NEGATIVE MICRO-EXTERNALITIES OF PURPOSE-BUILT STUDENT ACCOMMODATIONS

Abstract. The Dutch market for student housing has recently experienced a rise of Purpose-Built Student Accommodations (PBSA) as response to traditional forms of student housing. PBSA often correspond to the concentrated arrival of a few hundred students in a single real estate object, subsequently – intentionally or not – having impact on their direct surroundings and the neighbourhood. This paper provides insights into the external economic effects of PBSA on nearby residential property prices by applying a difference-in-difference specification of the hedonic pricing model. The initial results indicate a discount of 25.95% for properties in close proximity to PBSA, decaying concavely to near zero over a distance of about 70 meters. Trend variables indicate that these effects also decay over time, which can possibly be attributed to initial preconceptions towards PBSA that lose their strength over time. Alternative specifications illustrate that the magnitude of the results may vary by the use of different treatment radii and the scale of spatial fixed effects. The consistency of the sign and significance of the coefficients throughout the majority of specifications, including one where the control group was defined through coarsened exact matching, provide ample evidence for the existence of a negative relationship between PBSA and neighbouring house prices. The results may inform urban planners and policy makers that are responsible for the spatial allocation of PBSA.

Keywords. *PBSA, externalities, hedonic housing prices, studentification*

DANIËL SLOMP

Master Thesis Real Estate Studies

Final version January 2019

COLOFON

Title	Negative Micro-Externalities of Purpose-Built Student Accommodations
Version	Final
Author	Daniël Slomp
Student number	S2164868
E-mail	d.p.slomp@gmail.com
Primary supervisor	dr. M.N. (Michiel) Daams
Secondary supervisor	dr. F.J. (Frans) Sijtsma
Date	30-1-2019
Word count	10,300 (chapter 1-6)

University of Groningen
Faculty of Spatial Sciences

Disclaimer: "Master theses are preliminary materials to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the supervisor or research staff."

CONTENTS

COLOFON 2

CONTENTS 3

1. INTRODUCTION 4

  1.1 Motivation 4

  1.2 Literature review 4

  1.3 Research problem statement 5

  1.4 Structure of this thesis 6

2. THEORETICAL FRAMEWORK 7

  2.1 House price dynamics 7

  2.2 Academic origin and contemporary use of locational externalities 7

  2.3 Studentification and the housing market 8

  2.4 The (alternative) impact of PBSA 9

  2.5 Hypothesis 10

3. METHODOLOGY 11

  3.1 Hedonic Analysis 11

  3.2 Empirical model 11

4. DATA 13

  4.1 Study area and data cleaning 13

  4.2 Descriptive statistics 14

  4.3 Exploratory analysis into time-variant neighbourhood characteristics 16

5. RESULTS 19

  5.1 Empirical model 19

  5.2 Sensitivity analyses 21

  5.3 Matching of the target and control group 24

  5.4 Are results driven by saturation of students in the neighbourhood? 25

6. DISCUSSION & CONCLUSIONS 28

  6.1 Discussion 28

  6.2 Conclusions and recommendation 29

REFERENCES 31

APPENDICES 35

  Appendix A. Stata do file including data preparation 35

  Appendix B. Full estimation results 40

  Appendix C. Chow test for equality between sets of coefficients 42

1. INTRODUCTION

1.1 Motivation

The demand for student housing in the majority of Dutch university cities has been increasing since the end of the 20th century. To meet this demand, there existed a trend of ‘verkamering’ in the Netherlands. This refers to the splitting of single-family homes to be let to students. Both parents and private investors increasingly bought homes for their children or to let. Revenue for this kind of housing exceeded that of its previous function as a single-family home, which is why these kind of residential properties were undervalued in regards to its future purpose. As a result of this development, many neighbourhoods that previously did not contain any student households got disrupted in the eyes of the existing residents. Illustrating is the example of the Transvaal district in Leiden from an article of the Volkskrant. In the span of thirteen years, Leidenaar Pascal Groen saw more than half of the houses in his street split up to be let to students (Pols, 2017). This subsequently led to substantial nuisance for local residents, as these houses were not designed for groups of students.

In response to these developments, many municipalities came up with policies to limit the growth of this 'traditional' way of student housing. In order to continue to meet the demand for student housing though, ‘new’ initiatives as Purpose-Built Student Accommodations (PBSA) increasingly gained in popularity. This type of student housing distinguishes itself by often supplying additional services and facilities to tenants, accompanied by a higher rent. PBSA are usually also of a higher quality, located on different sites and on a larger scale. It is not uncommon for PBSA to result in the sudden arrival of hundreds of students, exerting their influence on the character of the neighbourhood. On a regular basis, local residents object to the new arrival of PBSA, though usually without success (Hayes, 2017, DeStentor, 2018.)

Despite the unwillingness of residents, there exist sound arguments to justify this form of student housing. It provides municipalities with control over the quality of living conditions and the amount of students living in a certain neighbourhood. The question arises whether these positive external effects outweigh the (perceived) negative effects of the arrival of students. Exploratory research into the economic effects of these kinds of developments may provide insight into this.

1.2 Literature review

Academic literature regarding student housing finds its popular origin in the early 2000s, with the explosive growth of the student population in the UK. In these years, the process of rapidly increasing levels of student concentration within a formerly traditional residential neighbourhood had been coined as *studentification*. Studies into the effects of this phenomenon found that it manifested itself through transformation of the physical, and subsequently, the social and economic realm (Kenyon, 1997; Hughes and Davis, 2002).

From a social point of view, studentification often came paired with exclusion and segregation. The study by Smith (2005) for example, saw studentification as a form of gentrification due to the increased social and spatial polarization through displacement of existing residents and due to the economic investment in property. The exception was that it also came coupled with a physical downgrading of the built environment, which is in contrast to other forms of gentrification. Comparatively, studies regarding “family flight” discuss how long-term residents vacate an area with the rise of student-households, as a result of negative effects (Sage et al., 2012).

From an economic point of view, however, studies have shown ambiguous results regarding the effects of studentification. Some studies initially claimed that an increase in the number of students would lead to impoverishment and neglect of neighbourhoods, with strong price decreases as result (Lofland, 1967; Kenyon, 1997). While other studies find evidence for an (initial) rise in house prices due to the growing interest of private investors (Rugg et al., 2000; Powell and Bark, 2008).

The rise of PBSA in the UK initially and in the Netherlands more recently, originated from the changing demand of students in housing, the increased interest from investors, and as policy makers felt the need to impose regulatory restrictions towards the traditional form of studentification: *Housing in Multiple Occupation* (Hubbard, 2009; Thomsen, 2004). Therefore, PBSA have become the backbone of many local authority policies seeking to disperse students in order to stimulate social mixing in studentified neighbourhoods and to pursue social diversity (Sage et al., 2012). The characteristics of PBSA as opposed to traditional forms of student housing lead to believe that there exist some fundamental differences in their external effects. While studies have been done into the social effects of PBSA, prior research into their economic effects has been limited to studentification in general. Furthermore, no academic precedent exists for research into the influence of PBSA on neighbouring property prices.

1.3 Research problem statement

The research aim of this study is to provide insights in the external effects of PBSA on residential neighbouring prices. The results of this thesis contribute to the broader literature on externalities and fill the gap on the differences between the economic effects of PBSA and traditional student housing. From this aim, the following research questions can be composed:

Central question: what are the external effects of PBSA on neighbouring residential property prices?

In order to formulate an adequate answer, this paper makes use of difference-in-difference variation on the hedonic pricing model. This is the same approach used in Van Duijn et al. (2014), though substituting industrial heritage transformations for PBSA developments. In short, the method compares residential transaction prices before and after completion of the PBSA, while controlling for various structural characteristics in addition to spatial and temporal fixed effects. In addition, it compares the transaction prices of homes within a treatment radius to a control group. This means that all transactions outside a range of 1 km are excluded from the analysis. Moreover, the study focuses exclusively on residential sales that occurred in six municipalities having research universities, during the period of 2008 through 2017.

1.4 Structure of this thesis

The remainder of this thesis is organized as follows. Chapter 2 describes the theoretical framework that serves as the foundation of the thesis. Chapter 3 continues by explaining the empirical methods applied in order to make adequate hedonic estimations. Chapter 4 continues with a description of the used datasets and required preparation in order to perform the hedonic analysis. The estimated results of the baseline and alternative specifications follow in chapter 5. Finally, chapter 6 concludes on the entirety of the thesis with a recap of the main conclusions in addition to a discussion on how the findings fit into existing academic literature.

2. THEORETICAL FRAMEWORK

What follows is a brief review about key-studies in the fields of urban economics, residential property valuation, and the geographies of studentification. The first paragraph provides a general introduction into the determinants of house prices. This is followed by an outline of the external influences of locational characteristics on property values. Next, an emphasis is made on the role of studentification and more specifically PBSA on the housing market. The chapter is concluded by listing the relevant hypotheses that guide the rest of the thesis.

2.1 House price dynamics

There exists broad consensus regarding the homogeneity of housing as a good. In other words, as each house is made of a unique set of characteristics, possible variations exists in its physical – such as house type, age, size and quality – and locational characteristics. These unique conditions affect the consumer’s willingness to pay for a certain house and subsequently the value of this property (Sirmans et al., 2005). Additionally, certain physical characteristics may have different values based on its geographical properties. A swimming pool or central air conditioning may for instance be of higher value in warmer climates. Moreover, the attributed value of the same given set of characteristics may vary due to unique utility functions of homebuyers (Sirmans et al., 2005). A certain homebuyer may for instance place a greater value on a fireplace than another.

These intertwining characteristics imply that valuation is not predefined, but inherently a fuzzy business. Nonetheless, many studies making use of hedonic modelling have (successfully) attempted to explain the value of housing by valuing these individual components (Wilkinson, 1973; Rosen, 1974; Li and Brown, 1980; Malpezzi, 1996; Van Duijn et al., 2016). In regards to the valuation of location characteristics specifically, there have been hundreds of studies that have contributed over the years. These vary for example from studies into the effect of natural space on nearby property prices (Daams et al., 2016) to that of wind turbines (Dröes and Koster, 2016).

2.2 Academic origin and contemporary use of locational externalities

As early as the 19th century, economists such as Ricardo (1817) and Von Thünen (1826) have made an attempt at defining the value of land based on its location. Von Thünen defined this relationship conceptually as a product of its distance to the centre of the town market (Von Thünen, 1826). His theory of the *bid-rent curve* can roughly be described as the varying maximum willingness between different types of users, to pay a rent premium for their proximity to the market. Over a century had passed when urban economists as Alonso (1964) and Mills (1967), built upon his location theory by replacing the town market by a central business district (CBD). In this *monocentric city model*, it were residents and firms that compete for land close to the heart of economic activity, instead of the different types of agricultural uses that defined location in Von Thünen’s (1826) model.

More recently, urban economists have stepped away from the view that location and land-use patterns are mostly determined by the production side of a city. Nowadays, theories prevail that explain this spatial allocation – of both firms and households alike – based on the presence of urban amenities such as parks, monuments, or even restaurants and theatres (Brueckner et al., 1999; Glaeser et al., 2006; Florida, 2001; Clark, 2002). This is based on a shared assumption that households try to maximise consumption of urban amenities that are considered to be favourable. In practice, not all households can benefit equally due to limitation of space surrounding positive amenities. As a consequence, more affluent households have better means to compete for homes that are in proximity to many of these amenities. Roughly speaking this implies that that there exists a certain willingness to pay for the proximity to an amenity, which is then reflected in property prices. This price increase can be seen as an (unintentional) external consequence of a specific environmental attribute, usually referred to as *externalities* in academic literature (Li and Brown, 1980; Malpezzi, 1996).

Equally important to positive amenities in price determination are external factors that negatively affect demand, and subsequently house prices. These are defined as negative externalities. Negative externalities often follow from amenities belonging to the group of *Not In My BackYard's* (NIMBY's). An example is that of traffic noise due to the presence of an airport or highway, which may impact transaction prices by a discount of up to 12 percent (Theebe, 2004). From a socio-economic perspective, negative externalities may for example arise from neighbourhood characteristics like the ethnic composition, share of poverty, (un)employment rates and amount of crowded dwellings (Cao and Cory, 1982). Especially the latter two are interesting for the research purposes of this thesis, as these correspond with characteristics of traditional student housing. The next paragraph will continue the discussion on this topic.

2.3 Studentification and the housing market

Traditionally, student housing in the UK was facilitated on university campuses. However, these could not keep up with the rapid growth of the student population since the beginning of the 90s. This led to the equally rapid student colonisation, and subsequently physical, social, economic and cultural transformation, or *studentification,* of established residential neighbourhoods in university towns and cities (Smith, 2005).

This growth was made possible by the profitability for private investors to buy-to-let to students, combined with limited regulation and low interest rates (Gibb, 2005; Leyshon and French, 2009; Smith and Holt, 2007). In these cases, houses that previously held a singly family household were bought to convert into Housing in Multiple Occupation (HMO) specifically aimed at students. The driving force behind this phenomenon is similar to that of gentrification, and described in academic literature as the *rent-gap* (Smith, 1987). In short, it describes the imbalance between the income that a property can generate in its current use and the income that can be achieved potentially. It is then the increased demand from students combined with the potential overvalue of letting to students through HMO which sparks investors interest.

Another argument as to why studentification would lead to a rise in residential property values is given by Smith (2005). This study concludes that studentification is a form of gentrification due to the increased social and spatial polarization through displacement of existing residents and due to the rising economic investment in properties. The exception is that it is often coupled with a physical downgrading of the built environment, which is in contrast to other forms of gentrification. Other important disparities of studentification versus gentrification are the displacement of the more affluent households instead of lower-income households. Groves et al. (2002) and Bark and Powell (2008) share the notion that increased studentification leads to property investment in these neighbourhoods.

Short-term problems associated with student HMO include issues with noise disturbances, street cleaning, car parking, and a rise in burglary rates (Hubbard, 2008). In the long-run, the persistent increase of the share of students in a neighbourhood leads to depopulation of families with children and broader seasonal depopulation during the holidays and the weekends. This phenomenon is therefore often associated with the abolishment of family-oriented amenities, leading to segregation of the neighbourhood. As a result, long-term residents thus increasingly vacate such an area, after which properties may then come available at a discount. Landlords subsequently capitalize by purchasing the vacated property to offer to other students (Sage et al., 2012). As this trend continues without regulatory intervention, this can lead to a self-sustaining cycle with further studentification as result. Lofland (1968) and Kenyon (1997) agreed on a similar notion, claiming that a rise of students in a neighbourhood would lead to impoverishment and neglect. This would result in strong price decreases of houses in what they called *student-* or *youth ghettos.* Bark and Powell (2008) on the other hand assert that this effect will stagnate and reverse when an area becomes saturated with student residents.

2.4 The (alternative) impact of PBSA

While literature on the effects of studentification is extensive, that regarding PBSA is not. Existing studies have mainly focused on the disparities between these types of student housing and preferences from the perspective of students. In general, differences are found in physical characteristics, regulations, and actors involved (Sage et al., 2013). Alternative impacts would especially arise in the case of urban regeneration processes, when PBSA developments occur on brownfield sites.

In practice, however, similarities are especially expected between HMO and PBSA in regards to negative effects. This statement follows from a survey study in Brighton, which concluded that local residents perceive PBSA negatively due to problems arising from litter, parking and noise-nuisance (Sage et al., 2013). Noise nuisance and anti-social behaviour occurring around the PBSA partly originates from the misuse of alcohol and drugs. Residents consume large quantities of alcohol in the PBSA before going out, resulting in nuisance when students leave and return to the complex. In addition, evidence is found for social withdrawal and sense of ‘imprisonment’ as natural reaction to retreat into homes from on-street disturbances.

Finally, it is easier for PBSA to be tied to a specific location in contrast to studentification of a neighbourhood. This is possible as PBSA usually form a single built entity, whiles a traditionally studentified neighbourhood through HMO is permeated by houses that cohesively affect the character of the neighbourhood. This also has implications from an analytical point of view. From these considerations we can conceptualise PBSA as an amenity in the neighbourhood that produces localized externalities, hence influencing neighbouring house prices as a locational characteristic.

2.5 Hypothesis

From the line of reasoning set out above, PBSA are conceptualized as localized amenities that may produce externalities. Expectations are mostly ambiguous in regards to how these externalities affect residential property prices. Price increases may result from (long-term) processes similar to gentrification or urban regeneration. A discount is also very likely, however, due to perceived negative effects arising from (short-term) noise nuisance and other anti-social behaviour. The ambiguity of the assumptions results in the following null hypothesis that is tested empirically:

H_0: The development of a PBSA does not have an effect on the sale price of nearby residential properties.

3. METHODOLOGY

In order to test the central hypothesis discussed in the previous section, this study makes use of a hedonic pricing model. The following paragraph starts with a brief introduction into the use of hedonic pricing models and their relevance in indirect valuation of property characteristics. Paragraph 3.2 subsequently describes the baseline specification of the difference-in-difference approach aimed at estimating the external effects of PBSA on surrounding residential property prices.

3.1 Hedonic Analysis

Hedonic price modelling for statistical economic analysis has been formalised in the study by Rosen (1974). This study provided a theoretical framework for measuring the marginal value of the characteristics of a heterogeneous good. The underlying assumption is that heterogeneous goods can be decomposed into individual characteristics with an implicit price attached to them. Consequently, even though the price of the heterogeneous good is known, the valuation of individual characteristics – at least those that are not traded on an explicit market – must be done indirectly.

The hedonic method is particularly relevant for economic property analysis, as the value of individual physical and locational components can often not be appropriately observed. Also, dwellings are heterogeneous goods that comprise of a bundle of attributes, each adding to the value of a property (Malpezzi et al., 1980). This is why there exists a wide range of studies that have attempted at measuring the added value of specific structural as well as locational characteristics of a dwelling through hedonic modelling (Palmquist, 1992; Palmquist, 2005; Sheppard, 1999; Sirmans et al., 2005; Dröes and Koster, 2016). The latter is the main focus in this thesis, as PBSA – like other localized externalities – would theoretically only affect those in proximity to the externality. Structural characteristics will not be forgotten, however; as these largely determine the price of a house and thus support in pinpointing the effect of the externality.

3.2 Empirical model

To determine if a relationship exists between PBSA and surrounding housing prices, the difference-in-difference specification of a hedonic regression model has been used. This method has proven its worth in studies by Schwartz et al. (2006) and Van Duin et al. (2014; 2016), providing a framework with which to estimate the differences between residential property values within a certain range of a PBSA (target group) and those that are not expected to be influenced by it (control group). Interaction variables are specifically added to the model to distinguish between observations receiving treatment and those that do not. Or in other words: between the effect after and prior to the completion of the PBSA. This is different from a 'naïve' linear approach, which would generally result in one coefficient; capturing the average effect of the externality depending on the average distances (Van Duijn et al., 2014).

Aside from the independent variables aimed at measuring the external effect of PBSA, other variables are added to control for variation in structural characteristics and (pre-)existing price trends. The latter is done similarly to the baseline specification of Schwartz et al. (2006) and Van Duijn et al. (2014, 2016), by including spatial and time fixed-effects variables. This means that observations are grouped based on their spatial and temporal coherence amongst each other in order to reduce omitted variable bias. In other words, year dummies are added to capture the trend of price increases over time. The resulting baseline specification of this research is then defined as:

$$\log(P_{int}) = \alpha + \sum_{s=1}^{S} \beta_s R_{itrs} + \sum_{s=1}^{S} \gamma_s R_{itrs} D_i + \sum_{s=1}^{S} \delta_s R_{itrs} D_i^2 + \sum_{k=1}^{K} \varphi_k X_{kit} + \sigma_t T_t + \vartheta_n N_n + \varepsilon_{it}$$

where P_{int} is the transaction price of residential property i located in postal code area n in year of sale t; α reflects a constant, R_{itrs} denotes a set of ring variables s which are dependent upon the location i, the year of sale t and the treatment radius r; D_i is the absolute distance between the property i and the PBSA, X_{kit} represents structural characteristics k of property i during transaction year t; dummies to control for time and spatial fixed effects have been added in T_t and N_n respectively, both taking the value one for year t and zip code area n, and zero otherwise; ε_{it} is the idiosyncratic error term and β_s, γ_s, δ_s, φ_k, σ_t and ϑ_n are the parameters estimated.

For *s*, three different ring variables are specified in order to capture the external effect of PBSA. Central to this thesis is *After*, which denotes the external effect after completion of a PBSA development and when students have taken residence. *Before* captures the external effect prior to completion. Finally, a trend variable is included for *s* that interacts with *After* to capture the possible linear decay of the effect over time. The value of this variable is calculated by the difference between the completion date of the PBSA and the date of sale of the residential property. No *Between* variable is included due to data incompleteness and difficulty in determining the construction commencement dates. As a result, the *Before* variable is susceptible to omitting possible external effects during construction, such as nuisance or anticipation (Van Duijn et al., 2016).

Finally, similar to the method used in Van Duijn et al. (2014, 2016), distance variables are included to capture the possible decay of the external effect over distance. Schwartz et al. (2006) referred to this as the distance gradient approach. Concretely, the set of ring variables are interacted with the absolute and quadratic distance to PBSA in meter. The latter is included in the model in order to determine if the decay is not linear, but rather takes a convex or concave form.

4. DATA

Chapter 4 starts with a concise description of the datasets and the preparation needed in order to conduct the hedonic analysis. The second paragraph continues with descriptive statistics illustrating the inherent differences that exist between observations in the target and control group. The chapter is concluded by an exploratory analysis into the observed differences of time-variant neighbourhood characteristics. These provide evidence to include additional exploratory model specifications that differentiate between over- and underrepresentation of these characteristics.

4.1 Study area and data cleaning

The data used in the hedonic analysis originates from three sources; NVM (The Dutch Association for Realtors), Savills Research and CBS (Statistics Netherlands). The first two sources combined provide information about recent residential sales and their spatial-temporal relationship to PBSA. CBS data enriches this dataset by adding, among other things, demographic and economic information about the neighbourhood. The remainder of this paragraph elaborates on the use of these sources and the preparation of the final dataset.

The entire database of the NVM covers 75% of all residential sales in the Netherlands (NVM, 2018). In general, this is not entirely distributed to researchers. Rather, a sample is provided based on spatial and temporal limitations imposed by the NVM; specifically to six municipalities and a ten-year period. This is not an issue, however, as PBSA developments are mostly limited to university cities – of which the Netherlands has twelve – and the majority have been realised in the last decade. The provided NVM dataset contains 61,212 transactions that have occurred in Groningen, Leiden, Wageningen, Delft, Eindhoven and Maastricht, between January 1st, 2008 and December 31st, 2017. The largest municipalities with a university, namely Amsterdam, Rotterdam and Utrecht, were deliberately excluded from the sample. This originates from the idea that these housing markets have been exposed to very high price increases in recent years, likely attributable to sentiment rather than market fundamentals and thus possibly affecting the empirical relationship between PBSA and property prices. The provided dataset includes a total of 70 variables. These contain information on the conditions of the transaction, such as the transaction price and date, in addition to a wide range of property characteristics including the location of the object. The selection of these variables has partly been based on studies by Daams et al. (2016) and van Duijn et al. (2016), who also made use of NVM datasets. An overview of these variables including descriptive statistics is listed in table II.

The dataset from Savills Research contains information about 64 unique PBSA developments that have been completed between 2007 and 2016 in the six municipalities mentioned above. Key variables are limited to the location of the PBSA, its construction year and the number of rooms. PBSA with fewer than 50 rooms were excluded from the analysis. The reason for this originates from the assumption that the external effects of such small scale developments would be negligible.

Both of these datasets have been geocoded through GIS-Software using RD-coordinates. This made it possible to compute the Euclidean distance between each house and the nearest PBSA, crucial in distinguishing between observations that fall within the target or control group. Key variables on the temporal interaction between both datasets were computed by comparing the transactions date of residential properties and the construction year of the PBSA closest to it. An overview of these new variables and their frequency is listed in table I.

Furthermore, spatial data from CBS has been used to provide an overview of the neighbourhoods in which the residential sales took place. On a yearly basis, CBS publishes *vierkantstatistieken* (literally square statistics) that divide the entire Netherlands in squares of 500 meter providing demographic and economic statistics. Again, GIS-software was used to merge the statistics of these units to the observations that fall within its geographical borders. As a result, house sales can be compared based on neighbourhood criteria, specifically on its share of young people, density, and average property values (WOZ)[1]. For the share of young people and population density it was of particular interest to obtain the value prior to completion of the PBSA. This indicates how the neighbourhood was before the arrival of students and if PBSA are built in different types of neighbourhoods. For this reason, CBS datasets between 2006 and 2016 have been used for these two variables. Property value data is only published sporadically, which is why the most recent dataset published in 2016 was used here.

Finally, the dataset has been trimmed by removing observations with missing values, outliers and implausible values on key variables[2]. In addition, some variables needed a log-transformation to satisfy the condition of normality. The full procedure of data preparation can be found in Appendix A[3]. Next paragraph continues with descriptive statistics of the variables used in the hedonic analysis.

4.2 Descriptive statistics

In order to perform the difference-in-difference approach, subsamples are created by assigning observations to a target and control group based on spatial and temporal criteria. In other words, house sales are divided by their distance to the nearest PBSA, if they occurred before or after realisation of the nearest PBSA and by additional interaction between these two. Table I illustrates the size of each subsample by listing frequencies of observations that meet these criteria. It also shows that 48.47% of the total observations (after data trimming) are excluded from the hedonic analysis as they occurred at a distance of over 1,000 meters. This makes it irrelevant for the scope of this thesis.

[1] WOZ refers to the valuation of immovable property act (*Wet Waardering Onroerende Zaken)*, which is assessed on a yearly basis for tax purposes.
[2] Total of 13,108 observations were dropped for implausibility in price or surface, in addition to missing values (11,713 observations for not having a bathroom).
[3] This also includes diagnostics that test for OLS assumptions of zero mean in errors, linearity in parameters, homoscedasticity, independence of error terms and normality of the error terms (Brooks and Tsolacos, 2010).

Table I | *Frequency table spatial and temporal criteria*

	Number of observations	*Percentage*	*Cum. percentage*
Temporal			
After	16,704	34.73%	34.73%
Before	31,399	65.27%	100.00%
Spatial			
0 – 100 meters (target)	588	1.22%	1.22%
100 – 1,000 meters (control)	24,200	50.31%	51.53%
> 1,000 meters	23,315	48.47%	100.00%
Interaction			
Target After	228	0.92%	0.92%
Target Before	360	1.45%	2.37%
Control After	8,707	35.13%	37.50%
Control Before	15,493	62.50%	100.00%

Table II | *Descriptive statistics*

		Target group *< 100 meters* (*N* = 588)		Control group *100 – 1000 meters* (*N* = 24,200)	
	Parameter	*Mean*	*Std. Dev.*	*Mean*	*Std. Dev.*
Transaction price	€	179,100.5	84,185.19	215,164.2	111,287.7
Distance to PBSA	meters	67.52	20.763	559.624	240.394
Structural characteristics					
Floor space	m^2	90.597	31.684	100.342	35.852
Quality Inside	1 = good	0.007	0.082	0.018	0.135
Quality Outside	1 = good	0.003	0.058	0.008	0.088
Balcony	1 = yes	0.643	0.480	0.457	0.498
Roof terrace	1 = yes	0.099	0.298	0.138	0.345
Central heating (CV)	1 = yes	0.951	0.217	0.933	0.249
Parking	1 = yes	0.134	0.341	0.179	0.384
Monument	1 = yes	0.060	0.237	0.030	0.170
Road Nuisance	1 = yes	0.114	0.318	0.055	0.229
Construction year					
1500-1905	1 = yes	0.044	0.206	0.091	0.288
1906-1930	1 = yes	0.160	0.367	0.180	0.384
1931-1944	1 = yes	0.107	0.310	0.142	0.349
1945-1959	1 = yes	0.111	0.314	0.085	0.279
1960-1970	1 = yes	0.163	0.370	0.176	0.381
1971-1980	1 = yes	0.150	0.357	0.094	0.292
1981-1990	1 = yes	0.066	0.249	0.069	0.253
1991-2000	1 = yes	0.097	0.296	0.077	0.266
>2001	1 = yes	0.102	0.303	0.086	0.281
Type of dwelling					
Simple	1 = yes	0.007	0.082	0.020	0.140
Single-family	1 = yes	0.077	0.266	0.292	0.455
Canal-house	1 = yes	-	-	0.010	0.098
Mansion	1 = yes	0.054	0.227	0.077	0.267
Farm	1 = yes	-	-	0.000	0.016
Bungalow	1 = yes	-	-	0.003	0.056
Villa	1 = yes	-	-	0.004	0.064
Country House	1 = yes	-	-	0.000	0.014
Estate	1 = yes	-	-	-	-
Downstairs house	1 = yes	0.121	0.326	0.105	0.306
Upstairs house	1 = yes	0.111	0.314	0.125	0.331
Maisonette	1 = yes	0.053	0.224	0.042	0.202
Flat (portiek)	1 = yes	0.270	0.445	0.215	0.411
Flat (galerij)	1 = yes	0.304	0.461	0.101	0.302
Hospice	1 = yes	-	-	0.000	0.021
Down- and upstairs house	1 = yes	0.053	0.224	0.042	0.202

**N* = number of observations

As discussed in 3.1., treatment occurs when both criteria *After* and *0 – 100 meters* are met. This is the key variable in the hedonic analysis, capturing the possible effect of the completion of a PBSA on nearby house prices. Only 228 observations meet these criteria, which may affect the reliability of the estimated results.

The descriptive statistics of the target and control group are stated separately in table II in order to make an adequate comparison. The figures show strong differences on virtually all variables. Most importantly, the mean transaction price is over €35,000 lower for observations that fall within the target group. This should not be misinterpreted by already attributing it to the proximity of PBSA. To begin with, houses that belong to the target group have on average 10% less floor space, while this is widely known to be an important determinant for house prices. Furthermore, houses in this group are on average of a lesser quality, have less parking and are exposed to more road nuisance. In addition, the type of dwellings within each group is substantially different. For example, the share of flats in the area receiving treatment is 57.4% versus 31.6% for the control group. These differences lead to believe that the target group is very different from the control group, although definitive conclusions cannot yet be made regarding the causal relationship between PBSA and neighbouring house prices. If these differences are not properly accounted for in the hedonic analysis, this could result in inconsistent estimates of the external effect due to omitted variable bias.

4.3 Exploratory analysis into time-variant neighbourhood characteristics

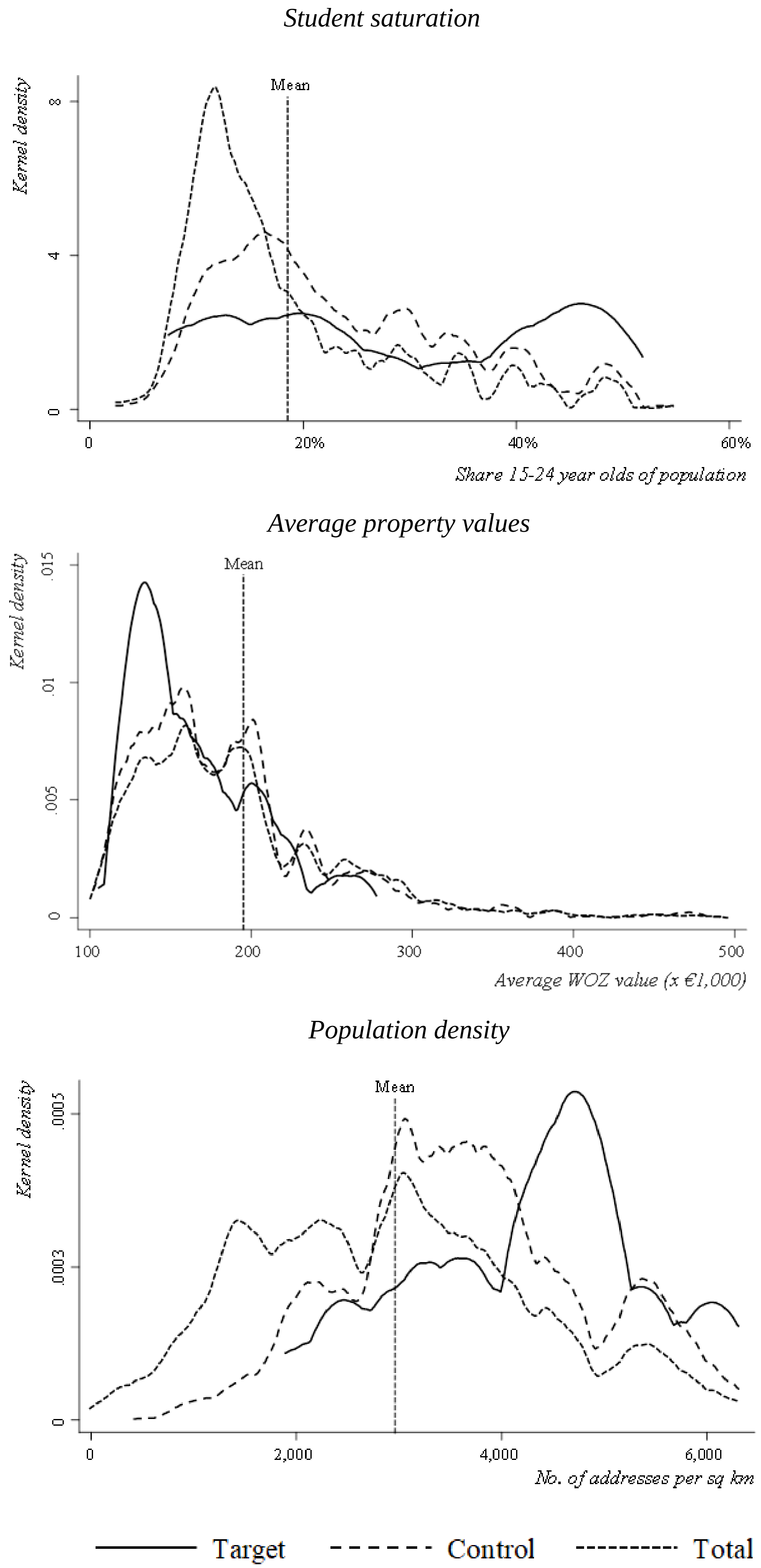
Another factor that accounts for the differences in transaction prices between the target and control group are local neighbourhood characteristics that are not fixed through time and fall outside of the scope of temporal fixed effects. Three criteria are used in this paper for comparison. More might be of importance, but as Van Duijn et al. (2016) discussed in their paper, it is difficult to cover all time-variant local characteristics as these are often not systematically registered. Time-variant neighbourhood characteristics are of importance as they provide insights in the type of neighbourhoods in which PBSA are developed. Moreover, results from the baseline specification could be (partly) driven by over or underrepresentation of these characteristics.

First of all, an area that is already *saturated* with students, a concept introduced by Bark and Powell (2008), might amplify or reduce the external effects of a PBSA. On the other hand, PBSA developments in poorer neighbourhoods might be seen as an extra negative impulse decreasing the willingness to buy a house there, or rather a revitalisation with a reversely increasing effect. Lastly, urban and buzzy environments area expected to be affected differently by the arrival of students than a formerly quiet and peaceful neighbourhood. Proxies have been used in order to measure the value of these criteria. These are respectively the share of young people (15-24) of the total population, average property values and number of addresses per square kilometre.

As discussed in section 4.1., each housing transaction is attributed a certain value for these three variables corresponding to its location within a unique geographic square of 500 meters. The distribution of observations based on these variables is visualised through kernel density estimations in figure I. These graphs distinguish between target group, control group and the total of observations, which includes those outside the boundaries of the control group. In addition, the mean value of the total of observations is indicated in each graph.

The graphs illustrate that PBSA are generally developed in areas with a substantially higher share of young people, lower average house prices and much higher population density. Observations in the control group show these same traits, though to a lesser extent. This gives substantial evidence to assume that the locations of PBSA are not random. This may thus affect the lower transaction prices of houses in the target group, but also the relationship between PBSA and house prices. Section 5.3. accounts for this through alternative specifications of the empirical model.

FIGURE I | Kernel density estimations of neighbourhood criteria



5. RESULTS

This chapter presents the regression results of the difference-in-difference approach which makes it possible to answer the central research question: *What are the external effects of PBSA on neighbouring residential property prices?* The chapter starts by an overview and interpretation of the estimation results for the baseline model, where the target group has been defined as being within 100 metres of a PBSA development. The next paragraph continues with sensitivity analyses that test the robustness of the model in the case of varying distances for the target group and varying scales of spatial fixed effects. This is followed by a section that accounts for the inherent differences between the target and control group, by using coarsened exact matching to find control observations that are similar to those in the target group. The chapter is concluded by regression estimates after splitting the dataset in subsamples based on its pre-existing student population. Additionally, the equality between these subsets is examined by a Chow (1960) test and its results are discussed.

5.1 Empirical model

Table III lists the baseline specification and illustrates the effect of gradually adding variables into the model on the coefficients, standard errors and R^2. Model (1) displays the results when the model is restricted to year fixed effects in addition to the temporal variables *After*, *Before* and *Trend After*. Model (2) adds structural characteristics, as well as the period in which the property has been built. Model (3) displays the preferred model, including spatial fixed effects that control for time-invariant neighbourhood characteristics. An R^2 of 0.86 for this model implies that the regression is able to explain around 86% of the variation in house prices.

The coefficient of the key variable *After* depicts the deviation in prices between the target area and control area after delivery of the PBSA. This coefficient is negative and significantly different from zero at the 1% level. This provides substantial evidence for a negative relationship between the realisation of a PBSA and neighbouring residential property prices. Concretely, the coefficient of *After* indicates that after completion of a PBSA, homes that are at a distance of 9 meter (the smallest distance in the sample) would sell at a discount of 20.8%[4] in regards to a comparable property in the control area. This is including the interaction with the distance variables that are both significant from zero at the 5% level. This implies that the negative external effects decay concavely over distance. Combining these coefficients implies that no further decay of the effect is noticeable after 71 meter and properties experience a 1.8%[5] discount on their transaction price as result of the presence of a PBSA. The effect of spatial decay is illustrated in figure II.

Like discussed in section 3.2, trend variables were included with the purpose of estimating the influence of time on the empirical relationship between PBSA and neighbouring property prices. All

[4] =(EXP-0.3005803 + 0.0079325 * 9 - 5.56E-05 * 9^2)-1)*100
[5] =(EXP-0.3005803 + 0.0079325 * 71 - 5.56E-05 * 71^2)-1)*100

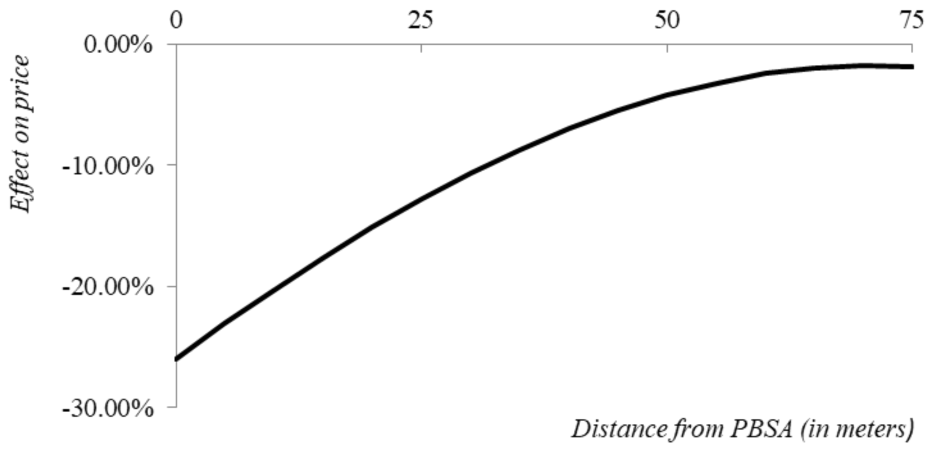
coefficients of the trend variables after completion of a PBSA are significant at the 10% or 5% level. Moreover, all are smaller than their ‘generic’ counterpart and inversely related. In other words, the coefficient of *Trend After* is smaller than *After*, and it is positive while the latter is negative. This means that the negative effects of the externality increasingly decay over time. This is interesting, as this would imply that people initially disregard PBSA resulting in a lower willingness to pay for nearby houses, but lose this negative tendency in only a few years. This could be attributed to certain prejudices that fade over time.

Table III | *Estimation results of the baseline specification*

	(1)		**(2)**		**(3)**	
Sample size	<1000 m		<1000 m		<1000 m	
Target area	0 – 100 m		0 – 100 m		0 – 100 m	
Control area	100 – 1000 m		100 – 1000 m		100 – 1000 m	
After	-0.75537***	(0.2479104)	-0.2261179	(0.1587595)	-0.3005803***	(0.1037431)
After Distance	0.01481*	(0.0088442)	0.0029481	(0.0050434)	0.0079325**	(0.0034748)
After Distance^2	-8.40e-05	(7.07e-05)	-7.62e-06	(3.79e-05)	-5.56e-05**	(2.71e-05)
Before	-0.93803***	(0.2807669)	-0.3436564*	(0.1496011)	-0.2166858**	(0.1039431)
Before Distance	0.01803**	(0.0090656)	0.0059702	(0.0047662)	0.0049313	(0.0034065)
Before Distance^2	-0.00010	(6.78e-05)	-2.4e-05	(0.0000355)	-2.7e-05	(2.59e-05)
Trend After	-0.16646	(0.1195111)	0.0052823	(0.0452695)	0.0623991**	(0.0313321)
Trend After Distance	0.00723*	(0.0040241)	0.0002148	(0.0014827)	-0.0018701*	(0.0010697)
Trend After Distance^2	-6.11e-05*	(3.14e-05)	-2.22e-06	(1.15e-05)	1.54e-05*	(8.62e-06)
Year Fixed effects (9)	*Yes*		*Yes*		*Yes*	
Structural characteristics (9)			*Yes*		*Yes*	
Building period dummies (8)			*Yes*		*Yes*	
ZIP4 Fixed effects (76)					*Yes*	
Observations	24,788		24,788		24,788	
Adjusted-R^2	0.0409		0.7868		0.8590	

Note: ***p<0.01, **p<0.05, *p<0.10. Robust standard errors in parentheses. Dependent variable is log (transaction price). An elaborate overview of the coefficients of the control variables for specification (3) can be found in Appendix B.

FIGURE II | *Distance decay of external effects PBSA after completion*



Isolated, however, the coefficients of *After* give a distorted image in regards to measuring the effect external effect of PBSA on property prices. The coefficient of *Before* also needs to be taken into account, as this captures the possible external effect of the site prior to completion of the PBSA. Of these coefficients, only *Before* is significant at a reliable interval (95%). The combination with the (non-significant[6]) interaction variables imply a discount of 16.0% at a distance of 9 meters compared to properties in the control group. It is then the difference between *After* and *Before* that reflects the external effect of the PBSA. This amounts to a discount of 4.8% for the property at 9 meters.

To conclude this section, I elaborate on the possible explanations for the negative external effect prior to completion of the PBSA. First, as commencement dates for construction of PBSA are unknown and therefore excluded from this paper, it is possible that the effects during the construction period are omitted within this coefficient. The cause may originate from anticipation effects, as discussed in the paper by Van Duijn et al. (2016) though substituting urban revitalisation for the arrival of students. Another cause may be nuisance from the construction site itself, although there is little academic precedence for the existence of such negative externalities. Secondly, it is possible that a significant majority of PBSA was developed on sites that were previously also regarded as *disamenities*. This is not unlikely, as PBSA are often developed in vacant or obsolete real estate, such as offices or industrial property. Would this be the case, then prices might have gone down regardless of the development of the PBSA. This would subsequently have implications for the validity, as the estimated coefficients would not represent the phenomenon that is under investigation. This emphasizes the importance of robustness tests that follow in the next paragraphs.

5.2 Sensitivity analyses

This section will discuss the sensitivity analyses that have been performed in order to test the robustness of the model. This is vital, as estimation results are heavily dependent on selected variables. First, changes in the treatment radius are reported in table IV. Next, the effects of varying scales of spatial fixed effects are displayed in table V. Both are discussed in the remainder of this paragraph.

The initial treatment radius of 100 meters was determined mainly on the assumption that the externality would be particularly apparent on the micro-scale. This is in line with Sage et al. (2013) and Hubbard (2008) which particularly associate perceived negative effects due to noise-nuisance. This is why two regressions have been performed under the same conditions as the preferred model (3) but with alternating distances below and above 100 meters. Model (4) reports the estimation results with a treatment radius of 50 meters. In this case, almost all coefficients become inverse, inflated and substantially less significant. There is a plausible explanation for this, as the results from model (3)

[6] These are included to compare the effect before and after the completion. However, non-significance also implies that no definitive conclusions can be made in regards to the possible distance decay of the *Before* coefficients.

already indicated that the effect cannot be fully captured within the 50 meters. On the other hand, estimation results for model (5) are slightly more significant than for model (3), though with smaller coefficients. This implies that the model with the target distance set at 250 meters is slightly better able at capturing the external effects, but also results in lower coefficients. This means that there is evidence that the effect may actually be lower than previously assumed in 5.1. Again, this should be nuanced as the differences are small. More importantly, however, is that it strengthens the validity of the model by providing further evidence for an empirical relationship between PBSA and neighbouring house prices.

Table IV | *Estimation results of alternative specification of distance to PBSA*

	(4)		**(5)**	
Sample size	<1000 m		<1000 m	
Target area	0 – 50 m		0 – 250 m	
Control area	50 – 1000 m		250 – 1000 m	
After	0.318005	(0.2765092)	-0.118335***	(0.0373065)
After Distance	-0.0352643*	(0.0206106)	0.0017986***	(0.0005452)
After Distance^2	-0.0006169*	(0.0003275)	-5.79e-06***	(1.80e-06)
Before	0.6232243	(0.4082011)	-0.0446397**	(0.0209488)
Before Distance	-0.0404282**	(0.0205763)	0.0004351	(0.002991)
Before Distance^2	0.005638**	(0.000257)	-7.27e-07	(09.79e-07)
Trend After	-0.033106	(0.0699054)	0.0247403**	(0.0123176)
Trend After Distance	0.0057353	(0.0059967)	-0.0003177*	(0.0001896)
Trend After Distance^2	-0.0001145	(0.0001018)	-1.12e-06*	(6.49e-07)
Year Fixed effects (9)	*Yes*		*Yes*	
Structural characteristics (9)	*Yes*		*Yes*	
Building period dummies (8)	*Yes*		*Yes*	
ZIP4 FE (76)	*Yes*		*Yes*	
Observations	24,788		24,788	
Adjusted-R^2	0.8597		0.8597	

Note: ***p<0.01, **p<0.05, *p<0.10. Robust standard errors in parentheses. Dependent variable is log (transaction price). An elaborate overview of the coefficients of the control variables for specification (4) and (5) can be found in Appendix B.

Furthermore, uncertainty exists in the appropriate scale of spatial fixed effects. Unobserved factors that are static across a certain spatial unit may cause omitted variable bias if these are not accounted for in the regression (Heintzelman & Tuttle, 2012). Omitted variable bias is best controlled at the smallest geographical grouping of the observations, though this also results in less variation within the groups. Concretely, taking a grouping or scale that is too small may cause the coefficients of the key variables to lose part of their statistical power. To illustrate the robustness of the initial findings, models (6) through (10) report the estimation results when restricted by respectively no spatial groupings down to street level fixed effects. A treatment radius of 250 meters was used for these estimations, as this specification showed to be slightly better capable at explaining the variance in transaction prices.

Models (6) through (8) illustrate how coefficients become smaller with each smaller spatial scale of fixed effect, thus reducing omitted variable bias. Model (9) and (10) report something remarkable, however, as coefficients seem to increase despite the increase in number of groupings for fixed effects. A possible explanation could be that street-level spatial fixed effects are better apt at capturing the externality of PBSA. Recall that problems with PBSA are likely to originate from noise-nuisance and anti-social behaviour, which is also likely to appear on the street level. What this also means is that when observations are grouped by street to control for time-invariant characteristics, variation of the important explanatory variables remains possible within these groups. The important coefficients of model (10) are all significant on the 1% level and the model explains 91.1% of the observed variation in transaction prices. .

Table V | *Estimation results of alternative specification based on spatial Fixed Effects (FE)*

	(6)	**(7)**	**(8)**	**(9)**	**(10)**
Sample size	<1000 m	<1000 m	<1000 m	<1000 m	<1000 m
Target area	0 – 250 m	0 – 250 m	0 – 250 m	0 – 250 m	0 – 250 m
Control area	250 – 1000 m	250 – 1000 m	250 – 1000 m	250 – 1000 m	250 – 1000 m
After	-0.169*** (0.0524)	-0.127395*** (-0.0385733)	-0.1183345*** (-0.0373065)	-0.1828624*** (-0.0381968)	-0.2689055*** (-0.0389908)
After Distance	0.00222*** (0.000739)	0.0021324*** (-0.0005621)	0.0017986*** (-0.0005452)	0.0023218*** (-0.0005509)	0.0037876*** (-0.0005431)
After Distance^2	-6.75e-06*** (2.41e-06)	-6.94e-06*** (-1.86e-06)	-5.79e-06*** (-1.80e-06)	-7.22e-06*** (-1.81e-06)	-0.0000116*** (-1.75e-06)
Before	0.135*** (0.0277)	-0.0819201*** (-0.0209363)	-0.0446397** (-0.0209488)	-0.1171361*** (-0.0210389)	-0.178559*** (-0.0285368)
Before Distance	0.00124*** (0.000392)	0.0009621** (-0.0003002)	0.0004351 (-0.0002991)	0.0011094*** (-0.0002976)	0.0023547*** (-0.0003636)
Before Distance^2	-2.82e-06** (1.27e-06)	-2.21e-06* (-9.86e-07)	-7.27e-07 (-9.79e-07)	-2.62e-06** (-9.75e-07)	-6.66e-06*** (-1.12e-06)
Trend After	0.0187 (0.0161)	0.0318296* (-0.012408)	0.0247403** (-0.0123176)	0.0186999 (-0.012246)	0.0417773*** (-0.0109235)
Trend After Distance	-0.000246 (0.000237)	-0.000404* (-0.0001922)	-0.0003177* (-0.0001896)	-0.0003057 (-0.0001855)	-0.0005887*** (-0.0001697)
Trend After Distance^2	7.98e-07 (8.02e-07)	1.40e-06* (-6.59e-07)	1.12e-06* (-6.49e-07)	1.11e-06 (-6.29e-07)	1.94e-06*** (-5.81e-07)
Year Fixed effects (9)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Structural characteristics (9)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Building period dummies (8)	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
District FE (42)		*Yes*			
ZIP4 FE (76)			*Yes*		
Neighbourhood FE (212)				*Yes*	
Street FE (2,423)					*Yes*
Observations	24,788	24,788	24,788	24,788	24,788
Adjusted-R^2	0.7866	0.8505	0.8597	0.8738	0.911

Note: ***p<0.01, **p<0.05, *p<0.10. Robust standard errors in parentheses. Dependent variable is log (transaction price). An elaborate overview of the coefficients of the control variables can be requested at the author.

5.3 Matching of the target and control group

Section 4.2 and 4.3 illustrated that there exist strong differences between the target and control group. If this imbalance is not properly addressed, this may lead to problems with causal inference (Holland, 1986). Or in other words; correlation between PBSA and a discount on nearby property prices in the last two sections may have wrongfully be interpreted as causation. To address this issue, many studies advocate the use of matching methods to make the treatment group as similar as possible to the control group (Ho et al., 2007). For example, the study by Van Duijn et al. (2016) – which applied a similar use of the differences-in-differences specification – made use of *propensity scores matching* method (PSM). Although this method has been very popular in academic research, relatively new studies have found evidence that PSM often results in contradictory results, with an increase of imbalance between the treatment and control group (King et al., 2016).

It is for this reason that I apply the *coarsened exact matching* (CEM) method as described by Iacus et al. (2012) to reduce the imbalance between the target and control group. The advantage of CEM compared to PSM is that no probit or logistic estimation step is required. Instead, this method matches observations with the same values for all the coarsened variables and puts them in a single stratum. The three neighbourhood characteristics of section 4.3 – student saturation, population density and average house value – have acted as the variables based on which the initial matching took place. Next, the large differences between the types of properties – especially in terms of amount of flats such as discussed in 4.2 – have been addressed. Observations that did not match were then excluded from the estimation, hence decreasing imbalance between the treatment and control group. In my study, CEM has been applied with the treatment radius set on 100 meters. Matching has subsequently been confined to the existing control group of up to 1,000 meters from PBSA. Moreover, determination of bin size was done through the Freedman-Diaconis rule, which makes use of the interquartile range thus making it less susceptible to outliers in the data.

The initial CEM procedure resulted in 19,313 observations not being matched[7]. Table VI reports the results of a similar specification to model (3), apart from having a different control area based on the CEM. Some results from estimation (11) are important to highlight in comparison to their counterparts from (3). The coefficients indicating the negative external effects appear to be stronger and more significant for both the *before* and *after* coefficients, including the distance interactions. Using the same log-transformation as in 5.1 yields that the closest home to a PBSA (D=9) would then sell at a discount of 20.01% before completion of the PBSA and at 26.29% after. The difference between these two also increases, indicating that the effect of the PBSA is actually stronger when controlling for neighbourhood characteristics. On the other hand, all coefficients for the trend variables become smaller and lose their significance.

[7] These all correspond to the control group. All observations (588) in the treatment group could be matched, however, thus not increasing possible bias due to decrease in sample size.

Finally, model (12) shows the results after including the property type matching criteria into the estimation. These results indicate a further increase of the coefficient of *after*, indicating that the effect becomes stronger when controlling for the type of property. In addition, all the *before* coefficients lose their significance. This implies that no definitive conclusions can be made in regards to the negative effect of the site prior to completion of PBSA. Conclusively, the alternative specifications with the control group through CEM procedure strengthens the validity of the external effect of PBSA on nearby property prices found in the main analysis.

Table VI | *Estimation results of the matching specification*

	(11)		**(12)**	
Sample size	<1000 m		<1000 m	
Target area	0 – 100 m		0 – 100 m	
Control area	CEM		CEM including property type	
After	-0.391***	(0.101)	-0.475***	(0.147)
After Distance	0.0102***	(0.00339)	0.0131***	(0.00482)
After Distance^2	-7.17e-05***	(2.66e-05)	-9.14e-05***	(3.72e-05)
Before	-0.273***	(0.101)	-0.117	(0.119)
Before Distance	0.00581*	(0.00325)	0.00163	(0.00372)
Before Distance^2	-3.10e-05	(2.46e-05)	-2.06e-06	(2.73e-05)
Trend After	0.0503	(0.0323)	0.0667	(0.0449)
Trend After Distance	-0.00154	(0.00112)	-0.00165	(0.00149)
Trend After Distance^2	-1.26e-05	(9.17e-06)	-1.13e-05	(1.18e-06)
Year Fixed effects (9)	*Yes*		*Yes*	
Structural characteristics (9)	*Yes*		*Yes*	
Building period dummies (8)	*Yes*		*Yes*	
ZIP4 FE (76)	*Yes*		*Yes*	
Observations	5,325		3,052	
Adjusted-R^2	0.8683		0.8407	

Note: ***p<0.01, **p<0.05, *p<0.10. Robust standard errors in parentheses. Dependent variable is log (transaction price). An elaborate overview of the coefficients of the control variables for specification (11) can be found in Appendix C.

5.4 Are results driven by saturation of students in the neighbourhood?

Paragraph 4.3 illustrated that PBSA are usually developed in areas with a high share of people between the ages of 15 through 24. Theory has suggested that saturation of students[8] within a neighbourhood may have a considerable impact. In addition, a parallel can be drawn with *carrying capacity* in environmental studies. Here, the concept indicates the maximum size of the population of a species that the environment can sustain based on the presence of resources. For my study, however, I define the concept as the capacity of a neighbourhood to sustain an overrepresentation of a subgroup within the population. If property prices would experience a higher negative external effect in a neighbourhood with an overrepresentation of the student population, this would then be conceptually defined as exceeding its carrying capacity.

[8] The share of young people (age 15-24) is used as proxy for the student population.

In order to explore if results are thus (partly) driven by the share of students in a neighbourhood, observations are grouped in below and above average[9] subsets for the share of 15-24 year olds. Subsequently, separate regressions have been performed under the same conditions as for model (5)[10]. The results of these separate estimations are reported in table VII and indicate stark differences between the subsets.

First and foremost, the negative external effect of the PBSA is much more apparent in a neighbourhood with a below average pre-existing student population. Compared to the pooled estimation of model (5), the coefficients for *after* become more than twice as large, while those for *before* stay approximately the same. The interpretation of the increased gap between *after* and *before* indicates a discount of 12.0% for model (13) versus 5.8% for model (5) (D=9).

On the other hand, neighbourhoods with above average student populations appear to lose virtually all negative external effects from the PBSA specifically. The same difference between *after* and *before* for the property closest to the PBSA (D=9) yields a discount of 1.0% for model (12). In addition, the Adjusted-R^2 indicates that both models seem slightly better able at explaining the variance in transaction prices.

Table VII | *Estimation results of alternative specifications based on neighbourhood characteristics*

Neighbourhood characteristic	**Student Saturation**	
	(12)	**(13)**
Subsample	>average	<average
Target area	0 – 250 m	0 – 250 m
Control area	250 – 1000 m	250 – 1000 m
After	-0.1043978** (-0.0517806)	-0.2216951*** (-0.0486971)
After*Distance	0.0013623* (-0.0007315)	0.0041315*** (-0.0007719)
After*Distance^2	-3.94e-06* (-2.38e-06)	-1.41e-05*** (-2.61e-06)
Before	-0.0890295*** (-0.0310164)	-0.0595172** (-0.0294402)
Before*Distance	0.0008764** (-0.000414)	0.0009785** (-0.0004813)
Before*Distance^2	-1.78e-06 (-1.31e-06)	-2.61e-06 (-1.65e-06)
Trend After	0.0224411 (-0.0147943)	0.0073444 (-0.0182255)
Trend After*Distance	-0.0002514 (-0.0002231)	-0.0001496 (-0.0003409)
Trend After*Distance^2	7.70e-07 (-7.62e-07)	8.47e-07 (-1.22e-06)
Year Fixed effects (9)	*Yes*	*Yes*
Structural characteristics (9)	*Yes*	*Yes*
Building period dummies (8)	*Yes*	*Yes*
Postal code (4-digit) FE (76)	*Yes*	*Yes*
Observations	14,579	10,209
Adjusted-R^2	0.8604	0.8629

Note: ***p<0.01, **p<0.05, *p<0.10. Robust standard errors in parentheses. Dependent variable is log (transaction price). An elaborate overview of the coefficients of the variables can be requested at the author.

[9] Average for total observations is 18.6% (15-24 year olds)

[10] Treatment radius of 250 meters is preferred, as as splitting the dataset resulted in very few observations in the case of a treatment radius of 100 meters, which could lead to bias in the estimations results.

The differences between the subsets are verified by the results of a Chow (1960) test for equality in intercepts and slopes of these subsamples. The *F*-statistic is significantly different from zero at the 1% level, meaning that the null hypothesis is rejected and that coefficients are not identical between the subsets. The full results of the Chow test can be found in Appendix D.

6. DISCUSSION & CONCLUSIONS

6.1 Discussion

Thus far, there was no academic precedence that aimed at estimating the value of possible externalities arising from PBSA through hedonic analysis. Academic relevance does not only arise from being the first of its kind, however. This method is also of interest as it provides new research insights into the capitalisation of (certain aspects of) studentification in residential property prices. This is party possible due to the localised aspect of PBSA, which enables to tie a substantial student population to a specific object in space, rather than estimating on the neighbourhood level.

Empirical research conducted in this thesis seems to validate the existence of negative external effects of PBSA developments on neighbouring residential property prices. This is mainly apparent on a micro scale, with an initial discount of 20.8% for homes closest to it. Evidence is also found for a relatively fast spatial decay, being concavely related to distance and thereby losing virtually all its effect on house prices after 71 meter. The performed sensitivity analyses illustrate that these figures are not set in stone. However, the negative relationship between PBSA and house prices is somewhat validated by consistency throughout the different models in this paper, particularly in regards to the sign of the coefficients at stake and their significance, and – to a lesser extent – the size of these coefficients.

An initial suggestion for the origin of the negative sign and short range would be nuisance from noise or other anti-social behaviour as discussed by Sage et al. (2013). The results of the trend variables, however, also indicate that the effect is decaying over time; losing virtually all its power in a span of approximately five years. This could imply several things. If the negative effects do originate from noise nuisance, it could be that these become less severe or are experienced differently. The plausibility of both is limited, although the latter may occur due to a change of the type of households that take residence compared to before the completion of PBSA. Another suggestion is the existence of an initial negative tendency towards the PBSA, originating from a negative attitude towards living nearby students but fading away as initial prejudices are not fulfilled.

Furthermore, the negative and significant *Before* coefficients implied that PBSA are developed on sites with inherently different characteristics, which may in advance already be portrayed as *disamenities*. This has also partly instigated the alternative explorations that provide evidence in regards to how certain time-variant preconditions of the neighbourhood affect the severity of the impact of PBSA on house prices. Amplification of the externality occurred for example in neighbourhoods with above average population density. In these neighbourhoods there is less room for a *buffer* between the PBSA and other properties, thus further strengthening the notion that the effect originates – at least to some extent – from noise nuisance as described by Sage et al. (2013). A counter intuition, however, is that high density neighbourhoods are usually also associated to more buzzy and lively environments, hence logically resulting in less of a contrast between the situation before and after the construction of a PBSA and arrival of students. On the other hand, poorer neighbourhoods – proxied by below average house values – experience a higher impact. To some extent this can be associated with the price decreases experienced in student- or youth ghetto's as described by Lofland (1967) and Kenyon (1997). But this development also fits within the statement of Smith (2005) that states that mainly less affluent households will take their residence in such neighbourhoods hence having a downward effect on residential property prices.

These results are also relevant for policy makers involved in the spatial allocation of PBSA. Macintyre (2003) stated for example that students expressed a desire to live in neighbourhoods already populated with students. The findings from 5.3 indicate that neighbourhoods saturated with students experience less of an impact from PBSA than those with below average share of youth. An answer could be to embrace this giving and concentrate PBSA development in some neighbourhoods, hence drawing out students from the neighbourhoods that experience problems. This process is also referred to as *destudentification* in academic literature (Sage et al, 2012).

The final academic parallel concerns that of the concept of NIMBY. Local opposition to PBSA and other forms of student housing are usually apt characterizations of NIMBY, as these people are likely to acknowledge the need for adequate and sufficient student housing, but merely advocate that these are situated in the wrong place (Hubbard, 2006). A lower house price is then caused by the negative willingness to pay for housing in the direct vicinity of PBSA, which can be conceptualized as a symptom of this NIMBY effect. An important concluding note is that the phenomenon of NIMBY in PBSA is not different from the traditional forms of studentification such as HMO. In the contrary, UK-based opposition to HMO such as the National HMO Lobby reflect a strong resemblance to other manifestations of NIMBY such as wind turbines (Alamel, 2018).

6.2 Conclusions and recommendation

This paper investigated the economic effects of PBSA. My analysis focused on possible externalities that would arise from the development of PBSA, subsequently influencing nearby residential property prices. Through a difference-in-difference specification of the hedonic pricing model, this study provides considerable evidence for the existence of such negative external effects on a micro-scale, based on which the null-hypothesis listed in paragraph 2.5 can be rejected, and thus providing an answer to the central research question; "*what are the external effects of PBSA on neighbouring residential property prices?*" However, the magnitude of this effect seems to be dependent on treatment radii and specific neighbourhood characteristics, increasing for instance in the case of below-average pre-existing student populations. In addition, the effect seems to fade over time. These conclusions lead to believe that the externalities are caused by short-term effects, such as noise-nuisance and anti-social behaviour, and that there exists a certain preconception of living near PBSA and the students that reside there.

Finally, although hedonic analysis remains enormously apt at showing that people attain a certain (negative) value to living nearby PBSA, it is limited in regards to revealing the exact underlying reasons. In order to further deepen our understanding of the driving forces behind the externalities of PBSA such as those that this study has revealed, I recommend the use of mixed methods in future research, including ‘softer’ methods such as case studies incorporating surveys in addition to quantitative analysis such as conducted in this study.

REFERENCES

Alamel, A. (2018). L'émergence des géographies étudiantes: une littérature anglophone substantielle, une recherche francophone à bâtir. Belgeo. *Revue belge de géographie*, (1).

Alonso, W. (1964). Location and land use. *Toward a general theory of land rent*. Cambridge: Harvard University Press.

Barke, M., & Powell, K. (2008). The impact of students on local house prices: Newcastle upon Tyne, 2000-2005. *Northern Economic Review*, 38, 39-60.

Bonger, S. (2016). *Discussie over aanpak studentenoverlast*. Retrieved on 11-5-2018 from https://www.delta.tudelft.nl/article/discussie-over-aanpak-studentenoverlast

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.

Cao, T. van, & Cory, D. C. (1982). Mixed land uses, land-use externalities, and residential property values: a reevaluation. *The Annals of Regional Science*, 16(1), 1-24.

Chow, G.C., (1960). Tests of equality between subsets of coefficients in two linear regressions. *Econometrica* 28 (3), 591–605.

Cheshire, P. & Sheppard, S. (1995). On the Price of Land and the Value of Amenities. Economica, vol. 62, pp. 247-267.

Clark, T. N., Lloyd, R., Wong, K. K., & Jain, P. (2002). Amenities drive urban growth. *Journal of urban affairs*, 24(5), 493-515.

Daams, M.N., Sijtsma, F.J. and van der Vlist, A.J., 2016. The effect of natural space on nearby property prices: Accounting for perceived attractiveness. *Land Economics*, 92(3), pp.389–410.

DeStentor. (2018). Rechter: studentenwoningen mogen in Deventer wijk Knutteldorp. *DeStentor*, 30-04-2018.

Dröes, M. & Koster, H. (2016). Renewable energy and negative externalities: the effect of wind turbines on house prices. *Journal of Urban Economics*, vol. 96, pp. 121-141

Duijn, M. van, Rouwendal, J. and Boersema, R. (2014). *Transformations of industrial heritage: Insights into external effects on house prices.*

Duijn, M. van, Rouwendal, J. and Boersema, R. (2016). Redevelopment of industrial heritage: Insights into external effects on house prices. *Regional Science and Urban Economics*, 57, pp.91–107.

Florida, R. (2006). The flight of the creative class: The new global competition for talent. *Liberal Education*, 92(3), 22-29.

Galati, G. & Teppa, F., Alessie, R., 2011. Macro and micro drivers of house price dynamics: An application to Dutch data. *DNB Working Paper*, 288, 1-35.

Gibb, K. and Nygaard, C. (2005) The impact of buy to let residential investment on local housing markets: evidence from Glasgow, Scotland, *European Journal of Housing Policy*, 5, pp. 301–326.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of economic geography*, 1(1), 27-50.

Groves, R., Revel, K., & Leather, P. (2003). The changing nature of the housing market in Selly Oak. *Birmingham, UK: Center for Urban and Regional Studies, University of Birmingham*.

Hayes, T. (2017). Eerste studentencomplex Brabantpark krijgt eindelijk groen licht. *BN De Stem*, 19-04-2017.

Heintzelman, M. & Tuttle, C. (2012). Values in the Wind: A Hedonic Analysis of Wind Power Facilities. *Land Economics*, 88 (3): 571– 88.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, 81(396), 945-960.

Hubbard, P. (2006). NIMBY by another name? A reply to Wolsink. *Transactions of the Institute of British Geographers*, 31(1), 92-94.

Hubbard, P. (2008). Regulating the social impacts of studentification: A Loughborough case study. *Environment and Planning A*, 40: 323-341.

Hubbard, P. (2009). Geographies of studentification and purpose-built student accommodation: leading separate lives?. *Environment and planning A*, 41(8), 1903-1923.

Hughes, D., & Davis, M. (2002). Student housing: a cautionary tale of one city. *The Journal of Social Welfare & Family Law*, 24(2), 135-155.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1), 1-24.

King, G., & Nielsen, R. (2016). Why propensity scores should not be used for matching. *Copy at http://j. mp/1sexgVw Download Citation BibTex Tagged XML Download Paper*, 378.

Kenyon, E.L. (1997). Seasonal sub-communities: the impact of student households on Residential Communities. *British Journal of Sociology*, 48: 286-301

Leyshon, A. and French, S. (2009) We all live in a 'Robbie Fowler house': the geographies of the buy-to-let market in the UK, *British Journal of Politics and International Relations*, 11, pp. 438–460.

Li, M. M., & Brown, H. J. (1980). Micro-neighbourhood externalities and hedonic housing prices. *Land economics*, 56(2), 125-141.

Lofland, J. (1968). The youth ghetto. *Journal of Higher Education*, 39(3): 121-143.

Malpezzi, S. (1996). Housing prices, externalities, and regulation in US metropolitan areas. *Journal of Housing Research*, 209-241.

NVM (2018). *About the NVM*. Retrieved on 22-12-2018 from https://www.nvm.nl/overnvm/about.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2), 197-210.

Palmquist, R. B. (1992). Valuing localized externalities. *Journal of urban economics*, 31(1), 59-68.

Palmquist, Raymond B. 2005. “Property Value Models.” In Handbook of Environmental Economics, Vol. 2, ed. Karl Goran-Mahler and Jeffrey R. Vincent, 763–819. Amsterdam: North Holland.

Pols, G., (2017). Studentensteden willen overlast door ‘verkamering’ straten aanpakken. *Volkskrant*, 10-07-2017.

Ricardo, D. (1817). On the principles of political economy and taxation.

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

Rugg, J., Rhodes, D., Jones, A. (2000). The nature and Impact of student demand on housing markets. Joseph Rowntree Foundation, 1-45.

Rugg, J., Rhodes, D., & Jones, A. (2002). Studying a niche market: UK students and the private rented sector. *Housing studies*, 17(2), 289-303.

Sage, J., Smith, D., & Hubbard, P. (2012). The rapidity of studentification and population change: there goes the (student) hood. *Population, Space and Place*, 18(5), 597-613.

Sage, J., Smith, D., & Hubbard, P. (2013). New-build studentification: A panacea for balanced communities?. *Urban Studies*, 50(13), 2623-2641.

Schwartz, A.E., Ellen, I.G., Voicu, I. and Schill, M.H., 2006. The external effects of placebased subsidized housing. *Regional Science and Urban Economics*, 36(6), pp.679–707.

Sheppard, S. (1999). Hedonic analysis of housing markets. *Handbook of Regional and Urban Economics*, vol. 3, pp. 1595-1635.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, 13(1), 1-44.

Smith, N. (1987). Gentrification and the rent gap. *Annals of the Association of American geographers*, 77(3), 462-465.

Smith, D. (2005). Patterns and processes of'studentification'in Leeds. *The Regional Review*, 12, 14-16.

Smith, D., Holt, L. (2007). Studentification and ‘apprentice’ gentrifiers within Britain’s provincial towns and cities: extending the meaning of gentrification. *Environment and Planning A*, 39: 142-161

Smith, D. P., & Hubbard, P. (2014). The segregation of educated youth and dynamic geographies of studentification. *Area*, 46(1), 92-100.

Theebe, M.A. (2004). Planes, Trains, and Automobiles: The Impact of Traffic Noise on House Prices. *Journal of Real Estate Finance and Economics*, vol. 28, no. 2, pp. 209-234.

Thünen, J.H. von. (1842). Der isolirte Staat. *Beziehung auf Landwirtschaft und Nationalökonomie.* Rostock, Leopold.

Thomsen, J. (2004). Home experiences in student housing: About institutional character and temporary homes. *Journal of Youth Studies*, 10(5): 577–596.

Wilkinson, R. (1973). House prices and the measurement of externalities. *Economic Journal*, vol. 83, pp. 72-86

APPENDICES

Appendix A. Stata do file including data preparation

```
*** // Import NVM Data // ***
import excel "X:\My Documents\Master thesis\Student housing\NVM data\Final Dataset.xls", sheet("Joined data Daniël Slomp") firstrow


*** // Generate variables // ***
/// dependent variable
gen logTPRICE = log(TRANSACTIEPRIJS)
// prijs per m2
gen PriceM2 = TRANSACTIEPRIJS/M2
/// variable generation
** Dummies target group
gen Target50m = Distance <= 50
gen Target100m = Distance <= 100
gen Target250m = Distance <= 250
** Dummies control group
gen Control50m = Target50m==0 & Distance <= 1000
gen Control100m = Target100m==0 & Distance <= 1000
gen Control250m = Target250m==0 & Distance <= 1000
** Dummies before & after
gen Trans_year = year(DATUM_AFMELDING)
gen after = Trans_year >= Constr_year
gen before = Trans_year < Constr_year
** Trend variables
gen Trend_before = Trans_year-Constr_year if before==1
gen Trend_after = Trans_year-Constr_year if after==1
/// Control variables
// location variables
gen logM2 = log(M2)
// gen house type dummies
keep if SOORTWONING==2 | SOORTWONING==5| SOORTWONING==6 | SOORTWONING==7| SOORTWONING==8 | SOORTWONING==9|
SOORTWONING==10 | SOORTWONING==11| SOORTWONING==12 | SOORTWONING==21| SOORTWONING==22| SOORTWONING==23 |
SOORTWONING==24| SOORTWONING==25 | SOORTWONING==26| SOORTWONING==27
qui tab SOORTWONING, gen(WTYPE)

// gen building period dummies
drop if BWPER==0
qui tab BWPER, gen(BWPER)
rename BWPER1 Bu1500_1905
rename BWPER2 Bu1906_1930
rename BWPER3 Bu1931_1944
rename BWPER4 Bu1945_1959
rename BWPER5 Bu1960_1970
rename BWPER6 Bu1971_1980
rename BWPER7 Bu1981_1990
rename BWPER8 Bu1991_2000
rename BWPER9 Bu2001_P
// gen physical dummies
gen Balcony = NBALKON >= 1
gen Terrace = NDAKTERRAS >= 1
gen CV = VERW == 2
gen BiKwGoed = ONBI<=4
gen BuKwGoed = ONBU<=4
gen Parking = PARKEER >=2
// transaction dummies
qui tab VERKOOPCOND, gen(TrType)
// Time fixed effects
gen post4 = substr(Postal,1,4)
destring post4, replace
// Time variant neighbourhood effects
gen street_busy = LIGDRUKW == 2
```

```
***// Descriptive statistics//***
// Descriptive statistics area characteristics of PBSA's
twoway kdensity Share15_24PBSA if Target100m==1 & _24Re>0 & _24Re<0.6 || kdensity _24Re if Control100m==1 & _24Re>0 & _24Re<0.6 || kdensity _24Re if _24Re>0 & _24Re<0.6, xline(0.1861865) legend(order(1 "Target" 2 "Control" 3 "Total")) name(StudSat, replace) title("KDE Student Saturation") xtitle("15-24 year olds as share of population") ytitle(Kernel density)
twoway kdensity WOZRe if Target100m==1 & WOZRe>100 & WOZRe<500 || kdensity WOZRe if Control100m==1 & WOZRe>100 & WOZRe<500 || kdensity WOZRe if WOZRe>100 & WOZRe<500, xline(194.6261) legend(order(1 "Target" 2 "Control" 3 "Total")) name(WOZ, replace) title("KDE Neighbourhood House Prices (WOZ)") xtitle("Average houseprice (x €1,000)") ytitle(Kernel density)
twoway kdensity OADRe if Target100m==1 || kdensity OADRe if Control100m==1 || kdensity OADRe, xline(2957.378) legend(order(1 "Target" 2 "Control" 3 "Total")) name(OAD, replace) title("KDE Population Density (OAD)") xtitle("No of addresses per sq km") ytitle(Kernel density)

// Trimming data
* Physical attributes
drop if mi(M2)
drop if M2<30 | M2>250
drop if NBADK==0
* Transaction attributes
drop if TRANSACTIEPRIJS<25000 | TRANSACTIEPRIJS>2000000
drop if PriceM2>7500

// Descriptive statistics
*summarize TRANSACTIEPRIJS TRANSACTIEPRIJS1 PriceM2 Distance TrType1 TrType2 TrType3 M2 BiKwGoed BuKwGoed Balcony Terrace CV Parking MONUMENT Bu1500_1905 Bu1906_1930 Bu1931_1944 Bu1945_1959 Bu1960_1970 Bu1971_1980 Bu1981_1990 Bu1991_2000 Bu2001_P WTYPE1 WTYPE2 WTYPE3 WTYPE4 WTYPE5 WTYPE6 WTYPE7 WTYPE8 WTYPE9 WTYPE10 WTYPE11 WTYPE12 WTYPE13 WTYPE14 WTYPE15 WTYPE16 if Target100m==1
*summarize TRANSACTIEPRIJS TRANSACTIEPRIJS1 PriceM2 Distance TrType1 TrType2 TrType3 M2 BiKwGoed BuKwGoed Balcony Terrace CV Parking MONUMENT Bu1500_1905 Bu1906_1930 Bu1931_1944 Bu1945_1959 Bu1960_1970 Bu1971_1980 Bu1981_1990 Bu1991_2000 Bu2001_P WTYPE1 WTYPE2 WTYPE3 WTYPE4 WTYPE5 WTYPE6 WTYPE7 WTYPE8 WTYPE9 WTYPE10 WTYPE11 WTYPE12 WTYPE13 WTYPE14 WTYPE15 WTYPE16 if Control100m==1
*summarize TRANSACTIEPRIJS TRANSACTIEPRIJS1 PriceM2 Distance TrType1 TrType2 TrType3 M2 BiKwGoed BuKwGoed Balcony Terrace CV Parking MONUMENT Bu1500_1905 Bu1906_1930 Bu1931_1944 Bu1945_1959 Bu1960_1970 Bu1971_1980 Bu1981_1990 Bu1991_2000 Bu2001_P WTYPE1 WTYPE2 WTYPE3 WTYPE4 WTYPE5 WTYPE6 WTYPE7 WTYPE8 WTYPE9 WTYPE10 WTYPE11 WTYPE12 WTYPE13 WTYPE14 WTYPE15 WTYPE16

*summarize TRANSACTIEPRIJS PriceM2 Distance TrType1 TrType2 TrType3 M2 BiKwGoed BuKwGoed Balcony Terrace CV Parking MONUMENT Bu1500_1905 Bu1906_1930 Bu1931_1944 Bu1945_1959 Bu1960_1970 Bu1971_1980 Bu1981_1990 Bu1991_2000 Bu2001_P WTYPE1 WTYPE2 WTYPE3 WTYPE4 WTYPE5 WTYPE6 WTYPE7 WTYPE8 WTYPE9 WTYPE10 WTYPE11 WTYPE12 WTYPE13 WTYPE14 WTYPE15 WTYPE16 if Target250m==1
*summarize TRANSACTIEPRIJS PriceM2 Distance TrType1 TrType2 TrType3 M2 BiKwGoed BuKwGoed Balcony Terrace CV Parking MONUMENT Bu1500_1905 Bu1906_1930 Bu1931_1944 Bu1945_1959 Bu1960_1970 Bu1971_1980 Bu1981_1990 Bu1991_2000 Bu2001_P WTYPE1 WTYPE2 WTYPE3 WTYPE4 WTYPE5 WTYPE6 WTYPE7 WTYPE8 WTYPE9 WTYPE10 WTYPE11 WTYPE12 WTYPE13 WTYPE14 WTYPE15 WTYPE16 if Control250m==1

* Dropping observations that don't belong in the regression due to being outside of control group
drop if Distance>1000

// Generating Interaction variables //
gen target1Distance = Target100*Distance
gen target1After = Target100*after
gen target1Before = Target100*before
gen target1AfterDistance = target1After*Distance
gen target1BeforeDistance = target1Before*Distance
gen target1Distance2 = Target100*(Distance^2)
gen target1AfterDistance2 = target1After*(Distance^2)
gen target1BeforeDistance2 = target1Before*(Distance^2)
gen targetTrend_after = Target100*Trend_after
gen targetTrend_afterDistance = Target100*Trend_after*Distance
gen targetTrend_afterDistance2 = Target100*Trend_after*(Distance^2)
*gen targetTrend_before = Target100*Trend_before
*gen targetTrend_beforeDistance = Target100*Trend_before*Distance
*gen targetTrend_beforeDistance2 = Target100*Trend_before*(Distance^2)


//Preparation for sensitivity analyses
gen target50Distance = Target50m*Distance
```

```
gen target50After = Target50m*after
gen target50Before = Target50m*before
gen target50AfterDistance = target50After*Distance
gen target50BeforeDistance = target50Before*Distance
gen target50Distance2 = Target50m*(Distance^2)
gen target50AfterDistance2 = target50After*(Distance^2)
gen target50BeforeDistance2 = target50Before*(Distance^2)
gen target50Trend_after = Target50m*Trend_after
gen target50Trend_afterDistance = Target50m*Trend_after*Distance
gen target50Trend_afterDistance2 = Target50m*Trend_after*(Distance^2)

gen target250Distance = Target250m*Distance
gen target250After = Target250m*after
gen target250Before = Target250m*before
gen target250AfterDistance = target250After*Distance
gen target250BeforeDistance = target250Before*Distance
gen target250Distance2 = Target250m*(Distance^2)
gen target250AfterDistance2 = target250After*(Distance^2)
gen target250BeforeDistance2 = target250Before*(Distance^2)
gen target250Trend_after = Target250m*Trend_after
gen target250Trend_afterDistance = Target250m*Trend_after*Distance
gen target250Trend_afterDistance2 = Target250m*Trend_after*(Distance^2)

**// Empirical Model //**
* Recode all missing variables to 0
mvencode _all, mv(0) override

set matsize 11000

sysdir set PLUS "X:\My Documents\Master thesis\Stata packages"
*ssc install outreg2

// Baseline model
reg logTPRICE target1After target1AfterDistance target1AfterDistance2 target1Before target1BeforeDistance target1BeforeDistance2
targetTrend_after targetTrend_afterDistance targetTrend_afterDistance2 i.Trans_year, robust
reg logTPRICE target1After target1AfterDistance target1AfterDistance2 target1Before target1BeforeDistance target1BeforeDistance2
targetTrend_after targetTrend_afterDistance targetTrend_afterDistance2 i.Trans_year i.SOORTWONING i.BWPER logM2 Balcony Terrace
BiKwGoed BuKwGoed Parking CV MONUMENT street_busy, robust
reg logTPRICE target1After target1AfterDistance target1AfterDistance2 target1Before target1BeforeDistance target1BeforeDistance2
targetTrend_after targetTrend_afterDistance targetTrend_afterDistance2 i.Trans_year i.SOORTWONING i.BWPER logM2 Balcony Terrace
BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4, robust

// OLS Assumptions
*vif
**jb logTPRICE
**predict r
**kdensity r, normal
**pnorm r

**estat hettest

//** Sensitivity analyses **//
reg logTPRICE target50After target50AfterDistance target50AfterDistance2 target50Before target50BeforeDistance
target50BeforeDistance2 target50Trend_after target50Trend_afterDistance target50Trend_afterDistance2 i.Trans_year i.SOORTWONING
i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4, robust
reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4, robust


// Spatial fixed effects
** No FE
*qui reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year, robust
*eststo m1
** District scale
```

```
*qui reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.obj_wijk_ID, robust
*eststo m2
** Postal code
*qui reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4, robust
*eststo m3
** Neighbourhood
*qui reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.obj_buurt_ID, robust
*eststo m4
** Street
*qui reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.obj_straatID, robust
*eststo m5
*estout m1 m2 m3 m4 m5, keep(target250After target250AfterDistance target250AfterDistance2 target250Before
target250BeforeDistance target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2)
cells(b(star) se(par)) stats(r2 df_r) delimiter(";") starl(* 0.1 ** 0.05 *** 0.01) style(tab)

// Chow-test for equality of coefficients between subsamples
// Generating subsamples based on area characteristics
gen OADSub = 0
gen StudSatSub = 0
gen WOZSub = 0

**summarize OADRe --> mean = 2957.378
**summarize _24Re if _24Re>0 --> mean = 0.1861865
**summarize WOZRe if WOZRe>0 & WOZRe<600 --> mean = 194.6261

replace OADSub = 1 if OADRe>2957.378
replace StudSatSub = 1 if _24Re>0.1861865
replace WOZSub = 1 if WOZRe>194.6261

count if OADSub==1 & Target100m==1 & after==1
count if OADSub==1 & Target250m==1 & after==1

count if StudSatSub==1 & Target100m==1 & after==1
count if StudSatSub==1 & Target250m==1 & after==1

count if WOZSub==1 &  Target100m==1 & after==1
count if WOZSub==1 & Target250m==1 & after==1

// Chow-test

reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if StudSatSub==1
reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if StudSatSub==0

**
* Student Saturation
* RSS = 661.604848 | RSS1 = 355.886109  | RSS2 = 284.624896

// Are results driven by the pre-existing student population?

reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance
target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year
i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if StudSatSub==1
eststo m8
```

```
reg logTPRICE target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2 i.Trans_year i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if StudSatSub==0
eststo m9

estout m8 m9, keep(target250After target250AfterDistance target250AfterDistance2 target250Before target250BeforeDistance target250BeforeDistance2 target250Trend_after target250Trend_afterDistance target250Trend_afterDistance2) cells(b(star) se(par)) stats(r2 df_r N) delimiter(";") starl(* 0.1 ** 0.05 *** 0.01) style(tab)

// Coarsened exact matching method
drop if _24Re<0
drop if WOZRe<0

cem OADRe _24Re WOZRe, treatment(Target100m) autocuts(fd)

reg logTPRICE target1After target1AfterDistance target1AfterDistance2 target1Before target1BeforeDistance target1BeforeDistance2 targetTrend_after targetTrend_afterDistance targetTrend_afterDistance2 i.Trans_year i.SOORTWONING i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if cem_matched==1

cem OADRe _24Re WOZRe SOORTWONING, treatment(Target100m) autocuts(fd)

reg logTPRICE target1After target1AfterDistance target1AfterDistance2 target1Before target1BeforeDistance target1BeforeDistance2 targetTrend_after targetTrend_afterDistance targetTrend_afterDistance2 i.Trans_year i.BWPER logM2 Balcony Terrace BiKwGoed BuKwGoed Parking CV MONUMENT street_busy i.post4 if cem_matched==1
```

Appendix B. Full estimation results

	(3)	(4)	(5)
Sample size	<1000 m	<1000 m	<1000 m
Target area	0 – 100 m	0 – 50 m	0 – 250 m
Control area	100 – 1000 m	50 – 1000 m	250 – 1000 m
After	-0.301***	0.318	-0.118***
	(0.104)	(0.276)	(0.0373)
After Distance	0.00793**	-0.0353*	0.00180***
	(0.00347)	(0.0206)	(0.000545)
After Distance^2	-5.56e-05**	0.000617*	-5.79e-06***
	(2.71e-05)	(0.000327)	(1.80e-06)
Before	-0.217**	0.623	-0.0446**
	(0.104)	(0.408)	(0.0209)
Before Distance	0.00493	-0.0404**	0.000435
	(0.00341)	(0.0206)	(0.000299)
Before Distance2	-2.69e-05	0.000564**	-7.27e-07
	(2.59e-05)	(0.000257)	(9.79e-07)
Trend After	0.0624**	-0.0331	0.0247**
	(0.0313)	(0.0699)	(0.0123)
Trend After Distance	-0.00187*	0.00574	-0.000318*
	(0.00107)	(0.00600)	(0.000190)
Trend After Distance^2	1.54e-05*	-0.000114	1.12e-06*
	(8.62e-06)	(0.000102)	(6.49e-07)
Single Family (ref: simple terraced house)	0.0963***	0.0962***	0.0963***
	(0.00804)	(0.00804)	(0.00805)
Canal house	0.251***	0.251***	0.251***
	(0.0163)	(0.0163)	(0.0162)
Mansion	0.231***	0.231***	0.231***
	(0.00949)	(0.00949)	(0.00949)
Farm	0.371**	0.371**	0.375**
	(0.148)	(0.148)	(0.148)
Bungalow	0.222***	0.222***	0.221***
	(0.0311)	(0.0311)	(0.0312)
Villa	0.368***	0.368***	0.368***
	(0.0240)	(0.0240)	(0.0239)
Countryhouse	0.218**	0.217**	0.218**
	(0.0850)	(0.0851)	(0.0853)
Downstairs house	0.0306***	0.0304***	0.0299***
	(0.00854)	(0.00853)	(0.00854)
Upstairs house	-0.0645***	-0.0645***	-0.0648***
	(0.00871)	(0.00871)	(0.00872)
Maisonette	-0.0636***	-0.0636***	-0.0634***
	(0.00954)	(0.00955)	(0.00955)
Flat (portiek)	-0.00396	-0.00430	-0.00440
	(0.00878)	(0.00877)	(0.00878)
Flat (galerij)	-0.0596***	-0.0594***	-0.0607***
	(0.00915)	(0.00914)	(0.00914)
Hospice	-0.292**	-0.292**	-0.293**
	(0.134)	(0.134)	(0.134)
Down- and upstairs house	0.00373	0.00332	0.00350
	(0.0197)	(0.0197)	(0.0198)
Surface (log m^2)	0.738***	0.738***	0.738***
	(0.00488)	(0.00488)	(0.00488)
Balcony (1=Yes)	0.0117***	0.0115***	0.0117***
	(0.00312)	(0.00312)	(0.00312)
Terrace (1=Yes)	0.0374***	0.0375***	0.0370***
	(0.00357)	(0.00357)	(0.00357)

BiKwGoed (1=Yes)	-0.133***	-0.133***	-0.133***
	(0.0105)	(0.0104)	(0.0104)
BuKwGoed (1=Yes)	-0.0525***	-0.0527***	-0.0531***
	(0.0181)	(0.0181)	(0.0180)
Parking (1=Yes)	0.112***	0.112***	0.112***
	(0.00371)	(0.00371)	(0.00371)
Central Heating (1=Yes)	0.0567***	0.0565***	0.0566***
	(0.00473)	(0.00473)	(0.00474)
Monumental building (1=Yes)	0.0633***	0.0634***	0.0628***
	(0.00761)	(0.00761)	(0.00761)
Street business (1= Busy road)	-0.0307***	-0.0303***	-0.0303***
	(0.00473)	(0.00472)	(0.00473)
Building period: 1906-1930 (ref: built in period	-0.0375***	-0.0374***	-0.0371***
1500-1905)	(0.00518)	(0.00518)	(0.00518)
Building period: 1931-1944	-0.0230***	-0.0231***	-0.0227***
	(0.00584)	(0.00584)	(0.00584)
Building period: 1945-1959	-0.140***	-0.139***	-0.139***
	(0.00678)	(0.00677)	(0.00677)
Building period: 1960-1970	-0.203***	-0.203***	-0.202***
	(0.00644)	(0.00644)	(0.00645)
Building period: 1971-1980	-0.128***	-0.128***	-0.128***
	(0.00656)	(0.00657)	(0.00656)
Building period: 1981-1990	-0.0729***	-0.0729***	-0.0721***
	(0.00680)	(0.00680)	(0.00678)
Building period: 1991-2000	0.0107	0.0109*	0.0118*
	(0.00656)	(0.00656)	(0.00656)
Building period: ≥ 2001	0.0565***	0.0562***	0.0573***
	(0.00682)	(0.00681)	(0.00681)
Time fixed effects	*Yes*	*Yes*	*Yes*
ZIP-code fixed effects	*Yes*	*Yes*	*Yes*
Constant	9.127***	9.125***	9.126***
	(0.0242)	(0.0242)	(0.0242)
Observations	24,788	24,788	24,788
Adjusted R-squared	0.8590	0.8597	0.8597

Appendix C. Chow test for equality between sets of coefficients

	Error Sum of Squares
	Student Saturation
Pooled	661.61
>Average	355.89
<Average	284.63
	F-Statistics
No. of regressors	125
F(125, 24538)	6.465***

Note: ***p<0.01, **p<0.05, *p<0.10.