club becomes part of the identity of a fan and can connect large groups of people (Stone, 2007). In the Netherlands, football is by far the most popular sport in terms of followers (Statista, 2019). It has more than 8 million fans, with on average 200.000 of them being present in a stadium each matchday and 2 million watching the games live at home (PWC, 2021). As football is often described as 'the most important of the least important things', stadiums facilitate this passion for all kinds of people, from businessmen in private skyboxes, to children in the stands. A substantial amount of Dutch stadium developments has either taken place, or is planned to happen, including renovations, redevelopments and even relocations. In addition to that, stadium developments are more often integrated into urban planning, and the economic aspects as well as the aesthetic side of stadiums are becoming increasingly important for cities and urban planners (Van Dam, 2000). The question arises whether football stadiums are a factor in the development of cities, and whether surrounding houses and companies benefit from a stadium in the vicinity. Therefore, the central research question of this thesis will be:

"To what extent do football stadiums create external effects as production and consumption amenities in urban regions?"

The determinants of urban development and growth have been widely investigated. According to Roback (1980), amenities like football stadiums can increase consumer utility or generate productive advantages, or both. Glaeser, Kolko and Saiz (2001) build on the increased consumer utility in their article 'Consumer City'. They state that the role of cities in providing consumption possibilities for their inhabitants is becoming more and more important, and cities that excel in this area have grown faster than other cities. According to this article, jobs follow people accordingly, so people supposedly allocate where the amenities are the best. It has even been found that within a city, rich people either live in the city centre or in the suburbs, depending on which area offers the highest amenity level (Brueckner, Thisse & Zenou, 1999). Amenities have become one of the factors that show a positive effect on people's willingness to live in a city, due to their contribution to an augmented quality of life (Mulligan & Carruthers, 2011) This increased demand for urban areas with high levels of recreational resources results in higher house prices (Feng & Humphreys, 2018; Wu, Adams & Plantinga, 2004).

In the Netherlands, the main research focus lies on the effect of other kinds of amenities than stadiums. It has been established that especially mobile workers, like high skilled employees active in the service sector, are particularly aiming to live in areas with high amenity levels rather than areas which are closer to their jobs (Van Oort, Weterings & Verlinde, 2003). With this in mind, it has been determined that investments in cultural amenities bring about price level increases in the surrounding housing market as well (Koster & Rouwendal, 2017). Furthermore, it has been specified that redevelopments of shopping centres also increase local house prices (Zhang, Van Duijn & Van der Vlist, 2020). Even natural amenities like a beautiful view, a house that borders water or an attractive landscape positively influence house prices on the Dutch housing market (Luttik, 2000). Football stadiums can be regarded as amenities as well, and as people are attracted by these, they could lead to higher demand for houses and thus higher housing prices in the areas surrounding the stadiums. Thus, the first sub question that will be used to answer the research question is whether the presence of a football stadium positively influences surrounding house prices, and if so, by how much. Next to this, the second sub question will look at whether the effect on house prices is a local, urban or regional effect.

Feng and Humphreys (2018) have already investigated the effect of the Nationwide Arena and the Crew Stadium, both in the United States, on house prices in the area around the stadiums. They found that house prices were significantly higher in close proximity to the facilities, and that there is a decay in this effect when moving further away from them. This could be an indicator that sport stadiums have the same effect on house prices in the Netherlands as well. Earlier on, Huang & Humphreys (2014) looked into 56 professional sport facilities. They show that in the United States, house prices in areas surrounding these facilities are generally higher than houses that were not close to such a facility. Closer to the Netherlands, Ahlfeldt & Maennig (2010), by examining the housing market in Berlin, found that their selected stadiums increase the prices of houses within 3000 metres as well. The increase due to stadiums is also confirmed in London, where Ahlfeldt & Kavetsos (2014) found the same effect in the housing market in proximity to the Wembley and Emirates Stadium.

Although quite a lot of articles have been written about cultural sites and shopping centres as amenities, these football stadiums have mostly been disregarded in the Netherlands. Considering the previously mentioned importance of football stadiums as amenities, research on this topic in the Dutch context is required. There is one master thesis in which the effect of stadium developments on house prices is investigated. It was found that during the developments, surrounding house prices relatively decreased, and after the developments, the author saw an increase in house prices (Bieze, 2021). Therefore, this study will be partly replicated, but different data will be used, and the focus will be on football stadiums only, instead of on sport stadiums in general. Moreover, the research will be expanded by splitting the effects of the stadiums into positive and negative externalities by means of proxy variables. According to Li and Brown (1980), the effect of amenities on house prices is the sum of these positive and negative externalities. Following this line of thinking, the third sub question will be asking whether the effects of Dutch football stadiums can be split into positive and negative externalities.

Another addition to the existing literature will be on the producer advantages side of amenities like football stadiums. As described by Van Dam (2000), football stadiums attract businesses in various ways. Not only do they offer special business seats or skyboxes, but multiple stadiums also have space for offices and facilities. Therefore, a football stadium can be a place for networking, but also provide working space for companies. Next to potential sponsor deals between businesses and football clubs, these factors have the potential to increase the producer advantages that football stadiums as amenities have to offer for surrounding firms. To operationalise this effect for Dutch stadiums, the fourth sub question focuses on whether the presence of a football stadium influences the number of jobs and the composition of jobs within the municipality.

Next to the new academic insights that this study brings, the results of this thesis will be useful for urban planning. If house prices are higher around football stadiums, this indicates that people value them. Increasing the number of stadiums, or upgrading the existing ones, can potentially increase the quality of life of people. Also, stadiums might have the potential to excel economic growth in certain cities and neighbourhoods, or even revive deprived areas. According to Mulligan and Carruthers (2011), urban planning should prioritise quality of life, and thus amenities like stadiums. Moreover, when the positive externalities are separated from the negative externalities, this might provide insightful information on to what extent both are present and might help in increasing the positive externalities while dealing with the negative ones.

Lastly, municipalities spend a lot of money on professional football clubs and their stadiums. KPMG (2003) shows that in the ten years prior to their report, all municipalities that include a professional football club have provided it with funds. Total spendings exceeded 300 million euros, of which around 80% is directed to the stadiums. For the United States, Carlino and Coulson (2004) have already indicated that high public expenditures on stadiums of National Football League clubs are worth it for society, but this thesis might help investigate whether this is true for Dutch football stadiums as well. The answers to the fourth sub question about the producer advantages of football stadiums as amenities might add valuable insights, because if football stadiums increase the number of jobs in an area, these economic advantages should be taken into consideration as well.

To summarise, the following sub-questions will be investigated to formulate an answer to the research question:

1. Does the presence of a football stadium positively influence surrounding house prices, and if so, by how much?
2. Is the potential stadium effect on house prices a local, urban or regional effect?
3. Can the effects of Dutch football stadiums be split into positive and negative externalities?

4. Does the presence of a football stadium influence the number of jobs and the composition of jobs within the municipality?

First, the literature behind the research question will be addressed in detail in the literature review section, and hypotheses will be formulated. Subsequently, the data sources and methods will be described, after which the results will be presented and discussed. These sections will be followed by a synthesis that compares the results to the examined literature. Lastly, there will be a conclusion that summarises the main findings, highlights potential limitations of this research and provides recommendations for future research.

Literature review

Football stadiums as amenities

A football stadium is a particular product, as its main value proposition is more of a service. It derives its main value from providing people with the opportunity to watch a football game. However, modern stadiums show an expanded focus on other functions as well (Paramio, Buraimo & Campos, 2008; Van Dam, 2000). When combining both cited articles, it becomes clear that football stadiums often include restaurants, bars and fanshops, which are open outside of matchdays as well. In addition to that, stadiums sometimes provide tours or even serve as event venues for concerts for instance. Also, the economic side of stadiums is becoming increasingly important. From business seats and skyboxes to office spaces and sponsorship deals, the commercial aspects of stadiums have never been more relevant.

Before investigating the external effects of football stadiums specifically, it is important to note that a football stadium is categorised as an amenity. An amenity can be defined as “something that helps to provide comfort, convenience, or enjoyment” (Merriam-Webster, n.d., Definition 1). This definition entails basic facilities like supermarkets and schools, but also leisure establishments such as restaurants, tourist attractions and sport establishments. Moreover, parking, beautiful sceneries or views, large forest areas or a pleasant climate can be included as well. Likewise, football stadiums are also facilities that can bring excitement or pleasure to people, so they are incorporated in the definition too. As pointed out in the Introduction, Roback (1980) concluded that amenities like football stadiums can potentially increase consumer utility and generate productive advantages.

The increased consumer utility has been operationalised by means of the first sub questions about house prices. In the existing literature, there is an ongoing debate about whether jobs follow people, or people follow jobs (Hoogstra, Van Dijk, & Florax, 2017; Partridge & Rickman, 2003; Steinnes, 1978). In other words: do people allocate where they can get a job, or do they choose an attractive place to live and find a job in that area? Partridge and Rickman (2003) and Porter (2015), amongst others, tend to be in favour of the first option. They either propose that labour demand is the most important for migration, or that to revitalise cities, jobs should be created which then automatically attract people. Opposed to this is the evidence that jobs follow people (Cooke, 1978; Hoogstra, Van Dijk, & Florax, 2017). The findings indicate that cities must attract people by providing amenities, and that jobs will come after that. Like indicated before, especially mobile workers, like high skilled employees active in the service sector, attach more value to amenities than distance to their job (Van Oort, Weterings, & Verlinde, 2003).

In this light, one of the first to link amenities to house prices is Roback (1980). Depending on the extent of which the amenities in an area have an influence on consumer utility and firm productivity, they have an impact on the rent and wages in that area. Given that amenities increase utility, rents will be higher in areas that contain a lot of amenities that have a beneficial impact on the productivity of firms, while there is no clarity in what direction wages will be influenced. These wages will be lower in the area when the amenities are generally unproductive, and in this case, it is unclear what happens to the rents. Glaeser, Kolko and Saiz (2001) elaborate on these findings, stating that cities with high level of amenities have grown faster than other cities. They argue that jobs follow people, so people allocate where the amenities are the best. Given that a football stadium increases the level of amenities in an area, it potentially attracts people and thus increases house prices.

Further research indicates that various amenities can have different ways in which they bring about an effect on people. Corney, Ives and Bekessy (2015) have elaborated on this in the light of natural amenities. For instance, these amenities might engage with people on the level of sensory perception. They might be aesthetically pleasing on the eye, or sun rays, wind and particular vegetation might address the sense of touch. Also, amenities are experienced in different ways by different people, based on cultural and social aspects, or knowledge and experience. Bringing this into a broader perspective, certain types of shops and tourist attractions might be pleasant on the eye, enjoyable, or bring comfort to people as well. This line of thinking can also be applied to football stadiums, where, according to Van Dam (2000), urban planners and developers have started to realise that stadiums can serve different purposes than just bringing joy to people by facilitating football matches. Van Dam (2000) presents multiple aspects and articles to support this statement. Stadiums have been researched as foundations for economic growth and redistribution (Coates, 2007), and it has been argued that professional football clubs, hosted by a stadium, can be a crucial factor in inter-city competition for business activity, inhabitants, and tourism (Fischer & Hamm, 2019). Moreover, football stadiums can be aesthetically pleasing and become urban landmarks (Churchman, 1995).

These factors can contribute to an augmented quality of life. Like already reasoned in the Introduction, this suggests that people are more inclined to live in areas with high levels of amenities (Mulligan & Carruthers, 2011). This conclusion potentially holds within cities as well, where it was found that rich people choose to live in either the centre or suburbs, whichever region outperforms the other in terms of quantity and quality of amenities (Brueckner, Thisse & Zenou, 1999). Following the simple model of supply and demand, the area with the most amenities should thus exhibit relatively higher house prices when taking into consideration the reasoning above. As it has been determined that a football stadium is an amenity, it can be expected that areas surrounding a stadium display relatively higher housing prices than other areas. This has been shown repeatedly in the existing literature, where findings from

Europe on one side, and the United States on the other side, have provided meaningful insights into the topic.

Europe versus the United States

Although findings from Europe and the United States mostly point to similar results, there are differences that should be highlighted. Firstly, it is important to note that almost all American research of stadium effects is on National Football League (NFL), National Basketball Association (NBA), Major League Baseball (MLB) or National Hockey League (NHL) stadiums, while European research mostly focusses on football stadiums due to the difference in popularity of the sports on both respective continents. This might make it harder to apply results found in the United States to our Dutch cases, as different sports might bring about different effects on people and thus house prices. However, as the research conducted in America is nevertheless investigating the effects of stadiums of the most popular sports, the results might still provide a good starting point.

Carlino and Coulson (2004) present one of the first American studies on this topic. They investigate the 60 most densely populated areas in the United States, exploring whether the house prices are related to whether the city has an NFL team or not. They find that hosting an NFL team in the city raises house prices by 8% on average in city centres, but outside of the centres, this effect decreases significantly. As hosting an NFL team is synonymous with having a stadium, the study proposes evidence that stadiums increase rents in surrounding areas in the cases investigated. Around 10 years later, Huang and Humphreys (2014) look at whether mortgage applications increased at 56 places where a new professional sport facility had opened. In the three years after the openings, mortgage applications became considerably higher, but when the researchers controlled for locational characteristics, the size of the effect declined drastically and even became insignificant. This points to the result that the mortgage applications would have been higher anyways, even without the introduction of the new professional sport facilities. Nonetheless, it is noted that the sport facilities are a part of the development plans for an area that could have been the reason for the increased mortgages. The authors conclude with indicating that location choice is important when investigating the effects of sport facilities.

The main results presented so far do not provide a clear view on the effect of sport stadiums on house prices yet, so more evidence is required. In a case study, Tu (2005) specifically looked at rents in areas near the FedEx Field, an American football stadium, prior to and after it was built. When the stadium was completed, the houses closest to the facility displayed the biggest price increases, with a decaying effect to up to 2.5 miles, after which the effect became negligible. Feng and Humphreys (2018), in another case study on the Nationwide arena and the Crew Stadium, generate similar results, with an aggregated increase of about 21 million dollars in house prices around the Nationwide Arena for example. These findings, in combination with the findings described in the previous paragraph, generally indicate a positive effect of stadiums on nearby house prices (Carlino & Coulson, 2004; Feng & Humphreys, 2018; Tu, 2005). This effect is limited to areas relatively close to the stadiums and does not reach far outside city centres in the United States.

In Europe, less research has been done into multiple stadiums, but case studies are the most common. Ahlfeldt and Maennig (2010) conducted such a case study in and around Berlin, Germany. Just like Tu (2005), their findings advocate a general positive effect of the stadiums on rents, this time up to 3 kilometres. However, the stadiums independently displayed various degrees of effects, which, according to the authors, is due to different ways in which the stadiums pose negative externalities on neighbouring areas and the extent to which these are dealt with. Ahlfeldt and Kavetsos (2014), in a rather similar case study, looked at the constructions of the Wembley and Emirates stadiums in London. Again, the findings imply that the stadiums increase house prices in neighbouring areas with a decaying effect. For the Wembley Stadium, the aggregated increase in property values was almost 2 billion pounds. Next to the effect on house prices, Zawadzki (2021) poses that technological innovations to existing stadiums also lead to social benefits. By researching the innovations Poland made to stadiums in 4 different cities to prepare them for the European Championship of 2012, the conclusion was drawn that these technological innovations led to a social benefit that was valued at around 18 million dollars by residents.

To summarise, both the American and European literature generally finds an increase in housing prices for areas which include a professional sports stadium, with a decaying effect over distance. In both general studies as case studies, similar conclusions are drawn. The big difference between stadiums in the United States and European stadiums is the magnitude of the effect on neighbouring house prices. In Europe, these aggregated effects were valued at a substantially higher level than in the United States. Relating these findings to the first sub question whether the presence of a football stadium positively influences surrounding house prices, and if so, by how much, it can be expected that people are willing to pay a sizable amount to live close to a football stadium. This means that football stadiums are predicted to have a positive effect on surrounding house prices. It is too early and premature, however, to speculate about a specific amount yet.

Local, urban and regional effects

As indicated above, the spatial extent to which football stadiums affect house prices slightly differs per article. For amenities in general, there is no consensus about this matter. Some amenities, like school quality, seem to have very local effects (Livy, 2017). On the contrary, landscape amenities are more likely to bring about an urban effect rather than a local effect (Waltert & Schläpfer, 2010). Furthermore, Ullman (1954) even suggests regional impacts due to amenities. Therefore, the effect seems to differ per amenity itself, something that is confirmed by Bartik and Smith (1987). They state that amenities can be characterised by their respective impacts, which can be regional, urban, local, or even on the level of blocks. Thus, empirical literature is needed to determine the reach of the effect of football stadiums on surrounding areas.

As described in the previous section, Carlino and Coulson (2004) indicate that the presence of an NFL team in the city raises house prices by 8% on average in city centres, but this effect fades outside of the centres. Where Tu (2005) in a case study of the FedEx Field posed that the houses closest to the facility displayed the biggest price increases, with a decaying effect to up to 2.5 miles, Feng and Humphreys (2018) used a 1-mile radius around the stadiums, but indicated that for at least one of the two cases, a 3-mile radius seems to be a good estimate of the effect. In Europe, Ahlfeldt and Maennig (2010) find a general positive effect of the stadiums on rents up to 3 kilometres, but this distance is dependent on how well negative externalities are dealt with. In a later study, Ahlfeldt and Kavetsos (2014) reported that the general evidence pointed to an effect up to 3-5 kilometres, and their empirical data confirmed a 3-kilometre range for the first stadium, and a 4-kilometre for the second one.

Mainly looking at the European findings as these are the closest to our Dutch research, we can expect an effect of professional football stadiums in a range of around 3-5 kilometres (Ahlfeldt & Kavetsos, 2014; Ahlfeldt & Maennig, 2010). When looking at Dutch cities, it can be expected that the stadium effects are therefore local and urban, but do not reach regional proportions. Therefore, in answer to the second sub question that investigates the spatial extent of the stadium effect, it can be hypothesised that the effect of stadiums on surrounding housing prices is an urban effect in the Netherlands. However, the effect is subject to the externalities the stadium produces.

Positive and negative externalities

That is because the range of the stadium effect is dependent on how well the negative externalities are countered (Ahlfeldt & Maennig, 2010). This suggests that football stadiums generate both positive and negative externalities. As mentioned before, Li and Brown (1980) state that the effect of amenities on house prices is the sum of these positive and negative externalities. They reason that the effect of an amenity consists of the positive effect of accessibility and the negative effect due to diseconomies, both decreasing with distance, but the latter declining faster than the former. Positive externalities can consist of an increased amount of sporters within an area with a professional football club (Pyun, Kim, Schlesinger & Matto, 2020). Another example is the pride or connection that local people experience due to the football club in their city or neighbourhood (Johnson & Whitehead, 2000). On the other side, negative externalities can consist of a variety of diseconomies like nuisance, pollution, reduced quality of views and congestion (Li & Brown, 1980; Tu, 2005).

The peak of the amenity effect on house prices is not found at houses right next to the amenity, but rather at houses that are slightly further away (Li & Brown, 1980). This indicates that houses close to the amenity potentially experience the diseconomies the most, while the houses at the peak are at the right balance between accessibility and diseconomies. Houses that are further from the peak are impacted less by the diseconomies, but experience lower degrees of accessibility. These statements are supported in the stadium related literature as well. It has been found that people who live very close to a proposed stadium location, often oppose the stadium, even though the region was in favour of the stadium (Ahlfeldt & Maennig, 2012; Davies, 2005). This is referred to as the 'not in my backyard' (NIMBY) effect. A recent case study presents empirical evidence for this effect regarding stadiums, even observing decreasing prices for houses located closest to the stadium (Hyun, 2021). These findings suggest that houses right next to football stadiums experience the most disturbances, while houses slightly further away are less impacted by this, but still profit from the high level of accessibility.

In answer to the third sub question about whether the effects of Dutch football stadiums can be split into positive and negative externalities, it is hypothesised that this can indeed be done. It is expected that there are both positive and negative externalities arising from a football stadium. Also, when linking the findings above to the previous section about the extent of the stadium effect, it is expected that the local impact is negative or lower than the urban impact due to the NIMBY effect. Houses right next to the stadium are expected to have lower increases than houses slightly further away, which are anticipated to be located at the peak of the effect. When moving further away from them, the effect will gradually decrease.

Productive advantages

Next to the consumer utility advantages of football stadiums that have been operationalised by using house prices, football stadiums have the potential to generate productive advantages for firms as well. As pointed out in the Introduction, these advantages might stem from amongst others business seats, skyboxes and stadium offices (Van Dam, 2000). These are not only used as working space, but networking is an important aspect as well. Contradictory to conference halls, meetings in football stadiums are more spontaneous. Whereas conversations in conference halls are often planned, encounters in a stadium are more casual. They are based on a shared interest instead of congruent job types, which means that the people that meet are more diverse. This may lead to cross- and spillovers. Moreover, football stadiums attract sponsors as well, which result in an increase in economic activity in the area. All these factors can lure companies towards the city or municipality in which the football stadium is located, leading to local development. This in turn would mean an increase in jobs and businesses in the area in which a football stadium is located.

However, the relevant literature indicates that there is no evidence to suggest that this happens. The substantial subsidies that stadiums receive are often labelled as unjustified when reviewing the economic effects (Baade & Dye, 1988; Coates, 2007). Even more so, the consensus is that football stadiums have no effect on local economies (Coates & Humphreys, 2000; Coates & Humphreys, 2003; Harger, Humphreys & Ross, 2016). Therefore, as an answer to the fourth sub question, it is hypothesised that a football stadium has no influence on the amount and composition of jobs.

Data

In this thesis, multiple data sources are used to create an integrated dataset. First, several data points on football stadiums are collected from sources like Transfermarkt, the Wijk- en buurtkaart 2020, National Institute for Public Health and the Environment (RIVM) and Nexis Uni. This data is then combined with three large datasets: the Kerncijfers wijken en buurten (KWB) 2020, a collection of transaction data of houses from the Dutch association of real estate agents (Nederlandse Coöperatieve Vereniging van Makelaars en Taxateurs, NVM), and a dataset originating from Stichting LISA. In this section, all data sources and procedures followed to create the integrated dataset will be discussed in detail.

The football stadiums that are used in this thesis can be found in Appendix A, along with additional information. These are all the Eredivisie (highest Dutch national league) stadiums of the season 2020/2021. Only Eredivisie stadiums are selected as these are generally much larger than the stadiums of Eerste Divisie clubs (second-highest league), which could lead to large differences in the possible external effects they might generate. Also, the 2020/2021 season is used because the other sources used in this research mostly provide data until 2020. In QGIS, an aerial picture provided by PDOK is used to pinpoint the location of the football stadiums on the map, and a new layer is created that contains a dot that is placed at the centre spot of the football field in each stadium. This layer is then set on top of the Wijk- en buurtkaart 2020, a map that is provided by Statistics Netherlands (CBS) that shows all the neighbourhoods. In this way, the neighbourhood in which a stadium is located can be determined, just like the adjacent neighbourhoods and the corresponding municipalities.

Especially the selection of adjacent neighbourhoods proved to be rather difficult, so eventually some criteria were developed to make the selection as accurate as possible. Examples of specific cases are described in Appendix B. Not every neighbourhood that directly bordered a neighbourhood with a stadium was marked as an adjacent neighbourhood. This mainly had to do with the position of a stadium within a neighbourhood. If a stadium was situated at the edge of a neighbourhood, the areas that border that neighbourhood on the other side might be that far away, that it would not make sense to mark that area as adjacent, so they were not included. An example of this is shown in Figure B.1. Another exclusion criteria is the existence of a border between the neighbourhood that hosts the stadium and its adjacent neighbourhood, of which an example is displayed in Figure B.2. Moreover, if a neighbourhood was not bordering a neighbourhood with a stadium, but it would still be logical to include it based on its distance to the stadium relative to other neighbourhoods that were included, it was nonetheless included. An example of this is provided in Figure B.3.

A specific case is the neighbourhood in which De Oude Meerdijk of FC Emmen is located. As Figure B.4 points out, the neighbourhood of the stadium is identified as ‘scattered houses around Emmen’. Therefore, the direct effect of the stadium in the neighbourhood itself has not been considered, as the houses are too scattered. Also, if an adjacent area is designed as an industrial site, it was excluded too like in the example in Figure B.5. More generally speaking, an adjacent neighbourhood would be included, except when the furthest point of the neighbourhood was that far from the stadium that it would make it illogical to select. This was often combined with the level houses present in the area relative to amongst others the level of industry or nature. In this way, the selection of adjacent neighbourhoods has been kept as consistent as possible.

Because the way of selecting adjacent neighbourhoods as described above might still be subjective or inconsistent, the selection is repeated by using a different method. For each neighbourhood, the geographical centroid is constructed in QGIS. Then, a circle with a radius of 1 kilometre is drawn around each stadium, and the neighbourhoods of which the centroids fall within these circles are identified as adjacent neighbourhoods. In this way, a consistent selection method is applied in which the centre of a neighbourhood should be close enough to a stadium to be chosen. A 1-kilometre distance is used as a maximum, as this distance is defined as the maximum distance to still be considered walking distance by the CBS. Also, by trying out different distances in QGIS, this distance resulted in the outcome that seemed most reasonable based on the extent to which the neighbourhoods were selected. The new selection of adjacent neighbourhoods is added to the dataset as well. However, during the analysis it became clear that the descriptive statistics remained the same for both methods, and the regression coefficient only changed by 0.001 without any change in significance, so the conclusion is drawn that the first method was robust. This method has been used in the thesis.

Also, several stadium-related variables are introduced. First, the stadium capacity is included, according to the numbers in Table A.1. Areas without a stadium have the value ‘0’ for capacity, and for the municipality of Rotterdam, the capacities of the Feyenoord and Sparta stadium have been added together. Next to this, two variables are added that serve as proxy variables for the positive and negative externalities of a stadium. These will be used to split the stadium effect. The positive externalities are the percentage of sporters in a municipality, measured as the percentage of people aged 18 or older that sport at least once per week. This data stems from the National Institute for Public Health and the Environment (RIVM) and is based on the Health Monitor (Gezondheidsmonitor) 2020 for adults, which uses statistics from Municipal Health Service (GGD’en), CBS and RIVM.

The negative externalities are the total number of negative articles about hooligans for each club. To compute this number, the search term ‘hooligans’ was used, followed by the name of the football club.

This name is kept as short as possible, without including city names in the search term (so for instance 'Ajax' instead of 'Ajax Amsterdam'). This was done because a test case of the first three stadiums of Table A.1 showed that this strategy resulted in the highest accuracy of newspaper articles. As articles often only mention the short name of a football club instead of its full name, the search became more complete in this way. The search is performed on Nexis Uni, for which several filters are used. The only newspaper included is the Algemeen Dagblad (AD), to rule out duplicate articles.

Also, it is important that only negative news is shown in the search results. That is, because the word 'hooligans' can also have a positive connotation. Football can be an important connector between fans that refer to themselves as hooligans in the good sense of the word. However, for the negative proxy variable, only negative articles are desired, as it is intended to measure the nuisance and other negative effects of football stadiums. To ensure this, the subject is set to "Crime, law enforcement & correction". In this way, only the negative articles about hooligans are shown without incorporating the positive ones in the search. The most exact option would have been to scan all the articles for each stadium systematically based on their sentiment, but this would require a full media study taking several months. As time is limited, this falls outside the scope of this thesis. Thus, it is chosen to filter the subject to "Crime, law enforcement & correction" as this is the most practical solution. This method is again validated by a scan of several stadiums, which proved that the search results became more accurate in terms of sentiment.

In the end, the number of articles that came up from the search term for each club became the value for the proxy variable. Some exceptions are made in specific cases. For Sparta, the name of the club that is used in the search term is 'Sparta Rotterdam', as just using 'Sparta' resulted in articles related to the Czech football club Sparta Prague. As a filter for Fortuna Sittard, the language is set to Dutch as the articles were otherwise often in German and related to the football club Fortuna Düsseldorf. Lastly, as both Sparta and Feyenoord are in Rotterdam, their amount of negative news articles were averaged relative to their stadium capacities. Areas without a stadium have the value of '0'.

The final stadium variable that is added, is the position of the club in the final ranking of the 2020/2021 Eredivisie season. Several categories were made, based on the respective sportive consequences of each position. The first category entails the first place, the champion of the league. The second category consists of the places 2, 3 and 4, which allowed these clubs to play European football in the next season. The third category includes the clubs that reached the domestic play-offs for European football, namely places 5, 6, 7 and 8, whereas the fourth category holds the 'middle' places 9, 10, 11 and 12. The teams that just escaped relegation on places 13, 14 and 15 are incorporated in category 5, while the team that finished 16^{th} and had to play play-offs for relegation is included in the sixth category. The relegated teams on places 17 and 18 make up the seventh category. The eighth category means that a neighbourhood or municipality did not include a stadium.

The information about the stadiums in neighbourhoods is then added to three main datasets, the first one being the Kerncijfers wijken en buurten (KWB) 2020, published by Statistics Netherlands (CBS). This dataset consists of key data about all neighbourhoods and municipalities within the Netherlands. It contains demographic data, as well as statistics about the housing market, amenities, and spatial structures. The main statistic for this thesis is the average property value (WOZ-waarde) per neighbourhood, which will be the dependent variable in the first model. Also, the 2020 data is selected as this is the most recent version that includes this variable. For each neighbourhood, it is indicated whether it accommodates a stadium or not, or whether it is adjacent to a neighbourhood that accommodates a stadium. Also, for each neighbourhood, there is a variable that expresses the presence of a stadium in its municipality. Dummies for provinces and the four biggest municipalities, that also include the four biggest cities, are added as well. The selection consists of the four largest, as their number of inhabitants is clearly above the other municipalities. All variables of the dataset are described in detail in Appendix C, Table C.1.

However, when a model was chosen to test the stadium effect on surrounding house prices, it was encountered that the coefficients did not present reliable estimates of the variables, there was high multicollinearity, and the R^2 could be improved. Therefore, it was decided to first establish a model that accurately describes the house prices without the stadiums, and then later add the stadiums to this model. To find this model, several changes to the data were made that are described in Appendix D. Table C.2 gives an overview of all the new and remaining variables.

For the splitting of the stadium effect into positive and negative externalities, the same data as above is used, but now on the level of municipalities as the positive proxy variable, the percentage of sporters, is only available per municipality and not for neighbourhoods. For each municipality, it is indicated whether there is a stadium present or not, and what the level of positive and negative externalities are. The control variables are the same as in the previous model to keep the analysis as consistent as possible.

The next dataset is a collection of transaction data of houses from the Dutch association of real estate agents (NVM). The NVM consists of more than 4000 real estate agents and appraisers, registering the transaction of houses for a considerable part of the Dutch housing market. For each transaction, several statistics are recorded, like the transaction price and characteristics of the house that was sold for the years 1985-2021. This data has then been aggregated per year to the level of municipalities by taking the average of the transaction prices and characteristics for each municipality. The transaction data provides a more accurate variable to measure house prices than the WOZ-waarde in the KWB, and the period of the collected data allows to control for time. However, this comes at the cost of losing a lot of control variables that were present in the KWB.

Again, for each municipality it is indicated whether it hosts a stadium or not, or whether the stadium was under construction during the year based on the information in Appendix A. Under construction means that a completely new stadium was being built and excludes redevelopments. For stadiums that were built after 1985 but were placed at the same place as the old stadium, it is indicated that there was a stadium for each year in the data. However, for stadiums that opened after 1985 that were placed at a new location, the stadium is only marked as present from the year of the opening of the stadium, and not before that. Again, a dummy variable for the four biggest municipalities is added, while there was already a variable indicating the province. Table C.3 provides a detailed description of all variables. For the variable *average surface*, all observations with value '0' were dropped, and for the variable *average volume*, all observations with value '99999' is these appeared to be unrealistic values.

The last dataset originates from Stichting Lisa, which holds a database of all businesses within the Netherlands. The database used in this thesis includes the number of jobs per municipality for each type of job categorised by the three-digit SBI-codes for the years 2000-2017. These codes are used to divide the jobs into categories, and the more digits the code contains, the higher the level of categorisation. The selected year for the data is 2017, as this is the most recent data. Also, when using multiple years, a lot of the control variables cannot be used anymore, as the availability for each year is limited.

Using a pivot table, the total amount of jobs per municipality is calculated, which becomes the first dependent variable. To look into the composition of jobs within municipalities and the stadium effect on this, a pivot table is also used to create the second dependent variable, which is the sum of sports-related jobs in each municipality (SBI code 931). It might be possible that a football stadium attracts or inspires others to create jobs that involve sports. For the third dependent variable, another pivot table is used to calculate the total amount of financial institutions, business services and consultancy firms (SBI codes 641, 642, 643, 649, 661, 662, 663, 692 and 702). These jobs are included, because these are commonly seen as jobs where a lot of money is involved, and which are typically the jobs that pay for expensive business seats and lounges with the purpose of networking. By adding these in the model, it can be seen whether stadiums attract these kinds of firms. Lastly, as a fourth dependent variable, the previous dependent variable will be narrowed down to firms related to accountancy and consultancy (SBI codes 692 and 702). This is because these types of firms are even more specifically focussed on networking at amongst others football stadium. Examples of companies that are included in these categories are McKinsey & Company, the Boston Consulting Group and PricewaterhouseCoopers.

Next to the LISA data itself, Kerncijfers wijken en buurten (KWB) 2017, published by Statistics Netherlands (CBS), is added to the dataset. Again, this dataset consists of key data about all neighbourhoods and municipalities within the Netherlands, but this time for the year 2017 to match the LISA data. Variables included are statistics describing the demographics, amenities, income, social security and spatial structures that can be used as control variables. Also, variables indicating the provinces and the biggest four municipalities are added again. Table C.4 contains a detailed description of all variables. Observations for the municipalities Midden-Groningen, Waadhoeke and Westerwolde were dropped as there was no match between the LISA dataset and the KWB 2017 due to reformations in the municipalities. Also, all observations for the municipality of Rozendaal were dropped, because there was no information on the number of sports-related jobs for the year 2017. Several other transformations can be found in Appendix D.

Methodology

In this research, several models will be used by using the different datasets (KWB 2020, NVM and LISA) First of all, the KWB 2020 dataset will be used to provide a first answer to the three sub questions about the effect of stadiums on house prices, formulated in the Introduction. The basis of the model is described in the article about hedonic prices by Rosen (1974). He reasons that differences in product prices are due to differences in product characteristics. As a result, a higher price for a product is the result of attributes that are valued higher. For amenities specifically, it has been argued by Cheshire and Sheppard (2017) that their value is incorporated into land prices, and that therefore, land prices can be used to make an approximation of the amenity effect. In consideration of both articles, a hedonic price model can be constructed to estimate the stadium effects. The regression equation is as follows:

$$\begin{aligned}\ln\,(average\ property\ value_i) &= \beta_0 + \beta_1 Neighbourhood\ with\ stadium + \ \beta_2 Adjacent\ neighbourhood \\ &+ \ \beta_3 Municipality\ with\ stadium_i + \ \beta X_i + \varepsilon_i\end{aligned}$$

The data will be used on the level of neighbourhoods. *Average property value$_i$* is a continuous variable that gives the average property value based on the WOZ-waarde for neighbourhood *i* in thousand euros, of which the natural logarithm is taken. *Neighbourhood with stadium* is a dummy variable that takes the value '1' if there is a stadium present in the neighbourhood, and '0' otherwise. *Adjacent neighbourhood* is also a dummy variable, this time taking the value '1' if the neighbourhood is adjacent to a neighbourhood with a stadium. More information on this variable can be found in the Data section and Appendix B. Next to that, *Municipality with stadium* is a dummy variable as well that takes the value '1' when a stadium is present in the municipality for neighbourhood *i*, and '0' if this is not the case. Furthermore, X_i presents a vector of control variables for each neighbourhood *i*. The control variables consist of data about demographics, the housing market, education, amenities, stadium characteristics and spatial factors. More details about the variables used can be found in Table C.2. Lastly, the regression coefficients are symbolised by means of β's, while ε_i represents the error term. Huber-White heteroskedasticity-robust standard errors are used in the analysis. Descriptive statistics can be found in Table E.1.

To split the stadium effect into positive and negative externalities, the proxy variables for both are used. This will be done on the level of municipalities, as the positive proxy variable, the percentage of sporters, is only available for each municipality and not for neighbourhoods. The regression equation used in the model is:

$$\ln (average\ property\ value_i) = \beta_0 + \beta_1 Municipality\ with\ stadium_i + \beta_2 Positive\ externalities + \beta_3 Positive\ externalities * Municipality\ with\ stadium + \beta_4 Negative\ externalities + \beta X_i + \varepsilon_i$$

Average property value$_i$ is a continuous variable that gives the average property value based on the WOZ-waarde for municipality *i* in thousand euros. *Municipality with stadium* is a dummy variable that takes the value '1' when a stadium is present in municipality *i*, and '0' if this is not the case. *Positive externalities* and *Negative* externalities represent their proxy variables respectively. An interaction term between the positive externalities and the variable indicating whether a municipality has a stadium is also included. This is not done for the negative externalities variable, because this already is kind of an interaction variable as it takes value '0' when there is no stadium in the municipality. Furthermore, X_i presents a vector of control variables for each municipality *i*. The control variables consist of data about demographics, the housing market, education, amenities, stadium characteristics and spatial factors. More details about the variables used can be found in Table C.2, as the same ones are used in this model but on a different level. This is done to keep the analysis as consistent as possible. Lastly, the regression coefficients are symbolised by means of β's, while ε_i represents the error term. Huber-White heteroskedasticity-robust standard errors are used in the analysis. Descriptive statistics can be found in Table E.2.

To investigate the stadium effects while controlling for time, the NVM dataset will be used in the second model. This data also provides a more accurate dependent variable for house prices, namely the average transaction price in a municipality for each year. Again, the hedonic pricing model as described above will be used. The regression equation for this model will be:

$$Average\ transaction\ price_{it} = \beta_0 + \beta_1 Stadium_t + \beta_2 Stadium\ under\ construction_t + \beta X_{it} + \varepsilon_{it}$$

Here, *Average transaction price$_{it}$* is a continuous variable for the average transaction price in municipality *i* for year *t*. *Stadium$_t$* is a binary variable taking one the value '1' if the municipality accommodates a stadium in year *t*, and '0' otherwise. Moreover, *Stadium under construction$_t$* is also a binary variable. It takes on value '1' when a stadium is being built in the municipality in the year *t*, and '0' when this is not the case. The year a stadium is being built is taken into consideration as well, as housing demand, and thus prices, in a neighbourhood might rise when people anticipate on the stadium being finished soon. Additionally, X_{it} presents a vector of control variables for each neighbourhood *i* per year *t*. The control variables consist mostly of house characteristics, a categorical variable to control for the years and spatial factors. The variables are described in more detail in Table C.3. Lastly, the regression coefficients are symbolised by means of β's, while ε_{it} represents the error term. Huber-White heteroskedasticity-robust standard errors are used in this analysis as well. The descriptive statistics are presented in Table E.3.

Finally, the third model allows to investigate the stadium effects on the number of jobs by using the LISA dataset in combination with the KWB 2017. To be able to investigate both the total amount of jobs, as well as the composition of jobs, four regressions will be run with four different dependent variables. The model will apply a linear regression with the following equations:

$$Jobs_i = \beta_0 + \beta_1 Stadium + \beta X_i + \varepsilon_i$$
$$Sports\ related\ jobs_i = \beta_0 + \beta_1 Stadium + \beta X_i + \varepsilon_i$$
$$Business\ related\ jobs_i = \beta_0 + \beta_1 Stadium + \beta X_i + \varepsilon_i$$
$$Network\ jobs_i = \beta_0 + \beta_1 Stadium + \beta X_i + \varepsilon_i$$

$Jobs_i$, $Sports\ related\ jobs_i$, $Business\ related\ jobs_i$ and $Network\ jobs_i$ are continuous variables representing the number of total jobs, sports-related jobs, business-related jobs, and jobs specifically involving networking in municipality *i* respectively. *Stadium* is a binary variable that takes value '1' if there is a stadium present in the municipality, and '0' otherwise. Just like before, X_i is a vector of control variables for the neighbourhoods, which consists of variable concerning demographics, income, social security, amenities and spatial factors. Table C.4 elaborates on all variables. Besides, the β's signify the coefficients, whereas ε_i stands for the error term. Again, Huber-White heteroskedasticity-robust standard errors are used in this analysis. The variables' descriptive statistics are displayed in Table E.4.

Results

To test the first two hypotheses, Table 1 is reviewed which contains the most important information about the regression results of the KWB 2020 dataset. The full results are shown in Table F.1. The dependent variable is the natural logarithm of the average property value (based on the WOZ-waarde), and the independent variables are an indicator of whether there is a stadium present in the neighbourhood, an indicator of whether the neighbourhood is adjacent to a neighbourhood that holds a stadium, and an indicator for the presence of a stadium within a neighbourhood's municipality.

Table 1: Main regression results for the models using the KWB 2020 dataset (CBS)

	(1) **ln(Average property value)**	**(2)** **ln(Average property value)**	**(3)** **ln(Average property value)**
Neighbourhood with stadium	0.065	0.064	0.064
	(0.061)	(0.061)	(0.061)
Adjacent neighbourhood		-0.022	-0.020
		(0.020)	(0.020)
Municipality with stadium			0.121***
			(0.034)
Constant	5.948***	5.948***	5.948***
	(0.081)	(0.081)	(0.081)
Observations	11,260	11,260	11,260
R^2	0.735	0.735	0.735

Note. This table shows the main results of three OLS regressions with the natural logarithm of the average property value as dependent variable, and having a stadium in the neighbourhood (1, 2 and 3), adjacent neighbourhoods (2 and 3) and municipalities with a stadium (3) as independent variables. The full results are available in Table F.1 Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

The first hypothesis says that it can be expected that people are willing to pay a sizable amount to live close to a football stadium, and thus that football stadiums are predicted to have a positive effect on surrounding house prices. In addition to that, it was hypothesised that the effect of stadiums on surrounding housing prices is an urban effect in the Netherlands. From Table 1, it becomes clear that the coefficients for neighbourhoods with a stadium and adjacent neighbourhoods are insignificant at the 10% confidence level, as their respective p-values are all larger than 0.1. This also holds when the second method of determining the adjacent neighbourhoods is used. However, the coefficient for municipalities with a stadium is significant at the 1% level, as the p-value is lower than 0.01. Given that the coefficient is positive, a municipality with a football stadium is associated with an average property value that is on average 12.1% higher than a municipality without a football stadium in the Dutch market, ceteris paribus. As this is quite much, the size is relevant enough.

This means that in these three models, no significant effect has been found of stadiums in house prices in the direct neighbourhood and adjacent neighbourhoods, but a significant effect has been found for stadiums within municipalities. Therefore, the first hypothesis is partially not rejected, as evidence has been found that people are willing to pay more to live in a municipality with a football stadium. The second hypothesis, stating that the stadium is an urban effect, is not rejected, as the stadium effect is significant at the level of municipalities. With regards to the stadium effect on the neighbourhood itself and the adjacent neighbourhoods, it becomes clear that the effect in the data is generally positive as the coefficients are positive. However, because these coefficients are not significant, it cannot be said with a sufficient degree of certainty that the effect is there.

For the third hypothesis regarding the splitting of the stadium effect, Table 2 provides the main regression results for the models that split the effect into positive and negative externalities. The full results are presented in Table F.2. Again, the dependent variable is the natural logarithm of the average property value (based on the WOZ-waarde). This time, the independent variables consist of an indicator for the presence of a stadium within a neighbourhood's municipality, and proxy variables for the positive and negative externalities. Also, an interaction term between the variables for the positive externalities and the presence of a stadium within a municipality is included. This has not been done for the negative externalities, because the negative externalities variable can already be treated as an interaction term, as it contains the number of negative articles for a municipality with a stadium and takes the value '0' otherwise.

Table 2: Main regression results for the models to split the stadium effect using the KWB 2020 dataset (CBS)

	(1) ln(Average property value)	(2) ln(Average property value)
Municipality with stadium	0.157***	0.219
	(0.046)	(0.196)
Positive externalities		0.011***
Percentage of sporters		(0.002)
Positive externalities * Municipality		-0.002
with a stadium		(0.004)
Negative externalities		0.002*
Amount of negative articles		(0.001)
Constant	6.098***	6.364***
	(0.480)	(0.471)
Observations	347	347
R^2	0.843	0.853

Note. This table shows the main results of two OLS regressions with the natural logarithm of the average property value as dependent variable, and having a stadium in the municipality (1 and 2), and positive and negative externalities (2) as independent variables. The full results are available in Table F.2. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Again, the stadiums seem to have a significant positive effect on the average property values in municipalities in the first model, as the p-value is lower than 0.01. In model 2, the effect becomes insignificant, which can be a sign that the effect is successfully split into positive and negative externalities. The negative externalities variable can be treated as an interaction term, as explained before. However, the negative externalities seem to have a positive effect on the average property value in a municipality, which is counterintuitive at first glance. The Discussion will elaborate further on this. The positive externalities of a stadium effect, represented by the interaction term, do not give a significant coefficient at the 10% level as the p-value is larger than 0.01. All in all, the third hypothesis is not rejected, as there is evidence that the stadium effects can be split. This is derived from the significant coefficient for the negative externalities. Even so, it should be noted that there are some clear issues with the way the externalities are split in this model. Again, the Discussion section will provide more detail on this matter.

To further investigate the stadium effect on house prices, the NVM dataset has been used additionally. This dataset provides more detailed information on house prices and characteristics of homes but includes less control variables. Table 3 provides the most important regression coefficients of the models, of which the complete overview is included in Table F.3.

Table 3: Main regression results for the models using the NVM dataset

	(1) Average transaction price	(2) Average transaction price
Stadium	-11,711.385	-12,882.814
	(11,986.908)	(11,944.847)
Stadium under construction		-92,978.824**
		(37,187.925)
Constant	-1.546e+07*	-1.545e+07*
	(9275744.829)	(9275286.128)
Observations	16,780	16,780
R-squared	0.048	0.048

Note. This table shows the main results of two OLS regressions with the average transaction price as dependent variable, and having stadium (1 and 2) and stadium under construction(2) as independent variables. The full regression results are presented in Table F.3. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

The coefficients for the stadium effect are insignificant as the p-value is larger than 0.1, meaning that a football stadium has no significant effect on house prices within the municipality in this model. On the other side, the coefficient for stadium under construction is significant at the 5% significance level, as the p-value is smaller than 0.05. The stadium under construction variable is a helpful addition to the previous analysis involving the negative externalities, as the construction captures a different negative aspect of a stadium. Next to riots, the construction of stadiums might also result in noise and other disturbances in an area. As the coefficient is negative, municipalities with stadium under construction in this sample have house prices that are on average 92,978.82 euros lower than in municipalities that do not have a stadium under construction. This is a solid indication that the construction of a stadium poses negative externalities on a municipality as well.

Next, the LISA dataset is analysed to test the fourth hypothesis, which states that a football stadium has no influence on the number and composition of jobs. Table 4 provides the most important results of this analysis, while Table F.4 gives the complete results. The dependent variables are the total amount of jobs to test the general influence of football stadiums on the job market, and sports-related jobs, business-related jobs and jobs specifically related to networking to assess the influence on the composition of jobs.

Table 4: Main regression results for the models using the LISA dataset combined with the KWB 2017 dataset (CBS)

	(1) **Jobs**	(2) **Sports-related jobs**	(3) **Business-related jobs**	(4) **Network jobs**
Stadium	5,118.266	92.406**	-32.841	22.563
	(3,190.925)	(38.752)	(398.196)	(179.960)
Constant	-1,159.113	-117.538**	313.193	28.778
	(2,784.167)	(50.044)	(319.994)	(160.857)
Observations	376	376	376	376
R-squared	0.987	0.942	0.983	0.985

Note. This table shows the main results of four OLS regressions with the total amount of jobs (1), sports-related jobs (2), business-related jobs (3) and network jobs (4) as dependent variables, and stadium (1, 2, 3 and 4) as independent variable. The full results are reported in Table F.4. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4 shows that the coefficient for stadium in the first model is not significant even at the 10% confidence level, as the p-value is larger than 0.1. This means that no significant effect of football stadiums on the total amount of jobs in a Dutch municipality has been found. Thus, the first part of the hypothesis that states that a football stadium has no effect on the total amount of jobs, is not rejected. In the data, the effect is positive on average, as a stadium is associated with an additional 5118 jobs, but this cannot be supported with sufficient degrees of certainty to also hold outside of the data.

In terms of job composition, Table 4 specifies that the stadium coefficient in the second model is significant at the 5% confidence level, as the p-value is smaller than 0.05. However, due to the small size of the coefficient, there is no evidence to suggest that the stadium effect is substantial. This will be elaborated on in the Discussion. Therefore, in combination with the insignificant stadium coefficient in model 3 where the p-value is smaller than 0.1, the last part of the third hypothesis, stating that stadiums have no influence on job composition, is not rejected. Also, when looking at the jobs that specifically involve networking, the same result is found. It is suggested by the coefficients that a municipality with a stadium has on average less business-related and more network jobs in our sample, but again these findings cannot be extrapolated.

Discussion

In the three models using the KWB 2020 dataset, no significant effect has been found of stadiums in house prices in the direct neighbourhood and adjacent neighbourhoods, but a significant effect has been found for stadiums within municipalities. Therefore, the first hypothesis is not rejected, as evidence has been found that people are willing to pay more to live close to a football stadium. The second hypothesis, stating that the stadium is an urban effect, is not rejected as well, as the stadium effect is significant at the level of municipalities. With regards to the stadium effect on the neighbourhood itself and the adjacent neighbourhoods, it becomes clear that the effect in the data is generally positive as the coefficients are positive. However, because these coefficients are not significant, it cannot be said with a sufficient degree of certainty that the effect is there.

On the contrary, the results obtained by analysing the NVM dataset show that the coefficients for the stadium effect are insignificant, meaning that a football stadium has no significant effect on house prices within the municipality in this model. One of the potential reasons for the difference in results between the models of the KWB 2020 dataset and the NVM dataset can be the amount of control variables, as the KWB 2020 contains some crucial controls that the NVM dataset lacks. Also, the R-squared of the NVM models are very low, indicating that they have less explanatory power in comparison to the KWB 2020 models. In combination with the fact that the KWB models contain more detail in terms of effect ranges, it is determined that more weight is put on these models. Therefore, the conclusions drawn in the first paragraph regarding the first and second hypothesis still prevail.

When testing for hypothesis 3 concerning the splitting of the stadium effect, the coefficient for the negative externalities is significant. However, the negative externalities seem to have a positive effect on the average property value in a municipality, which is counterintuitive at first glance. This effect might be due to a media bias towards the bigger clubs. Riots and conflicts by supporters of bigger clubs like Ajax, PSV or Feyenoord might sooner be the subject of an article in a Dutch national newspaper like AD, compared to fights between smaller clubs. Therefore, bigger clubs are likely to have a relatively higher negative externalities score than smaller clubs, so part of the negative externalities coefficient contains the stadium effect for bigger clubs. As this effect might be larger than the effect for small clubs, the negative externalities coefficient has become positive.

On the other side, the coefficient for stadium under construction in the models using the NVM dataset is significant at the 5% significance level. Stadiums under construction add to the previous analysis involving the negative externalities, as the construction captures a different negative aspect of a stadium. Next to riots, the construction of stadiums might also result in noise and other disturbances in an area. As the coefficient is negative, municipalities with a stadium under construction in the NVM sample have house prices that are on average 92,978.82 euros lower than in municipalities that do not have a stadium under construction. This is a solid indication that the construction of a stadium poses negative externalities on a municipality as well. Additionally, the interaction term's coefficient for the positive externalities is insignificant. This leads to the conclusion that there is no evidence pointing towards positive externalities arising from football stadiums in our sample.

In the end, the third hypothesis is not rejected, as there is evidence that the stadium effects can be split. This is derived from the significant coefficient for the negative externalities when it comes to the number of negative articles and stadiums under construction. Even so, it should be noted that there are some clear issues with the way the externalities are split in this model. Especially the selected proxy variables might not be suitable for the split that was intended. It proved to be very hard to measure and quantify the positive and negative externalities that may arise due to the presence of a stadium. Therefore, the attempt to split the stadium effect is not completely successful, as the aim was to create a clear split between the positive and negative externalities. Different proxy variables or different measurement techniques might result in more reliable results for this model.

Lastly, the fourth hypothesis that states that a football stadium has no effect on the total amount of jobs, is not rejected. No significant effect of football stadiums on the total number of jobs has been found. The same is true for the composition of jobs. The only significant coefficient suggests that a municipality with a football team has on average 92 more sports-related jobs than a municipality without a stadium. Although this might seem like a rather large amount, especially since Table E.4 indicates that a municipality on average has 163 sports-related jobs, it is not unlikely that the additional 92 jobs are solely jobs related to the stadium. It is not unthinkable that 92 people work at a professional football stadium, so the additional jobs are not visible in the rest of the municipality. Therefore, in combination with the other insignificant coefficients, the last part of the third hypothesis, stating that stadiums have no influence on job composition, is not rejected. Also, when looking at the jobs that specifically involve networking, the same result is found.

Synthesis

In this section, the differences between the results obtained in the empirical analysis and the analysed literature will be compared. For the first part of the thesis, some major differences are found. First, almost every paper has found a significant positive effect of football stadiums on house prices in areas that would count as the stadium neighbourhood itself and the adjacent neighbourhoods. In the United States, although most research was done on other sport stadiums than football (soccer) stadiums, both Carlino and Coulson (2004) in a general study involving 60 areas, as Feng and Humphreys (2018) and Tu (2005) in specific case studies, concluded that house prices are generally higher near stadiums. In Europe, Ahlfeldt and Kavetsos (2014) and Ahlfeldt and Maennig (2010), by means of case studies as well, came to the same conclusion, even reporting substantially higher aggregated effects than in the United States. These findings are in contrast with the results reported in this thesis, where analyses with the KWB 2020 dataset and the NVM dataset suggest no local effects of football stadiums on surrounding house prices. The expected effect with a range of around 3-5 km, based on the papers by Ahlfeldt and Kavetsos (2014) and Ahlfeldt Maennig (2010), could not be replicated in this study, as significant local effects were not observed.

One of the reasons for this contradiction might be the different research scopes deployed. All European papers that are discussed use case studies, just like the American paper by Tu (2005). These case studies often focus on the bigger stadiums. Tu (2005) investigated the FedEx Field, Ahlfeldt and Maennig (2010) two stadiums that were part of Berlin's bid for the Olympics, and Ahlfeldt and Kavetsos (2014) investigated the Wembley and Emirates stadiums in London. Therefore, it might be possible that there is an effect of football stadiums, but only if the stadium is large enough. When considering all stadiums, the coefficients might become insignificant due to a subgroup of stadiums that have no effect. As a result, just like in the study by Carlino and Coulson (2004) in which multiple venues were investigated, the results can become inconclusive. In addition to that, stadiums outside the Netherlands might be situated in the middle of residential areas, whereas Dutch stadiums are often located on the edge of cities. Foreign stadiums might therefore have a bigger effect on house prices in other countries and have an insignificant effect in the Netherlands. Moreover, some regions in the Netherlands are very densely populated, which can lead to different outcomes as well.

Furthermore, reverse causality can play a role in the models. As indicated in the Literature review, there is an ongoing discussion about whether jobs follow people or people follow jobs (Hoogstra, Van Dijk, & Florax, 2017; Partridge & Rickman, 2003; Steinnes, 1978). It is possible that a football stadium attracts people who will start living in that area, but there is also a possibility that the stadium has been built in places where already a lot of people lived. This bias can result in coefficients that are different in size or significance than what they are supposed to be.

However, just like the literature reported, stadiums do seem to influence house prices in the municipalities. Another similarity between the articles cited and this thesis can be found in the stadiums under construction. In the master thesis by Bieze (2021), it was found that surrounding house prices decreased during the construction of a stadium, and our significant coefficient for stadium under construction in the second model of the NVM dataset confirms this finding. Nonetheless, the local increase in house prices that Bieze (2021) found after the stadiums were finished, is not present in our analysis. Concerning the splitting of the stadium effects, the literature suggests that it might be possible, and this has also been shown in this thesis. However, as the results turned out to be counter intuitive, further research is needed to disentangle hypothesised positive and negative externalities.

Where for first part of this Synthesis quite a lot of differences between the prevailing literature and the thesis are reported, more similarities are found when it comes to the stadium effects on the number of jobs in an area. The prevailing literature reveals that football stadiums have no effect on local economies (Coates & Humphreys, 2000; Coates & Humphreys, 2003; Harger, Humphreys, & Ross, 2016). Accordingly, the fourth analysis with the LISA dataset suggests that a football stadium in the Netherlands has no significant effect on the number of jobs in the municipality where it is located. Nonetheless, it should be noted that the setup of the research differs from the existing literature. Where the discussed papers focus on local economies in general, the number of new businesses founded and inhabitants' income, this thesis uses the number of jobs as a dependent variable. It might be possible that a football stadium does not lead to more businesses being opened or to higher incomes in a municipality, but that existing businesses can expand and thrive due to the presence of the stadium. Therefore, there are more jobs available despite not influencing the other factors of the local economy. This can be a reason for the generally positive, although insignificant, effect in the models. Because of that, a broad overview including all different aspects of stadium effect might be necessary to be able to draw a clear conclusion.

The results of the last part of the analysis are also in line with the expectations that were created by the existing literature: football stadiums have arguably no effect on the composition of jobs within a municipality. Whether it involves sports-related, business-related or network jobs, the data gives no cause to assume that a football stadium influences a specific sector. As reasoned above, there is a chance that they do not increase the number of businesses or income in general, but that existing companies benefit from a stadium. This would lead to a general increase in the job market, keeping the job composition stable. All in all, the results in this thesis when it comes to local economies are in line with the existing literature.

Again, one thing that should be noted for the last models is the issue of reverse causality. The debate of whether jobs follow people or people follow jobs specifically pertains to these models. There is no consensus on whether a stadium as an amenity attracts jobs to an area, or whether the stadiums are built wherever there is a large job market. Intuition suggests the latter is at least partly true, so caution is required when interpreting the size and significance of the coefficients in the models presented. Moreover, omitted variable bias is also likely to play a role. This will be further explored in the Limitations section.

Concluding remarks

Conclusion

In this thesis, the external effects of football stadiums as consumption and production amenities in urban regions is investigated. In terms of consumption amenities, house prices were collected, whereas for the stadium as production amenity, the total amount and composition of jobs were obtained. The prevailing literature suggests that football stadiums have a local and urban effect on house prices within a range of around 3-5 km. Also, it is expected that the stadium effect on house prices could be split into positive and negative externalities. Lastly, the literature suggests that there is no effect of football stadiums on local economies, meaning that there is no relationship between the amount and composition of jobs within an area.

The results found in this thesis show some similarities with the literature, but also major differences are found. In answer to the first sub question about whether people are willing to pay extra to live close to a stadium, a significant positive coefficient of football stadiums on house prices is found on the level of municipalities. Therefore, the results point to a potential positive effect of football stadiums on house prices. However, whether a football stadium causes the house prices to be higher remains debatable. Due to issues like reversed causality, it cannot be stated with certainty that football stadiums generally increase house prices in a municipality.

On the contrary, no significant coefficient has been found for adjacent neighbourhoods and the neighbourhood itself. The results suggest that the local stadium effect is generally positive due to the positive sign of the coefficients, but their insignificance make it impossible to state that there is an effect with a high degree of certainty. Therefore, concerning the second sub question about the extent of the stadium effect, it must be noted that for the Dutch market, there are indications that there might be a positive effect, but there is too much uncertainty yet. For the first two sub questions, additional research is needed to arrive at a clear conclusion. When looking at the third sub question regarding the splitting of the stadium effect, the results pointed towards a possibility to do so, but not in the way that this thesis tried to accomplish. It proved to be rather hard to split the effect correctly into a positive and negative effect, mainly due to difficulties with measuring the respective externalities. Further research is needed to potentially achieve the splitting of the stadium effect.

Finally, the results, in accordance with the prevailing literature, imply that a football stadium has no significant effect on the total amount of jobs in a municipality. The positive coefficient points to a generally positive effect, but this cannot be supported with enough degrees of certainty. When it comes to job composition, the results also indicate that there is no stadium effects. This is the same for sports- and business-related jobs, as well as for jobs specifically involving networking. However, it should be noted again that when interpretating the coefficients in the fourth analysis, causation remains questionable due to the possible issue of reversed causality amongst other. Once more, additional research is required to make well-founded statements.

In conclusion, professional football stadiums in the Netherlands are related to several external effects as production and consumption amenities. Football stadiums are associated with higher house prices within a municipality, but it is too premature to talk about causation. In the data, indications for positive local effects are found as well, but these cannot be supported with a sufficient level of certainty. Moreover, the results imply that there is no significant effect of football stadiums on the total amount and composition of jobs within a municipality. Also, an attempt has been made to split the stadium effect into positive and negative externalities, but this was not successful. No clear distinction with accurate results was found, but there are indications that it should be possible when applying the right methods. Therefore, future research is needed.

However, it should also be noted that causal interpretations are not possible, as the models might suffer from several problems like reverse causality. It remains questionable whether stadiums can increase house prices and attract jobs, or whether stadiums are built in places where a lot of people already live, or where there are many jobs and thriving businesses. In addition to that, the modifiable areal unit problem and omitted variable bias are expected to play a crucial role. They will be expanded on in the next section.

Limitations

As described, the results found in this research sometimes differ substantially from the expectations risen by the prevailing literature. This could be due to some limitations. First, the way in which the potential areas on which a football stadium has an effect were chosen, might lead to differing results. Instead of looking into various rings with selected radiuses of stadium effects, the neighbourhood of the stadium itself, adjacent neighbourhood and the municipality was chosen for the KWB 2020 dataset. This might lead to atypical results, as it was often seen that a neighbourhood in which the stadium was located, did not contain a lot of houses. More often, the neighbourhood would consist of the stadium only, or included mostly industries or nature. The direct effect of the football stadium on house prices in its own neighbourhood is thus impossible to establish sometimes. Also, a fair share of Dutch stadiums has been built on the border of municipalities or cities, whereas some are in the middle of residential areas. This might lead to unequal comparisons between stadiums.

In addition to that, the way in which adjacent neighbourhoods had to be selected proved to be rather difficult as well. Despite the selection criteria that were formed, it might be possible that the selection was not always consistent, and some stadiums simply were surrounded by way more houses than other stadiums. Even though the same results were found when deploying a different method of appointing adjacent neighbourhoods, there is room for improvement. Whereas the first method by manually picking the neighbourhood might be inconsistent, the second method by using the radius might include areas without houses or areas that are clearly separated from stadiums due to natural borders. For future research, it would be best to use a combination of both methods. At first, construct rings around the stadium, after which a manual selection should be made based on a vast set of criteria. In combination with the paragraph above, it might be best in general to study each stadium as a different case with methods specifically adjusted to each stadium.

Also, there could be too little observations for neighbourhoods with stadiums compared to the large number of total neighbourhoods. This could lead to an insignificant effect of stadiums in both the neighbourhood itself and the adjacent neighbourhoods. As the number of municipalities with a stadium relative to total municipalities was rather large, this is possibly a reason that there was an effect found for stadiums in municipalities. For future research, it might be better to, for each neighbourhood with a stadium, pick a counterfactual neighbourhood that is almost the same, except for the presence of the stadium. In that way, the stadium effect is properly isolated from other potential factors that influence the house prices.

Moreover, the place in the neighbourhood at which the stadium is located might make a big difference. If the stadium is located at the far left of a large neighbourhood, it might have a different effect on the houses close by, than on houses located at the far right of that same neighbourhood. For other stadiums, these houses on the right could have counted as adjacent neighbourhoods, or in extreme cases even could not have been considered at all. However, they are included for this stadium, as they are located in the same neighbourhood and cannot be excluded. Also, the location of a stadium within a city might make a difference. Houses in the city centre are generally more expensive than houses on the border of cities. As stadiums are often located on the edge or even outside of the city, this could lead to lower house prices in neighbourhoods with a stadium compared to other neighbourhoods. However, this effect is thus not solely due to the presence of a stadium, but also because of the location of the neighbourhood within the city.

In addition to that, neighbourhoods have different sizes in each municipality. It could thus be that for example for the Ajax stadium in Amsterdam, the size of the neighbourhood itself and adjacent neighbourhoods amounted to a range of 2km, while for the RKC stadium in Waalwijk, the range totalled up to 4km. This makes the comparisons rather odd, and combined with the insights above, these factors could have led to the insignificant effects that were found. On top of that, different stadiums might bring different effects, as discussed in the Synthesis. However, by researching the general stadium effects, this has not been considered. Again, this suggests that for future research, case studies seem important to be taken into consideration.

Most of the limitations above are part of one bigger issue: the modifiable areal unit problem (MAUP). This problem arises when data is aggregated to certain areas with artificial borders, like the neighbourhoods or municipalities in this thesis. Fotheringham and Wong (1991) indicate that the MAUP negatively affects the reliability of the results in all kinds of studies involving aggregated geographical data. They have even found that in multivariate analysis, different zoning techniques lead to different results in terms of size and significance. The MAUP is therefore expected to be a problem in this thesis as well. The neighbourhood in which a stadium is located is demarcated by artificial borders, and drawing these borders differently can potentially lead to different results. Also, the place of a stadium in a neighbourhood, as discussed in the previous paragraph, is determined by these borders, just like the number and types of houses, amenities, sceneries and many more factors that potentially influence housing prices or stadium effects. It is possible that when drawing the borders differently, the stadium's neighbourhoods contain different characteristics leading to different results.

Also, the dependent variable is a limitation as well. For average property value, the WOZ-waarde was taken, but the WOZ is not limited to houses and includes business establishments and other properties. Although it has been tried to limit this by manually selecting adjacent neighbourhoods, it cannot be ruled out that this has played a role, especially since the neighbourhoods with stadiums often did not include many houses. Additionally, the WOZ-waarde is the value of a property according to the law of the WOZ, and does not necessarily equal the transaction price, which would be more accurate. Also, WOZ-waarde per m^2 would have controlled for the size of houses, as this might differ substantially depending on the area they are located. However, this was not available. Therefore, it has been tried to solve this by using a wide range of control variables, but it is not certain that it eliminated the effect.

Therefore, the NVM dataset provided a more accurate dependent variable, as transaction prices accurately project the prices people were willing to pay for a house on the market. Also, it included control variables for the surface and volume of the house, as well as for other characteristics. However, the data was not available at the level of neighbourhoods, and it did not include as many control variables as the KWB 2020 dataset. As neighbourhood characteristics could not be considered, different effects might have played a role that were not taken into account. Consequently, omitted variable bias is likely to have played a role.

For the LISA dataset, a lot of control variables could be added and the number of jobs was a clear dependent variable. However, especially for the business-related jobs, the selection of SBI-codes might be too arbitrary, leading to insignificant results. Also, the effects of stadiums on the composition of jobs can be more accurately researched by using four- or five-digit SBI codes instead of the three-digit ones that were available now. Lastly, as discussed before, reverse causality might have played a major role, specifically in these models.

In general, a limitation for all three datasets and analyses is the potential omitted variables bias that could play a role in the regressions. The number of jobs in an area and, maybe even more so, house prices are the results of a very large number of factors. Some can be very obvious, but not all determinants might be clear, or some could even be unknown. It is therefore possible that important variables have been left out. As a result, the independent variables potentially correlate with the error term, leading to biased estimators, as the zero-conditional man assumption of OLS is violated. In addition to that, reverse causality might be a problem in all models, especially for models using the KWB (2017 and 2020) and LISA datasets. Amenities might be a good example. Do more amenities result in higher house prices or more jobs, or are there more amenities in an area with a lot of jobs, where people live and increase house prices for that reason? Due to this issue, the estimators might be biased. Lastly, the coronavirus pandemic might have led to unusual values for certain variables, which could have affected the results. This mainly pertains to the NVM dataset, as this dataset included the years 2020 and 2021. The KWB 2020 dataset is mostly unaffected by the pandemic, as the reference date for all variables is January 1, 2020. At this time, the covid pandemic had not yet reached the Netherlands.

Recommendations

A suggestion for future research is to collect a lot of additional data, making the analyses more accurate. In the best case, transaction data on the neighbourhood level would be taken, and there should be a lot of variables to control for. Basically, a combination of the transaction price and housing characteristics of the NVM dataset, this time available on neighbourhood levels, combined with the neighbourhood characteristics in the KWB 2020 dataset seems like a big step in the right direction. Another suggestion is the use of case studies, or to at least look at each stadium in more depth.

These case studies might also reduce the impact of the MAUP. Dark and Bram (2007) give an overview of multiple solutions for the MAUP, and conclude that it a different approach might be required for each study, depending on the topic and analysis. Their list of solutions includes simply ignoring the problem (based on Openshaw (1984)), or using object-specific spatial measures to downplay the MAUP (based on Hay, Marceau, Dube and Bouchard (2001)). Also, Jelinski & Wu (1996) are mentioned that propose the analysis of individual entities. For the stadium effect, case studies might provide the best solution in combination with using the solutions mentioned above. A study that focuses on each stadium specifically and analyses the surroundings of each stadium, might lead to the most accurate results. For each stadium, the possible extent of the effects can be determined and an optimal zoning strategy can be applied by closely studying the surrounding areas and object around football stadiums.

Furthermore, different methodologies might be used to solve potential problems like omitted variable bias or reversed causality. Over the past years, there is little variation in stadiums and their locations. It has been a relatively long time ago since the last big projects were realised, and changes in location are rare. However, new stadiums being built are crucial for the research, as it comes closer to an experiment than the current method. Whenever a stadium is built in a neighbourhood or municipality, you are able to analyse the change in house prices more accurately. More concrete, it is easier to test whether house prices have gone up since the arrival of the new stadium. Mainly the set-up of a repeated sales method is likely to give accurate results, but also a regression continuity design could be helpful. Lastly, an instrumental variable might be a good suggestion. A different methodology for the splitting of the effect is required as well. Different proxy variables may be used for the positive externalities, whilst a different approach is advised for the negative externalities. Instead of only computing the negative articles for the municipalities with football stadiums, a full media analysis that focuses on the sentiment of each article for all municipalities is suggested. In this way, the negative externalities can more accurately be estimated and the stadium effect can be singled out.

Also, when investigating the external effects of stadiums, all potential effects should be considered. Whereas a lot of academic papers focus solely on house prices or local economies, an overview of effects ranging from house prices to jobs, to intangible benefits and many more effects should all be considered to arrive at a definitive conclusion about the external effects of football stadiums. A general overview of all external effects would provide very meaningful insight of how a football stadium influences the areas around it. This study can then be extended to multiple countries, which presents even more insight into the different stadium effects.

In terms of practical applications, both the literature and the empirical analysis suggest that amenities, specifically stadiums in this thesis, result in higher house prices. Local policy should be aimed at attracting people by means of amenities like football stadiums. From this follows that it is important that football clubs receive funding to ensure that they build proper stadiums. However, the results offer no certainty yet. Especially in the Netherlands, additional research is needed on the size and scope of stadium effects. Specifically, investigating the positive and negative externalities is expected to be crucial, as it is likely that dealing with them largely determines the size of the stadium effects. Lastly, it is implied that football stadiums have no effect on local economies, so based on the current results, they should not be used by the local authorities as tools to attract new companies or create more jobs.

Bibliography

Ahlfeldt, G. M., & Kavetsos, G. (2014). Form or function?: the effect of new sports stadia on property prices in London. *Journal of the Royal Statistical Society: series A (statistics in society)*, *177*(1), 169-190. https://doi.org/10.1111/rssa.12006

Ahlfeldt, G. M., & Maennig, W. (2010). Impact of sports arenas on land values: evidence from Berlin. *The Annals of Regional Science*, *44*(2), 205-227. https://doi.org/10.1007/s00168-008-0249-4

Ahlfeldt, G. M., & Maennig, W. (2012). Voting on a NIMBY facility: Proximity cost of an "iconic" stadium. *Urban Affairs Review*, *48*(2), 205-237. https://doi.org/10.1177/1078087411423644

Baade, R. A., & Dye, R. F. (1988). An analysis of the economic rationale for public subsidization of sports stadiums. *The Annals of Regional Science*, *22*(2), 37-47. https://doi.org/10.1007/BF01287322

Bartik, T. J., & Smith, V. K. (1987). Urban amenities and public policy. In E.S. Mills (ed.), *Handbook of regional and urban economics* (pp. 1207-1254). Amsterdam: Elsevier. https://doi.org/10.1016/S1574-0080(87)80017-2

Bieze, G. (2021). *Stadium development in the Netherlands: the effect on surrounding house prices* [Master's thesis, Rijksuniversiteit Groningen]. Student Theses Faculty Spatial Sciences. https://frw.studenttheses.ub.rug.nl/id/eprint/3713

Brueckner, J. K., Thisse, J. F., & Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European economic review*, *43*(1), 91-107. https://doi.org/10.1016/S0014-2921(98)00019-1

Carlino, G., & Coulson, N. E. (2004). Compensating differentials and the social benefits of the NFL. *Journal of Urban Economics*, *56*(1), 25-50. https://doi.org/10.1016/j.jue.2004.03.001

Cheshire, P., & Sheppard, S. (1995). On the Price of Land and the Value of Amenities. *Economica, 62*(246), 247–267. https://doi.org/10.2307/2554906

Churchman, C. (1995). Sports Stadia and the Landscape: A Review of the Impacts and Opportunities arising as a Result of the Current Redevelopment of Football Grounds. *Built Environment (1978-)*, *21*(1), 6-24. https://www.jstor.org/stable/23287711

Coates, D. (2007). Stadiums and arenas: Economic development or economic redistribution? *Contemporary economic policy, 25*(4), 565-577. https://doi.org/10.1111/j.1465-7287.2007.00073.x

Coates, D., & Humphreys, B. R. (2000). The stadium gambit and local economic development. *Regulation, 23*(2), 15-20. https://heinonline.org/HOL/P?h=hein.journals/rcatorbg23&i=93

Coates, D., & Humphreys, B. R. (2003). Professional sports facilities, franchises and urban economic development. *Public Finance and Management*, *3*(3), 335-357. https://www.researchgate.net/publication/23693796_Professional_Sports_Facilities_Franchises_and_Urban_Economic_Development

Cooke, T. W. (1978). Causality reconsidered: a note. *Journal of Urban Economics*, *5*(4), 538-542. https://doi.org/10.1016/0094-1190(78)90009-8

Corney, H., Ives, C. D., & Bekessy, S. (2015). Amenity and ecological management: A framework for policy and practice. *Ecological Management & Restoration*, *16*(3), 199-205. https://doi.org/10.1111/emr.12177

Dark, S. J., & Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, *31*(5), 471-479. https://doi.org/10.1177/0309133307083294

Davies, L. E. (2005). Not in my back yard! Sports stadia location and the property market. *Area*, *37*(3), 268-276. https://doi.org/10.1111/j.1475-4762.2005.00630.x

Feng, X., & Humphreys, B. R. (2018). Assessing the economic impact of sports facilities on residential property values: A spatial hedonic approach. *Journal of Sports Economics*, *19*(2), 188-210. https://doi.org/10.1177/1527002515622318

Fischer, C., & Hamm, R. (2019). Football clubs and regional image. *Review of Regional Research*, *39*(1), 1-23. https://doi.org/10.1007/s10037-018-00129-5

Fotheringham, A. S., & Wong, D. W. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and planning A*, *23*(7), 1025-1044. https://doi.org/10.1068/a231025

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of economic geography*, *1*(1), 27-50. https://doi.org/10.1093/jeg/1.1.27

Harger, K., Humphreys, B. R., & Ross, A. (2016). Do new sports facilities attract new businesses? *Journal of Sports Economics, 17*(5), 483-500. https://doi.org/10.1177/1527002516641168

Hay, G. J., Marceau, D. J., Dube, P., & Bouchard, A. (2001). A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology*, *16*(6), 471-490. https://doi.org/10.1023/A:1013101931793

Hoogstra, G. J., Van Dijk, J., & Florax, R. J. G. M. (2017). Do jobs follow people or people follow jobs? A meta-analysis of Carlino–Mills studies. *Spatial Economic Analysis*, *12*(4), 357-378. https://doi.org/10.1080/17421772.2017.1340663

Huang, H., & Humphreys, B. R. (2014). New sports facilities and residential housing markets. *Journal of Regional Science*, *54*(4), 629-663. https://doi.org/10.1111/jors.12120

Hyun, D. (2021). Proud of, but too close: the negative externalities of a new sports stadium in an urban residential area. *The Annals of Regional Science*, *68*(3), 615-633. https://doi.org/10.1007/s00168-021-01095-6

Jelinski, D. E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape ecology*, *11*(3), 129-140. https://doi.org/10.1007/BF02447512

Johnson, B. K., & Whitehead, J. C. (2000). Value of public goods from sports stadiums: The CVM approach. *Contemporary economic policy*, *18*(1), 48-58. https://doi.org/10.1111/j.1465-7287.2000.tb00005.x

Koster, H. R., & Rouwendal, J. (2017). Historic amenities and housing externalities: evidence from the Netherlands. *The Economic Journal*, *127*(605), F396–F420. https://doi.org/10.1111/ecoj.12477

KPMG. (2003). *De gemeente als twaalfde man.* http://docplayer.nl/27679549-De-gemeente-als-twaalfde-man.html

Li, M. M., & Brown, H. J. (1980). Micro-neighborhood externalities and hedonic housing prices. *Land economics*, *56*(2), 125-141. https://doi.org/10.2307/3145857

Livy, M. R. (2017). The effect of local amenities on house price appreciation amid market shocks: The case of school quality. *Journal of Housing Economics*, *36*, 62-72. https://doi.org/10.1016/j.jhe.2017.02.003

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Landscape and urban planning*, *48*(3-4), 161-167. https://doi.org/10.1016/S0169-2046(00)00039-6

Merriam-Webster. (n.d.). Amenity. In *Merriam-Webster.com dictionary*. Retrieved July 25, 2022, from https://www.merriam-webster.com/dictionary/amenity

Mulligan, G. F., & Carruthers, J. I. (2011). Amenities, quality of life, and regional development. In R. W. Marans, & R. J. Stimson (Eds.), *Investigating quality of urban life* (pp. 107-133). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-1742-8_5

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment and planning A*, *16*(1), 17-31. https://doi.org/10.1068/a160017

Paramio, J. L., Buraimo, B., & Campos, C. (2008). From modern to postmodern: The development of football stadia in Europe. *Sport in society*, *11*(5), 517-534. https://doi.org/10.1080/17430430802196520

Partridge, M. D., & Rickman, D. S. (2003). The waxing and waning of regional economies: The chicken–egg question of jobs versus people. *Journal of Urban Economics*, *53*(1), 76-97. https://doi.org/10.1016/S0094-1190(02)00501-6

Porter, M. E. (1995). The competitive advantage of the inner city. *Harvard Business Review*, *73*(3), 55-71. http://dx.doi.org/10.4324/9780429261732-38

PWC. (2021). *De maatschappelijke impact van betaald voetbal in Nederland.* https://www.knvbexpertise.nl/downloads/bestand/354/de-maatschappelijke-impact-van-betaald-voetbal

Pyun, H., Kim, J., Schlesinger, T., & Matto, L. (2020). Positive externalities from professional football clubs in the Metropolitan Rhine-Ruhr, Germany: Trickle-down effects associated with promotion and relegation. *Sustainability*, *12*(20), 8638. https://doi.org/10.3390/su12208638

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? *Journal of targeting, measurement and analysis for marketing*, *17*(2), 139-142. https://doi.org/10.1057/jt.2009.5

Roback, J. A. (1980). *The value of local urban amenities; theory and measurement* (Publication No. 8028928) [Doctoral dissertation, University of Rochester]. ProQuest Dissertations & Theses Global. https://www.proquest.com/dissertations-theses/value-local-urban-amenities-theory-measurement/docview/303092038/se-2

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), 34-55. https://doi.org/10.1086/260169

Statista. (2019). *Sports with the most followers through media and events in the Netherlands in 2019*. https://www.statista.com/statistics/1172147/most-popular-sports-in-media-and-events-in-the-netherlands/

Steinnes, D. N. (1978). Causality and migration: A statistical resolution of the "chicken or egg fowl-up". *Southern Economic Journal*, *45*(1), 218-226. https://doi.org/10.2307/1057628

Stone, C. (2007). The role of football in everyday life. *Soccer & society*, *8*(2-3), 169-184. https://doi.org/10.1080/14660970701224319

Tu, C. C. (2005). How does a new sports stadium affect housing values? The case of FedEx field. *Land Economics*, *81*(3), 379-395. http://dx.doi.org/10.3368/le.81.3.379

Ullman, E. L. (1954). Amenities as a factor in regional growth. *Geographical Review*, *44*(1), 119-132. https://doi.org/10.2307/211789

Van Dam, F. (2000). Refurbishment, redevelopment or relocation? The changing form and location of football stadiums in the Netherlands. *Area*, *32*(2), 133-143. https://doi.org/10.1111/j.1475-4762.2000.tb00124.x

Van Oort, F., Weterings, A., & Verlinde, H. (2003). Residential amenities of knowledge workers and the location of ICT-Firms in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie*, *94*(4), 516-523. http://dx.doi.org/10.1111/1467-9663.00278

Waltert, F., & Schläpfer, F. (2010). Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies. *Ecological Economics*, *70*(2), 141-152. https://doi.org/10.1016/j.ecolecon.2010.09.031

Wu, J., Adams, R. M., & Plantinga, A. J. (2004). Amenities in an urban equilibrium model: Residential development in Portland, Oregon. *Land Economics*, *80*(1), 19-32. http://dx.doi.org/10.2307/3147142

Zawadzki, K. M. (2021). Social perception of technological innovations at sports facilities: justification for financing ‘white elephants’ from public sources? The case of Euro 2012 Stadiums in Poland. *Innovation: The European Journal of Social Science Research*, *35*(2), 346-366. https://doi.org/10.1080/13511610.2021.1937070

Zhang, S., van Duijn, M., & van der Vlist, A. J. (2020). The external effects of inner-city shopping centers: Evidence from the Netherlands. *Journal of Regional Science*, *60*(4), 583-611. https://doi.org/10.1111/jors.12473

Appendices

Appendix A – Stadium data

Table A.1: Football stadiums of the Eredivisie teams of the 2020/2021 season

Nr	Stadium	Club	City	Start of construction	Opening	After 1986	New location	Capacity
1	Johan Cruijff ArenA	Ajax	Amsterdam	1993	1996	Yes	Yes	55,600
2	Stadion Feijenoord "De Kuip"	Feyenoord	Rotterdam	1935	1937	No	-	47,500
3	Philips Stadion	PSV	Eindhoven	1910	1910	No	-	35,000
4	De Grolsch Veste	FC Twente	Enschede	1997	1998	Yes	Yes	30,205
5	Abe Lenstra Stadion	sc Heerenveen	Heerenveen	1993	1994	Yes	Yes	27,224
6	Stadion Galgenwaard	FC Utrecht	Utrecht	1936	1936	No	-	23,750
7	Euroborg	FC Groningen	Groningen	2004	2006	Yes	Yes	22,555
8	GelreDome	Vitesse	Arnhem	1996	1998	Yes	Yes	21,248
9	AFAS Stadion	AZ	Alkmaar	2004	2006	Yes	Yes	19,478
10	Cars Jeans Stadion	ADO Den Haag	Den Haag	2005	2007	Yes	Yes	15,000
11	Koning Willem II Stadion	Willem II	Tilburg	1993	1995	Yes	No	14,700
12	MAC³PARK stadion	PEC Zwolle	Zwolle	2007	2009	Yes	No	14,000
13	Fortuna Sittard Stadion	Fortuna Sittard	Sittard	1999	1999	Yes	Yes	12,500
14	Erve Asito	Heracles Almelo	Almelo	1999	1999	Yes	Yes	12,080
15	Sparta-stadion Het Kasteel	Sparta Rotterdam	Rotterdam	1916	1916	No	-	11,026
16	De Oude Meerdijk	FC Emmen	Emmen	1977	1977	No	-	8,600
17	Covebo Stadion - De Koel	VVV-Venlo	Venlo	1971	1972	No	-	8,000
18	Mandemakers Stadion	RKC Waalwijk	Waalwijk	1996	1996	Yes	No	7,508

The capacity numbers are retrieved from *Transfermarkt.nl*, specifically from the websites https://www.transfermarkt.nl/eredivisie/stadien/wettbewerb/NL1 and https://www.transfermarkt.nl/keuken-kampioen-divisie/stadien/wettbewerb/NL2/plus/. The years of opening and start of construction are found in newspaper articles accessed through *Nexis Uni*, or found on Wikipedia or specific websites related to the stadiums or their football clubs. The links match with the numbers of the stadiums in the table:

1. https://advance.lexis.com/document/?pdmfid=1516831&crid=b7c023e8-dc54-4b04-aa12-7c6f33e43899&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A63C5-N3B1-JBKF-M13R-00000-00&pdcontentcomponentid=316966&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=417c8b9d-79e0-487e-901c-c2a7ccd14ca4
2. https://www.dekuip.nl/stadion-feijenoord/historie#:~:text=Tien%20maanden%20na%20het%20slaan,op%2023%20juli%201936%20opgeleverd.
3. https://nl.wikipedia.org/wiki/Philips_Stadion
4. https://advance.lexis.com/document/?pdmfid=1516831&crid=d64ed4fb-a777-44a2-b319-f4a1e9600a11&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3TK6-G1W0-009Y-90BT-00000-00&pdcontentcomponentid=149018&pdteaserkey=sr1&pditab=allpods&ecomp=szznk&earg=sr1&prid=93a02264-b184-41dd-b2ce-366fc9a070b7 & https://advance.lexis.com/document/?pdmfid=1516831&crid=ebce0f4f-5baa-4530-b881-8dc14b829b1b&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A48MV-YM30-0150-X1HS-00000-00&pdcontentcomponentid=294298&pdteaserkey=sr9&pditab=allpods&ecomp=szznk&earg=sr9&prid=55ed8876-c7c4-4753-b3ec-cb87ba694107
5. https://advance.lexis.com/document/?pdmfid=1516831&crid=72f84df0-ef7c-4d4c-9156-cb8f0b710682&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A48MV-S4Y0-0150-X29X-00000-00&pdcontentcomponentid=294298&pdteaserkey=sr3&pditab=allpods&ecomp=szznk&earg=sr3&prid=a98f1146-0ab8-4e5a-88a7-0478dc119ac0

6. https://advance.lexis.com/document/?pdmfid=1516831&crid=3f1ad1b5-cb48-40ca-a32b-0cb5b0069381&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A5WTC-7261-JC8X-608D-00000-00&pdcontentcomponentid=294313&pdteaserkey=sr6&pditab=allpods&ecomp=szznk&earg=sr6&prid=254835bc-973e-4b8a-95b9-ad46851062a0
7. https://advance.lexis.com/document/?pdmfid=1516831&crid=5fdcf755-abd2-45bf-ad20-70ce32e4f625&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4HH8-VGH0-TWS7-139J-00000-00&pdcontentcomponentid=277856&pdteaserkey=sr1&pditab=allpods&ecomp=szznk&earg=sr1&prid=08be9dc8-6b22-480d-8f48-7b9d688fcc05 & https://advance.lexis.com/document/?pdmfid=1516831&crid=e033e5c1-8b85-48b7-8f71-3539de22b20f&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4K1P-4Y60-TX55-02WC-00000-00&pdcontentcomponentid=277856&pdteaserkey=sr3&pditab=allpods&ecomp=szznk&earg=sr3&prid=9c414b11-3091-4fc3-af6a-73de923d96df
8. https://advance.lexis.com/document/?pdmfid=1516831&crid=fa508646-7ade-41b6-80e2-24e4f6742f44&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A48MW-1650-0150-W0G1-00000-00&pdcontentcomponentid=259064&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=7e9a2fa1-c1e7-4361-b922-e59592b5857b & https://advance.lexis.com/document/?pdmfid=1516831&crid=7f613ee6-3a50-43ec-a6b7-f8298cab54cf&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3TDS-1W00-009Y-952K-00000-00&pdcontentcomponentid=149018&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=b2206330-8af0-4ff1-a5d1-f899b0e4ed07
9. https://advance.lexis.com/document/?pdmfid=1516831&crid=737287aa-c853-4c1b-802b-6987ba4cf03e&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4KK0-K860-TWS7-131N-00000-00&pdcontentcomponentid=277856&pdteaserkey=sr9&pditab=allpods&ecomp=szznk&earg=sr9&prid=7a162da5-850f-4c80-a936-87bad8456366 & https://advance.lexis.com/document/?pdmfid=1516831&crid=caa010f1-5b0f-453b-9f13-e78b1e8ce597&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4F29-7XM0-00J5-K1F8-00000-00&pdcontentcomponentid=168873&pdteaserkey=sr4&pditab=allpods&ecomp=szznk&earg=sr4&prid=681b6c3f-e94a-4d53-9902-d3d922277bd4
10. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=eee068a7-3c2b-4b2f-807f-e9b4a1c958df&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4P94-HJ90-TXJ3-V416-00000-00&pdcontentcomponentid=259064&pdteaserkey=sr0&pditab=allpods&ecomp=rz2yk&earg=sr0&prid=14b91bde-286d-429d-9455-df886f335799 & https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=3c5ab34c-93ec-4bb9-a11c-8d5e8ad24967&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A4HMT-C0K0-00J5-K3WW-00000-00&pdcontentcomponentid=168873&pdteaserkey=sr5&pditab=allpods&ecomp=rz2yk&earg=sr5&prid=4a9921fd-bcd6-402d-9730-238ac6372611
11. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=320febd6-6069-4d9b-9ec4-97a28d50b937&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A64B4-HV71-JC8X-63BK-00000-00&pdcontentcomponentid=149022&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=c256b109-66fa-4d1e-b1e3-e76190d31d69 & https://www.willem-ii.nl/stadion
12. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=2f2a1f62-35a7-414b-b5db-d71213f6d7ff&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A5WVH-9J61-DY0X-92N9-00000-00&pdcontentcomponentid=467102&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=3c252e44-11a7-461e-9081-fa765841b18c
13. https://www.fortunasittard.nl/fortuna-sittard-stadion/
14. https://www.heracles.nl/stadion/ & https://nos.nl/artikel/183561-nieuw-stadion-heracles-krijgt-vorm
15. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=d30de806-8d7d-4dfe-b3c8-be7ac82b0383&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A5KWK-VYJ1-JC8X-638N-00000-00&pdcontentcomponentid=294311&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=3d7ea5c4-77f1-4913-bb9b-c65b4e6d9301
16. https://nl.wikipedia.org/wiki/De_Oude_Meerdijk
17. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=d894a1b7-c876-4350-ad68-239af40d6b51&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A6510-K001-JCWP-93NR-00000-00&pdcontentcomponentid=475365&pdteaserkey=sr0&pditab=allpods&ecomp=rz2yk&earg=sr0&prid=fdef0179-1194-4734-87e9-af6a3c99061d & https://nl.wikipedia.org/wiki/Covebo_Stadion_-_De_Koel_-
18. https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=3a138891-4687-450c-9b7a-00af89253d62&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3SKN-2150-009Y-619J-00000-00&pdcontentcomponentid=149022&pdteaserkey=sr0&pditab=allpods&ecomp=szznk&earg=sr0&prid=01314248-c3a3-4ee7-b37b-f850cf9a6813 & https://advance-lexis-com.eur.idm.oclc.org/document/?pdmfid=1516831&crid=5654ef55-0417-45bd-a937-356ee8dcedcb&pddocfullpath=%2Fshared%2Fdocument%2Fnews%2Furn%3AcontentItem%3A3SKN-2CS0-009Y-63NF-00000-00&pdcontentcomponentid=149022&pdteaserkey=sr2&pditab=allpods&ecomp=szznk&earg=sr2&prid=87e8ca14-117c-4e50-89f9-52e3f2e4bd7f

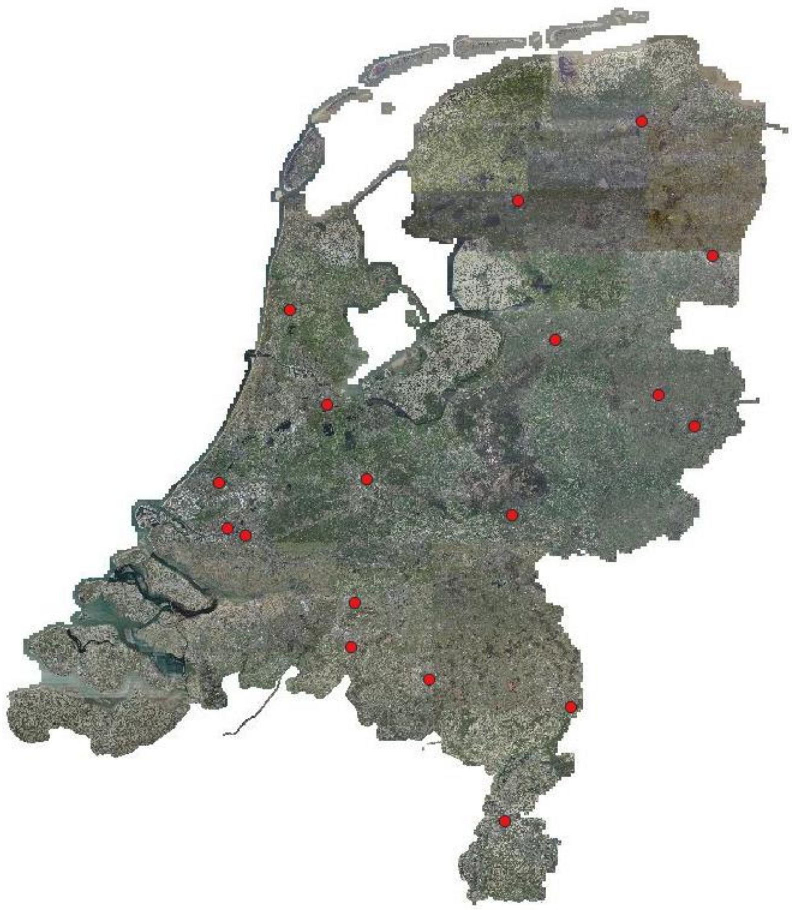

Figure A.1: Football stadiums of the Eredivisie teams of the 2020/2021 season on a map of the Netherlands

Appendix B – Neighbourhood selection

This appendix includes several examples of how the neighbourhoods were selected.

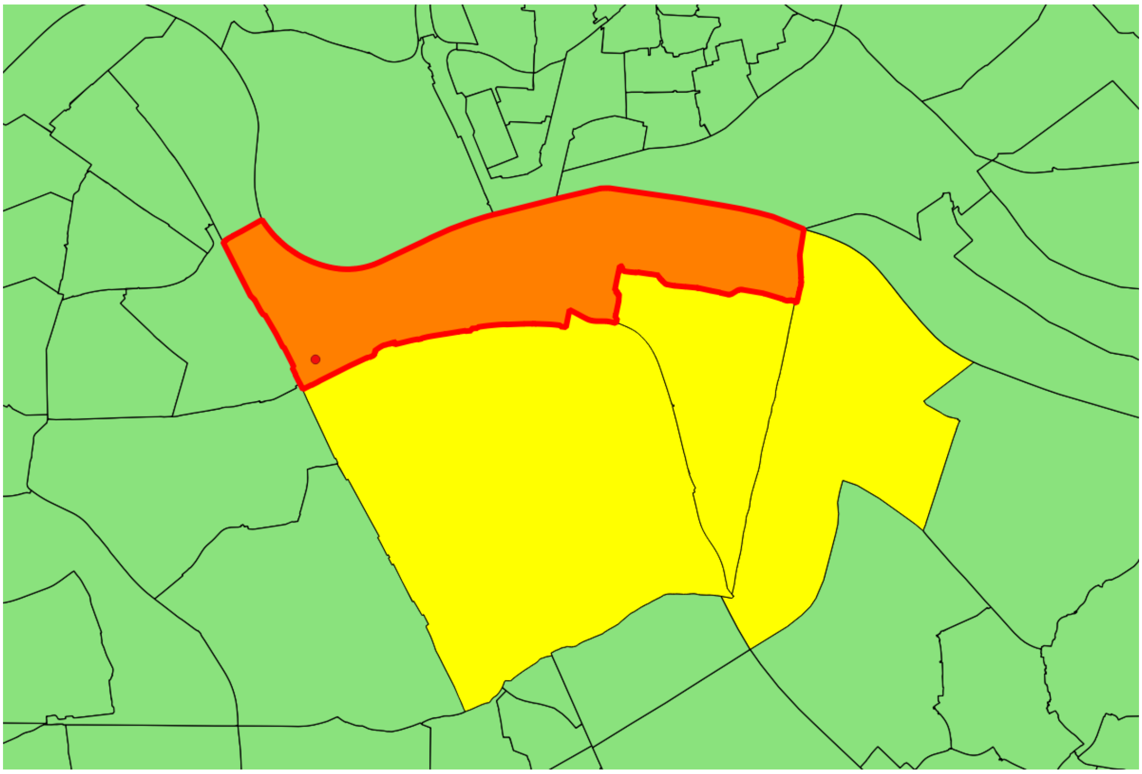

Figure B.1: Snapshot of the neighbourhoods surrounding Stadion Feijenoord "De Kuip" in Rotterdam

Figure B.1 shows the neighbourhood that accommodates Stadion Feijenoord "De Kuip" in red, with three yellow neighbourhoods used as an example. The stadium is located to the far left of the neighbourhood, meaning that it is actually quite far away from the yellow neighbourhood on the far left. Therefore, although the far right neighbourhood is bordering the red neighbourhood, it is not selected in our sample as it is too far away from the actual stadium, while the other two neighbourhoods are included as they are closer to the stadium.

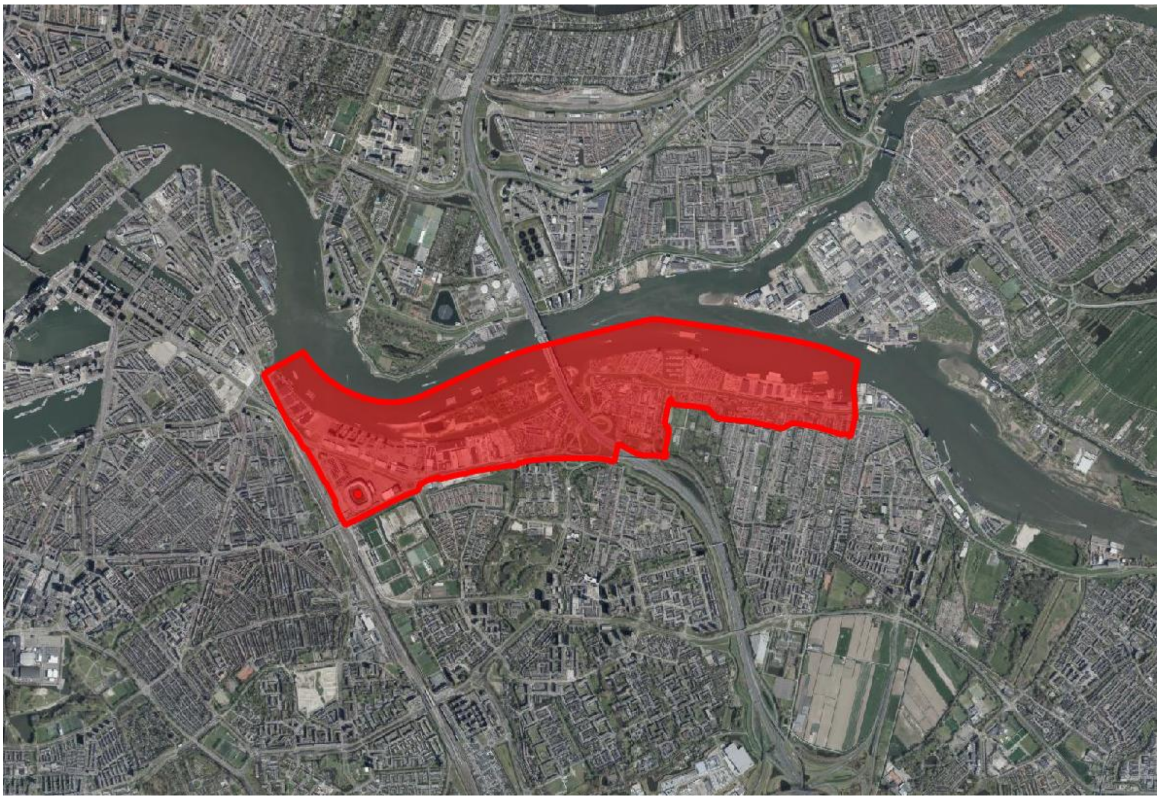

Figure B.2: Aerial photo of the neighbourhoods surrounding Stadion Feijenoord "De Kuip" in Rotterdam

Figure B.2 presents the same geographical location as Figure B.1, but now as an aerial photo. From the picture, it becomes clear that the north side of the neighbourhood that includes Stadion Feijenoord "De Kuip" is bordered by a river. This border is ought to be that big, that it significantly decreases the stadium effect on the northern adjacent areas, that these are not included in the sample, especially since there are not many bridges etc. The same problem was found for neighbourhoods that were separated by a train track, large highway or natural areas without (many) connections to the stadium, so these were also excluded.

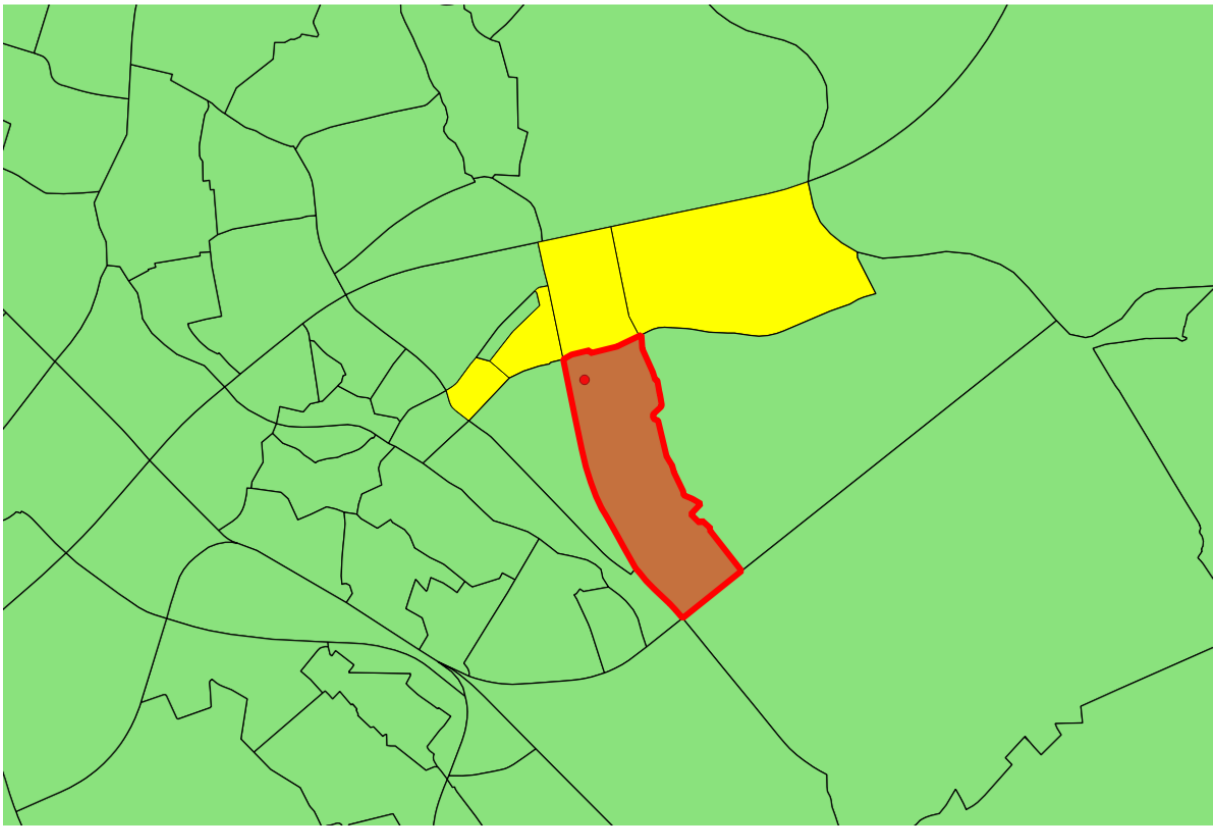

Figure B.3: Snapshot of the neighbourhoods surrounding the MAC³PARK stadion in Zwolle.

In this case, the yellow neighbourhood that is the furthest to the left is not adjacent to the red neighbourhood that includes the MAC³PARK stadion. However, as the yellow neighbourhood on the far right is included, the neighbourhood on the left should be included as well, based on its relative distance to the stadium compared to the neighbourhood on the right.

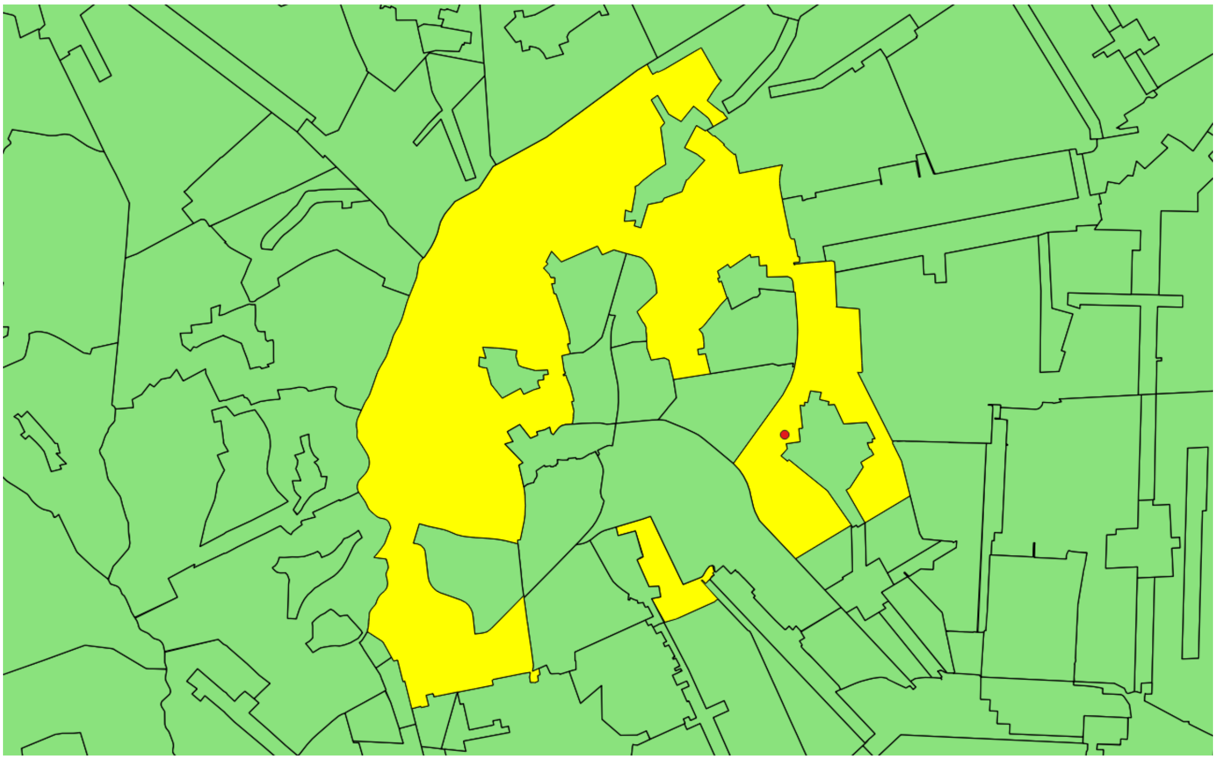

Figure B.4: Snapshot of the neighbourhoods surrounding De Oude Meerdijk in Emmen

Figure B.4 shows that the neighbourhood in which De Oude Meerdijk of FC Emmen is located, is scattered all over the municipality. Therefore, the effect on that neighbourhood is not representative as the stadium effect and is thus not included in the sample.

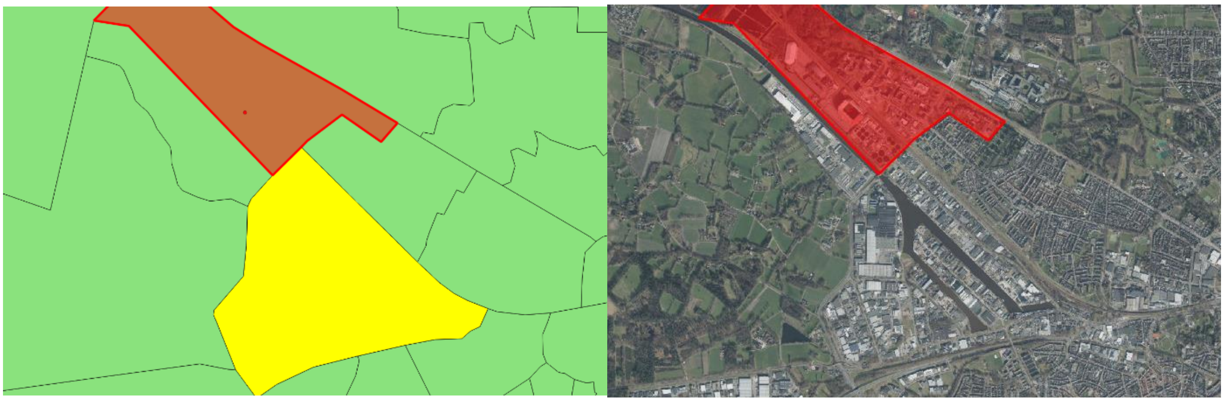

Figure B.5: Snapshot and aerial photo of the neighbourhoods surrounding De Grolsch Veste in Enschede

The yellow neighbourhood is adjacent to the red neighbourhood that hosts De Grolsch Veste, but as the aerial photo shows, the area is used as an industrial site. The name confirms this, as the neighbourhood is called ‘Industrie- en havengebied’, which translates to area for industry and port. As there are (almost) no houses in these area, the stadium effect on house prices will not be measured in these areas, but only the value of the company buildings. This is not included in the scope of the investigation, so the neighbourhood is excluded from the sample.

Appendix C – Description of variables

Table C.1: Description of variables for the originally used KWB 2020 dataset (CBS)

Variable	Description
Dependent variable	
Average property value	Average property value based on the 'WOZ waarde' in thousand euros.
Independent variables	
Neighbourhood with stadium	Dummy variable that takes the value '1' if there is a stadium present in the neighbourhood, and ‘0’ otherwise
Adjacent neighbourhood	Dummy variable that takes the value '1' if the neighbourhood is adjacent to a neighbourhood with a stadium, and ‘0’ otherwise. More information can be found in the Data section and in Appendix B.
Municipality with stadium	Dummy variables that takes the value '1' if there is a stadium present in the municipality, and '0' otherwise
Control variables	
Demographics	
Inhabitants	Amount of inhabitants
Male	Amount of male inhabitants
Female	Amount of female inhabitants
Age 0-14	Amount of inhabitants aged 0 until 15
Age 15-24	Amount of inhabitants aged 15 until 25
Age 25-44	Amount of inhabitants aged 25 until 45
Age 45-64	Amount of inhabitants aged 45 until 65
Age 65+	Amount of inhabitants aged 65 or older
Unmarried	Amount of inhabitants that are unmarried
Married	Amount of inhabitants that are married
Divorced	Amount of inhabitants that are divorced
Widowed	Amount of inhabitants that are widows
Western	Amount of inhabitants with a Western migration background
Non-Western	Amount of inhabitants with a non-Western migration background
Moroccan	Amount of Moroccan inhabitants
Antillean and Aruban	Amount of Antillean and Aruban inhabitants
Surinamese	Amount of Surinamese inhabitants
Turkish	Amount of Turkish inhabitants
Other non-Western	Amount of other non-Western inhabitants
Birth rate	Amount of births (total)
Death rate	Amount of deaths (total)
Households	Amount of households
Singe-person households	Amount of single-person households
Households without children	Amount of households without children
Households with children	Amount of households with children
Average household size	Average size of households

Housing market	
Housing stock	Housing stock within a neighbourhood
Single-family housing	Percentage of single-family homes. A single-family home is a house that is a separate building on its own.
Multi-family housing	Percentage of multi-family homes. A multi-family home is a home that is with multiple other homes or businesses in one building.
Inhabited	Percentage of houses that are inhabited
Uninhabited	Percentage of houses that are uninhabited
Owner-occupied houses	Percentage of houses that are owner-occupied houses
Rental houses	Percentage of houses that are rental houses
In possession of housing associations	Percentage of houses owned by housing associations
In possession of other landlords	Percentage of houses owned by other landlords
Ownership unknown	Percentage of houses of which the ownership is unknown
Year of construction before 2000	Percentage of houses with the year of construction before 2000
Year of construction after 2000	Percentage of houses with the year of construction from 2000 onwards
Education	
Low-skilled	Amount of people that are low-skilled
Medium-skilled	Amount of people that are medium-skilled
High-skilled	Amount of people that are high-skilled
Business	
Business establishments	Amount of business establishments
Amenities	
Distance to general practice	Average road distance of all inhabitants to the nearest general practice
Distance to big supermarket	Average road distance of all inhabitants to the nearest big supermarket. A big supermarket should have multiple types of daily products and a surface of at least 150 m^2.
Distance to daycare	Average road distance of all inhabitants to the nearest daycare. A daycare is a place where people can bring their children aged 0-4 for at least one part of the day each for throughout the whole year.
Distance to school	Average road distance of all inhabitants to nearest school.
Schools within 3km	Amount of schools within a range of 3 kilometres. The range takes road distance as scale.
Spatial factors	
Surrounding address density	Calculates the amount of addresses within a radius of 1 kilometre around an address, divided by the surface of the circle. The variable shows the average surrounding address density taking into account all addresses in the area.
Province	Categorical variable indicating the province

G4	Dummy variable that takes value '1' if the neighbourhood or municipality itself is part of the four biggest municipalities that also contain the largest four Dutch cities: Amsterdam, Rotterdam, 's-Gravenhage or Utrecht

Note. Almost all definitions are retrieved from and available on the website of Statistics Netherlands (CBS).

Table C.2: Description of used variables of the KWB 2020 dataset for the final model (CBS)

Variable	**Description**
Dependent variable	
Average property value	Average property value based on the 'WOZ waarde' in thousand euros.
Independent variables	
Neighbourhood with stadium	Dummy variable that takes the value '1' if there is a stadium present in the neighbourhood, and ‘0’ otherwise
Adjacent neighbourhood	Dummy variable that takes the value '1' if the neighbourhood is adjacent to a neighbourhood with a stadium, and ‘0’ otherwise. More information can be found in the Data section and in Appendix B.
Municipality with stadium	Dummy variables that takes the value '1' if there is a stadium present in the municipality, and '0' otherwise
Control variables	
Demographics	
Working age population	Percentage of total inhabitants aged between 25-64
Labour migration	Percentage of total inhabitants with a Morrocan, Antillean/ Aruban, Surinamese or Turkish background
Other non-Western	Percentage of other non-Western inhabitants
Average household size	Average size of households
Housing market	
Housing stock	Housing stock within a neighbourhood
Inhabited	Percentage of houses that are inhabited
Uninhabited	Percentage of houses that are uninhabited
Owner-occupied houses	Percentage of houses that are owner-occupied houses
Rental houses	Percentage of houses that are rental houses
Year of construction before 2000	Percentage of houses with the year of construction before 2000
Year of construction after 2000	Percentage of houses with the year of construction from 2000 onwards
Education	
Low-skilled	Percentage of total inhabitants that are low-skilled
Medium-skilled	Percentage of total inhabitants that are medium-skilled
High-skilled	Percentage of total inhabitants that are high-skilled
Amenities	

Distance to amenities	Average road distance of all inhabitants to the nearest amenities (average is taken for distance to nearest general practice, large supermarket, daycare and school)
Stadium variables	
Position	For every position category, a dummy variable is included like explained in the Data section
Positive externalities	Percentage of people older than 18 that sports at least every week, as explained in the Data section
Negative externalities	Amount of negative newspaper articles when using the search term ‘hooligans’, as explained in the Data section
Spatial factors	
Surrounding address density	Calculates the amount of addresses within a radius of 1 kilometre around an address, divided by the surface of the circle. The variable shows the average surrounding address density taking into account all addresses in the area.
Province	Categorical variable indicating the province
G4	Dummy variable that takes value '1' if the neighbourhood or municipality itself is part of the four biggest municipalities that also contain the largest four Dutch cities: Amsterdam, Rotterdam, 's-Gravenhage or Utrecht

Note. Almost all definitions are retrieved from and available on the website of Statistics Netherlands (CBS).

Table C.3: Description of variables for the NVM dataset

Variable	Description
Dependent variable	
Average transaction price	Average transaction price in municipality per year in euros, rounded to cents
Independent variables	
Stadium	Dummy variable that takes the value '1' if there is a stadium present in the municipality during the year, and '0' otherwise
Stadium under construction	Dummy variable that takes the value '1' if there is a stadium in the municipality that is under construction during that specific year, based on Table A.1, and '0' otherwise.
Control variables	
Average surface	Average surface in m^2, rounded to two decimal places
Average volume	Average volume in m^3, rounded to two decimal places
Share of apartments	Share of apartments in the municipality, rounded to three decimal places
Built before 1945	Share of homes built before 1945 (pre-WWII homes), rounded to three decimal places
Built between 1945-1970	Share of homes built between 1945 and 1970 (post WWII rebuilding), rounded to three decimal places
Built between 1971-1990	Share of homes built between 1971 and 1990 (oil crisis and material shortages), rounded to three decimal places
Built between 1991-2000	Share of homes built between 1991 and 2000 (large suburban expansion), rounded to three decimal places
Built after 2000	Share of homes built after 2000 (urban renewal), rounded to three decimal places
Year	Categorical variable indicating the year of the transaction
Province	Categorical variable indicating the province
G4	Dummy variable that takes value '1' if the municipality is part of the four biggest municipalities that also contain the largest four Dutch cities: Amsterdam, Rotterdam, 's-Gravenhage or Utrecht

Table C.4: Description of variables for the LISA dataset combined with the KWB 2017 dataset (CBS)

Variable	**Description**
Dependent variables	
Jobs	Total amount of jobs per municipality
Sports-related jobs	Total amount of sports-related jobs (SBI code 931)
Business-related jobs	Total amount of business-, consultancy- and finance-related jobs (SBI codes 641, 642, 643, 649, 661, 662, 663, 692 and 702)
Network jobs	Total amount of jobs that specifically involve networking (SBI codes 692 and 702)
Independent variable	
Stadium	Dummy variable that takes the value '1' if there is a stadium present in the municipality, and '0' otherwise
Control variables	
Demographics	
Inhabitants	Amount of inhabitants
Male	Amount of male inhabitants
Female	Amount of female inhabitants
Age 0-14	Amount of inhabitants aged 0 until 15
Age 15-24	Amount of inhabitants aged 15 until 25
Age 25-44	Amount of inhabitants aged 25 until 45
Age 45-64	Amount of inhabitants aged 45 until 65
Age 65+	Amount of inhabitants aged 65 or older
Western	Amount of inhabitants with a Western migration background
Non-Western	Amount of inhabitants with a non-Western migration background
Moroccan	Amount of Moroccan inhabitants
Antillean and Aruban	Amount of Antillean and Aruban inhabitants
Surinamese	Amount of Surinamese inhabitants
Turkish	Amount of Turkish inhabitants
Other non-Western	Amount of other non-Western inhabitants
Amenities	
Distance to general practice	Average road distance of all inhabitants to the nearest general practice
Distance to big supermarket	Average road distance of all inhabitants to the nearest big supermarket. A big supermarket should have multiple types of daily products and a surface of at least 150 m^2.
Distance to daycare	Average road distance of all inhabitants to the nearest daycare. A daycare is a place where people can bring their children aged 0-4 for at least one part of the day each for throughout the whole year.
Distance to school	Average road distance of all inhabitants to nearest school.
Schools within 3km	Amount of schools within a range of 3 kilometres. The range takes road distance as scale.

Spatial factors	
Surrounding address density	Calculates the amount of addresses within a radius of 1 kilometre around an address, divided by the surface of the circle. The variable shows the average surrounding address density taking into account all addresses in the area.
Surface	Total surface in hectares
Land surface	Land surface in hectares
Water surface	Water surface in hectares
Province	Categorical variable indicating the province
G4	Dummy variable that takes value '1' if the neighbourhood or municipality itself is part of the four biggest municipalities that also contain the largest four Dutch cities: Amsterdam, Rotterdam, 's-Gravenhage or Utrecht

Note. Almost all definitions are retrieved from and available on the website of Statistics Netherlands (CBS).

Appendix D – Changes to dataset

Changes to KWB 2020 dataset

As described in the Data section, the primary model that was chosen did not present reliable estimates of the variables. Also, a lot of variables were strongly correlated, and the R^2 could be improved. Therefore, it was decided to use a different model. The descriptive statistics and regression results of the primary model are presented in Table D.1 and D.2 respectively.

Table D.1: Descriptive statistics for the primary model with the KWB 2020 dataset (CBS)

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variable					
Average property value	11961	315.44	140.42	19.00	2065.00
Independent variables					
Neighbourhood with stadium	13807	0.00	0.04	0.00	1.00
Adjacent neighbourhood	13808	0.00	0.07	0.00	1.00
Municipality with stadium	13808	0.14	0.35	0.00	1.00
Control variables					
Demographics					
Inhabitants	13808	1259.97	1674.82	0.00	28870.00
Male	13808	625.62	824.71	0.00	13615.00
Female	13808	633.71	851.37	0.00	15250.00
Age 0-14	13808	197.45	285.95	0.00	4895.00
Age 15-24	13808	155.25	228.93	0.00	3315.00
Age 25-44	13808	312.05	470.98	0.00	7785.00
Age 45-64	13808	350.26	453.11	0.00	6880.00
Age 65+	13808	245.70	340.99	0.00	7920.00
Unmarried	13808	613.36	880.78	0.00	14710.00
Married	13808	485.85	625.51	0.00	9600.00
Divorced	13808	98.23	150.06	0.00	3070.00
Widowed	13808	61.89	95.47	0.00	2195.00
Western	13808	132.39	230.28	0.00	4060.00
Non-western	13808	173.21	507.84	0.00	11665.00
Moroccan	13808	29.62	123.69	0.00	2970.00
Antillean and aruban	13808	11.94	45.35	0.00	1840.00
Surinamese	13808	25.81	110.25	0.00	2430.00
Turkish	13808	30.21	125.09	0.00	4090.00
Other non-western	13808	75.70	173.03	0.00	3420.00
Birth rate	13808	11.19	18.48	0.00	315.00
Death rate	13808	11.23	21.34	0.00	520.00
Households	13808	577.87	805.04	0.00	14035.00
Single-person households	13808	223.04	385.18	0.00	6460.00
Households without children	13808	166.82	214.87	0.00	3455.00
Households with children	13808	189.42	263.68	0.00	4490.00
Average household size	13462	2.30	0.46	1.00	9.00

Housing market					
Housing stock	13808	571.54	781.20	0.00	14222.00
Single-family housing	12271	76.78	27.37	0.00	100.00
Multi-family housing	12271	23.22	27.37	0.00	100.00
Inhabited	12271	94.33	7.20	0.00	100.00
Uninhabited	12271	5.67	7.20	0.00	100.00
Owner-occupied houses	12265	68.15	22.68	0.00	100.00
Rental houses	12265	31.43	22.60	0.00	100.00
In possession of housing associations	12265	18.53	21.08	0.00	100.00
In possession of other landlords	12265	12.90	13.73	0.00	100.00
Ownership unknown	12265	0.38	1.71	0.00	45.00
Year of construction before 2000	12271	82.58	23.31	0.00	100.00
Year of construction after 2000	12271	17.42	23.31	0.00	100.00
Education					
Low-skilled	11817	305.60	413.38	0.00	7900.00
Medium-skilled	11822	458.47	544.01	0.00	9100.00
High-skilled	11817	340.86	472.80	0.00	7950.00
Business					
Business establishments	13808	124.54	160.83	0.00	4010.00
Amenities					
Distance to general practice	13138	1.68	1.42	0.00	11.30
Distance to big supermarket	13138	1.54	1.34	0.00	11.20
Distance to daycare	13138	1.12	0.99	0.00	9.80
Distance to school	13138	1.16	0.94	0.10	10.00
Schools within 3km	13138	7.36	7.65	0.00	60.70
Spatial factors					
Surrounding address density	13739	1192.43	1499.67	0.00	12421.00
G4	13808	0.06	0.23	0.00	1.00

Table D.2: Regression results for the primary models using the KWB 2020 dataset (CBS)

	(1) **Average property value**	**(2)** **Average property value**	**(3)** **Average property value**
Neighbourhood with stadium	2.220	2.463	-0.643
	(21.847)	(21.870)	(21.971)
Adjacent neighbourhood		9.818	7.326
		(9.590)	(9.775)
Municipality with stadium			4.762
			(3.114)
Inhabitants	-0.326	-0.325	-0.344
	(0.388)	(0.388)	(0.388)
Male	0.085	0.084	0.104

	(0.293)	(0.293)	(0.293)
Female	0.130	0.130	0.154
	(0.294)	(0.294)	(0.295)
Age 0-14	0.180	0.180	0.188
	(0.126)	(0.126)	(0.125)
Age 15-24	-0.008	-0.008	0.001
	(0.122)	(0.122)	(0.122)
Age 25-44	-0.147	-0.147	-0.136
	(0.121)	(0.121)	(0.121)
Age 45-64	0.125	0.125	0.136
	(0.122)	(0.122)	(0.122)
Age 65+	0.156	0.157	0.166
	(0.122)	(0.122)	(0.122)
Unmarried	0.215	0.215	0.204
	(0.209)	(0.209)	(0.209)
Married	0.132	0.132	0.122
	(0.209)	(0.209)	(0.209)
Divorced	0.057	0.056	0.044
	(0.210)	(0.210)	(0.210)
Widowed	0.112	0.111	0.098
	(0.212)	(0.212)	(0.212)
Western	0.014	0.015	0.015
	(0.013)	(0.013)	(0.013)
Non-Western	0.335**	0.333**	0.331**
	(0.156)	(0.156)	(0.156)
Moroccan	-0.343**	-0.341**	-0.337**
	(0.156)	(0.156)	(0.156)
Antillean and Aruban	-0.132	-0.132	-0.130
	(0.160)	(0.160)	(0.160)
Surinamese	-0.426***	-0.424***	-0.420***
	(0.157)	(0.157)	(0.157)
Turkish	-0.343**	-0.341**	-0.339**
	(0.156)	(0.156)	(0.156)
Other non-Western	-0.261*	-0.259*	-0.257
	(0.157)	(0.157)	(0.157)
Birth rate	-0.270*	-0.266*	-0.261*
	(0.148)	(0.148)	(0.148)
Death rate	-0.031	-0.032	-0.035
	(0.118)	(0.118)	(0.118)
Households	0.580**	0.578**	0.554**
	(0.227)	(0.227)	(0.229)
Single-person households	-0.432*	-0.431*	-0.407*
	(0.227)	(0.227)	(0.229)
Households without children	-0.425*	-0.423*	-0.402*
	(0.227)	(0.227)	(0.229)
Households with children	-0.510**	-0.509**	-0.488**
	(0.228)	(0.228)	(0.229)
Average household size	92.704***	92.722***	92.929***
	(5.126)	(5.126)	(5.132)
Housing stock	-0.127***	-0.127***	-0.127***
	(0.023)	(0.023)	(0.023)
Single-family housing	-145.709***	-145.765***	-146.075***
	(30.163)	(30.159)	(30.237)
Multi-family housing	-145.833***	-145.890***	-146.197***
	(30.162)	(30.157)	(30.235)
Inhabited	-2.931***	-2.931***	-2.939***
	(0.345)	(0.345)	(0.345)

Owner-occupied houses	5.362*	5.338*	5.412*
	(2.858)	(2.858)	(2.858)
Rental houses	0.131	0.118	0.187
	(2.829)	(2.829)	(2.830)
In possession of housing associations	3.152*	3.140*	3.144*
	(1.648)	(1.648)	(1.648)
In possession of other landlords	4.526***	4.514***	4.514***
	(1.657)	(1.657)	(1.658)
Ownership unknown	5.073*	5.052*	5.114*
	(2.935)	(2.936)	(2.936)
Year of construction before 2000	19.651***	19.906***	21.565***
	(6.769)	(6.774)	(6.829)
Year of construction after 2000	20.018***	20.274***	21.931***
	(6.766)	(6.771)	(6.826)
Low-skilled	-0.001	-0.002	-0.003
	(0.027)	(0.027)	(0.027)
Medium-skilled	-0.105***	-0.105***	-0.106***
	(0.027)	(0.027)	(0.027)
High-skilled	0.033	0.032	0.031
	(0.027)	(0.027)	(0.027)
Business establishments	0.160***	0.160***	0.161***
	(0.016)	(0.016)	(0.016)
Distance to general practice	-3.837***	-3.839***	-3.809***
	(0.953)	(0.953)	(0.953)
Distance to big supermarket	2.479**	2.493**	2.485**
	(1.137)	(1.137)	(1.137)
Distance to daycare	4.271***	4.275***	4.326***
	(1.595)	(1.595)	(1.594)
Distance to school	20.319***	20.309***	20.248***
	(1.739)	(1.739)	(1.738)
Schools within 3km	1.848***	1.850***	1.842***
	(0.409)	(0.409)	(0.409)
Surrounding address density	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Provinces			
Flevoland	36.903***	36.988***	37.784***
	(8.135)	(8.137)	(8.169)
Friesland	-20.849***	-20.860***	-20.477***
	(3.992)	(3.992)	(4.006)
Gelderland	54.077***	54.105***	54.571***
	(4.082)	(4.083)	(4.103)
Groningen	-36.109***	-36.084***	-36.287***
	(4.740)	(4.741)	(4.725)
Limburg	23.017***	22.952***	22.861***
	(4.320)	(4.322)	(4.336)
Noord-Brabant	83.671***	83.649***	83.580***
	(4.119)	(4.120)	(4.134)
Noord-Holland	123.081***	123.101***	123.627***
	(6.431)	(6.432)	(6.449)
Overijssel	20.744***	20.714***	20.518***
	(4.039)	(4.040)	(4.050)
Utrecht	105.809***	105.820***	106.593***
	(6.433)	(6.433)	(6.474)
Zeeland	-12.981**	-12.979**	-12.466**
	(5.770)	(5.771)	(5.778)
Zuid-Holland	68.791***	68.839***	69.510***
	(5.225)	(5.226)	(5.247)

G4	76.388***	76.302***	72.394***
	(8.125)	(8.129)	(8.406)
Constant	12,401.482***	12,383.919***	12,241.336***
	(3,086.653)	(3,086.429)	(3,094.873)
Observations	11,385	11,385	11,385
R^2	0.567	0.567	0.567

Note. This table shows the results of three OLS regressions with the average property value as dependent variable, and having a stadium in the neighbourhood (1, 2 and 3), adjacent neighbourhoods (2 and 3) and municipalities with a stadium (3) as independent variables. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

To improve this model, a new model was developed that initially does not take into account the stadium effect. In that way, it was attempted to create a model that realistically predicts house prices, before adding the stadium effect to it. First of all, the amount of men and women in the dataset was deleted as it had no significant effect on house prices in the primary model, and there is no reason to assume that a higher share of men or women in a neighbourhood has an influence on the house prices. This deletion did not lead to major changes in the coefficients of the primary model. Also, instead of using the absolute values, the demographic and education variables were taken as a share of the total amount of inhabitants, or as a share of the total amount of households. Then, the top 1% of property values was deleted from the dataset, as this variable contained some very high outliers. These outliers are probably expensive villas in rich neighbourhoods, and these are usually not representative of neighbourhoods in which a football stadium could be placed.

After this, all observations with value '0' for either the amount of inhabitants, the housing stock or the amount of households was deleted. This is because if one of these is 0, there will be no houses or people living in that neighbourhood, which makes it irrelevant to examine the house prices in that neighbourhood. Next, instead of using the total amount, the variables concerning age, ethnicity and marital status were expressed as a percentage of the inhabitants in the neighbourhoods. Just like this, the variables for single-person households, households with and households without children were expressed as percentages of total households.

Subsequently, a correlation table was constructed using pairwise correlations to track down any cases of multicollinearity. Every correlation that was higher than 0.7 or lower than -0.7 between two independent was highlighted, as a correlation of more than 0.7 or less than -0.7 points to a strong correlation between the two variables (Ratner, 2009). This correlation was then extracted to a smaller table, including the correlated variables and the dependent variables. On this basis, a selection has been made of which variables to keep and which variables to exclude from the model. The separate smaller tables with correlation coefficients are presented below, with an explanation in which the choices made are clarified.

Table D.3: Correlation table for several variables strongly correlated to the amount of inhabitants

Variables		**1**	**2**	**3**	**4**	**5**	**6**
1	Average property value	1.00					
2	Inhabitants	-0.25*	1.00				
3	Birth rate	-0.22*	0.91*	1.00			
4	Death rate	-0.25*	0.74*	0.61*	1.00		
5	Households	-0.27*	0.98*	0.88*	0.74*	1.00	
6	Housing stock	-0.28*	0.98*	0.88*	0.77*	0.99*	1.00

Note. * $p < 0.05$

In Table D.3, it can be seen that almost all of the presented variables are correlated with each other, but specifically strongly with inhabitants. In the end, inhabitants, birth rate, death rate, and households were dropped because the housing stock has the strongest correlation with the dependent variable compared to the variables that were dropped.

Table D.4: Correlation table for several housing characteristics

Variables		**1**	**2**	**3**	**4**	**5**
1	Average property value	1.00				
2	Owner-occupied houses	0.46*	1.00			
3	Rental houses	-0.46*	-1.00*	1.00		
4	Single-family housing	0.24*	0.75*	-0.74*	1.00	
5	Multi-family housing	-0.24*	-0.75*	0.74*	-1.00*	1.00

Note. * $p < 0.05$

Owner-occupied houses and rental houses, just like single-family and multi-family housing, have a rounded correlation of -1.00 according to Table D.4, which is logical as they are percentages that add up to 100%. Therefore, only one of each pair will possibly end up in the regression. However, single-family housing and multi-family housing were dropped as they have a strong correlation with owner-occupied housing and rental housing, and the last two have highest correlation with the dependent variable.

Table D.5: Correlation table for several household characteristics

Variables		**1**	**2**	**3**	**4**
1	Average property value	1.00			
2	Average household size	0.42*	1.00		
3	Single-person households	-0.37*	-0.77*	1.00	
4	Households with children	0.32*	0.73*	-0.42*	1.00

Note. * $p < 0.05$

Table D.5 shows that the average household size, as could be expected, is correlated with single-person households and households with children, so single-person households and households with children were dropped as these are the least strongly correlated with the dependent variable. Also, there was a higher level of significance for average household size in the primary model. In addition to this, households without children is dropped as this is the same categorisation as the other dropped variables and does not show a very strong correlation with the dependent variable.

Table D.6: Correlation of housing characteristics and married

Variables		1	2	3	4
1	Average property value	1.00			
2	Owner-occupied houses	0.46*	1.00		
3	Rental houses	-0.46*	-1.00*	1.00	
4	Married	0.30*	0.72*	-0.72*	1.00

Note. * $p < 0.05$

It becomes clear from Table D.6 that there is a strong correlation between married on one side, and owner-occupied and rental houses on the other side. The proportion of married people has the smallest effect on the dependent variable, so this variable is dropped. Also, unmarried, divorced and widowed were dropped, as most all showed a rather high correlation as well and are related. They were insignificant in the primary model too. On top of that, there is no logical reason, other than controlled for, why marital status should have an effect on house prices.

Table D.7: Correlation table for housing stock and business establishments

Variables		1	2	3
1	Average property value	1.00		
2	Housing stock	-0.28*	1.00	
3	Business establishments	-0.03*	0.80*	1.00

Note. * $p < 0.05$

The number of business establishments is strongly correlated with the housing stock, and business establishments has a very weak correlation with the dependent variable, as can be seen in Table D.7. Also, as discussed in the Literature review, there is an ongoing discussion about whether jobs follow people or people follow jobs, so the amount of business establishments could have the problem of reverse causality. It is therefore decided to drop the number of business establishments in the model.

Table D.8: Correlation table for surrounding address density and schools within 3km

Variables		1	2	3
1	Average property value	1.00		
2	Surrounding address density	-0.14*	1.00	
3	Schools within 3km	-0.11*	0.91*	1.00

Note. * p < 0.05

Here, the amount of schools within 3km is dropped, as this is strongly correlated with the surrounding address density and has the relatively weaker correlation with the dependent variable, as indicated in Table D.8.

Table D.9: Correlation table for several variables related to ethnicity

Variables		1	2	3	4	5	6	7	8
1	Average property value	1.00							
2	Western	-0.08*	1.00						
3	Non-Western	-0.28*	0.30*	1.00					
4	Moroccan	-0.17*	0.17*	0.70*	1.00				
5	Antillean and Aruban	-0.25*	0.18*	0.57*	0.31*	1.00			
6	Surinamese	-0.11*	0.18*	0.67*	0.37*	0.50*	1.00		
7	Turkish	-0.26*	0.20*	0.69*	0.53*	0.31*	0.29*	1.00	
8	Other non-Western	-0.25*	0.31*	0.84*	0.39*	0.42*	0.46*	0.37*	1.00

Note. * p < 0.05

From Table D.9, it becomes clear that especially non-Western is quite strongly correlated with most variables about ethnicity, with a correlation of above 0.7 for Moroccan and other non-Western. In combination with the fact that Non-Western entails all the mentioned ethnicities, and Western has a very weak correlation with the dependent variable, it is decided to drop both Western and non-Western. As there is also, although not the strongest, correlation between the other ethnicity variables, the share of Moroccan, Antillean and Aruban, Surinamese and Turkish inhabitants are joint into a variable named labour migration.

Table D.10: Correlation table for the housing ownership variables

Variables		1	2	3	4	5	6
1	Average property value	1.00					
2	Owner-occupied houses	0.46*	1.00				
3	Rental houses	-0.46*	-1.00*	1.00			
4	In possession of housing associations	-0.54*	-0.80*	0.81*	1.00		
5	In possession of other landlords	0.09*	-0.41*	0.40*	-0.21*	1.00	
6	Ownership unknown	0.01	-0.07*	-0.01	-0.06*	0.08*	1.00

Note. * p < 0.05

As can be seen in Table D.10, the variable representing the percentage of houses in possession of housing associations is strongly correlated with both the share of owner-occupied houses and the share of rental houses. As argued before, owner-occupied houses and rental houses have a correlation of 1.00 rounded of, so only one included in regression. Despite the slightly stronger correlation between houses in possession of housing associations and the dependent variable in comparison to the correlation between owner-occupied and rental houses with the dependent variable, only the owner-occupied houses and rental houses variables are kept. This has also to do with the other two variables describing ownership, which are very weakly correlated or have no significant correlation with the dependent variable. Therefore, variables 2 and 3 in Table D.10 are kept, while ownership variables 4, 5 and 6 are dropped.

Table D.11: Correlation table for the amenity variables

Variables		**1**	**2**	**3**	**4**	**5**
1	Average property value	1.00				
2	Distance to general practice	0.23*	1.00			
3	Distance to big supermarket	0.26*	0.82*	1.00		
4	Distance to daycare	0.32*	0.70*	0.71*	1.00	
5	Distance to school	0.34*	0.66*	0.66*	0.86*	1.00

Note. * $p < 0.05$

All remaining variables for amenities are (almost) strongly correlated to each other, as becomes evident from Table D.11. Thus, the average of them is taken in a new variable named ‘distance to amenities’. The separate variables are dropped.

Table D.12: Correlation table for several housing characteristics

Variables		**1**	**2**	**3**	**4**	**5**	**6**
1	Inhabited	1.00					
2	Uninhabited	-1.00*	1.00				
3	Owner-occupied houses	0.10*	-0.10*	1.00			
4	Rental houses	-0.08*	0.08*	-1.00*	1.00		
5	Year of construction before 2000	0.03*	-0.03*	0.00	0.00	1.00	
6	Year of construction after 2000	-0.03*	0.03*	0.00	0.00	-1.00*	1.00

Note. * $p < 0.05$

The variables inhabited and uninhabited, owner-occupied houses and rental houses, and year of construction before and year of construction after 2020 are each perfectly correlated to each other when rounded to two decimals. However, the model does not leave one of the options out by itself. Given that each pair of variables should represent percentages that add up to 100% together, one of each pair is left out in the regression equation.

Now that the biggest problems regarding multicollinearity have been dealt with, the natural logarithm will be taken of the variables average property value, housing stock and average household size. The variables have been selecting on the basis of several trial regressions with using different variables to take the ln of, after which the variables were choses that resulted in the highest adjusted R^2. Later on, the natural logarithm is also taken for the distance to amenities and stadium capacity variables. Finally, the age groups between 25-44 and 45-64 have been joint together in a new variable called 'working age population'. Although this resulted in a slightly lower adjusted R^2, it changed the variable other non-Western in a logical way, so the change was kept. All other age groups were dropped from the dataset.

Lastly, when the stadium variables are included, another pairwise correlation table is constructed to look for multicollinearity (correlation that is higher than 0.7 or lower than -0.7). The potential problematic variables are shown in Table D.13.

Table D.13: Correlation table for the potential problematic stadium variables

Variables		**1**	**2**	**3**	**4**	**5**
1	Average property value	1.00				
2	Municipality with stadium	-0.03*	1.00			
3	Capacity	0.06*	0.83*	1.00		
4	Position category 1	0.14*	0.45*	0.78*	1.00	
5	Position category 8	0.03*	-1.00*	-0.83*	-0.45*	1.00

Note. * $p < 0.05$

As becomes evident from Table D.13, the capacity is strongly correlated with the municipality with a stadium variable, and with both position variables. It is therefore decided to drop the capacity variable. Also, the position category 8 is perfectly correlated with the municipality with a stadium variable, so this category will be dropped and will thus be the reference category.

Appendix E – Descriptive statistics

Table E.1: Descriptive statistics for the KWB 2020 dataset on neighbourhood level (CBS)

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variable					
Average property value	11821	307.71	118.49	19.00	766.00
Independent variables					
Neighbourhood with stadium	12989	0.00	0.04	0.00	1.00
Adjacent neighbourhood	12990	0.01	0.07	0.00	1.00
Municipality with stadium	12990	0.14	0.34	0.00	1.00
Control variables					
Demographics					
Working age population	12990	0.52	0.10	0.00	2.00
Labour migration	12990	0.04	0.08	0.00	1.00
Other non-western	12990	0.04	0.06	0.00	1.00
Average household size	12990	2.30	0.43	1.00	6.00
Housing market					
Housing stock	12990	603.23	792.67	1.00	14222.00
Inhabited	12127	94.41	7.08	0.00	100.00
Uninhabited	12127	5.59	7.08	0.00	100.00
Owner-occupied houses	12121	68.09	22.66	0.00	100.00
Rental houses	12121	31.50	22.59	0.00	100.00
Year of construction before 2000	12127	82.57	23.30	0.00	100.00
Year of construction after 2000	12127	17.43	23.30	0.00	100.00
Education					
Low-skilled	11681	0.21	0.08	0.00	1.00
Medium-skilled	11686	0.34	0.11	0.00	7.00
High-skilled	11682	0.23	0.12	0.00	5.00
Amenities					
Distance to amenities	12915	1.37	1.05	0.15	10.48
Stadium variables					
Position category 1	12990	0.03	0.17	0.00	1.00
Position category 2	12990	0.02	0.14	0.00	1.00
Position category 3	12990	0.02	0.15	0.00	1.00
Position category 4	12990	0.02	0.12	0.00	1.00
Position category 5	12990	0.02	0.15	0.00	1.00
Position category 6	12990	0.01	0.08	0.00	1.00
Position category 7	12990	0.02	0.13	0.00	1.00
Position category 8 (no stadium)	12990	0.86	0.35	0.00	1.00
Spatial factors					
Surrounding address density	12990	1197.18	1483.29	2.00	12421.00

G4	12990	0.05	0.23	0.00	1.00

Table E.2: Descriptive statistics for the KWB 2020 dataset on municipality level (CBS)

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variable					
Average property value	347	271.45	58.13	138.00	440.00
Independent variables					
Municipality with stadium	349	0.05	0.22	0.00	1.00
Positive externalities	349	48.01	5.11	33.00	60.20
Negative externalities	349	1.34	8.43	0.00	89.00
Control variables					
Demographics					
Working age population	349	0.51	0.02	0.45	0.61
Labour migration	349	0.04	0.04	0.00	0.26
Other non-western	349	0.04	0.02	0.01	0.19
Average household size	349	2.25	0.18	1.70	3.30
Housing market					
Housing stock	349	22480.63	36060.72	579.00	447351.00
Inhabited	349	95.69	2.51	76.00	98.00
Uninhabited	349	4.31	2.51	2.00	24.00
Owner-occupied houses	349	64.89	8.56	29.00	78.00
Rental houses	349	34.71	8.52	20.00	70.00
Year of construction before 2000	349	83.15	5.79	56.00	95.00
Year of construction after 2000	349	16.85	5.79	5.00	44.00
Education					
Low-skilled	349	0.21	0.03	0.12	0.31
Medium-skilled	349	0.33	0.03	0.21	0.43
High-skilled	349	0.21	0.06	0.08	0.43
Amenities					
Distance to amenities	349	0.95	0.28	0.45	1.92
Stadium variables					
Position category 1	349	0.00	0.05	0.00	1.00
Position category 2	349	0.01	0.09	0.00	1.00
Position category 3	349	0.01	0.09	0.00	1.00
Position category 4	349	0.01	0.11	0.00	1.00
Position category 5	349	0.01	0.09	0.00	1.00
Position category 6	349	0.00	0.05	0.00	1.00
Position category 7	349	0.01	0.08	0.00	1.00
Position category 8 (no stadium)	349	0.95	0.22	0.00	1.00

Spatial factors					
Surrounding address density	349	1160.68	779.53	208.00	6074.00
G4	349	0.01	0.11	0.00	1.00

Table E.3: Descriptive statistics for the NVM dataset

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variable					
Average transaction price	16780.00	237072.12	638201.02	30479.00	39027240.00
Independent variables					
Stadium	16780.00	0.03	0.17	0.00	1.00
Stadium under construction	16780.00	0.00	0.03	0.00	1.00
Control variables					
Average surface	16780.00	117.29	27.55	0.49	341.00
Average volume	16780.00	7498.84	17692.43	160.00	99733.21
Share of apartments	16780.00	0.10	0.14	0.00	1.00
Built before 1945	16780.00	0.24	0.16	0.00	1.00
Built between 1945-1970	16780.00	0.24	0.11	0.00	1.00
Built between 1971-1990	16780.00	0.37	0.15	0.00	1.00
Built between 1991-2000	16780.00	0.10	0.09	0.00	1.00
Built after 2000	16780.00	0.04	0.06	0.00	1.00
Year	16780.00	2003.11	10.63	1985.00	2021.00
G4	16780.00	0.01	0.09	0.00	1.00

Table E.4: Descriptive statistics for the LISA dataset combined with the KWB 2017 dataset (CBS)

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Dependent variable					
Jobs	376	22384.70	46030.69	570.00	637039.00
Sports-related jobs	376	163.38	288.04	4.00	3477.00
Business-related jobs	376	1281.57	4808.63	1.00	82289.00
Network jobs	376	828.56	2410.04	1.00	38269.00
Independent variables					
Stadium	376	0.05	0.21	0.00	1.00
Control variables					
Demographics					
Inhabitants	376	44995.17	69337.21	941.00	844947.00
Male	376	22321.70	34279.69	483.00	418127.00
Female	376	22673.47	35063.42	458.00	426820.00
Age 0-14	376	7329.97	11126.07	110.00	126764.00
Age 15-24	376	5541.82	9480.76	92.00	109253.00
Age 25-44	376	11118.45	22424.53	186.00	299004.00
Age 45-64	376	12694.93	17302.53	292.00	207391.00
Age 65+	376	8309.99	9675.77	205.00	102535.00
Western	376	4467.00	10655.90	68.00	147614.00
Non-western	376	5760.24	22764.83	17.00	296073.00
Moroccan	376	1039.07	5088.63	0.00	75758.00
Antillean and aruban	376	404.95	1639.66	1.00	24362.00
Surinamese	376	926.46	5118.58	0.00	65468.00
Turkish	376	1061.45	4171.18	0.00	47772.00
Other non-western	376	2328.31	7390.33	11.00	99410.00
Amenities					
Distance to general practice	376	1.14	0.39	0.50	2.80
Distance to big supermarket	376	1.05	0.38	0.40	2.50
Distance to daycare	376	1.03	1.25	0.40	22.70
Distance to school	376	0.76	0.18	0.40	1.70
Schools within 3km	376	6.57	4.45	0.90	34.80
Spatial factors					
Surrounding address density	376	1107.03	758.73	155.00	6004.00
Surface	376	10750.21	11486.74	784.00	83871.00
Land surface	376	8675.69	7622.92	703.00	46005.00
Water surface	376	2074.48	6634.73	0.00	59054.00
G4	376	0.01	0.10	0.00	1.00

Appendix F – Regression results

Table F.1: Regression results for the models using the KWB 2020 dataset (CBS)

	(1) ln(Average property value)	(2) ln(Average property value)	(3) ln(Average property value)
Neighbourhood with stadium	0.065	0.064	0.064
	(0.061)	(0.061)	(0.061)
Adjacent neighbourhood		-0.022	-0.020
		(0.020)	(0.020)
Municipality with stadium			0.121***
			(0.034)
Working age population	-0.674***	-0.674***	-0.673***
	(0.058)	(0.058)	(0.058)
Labour migration	-0.668***	-0.667***	-0.667***
	(0.044)	(0.044)	(0.044)
Other non-Western	-0.322***	-0.322***	-0.322***
	(0.100)	(0.100)	(0.100)
ln(Average household size)	0.645***	0.645***	0.645***
	(0.023)	(0.023)	(0.024)
ln(Housing stock)	-0.055***	-0.055***	-0.055***
	(0.002)	(0.002)	(0.002)
Inhabited	-0.006***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)
Owner-occupied houses	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)
Year of construction before 2000	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Low-skilled	-0.189***	-0.189***	-0.190***
	(0.055)	(0.055)	(0.055)
Medium-skilled	-0.274***	-0.274***	-0.274***
	(0.055)	(0.055)	(0.055)
High-skilled	0.949***	0.950***	0.949***
	(0.056)	(0.056)	(0.056)
ln(Distance to amenities)	0.092***	0.092***	0.092***
	(0.005)	(0.005)	(0.005)
Position			
Position category 1	0.463***	0.463***	0.343***
	(0.026)	(0.026)	(0.026)
Position category 2	-0.001	0.000	-0.121***
	(0.013)	(0.013)	(0.037)
Position category 3	0.171***	0.173***	0.052
	(0.020)	(0.020)	(0.033)
Position category 4	-0.032**	-0.030**	-0.152***
	(0.013)	(0.013)	(0.036)
Position category 5	-0.034***	-0.033***	-0.154***
	(0.012)	(0.012)	(0.036)
Position category 6	-0.102***	-0.102***	-0.223***
	(0.021)	(0.021)	(0.041)
Position category 7	0.130***	0.131***	0.010
	(0.017)	(0.017)	(0.033)
Provinces			
Flevoland	0.132***	0.132***	0.132***
	(0.018)	(0.018)	(0.018)
Friesland	-0.080***	-0.080***	-0.080***
	(0.012)	(0.012)	(0.012)
Gelderland	0.205***	0.205***	0.205***

	(0.011)	(0.011)	(0.011)
Groningen	-0.186***	-0.186***	-0.186***
	(0.014)	(0.014)	(0.014)
Limburg	0.072***	0.072***	0.072***
	(0.012)	(0.012)	(0.012)
Noord-Brabant	0.287***	0.287***	0.287***
	(0.011)	(0.011)	(0.011)
Noord-Holland	0.333***	0.333***	0.333***
	(0.013)	(0.013)	(0.013)
Overijssel	0.096***	0.096***	0.096***
	(0.012)	(0.012)	(0.012)
Utrecht	0.353***	0.353***	0.353***
	(0.013)	(0.013)	(0.013)
Zeeland	-0.020	-0.020	-0.020
	(0.016)	(0.016)	(0.016)
Zuid-Holland	0.252***	0.252***	0.252***
	(0.012)	(0.012)	(0.012)
Surrounding address density	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
G4	0.000	0.000	-0.001
	(0.021)	(0.021)	(0.021)
Constant	5.948***	5.948***	5.948***
	(0.081)	(0.081)	(0.081)
Observations	11,260	11,260	11,260
R^2	0.735	0.735	0.735

Note. This table shows the results of three OLS regressions with the natural logarithm of the average property value as dependent variable, and having a stadium in the neighbourhood (1, 2 and 3), adjacent neighbourhoods (2 and 3) and municipalities with a stadium (3) as independent variables. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table F.2: Regression results for models to split the stadium effect using the KWB 2020 dataset (CBS)

	(1) ln(Average property value)	(2) ln(Average property value)
Municipality with stadium	0.157***	0.219
	(0.046)	(0.196)
Positive externalities		0.011***
Percentage of sporters		(0.002)
Positive externalities * Municipality		-0.002
with a stadium		(0.004)
Negative externalities		0.002*
Amount of negative articles		(0.001)
Working age population	0.010	0.403
	(0.704)	(0.696)
Labour migration	-0.651*	-0.484
	(0.337)	(0.333)
Other non-Western	-0.133	-0.111
	(0.739)	(0.689)
ln(Average household size)	1.034***	0.918***
	(0.213)	(0.207)
ln(Housing stock)	-0.015	-0.017
	(0.011)	(0.010)
Inhabited	-0.017***	-0.020***
	(0.003)	(0.003)
Owner-occupied houses	0.002	0.001
	(0.002)	(0.002)
Year of construction before 2000	-0.002	-0.001
	(0.001)	(0.001)
Low-skilled	0.195	-0.197
	(0.765)	(0.754)
Medium-skilled	-0.417	-1.199*
	(0.651)	(0.625)
High-skilled	2.326***	1.165
	(0.724)	(0.720)
ln(Distance to amenities)	0.068**	0.067**
	(0.034)	(0.033)
Position		
Position category 1	0.355***	0.329***
	(0.060)	(0.074)
Position category 2	-0.200***	-0.210***
	(0.060)	(0.068)
Position category 3	-0.024	-0.032
	(0.057)	(0.049)
Position category 4	-0.192***	-0.186***
	(0.049)	(0.049)
Position category 5	-0.154***	-0.119**
	(0.048)	(0.052)
Position category 6	-0.202***	-0.174***
	(0.054)	(0.062)
Provinces		
Flevoland	0.053	0.028
	(0.051)	(0.052)
Friesland	0.003	-0.016
	(0.039)	(0.038)
Gelderland	0.153***	0.104***
	(0.031)	(0.033)

	(1)	(2)
Groningen	-0.164***	-0.144***
	(0.037)	(0.038)
Limburg	0.038	0.008
	(0.033)	(0.032)
Noord-Brabant	0.213***	0.135***
	(0.030)	(0.033)
Noord-Holland	0.279***	0.208***
	(0.035)	(0.038)
Overijssel	0.062**	0.023
	(0.031)	(0.032)
Utrecht	0.257***	0.209***
	(0.034)	(0.035)
Zeeland	-0.029	-0.039
	(0.034)	(0.033)
Zuid-Holland	0.209***	0.168***
	(0.033)	(0.033)
Surrounding address density	-0.000	-0.000
	(0.000)	(0.000)
G4	-0.070	-0.187*
	(0.066)	(0.102)
Constant	6.098***	6.364***
	(0.480)	(0.471)
Observations	347	347
R^2	0.843	0.853

Note. This table shows the results of two OLS regressions with the natural logarithm of the average property value as dependent variable, and having a stadium in the municipality (1 and 2), and positive and negative externalities (2) as independent variables, as well as an interaction term between the positive externalities and having a stadium in the municipality. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table F.3: Regression results for the models using the NVM dataset

	(1) Average transaction price	(2) Average transaction price
Stadium	-11,711.385	-12,882.814
	(11,986.908)	(11,944.847)
Stadium under construction		-92,978.824**
		(37,187.925)
Average surface	4,123.729***	4,122.479***
	(900.678)	(900.564)
Average volume	4.717***	4.712***
	(1.020)	(1.019)
Share of apartments	46,176.785	47,580.462
	(33,096.764)	(33,271.153)
Built before 1945	15057509.281	15053584.637
	(9209529.810)	(9209105.261)
Built between 1945-1970	14991242.086	14987204.419
	(9204512.148)	(9204073.905)
Built between 1971-1990	14947542.053	14943546.258
	(9212504.235)	(9212070.823)
Built between 1991-2000	14658962.150	14654636.710
	(9127157.704)	(9126679.402)
Built after 2000	14830133.174	14826347.886
	(9223577.425)	(9223173.360)
Years		
1986	2,005.638	2,015.390
	(5,100.032)	(5,101.849)
1987	12,905.289**	12,899.468**
	(5,394.128)	(5,395.712)
1988	14,255.056**	14,242.972**
	(5,552.612)	(5,554.404)
1989	18,608.882***	18,594.753***
	(5,234.475)	(5,236.090)
1990	17,391.054***	17,372.971***
	(5,030.441)	(5,032.514)
1991	26,184.987***	26,157.282***
	(5,180.496)	(5,181.919)
1992	32,497.240***	32,472.692***
	(5,209.929)	(5,211.548)
1993	48,005.964***	48,396.576***
	(5,857.890)	(5,884.207)
1994	60,865.503***	61,062.149***
	(6,695.025)	(6,708.970)
1995	84,580.000***	84,718.348***
	(9,741.938)	(9,756.266)
1996	106,339.954***	106,438.932***
	(12,665.145)	(12,679.508)
1997	126,405.361***	126,711.124***
	(14,701.109)	(14,739.758)
1998	165,113.594***	165,021.123***
	(32,118.082)	(32,109.683)
1999	224,112.841***	224,027.166***
	(33,492.765)	(33,490.211)
2000	245,235.150***	245,147.302***
	(26,711.547)	(26,707.090)
2001	243,345.390***	243,261.947***
	(21,505.931)	(21,498.824)
2002	313,992.022***	313,909.061***

	(37,536.427)	(37,533.002)
2003	273,384.712***	273,292.446***
	(28,530.297)	(28,524.651)
2004	304,452.841***	304,756.895***
	(37,084.313)	(37,111.948)
2005	406,550.262***	407,053.102***
	(101,968.509)	(102,083.414)
2006	381,404.932***	381,509.822***
	(77,751.053)	(77,756.421)
2007	316,661.352***	316,557.397***
	(42,440.765)	(42,441.914)
2008	284,019.821***	283,904.762***
	(25,387.055)	(25,380.904)
2009	377,182.068***	377,046.417***
	(63,681.996)	(63,681.553)
2010	254,206.780***	254,072.807***
	(17,423.742)	(17,412.525)
2011	370,160.045***	370,020.407***
	(97,262.384)	(97,266.897)
2012	234,044.431***	233,912.709***
	(18,966.966)	(18,955.378)
2013	229,942.458***	229,800.058***
	(19,543.704)	(19,531.219)
2014	229,166.076***	229,019.558***
	(19,569.511)	(19,557.082)
2015	235,700.530***	235,548.836***
	(19,518.580)	(19,506.032)
2016	249,389.754***	249,239.367***
	(19,307.971)	(19,296.327)
2017	267,930.058***	267,787.412***
	(18,298.400)	(18,288.117)
2018	294,003.283***	293,869.742***
	(17,830.667)	(17,821.944)
2019	306,866.274***	306,738.750***
	(17,280.829)	(17,271.466)
2020	344,865.710***	344,740.813***
	(17,886.216)	(17,877.251)
2021	417,902.592***	417,756.704***
	(18,161.977)	(18,152.151)
Provinces		
Drenthe	-30,316.292**	-30,257.257**
	(14,886.550)	(14,891.906)
Flevoland	65,169.836*	65,192.737*
	(34,030.667)	(34,033.750)
Friesland	-21,033.993	-20,928.287
	(18,787.519)	(18,798.057)
Gelderland	3,277.352	3,289.266
	(13,723.605)	(13,725.059)
Groningen	-53,142.867**	-52,986.937**
	(21,973.490)	(21,985.171)
Limburg	-72,354.477***	-72,352.481***
	(11,977.727)	(11,978.009)
Noord-Holland	115,643.959***	115,626.821***
	(37,574.353)	(37,573.032)
Overijssel	14,850.459	15,007.765
	(22,133.177)	(22,142.871)
Utrecht	82,313.303***	82,083.968***

	(1)	(2)
	(21,869.230)	(21,865.860)
Zeeland	-18,289.271	-18,376.169
	(19,744.346)	(19,737.392)
Zuid-Holland	39,963.438*	39,745.799*
	(21,221.908)	(21,209.477)
G4	-11,879.599	-8,575.515
	(24,379.919)	(24,088.319)
Constant	-1.546e+07*	-1.545e+07*
	(9275744.829)	(9275286.128)
Observations	16,780	16,780
R-squared	0.048	0.048

Note. This table shows the results of two OLS regressions with the average transaction price as dependent variable, and having stadium (1 and 2) and stadium under construction(2) as independent variables. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table F.4: Regression results for the models using the LISA dataset combined with the KWB 2017 dataset (CBS)

	(1) Jobs	(2) Sports-related jobs	(3) Business-related jobs	(4) Network jobs
Stadium	5,118.266	92.406**	-32.841	22.563
	(3,190.925)	(38.752)	(398.196)	(179.960)
Inhabitants	-0.158	0.006	0.226**	0.163***
	(1.221)	(0.011)	(0.106)	(0.047)
Male	0.306	0.014	-0.487**	-0.310***
	(1.337)	(0.017)	(0.190)	(0.087)
Age 0-14	-0.511	-0.017*	-0.118	0.064
	(1.291)	(0.010)	(0.116)	(0.052)
Age 15-24	0.002	-0.012	-0.330**	-0.142***
	(1.252)	(0.010)	(0.130)	(0.054)
Age 25-44	1.737***	-0.001	0.376***	0.094**
	(0.550)	(0.008)	(0.088)	(0.037)
Age 45-64	0.597	-0.012	-0.019	-0.073
	(1.937)	(0.015)	(0.161)	(0.064)
Western	-0.119	-0.012***	-0.059	0.011
	(0.316)	(0.004)	(0.047)	(0.020)
Non-western	2.215*	0.016	0.679***	0.393***
	(1.221)	(0.010)	(0.163)	(0.081)
Moroccan	-1.133	-0.020	-0.321	-0.307***
	(1.064)	(0.013)	(0.199)	(0.090)
Antillean and Aruban	-3.261*	-0.005	-1.453***	-0.768***
	(1.912)	(0.023)	(0.324)	(0.141)
Surinamese	-4.396*	-0.008	-1.008***	-0.569***
	(2.273)	(0.020)	(0.251)	(0.129)
Turkish	-3.454**	-0.033***	-1.010***	-0.516***
	(1.512)	(0.012)	(0.191)	(0.095)
Distance to general practice	598.942	12.477	-89.433	-62.561
	(695.481)	(15.062)	(92.488)	(46.741)
Distance to big supermarket	-1,542.137*	-8.518	-116.006	-43.363
	(916.155)	(16.266)	(127.382)	(62.068)
Distance to daycare	-2.210	3.077*	-30.938**	-10.987
	(93.742)	(1.635)	(15.489)	(7.983)
Distance to school	395.442	63.442***	-100.202	12.906
	(1,546.538)	(22.482)	(196.154)	(93.353)
Schools within 3km	111.215	-0.785	-36.848	-18.232
	(218.164)	(2.464)	(24.122)	(13.101)
Surrounding address density	-3.078*	0.008	-0.338*	-0.156
	(1.829)	(0.021)	(0.201)	(0.095)
Surface	0.000	-0.001	-0.001	0.002
	(0.034)	(0.001)	(0.005)	(0.003)
Land surface	0.010	-0.000	0.022**	0.011*
	(0.081)	(0.001)	(0.011)	(0.006)
Provinces				
Flevoland	-1,936.014	36.255	-217.870	-185.327
	(3,925.854)	(51.570)	(466.706)	(219.779)
Friesland	1,622.638	25.764	365.117*	131.690
	(1,879.728)	(31.622)	(203.420)	(101.596)
Gelderland	2,421.407	56.930*	570.803***	400.372***
	(1,642.983)	(31.150)	(153.565)	(98.403)
Groningen	1,853.807	42.145	414.620**	248.937**
	(1,755.940)	(30.363)	(174.633)	(107.735)
Limburg	920.169	52.594*	209.157	146.886

	(1)	(2)	(3)	(4)
	(1,885.842)	(30.975)	(183.446)	(108.872)
Noord-Brabant	2,156.065	28.868	515.673***	356.526***
	(1,710.027)	(30.646)	(158.850)	(103.992)
Noord-Holland	1,413.791	68.879**	609.512***	267.933**
	(2,202.841)	(29.901)	(181.725)	(108.082)
Overijssel	2,984.692	25.425	660.471***	256.662**
	(1,895.754)	(33.287)	(173.815)	(102.164)
Utrecht	2,239.957	85.695**	720.041***	524.957***
	(1,823.369)	(37.795)	(190.955)	(115.019)
Zeeland	2,806.409	28.740	254.750	72.024
	(2,094.620)	(32.614)	(196.566)	(122.770)
Zuid-Holland	1,162.503	26.901	561.420***	284.937***
	(1,860.140)	(31.402)	(177.521)	(105.532)
G4	-5,287.428	-456.172**	-2,823.003	-795.174
	(18,600.372)	(202.852)	(2,474.290)	(980.854)
Constant	-1,159.113	-117.538**	313.193	28.778
	(2,784.167)	(50.044)	(319.994)	(160.857)
Observations	376	376	376	376
R-squared	0.987	0.942	0.983	0.985

Note. This table shows the results of four OLS regressions with the total amount of jobs (1), sports-related jobs (2), business-related jobs (3) and network jobs (4) as dependent variables, and stadium (1, 2, 3 and 4) as independent variable. Also, several control variables are included. The values represent the regression coefficients. Standard errors are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.