The selling process of houses around the Groningen gas field

A Two-Stage Least Squares (2SLS) regression analysis of the influence of earthquake risk on the time-on-the-market (TOM)

Jelte van den Broek
s2352265

Master’s thesis
Research Master in Spatial Sciences

May 2018

Faculty of Spatial Sciences
University of Groningen

Supervisors:
prof. dr. ir. G.R.W. de Kam (Faculty of Spatial Sciences, University of Groningen)
prof. dr. J.P. Elhorst (Faculty of Economics and Business, University of Groningen)

Preface

This thesis is my final work of the Research Master in Spatial Sciences at the University of Groningen. I hope that this research is able to provide more insight into the selling process of houses around the Groningen gas field. The positionality of the researcher can influence the chosen empirical strategy or the outcomes of the study. I live in Groningen which is close to the earthquake area; furthermore, I interviewed home-owners in financial problems due to the earthquake damage for another research project. To limit the influence of these factors on the strategy or outcomes, this research is transparent about the choices made during the research process. The decisions regarding model specifications and included variables are made based on theory or earlier empirical work. The supervisors are also regularly consulted about the statistical modeling. Regarding my philosophical positioning, I tend towards positivism because of the search for general patterns in the housing market. However, I do understand the limitations to generalization in statistical research, qualitative studies are needed to gain insight into the experiences of seller and buyers and to fully understand the mechanisms of the real estate market. I chose a quantitative approach to be able to investigate a variety of factors influencing the time-on-the-market (TOM) and to highlight the specific effect of earthquakes. I hope my thesis can improve the understanding of the relationship between earthquakes and the housing market around the Groningen gas field.

I want to thank my supervisors George de Kam and Paul Elhorst for their valuable support and feedback during the empirical analysis and the writing of this thesis. I also want to express my gratitude to my friends and family for their continued help and support. Finally, I would like to thank the NVM for providing the data necessary for the statistical modeling. I appreciate all your contributions during the process leading to this thesis.

Summary

Earthquakes around the Groningen gas field are associated with damage to properties and a lower quality of life, causing difficulties for sellers to find a suitable buyer. Besides having to accept a lower selling price, sellers often face a lengthy selling process. This Master's thesis focuses on the influence of earthquake risk on the time-on-the-market (TOM) of house sales around the Groningen gas field. The literature review shows an array of factors affecting TOM: seller characteristics, structural and locational attributes, and market conditions. Furthermore, TOM is simultaneously determined with selling price in the selling process. Earthquake risk has both a spatial dimension since it is connected to a certain region and a temporal dimension because multiple earthquake take place over time. Two approaches are used to measure earthquake risk: a Difference-In-Difference (DID) technique comparing a risk and a reference area before and after a major earthquake and a variable accumulating the Peak Ground Velocity (PGV) at the house location of all previous earthquakes. Employing NVM data on housing transactions in the Northern Netherlands from 2003 until 2014, the final result in both approaches is a Two-Stage Least Squares (2SLS) regression model, including the earthquake indicator, selling price as an endogenous explanatory variable, structural and locational attributes, and spatial and temporal fixed effects. The DID model shows that TOM in the risk area, being neighborhoods with damaged houses, is 6.2% higher after the Huizinge earthquake of 2012 compared to similar neighborhoods surrounding the gas field, while it used to be 8.7% lower. The PGV model indicates that an increase of 10% in PGV causes a rise of 0.5% in TOM. Earthquakes start to have an impact after a PGV of 0.7 m/s. This thesis provided more insight into the housing market dynamics around the Groningen gas field by showing that the risk of earthquakes appears to increase TOM.

Keywords: earthquakes, Groningen gas field, selling process, time-on-the-market, Two-Stage Least Squares regression.

Table of Contents

1 Introduction

1.1 Gas extraction, earthquakes and the housing market

The northeast of the Netherlands contains the largest natural gas field of Europe, where gas extraction is taking place since 1963 by the Dutch Petroleum Company (NAM) (Bosker et al., 2016; Whaley, 2009). The large-scale and long-term gas production has caused soil subsidence and the frequent occurrence of earthquakes in the surrounding region (see Figure 1.1). Houses have sustained earthquake damage and are at risk of future damage (Koster & Van Ommeren, 2015). Compensation schemes exist for incurred earthquake damage; however, they can impose high transaction costs to households and granted budgets do not always enable adequate repairs (De Kam & Spijkerboer, 2015; Van der Voort & Vanclay, 2015). Furthermore, earthquake damage causes feelings of unsafety and health problems, thereby decreasing the quality of life in the area around the gas field (Boelhouwer et al., 2016; Postmes et al., 2017). These negative developments set in motion by earthquakes can be related to a declining trend in local property values (Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster, 2016).

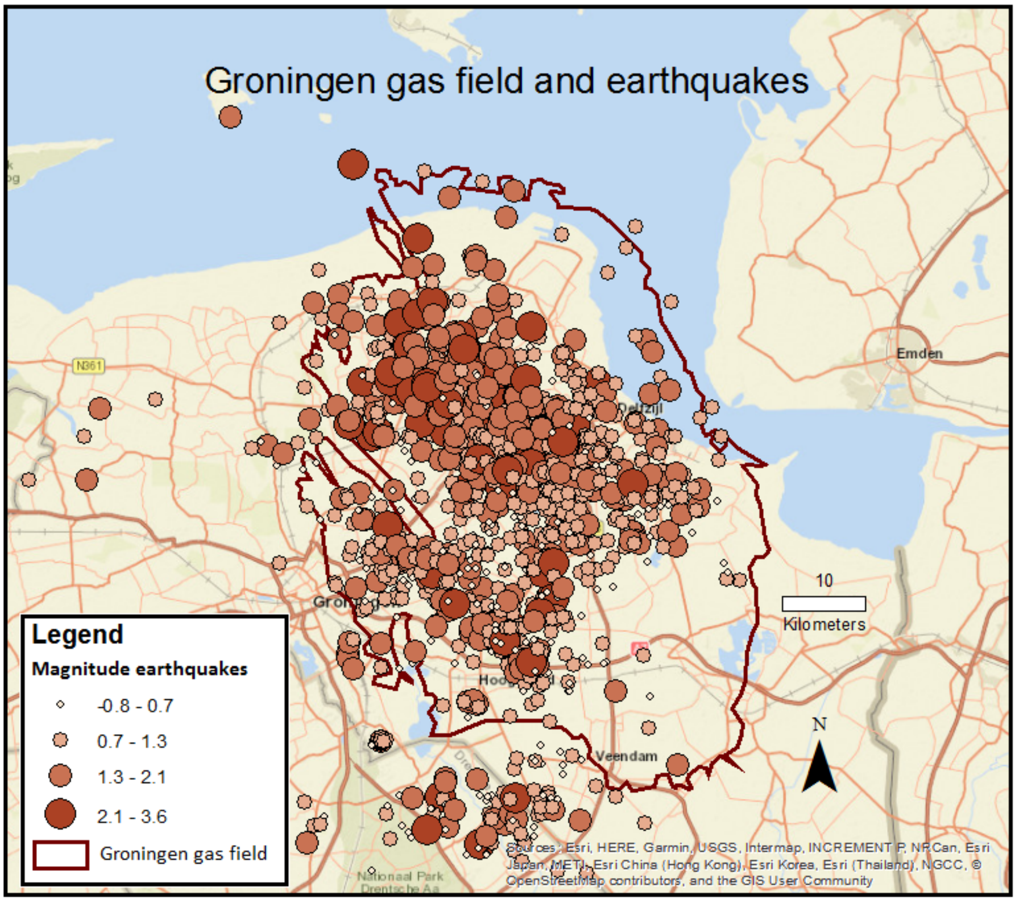


Figure 1.1: The location of the Groningen gas field and occurring earthquakes from 1986 to February 2018. Source: created in ESRI ArcGIS, earthquake map from Groninger Bodem Beweging, KNMI, NAM Platform, Rijksuniversiteit Groningen, retrieved from ArcGIS Online.

Buyers appear to be less willing to buy houses in a region characterized by earthquake risk. Therefore, sellers often have to decrease their listing price or accept a lower selling price to ensure a house sale, otherwise they face a lengthy selling process (De Kam & Mey, 2017). A lower selling price can be problematic for the considerable share of home-owners around the gas field with a mortgage debt higher than the value of their house, a situation caused by decreasing property values (De Kam et al., 2018). The difficulties experienced by sellers around the Groningen gas field cause the time-on-the-market (TOM), being the time between offering the house to the market and the house sale, to be longer compared to similar regions without earthquakes (CBS, 2017b). However, the selling process is also influenced by busts in the housing market or population decline characterizing many regions around the gas field (Boelhouwer et al., 2016; CBS, 2017b; De Kam & Mey, 2017; Koster, 2016). Taking into account other factors affecting the selling process, this Master's thesis aims to gain insight into the effect of earthquakes on TOM.

1.2 Time-on-the-market modeling

While a significant strand of research focuses on modeling housing prices around the Groningen gas field (e.g. Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster, 2016), a comprehensive regression model on TOM appears to be missing. The urgency to sell can differ between sellers (De Kam & Mey, 2017; Evans, 2004); however, the seller usually tries to combine achieving a high transaction price with selling the house as quickly as possible, trying to minimize TOM (Dubé & Legros, 2016; Yavas & Yang, 1995). A considerable methodological challenge in modeling TOM is the interrelationship with selling price, both are simultaneously determined in the selling process. A higher transaction price usually requires a longer TOM and vice versa. This trade-off causes endogeneity if selling price is included in a TOM model. A common solution is a two-stage approach using instrumental variables, a Two-Stage Least Squares (2SLS) regression model (Brooks & Tsolacos, 2010; Dubé & Legros, 2016; Knight, 2002). Besides the price and TOM trade-off, structural and locational attributes of the house, motivation and characteristics of the seller, and market conditions are affecting the length of the selling process (Anglin et al., 2003; De Kam & Mey, 2017; Dubé & Legros, 2016; Knight, 2002; Springer, 1996).

The second methodological challenge is finding a suitable earthquake impact indicator. Hedonic models developed in earlier studies employed a variety of indicators, ranging from physical damage to the property, the percentage of damaged houses in the surrounding area, being located in an area at risk of earthquakes to the intensity of earthquakes at the house location (Atlas voor Gemeenten, 2017; Duran & Elhorst, 2017; Francke & Lee, 2014; Koster, 2016). It is already shown that areas with a high percentage of damaged houses have a longer TOM than nearby areas with similar locational attributes but without earthquake damage (CBS, 2017b). Earthquakes might affect the risk perception of buyers; it could create a negative image of the area around the gas field. Comparing risk and reference areas to filter out the effect of earthquakes is based on a Difference-In-Difference (DID) technique. One approach to identify a risk area is to use the percentage of damaged houses in a region (CBS, 2017b). Besides the spatial dimension of earthquake risk, buyers could be more aware of earthquake hazards after a major earthquake, thereby adding a temporal dimension (Atlas voor Gemeenten, 2017; Beron et al., 1997; Duran & Elhorst, 2017). A DID approach is able to compare areas before and after an event, for example, a major earthquake (Schwartz et al., 2006; Van Duijn et al., 2016). However, the temporal effect might be more complex since many earthquakes occur around the gas field and differences also exist within the risk area. Koster (2016) employs the percentage of damaged houses in the surrounding ZIP code area to measure earthquake risk. Unfortunately, the employed dataset does not allow this thesis to include the variable of Koster (2016) in the analysis. The reporting of earthquake damage mainly started after the 2012 Huizinge earthquake. Therefore, the damage percentages might not represent the effect of earthquakes on the housing market before 2012. It might be more suitable to use the accumulated earthquake intensity at the house location which can be done using the Peak Ground Velocity (PGV) of an earthquake (Duran & Elhorst, 2017; Koster & Van Ommeren, 2015). It is interesting to use both a DID and a PGV approach to include earthquake risk in a regression model estimating TOM.

1.3 Research questions

Houses around the Groningen gas field are associated with earthquake risk. The property is at risk of future damage and compensation schemes often offer too low repair budgets or impose high transactions costs (De Kam & Spijkerboer, 2015; Koster & Van Ommeren, 2015; Van der Voort & Vanclay, 2015). Furthermore, the quality of life has declined due to feelings of unsafety (Boelhouwer et al., 2016; Postmes et al., 2017). Therefore, sellers encounter difficulties with finding a suitable buyer. They often have to settle for lower selling prices and face a lengthy selling process (De Kam & Mey, 2017). This thesis focuses on the relationship between earthquakes and the length of the selling process or TOM, taking into account other factors influencing TOM. The societal relevance can be found in addressing the often problematic situation of sellers around the Groningen gas field, while the academic relevance of this study is increasing the understanding of the role of earthquakes in housing market dynamics and tackling methodological challenges in modeling TOM and earthquake impact. This thesis is structured around the following main research question:

To what extent do earthquakes influence the time-on-the-market of house sales around the Groningen gas field?

The first step of this thesis is to gain insight into a variety of factors playing a role in the selling process. In order to find the specific effect of earthquakes on TOM in a regression model, it is crucial to control for other aspects influencing TOM. Furthermore, it is useful to further explore the spatial and temporal dimension of the impact of earthquakes on the housing market. Theories and earlier empirical work give insight into the selling process and the connection to earthquakes. The literature review creates the foundation for the empirical modeling and is covered by the first two sub questions:

1. Which factors influence the selling process of houses?
2. How are earthquakes affecting the housing market?

The second step of this thesis is to estimate a 2SLS regression model including the relationship between earthquakes and TOM. The risk of earthquakes around the Groningen gas field causes difficulties for sellers to find a buyer. The majority of buyers relocates within the region and has detailed knowledge about the situation in the area, although the perception of the size of the risk area can differ between buyers (De Kam & Mey, 2017). Besides the spatial aspect, there is a time dimension to earthquake risk since locations can be hit by multiple earthquakes, especially a major earthquake can create a negative image of the region (Atlas voor Gemeenten, 2017; Duran & Elhorst, 2017). Measuring earthquake risk is done using both a DID and a PGV approach. The DID model identifies a risk area based on the percentage of damaged houses in the neighborhood and compares it to a reference area before and after the Huizinge earthquake of 2012 and the Middelstum earthquake of 2006. Neighborhoods with damaged houses are assumed to have a negative image regarding earthquake risk. However, it has to be noted that the reporting of earthquake damage mainly took place after 2012 and that the size of the risk area might have changed over the years. The risk area could have been smaller around the 2006

Middelstum earthquake. However, the DID model on the Middelstum earthquake employs the same risk areas as the Huizinge earthquake model since a smaller risk area might exclude regions at risk of earthquakes in later years. The PGV approach uses a variable measuring the accumulated PGV at the house location before the sale. Both approaches estimate a 2SLS model to take into account the simultaneity between selling price and TOM. Furthermore, the models include variables representing structural and locational attributes and market conditions. The modeling of the spatial and temporal aspect of earthquake risk and its effect on TOM is covered by the following sub questions:

3. To what extent does the location in an area at risk of earthquakes affect the time-on-the-market of houses around the Groningen gas field?
4. How does a major earthquake influence the time-on-the-market of houses around the Groningen gas field?

1.4 Section outline

In chapter 2, earlier theoretical and empirical work on TOM is discussed. It gives insight into factors affecting TOM and presents findings of earlier studies on the connection between earthquakes and the housing market, thereby covering the first two sub questions. Chapter 3 discusses the methodology, describing the dataset, the employed variables and the statistical model. Chapter 4 discusses the findings of the statistical analyses. The results of the regression models are interpreted here which can be used to answer sub questions 3 and 4. Chapter 5 presents the conclusions of this study and presents an answer to the main research question. The findings are also connected to policy recommendations and suggestions are given for further research.

2 Theoretical framework

2.1 Selling process

The process of selling a house generally starts with determining the appropriate listing price, where most sellers contact a broker for professional support. Buyers compare this price to the price they are willing to pay for the concerned house, and to other properties they are considering during their search. The buyer can proceed by making a bid, thereby starting the bargaining process. The house is sold if the seller accepts, although it is also possible to reject the bid or make a counter-offer. In the latter situation, the buyer can then chose to accept, reject or make a counter-offer which leads to a continuing bargaining process. The selling process is finalized if a deal is concluded between buyer and seller, resulting in an agreed transaction price of the house. Without a deal, the house remains active on the market and the seller continues the search for a buyer (Anglin et al., 2003; Dubé & Legros, 2016; Evans, 2004). The search and bargaining process plays an important role because the housing market does not determine a fixed price, caused by the fact that it is an inefficient and imperfect market (Evans, 2004).

In an efficient market, the information available will be fully capitalized into the prices of traded goods. However, participants in the property market are usually not completely aware of changing market conditions; they only observe the direction of price changes or do not respond immediately. Therefore, it is questionable whether the property market is economically efficient (Evans, 2004). The characteristics of an efficient market are reflected in the basic model of supply and demand in economics, being the perfect market. This model has three basic assumptions (Evans, 2004):

- Many buyers and sellers.
- A homogeneous product.
- The participants in the market have full information on product prices.

The model of the perfect market mainly applies to an explicit market, where the good itself is actually being traded, usually in a marketplace. However, the property market is an implicit market since the location of houses is fixed and only their characteristics are traded (Evans, 2004). Rosen (1974) states that implicit or hedonic prices can be connected to product attributes using observed prices of the differentiated products. Hedonic modelling can give insight into the valuation of property characteristics, often divided into structural, locational and market attributes (Daams et al., 2016; Livy & Klaiber, 2016; Schwartz et al., 2006). However, the maximum explained variation appears to be around 90 percent, meaning determining accurate property prices is difficult and professional brokers are usually only able to set a certain price range (Evans, 2004).

The main reason for the inefficiency and imperfection of the real estate market is the absence of a homogeneous product. Real estate is fixed in location and even for identical properties a price adjustment has to be made for locational differences. It is even doubtful if truly identical properties exist since they are varying bundles of characteristics (Evans, 2004). Furthermore, buyers and sellers only trade on the market infrequently and they are facing search costs, limiting them in acquiring full information on alternatives, the value of property attributes, or the influence of market conditions on relative prices. The product heterogeneity, search costs and a lack of information cause the number the number of buyers and sellers to be limited, leaving room for negotiation on the selling price (Evans, 2004; Knight, 2002; Yavas & Yang, 1995). It takes time to match a buyer and seller and to negotiate a final transaction price for the property. In general, the target of achieving the highest possible selling price is combined with shortening the time-on-the-market (TOM) (Dubé & Legros, 2016; Yavas & Yang, 1995). This Master's thesis focuses on explaining TOM.

2.2 Selling price and time-on-the-market trade-off

Sellers appear to face a trade-off between maximizing selling price and minimizing TOM, being a simultaneous optimization problem (Dubé & Legros, 2016). The chosen listing price plays an important role in determining both targets. This can be illustrated using search theory; an approach from labor economics used to analyze markets where a buyer cannot immediately find a seller (Boeri & Van Ours, 2013; Knight, 2002). The listing price signals the seller's reservation price, being the minimum price the seller intends to receive for the property. The buyer compares the listing price to his or her own reservation price, being the maximum price the buyer is willing to pay for a house based on a valuation of the property characteristics and prices of similar properties. The height of the listing price determines the arrival rate of potential buyers and bid distributions. A lower listing price increases the arrival rate and enables the seller to realize a quick sale, but the final price is expected to be lower. On the other hand, a higher listing price causes a lower arrival rate of potential buyers; however, it increases the probability of finding a buyer with a higher reservation price (Knight, 2002; Yavas & Yang, 2002). The effect might be offset by negative herding, causing houses that are on the market for a considerable time to become stigmatized (Taylor, 1999). De Kam and Mey (2017) found evidence for this effect around the Groningen gas field, where sellers get stuck in the market because they do not want to lower their listing price. Anglin et al. (2003) measure the relative height of the listing price using the degree of overpricing (DOP): the percentage deviation of the chosen listing price from a typical listing price for a house with certain attributes under particular market conditions. The estimated hazard model indicated that a higher DOP increases TOM, where the hedonic model also shows an increase in selling price (Anglin et al., 2003). In general, models estimating TOM have a relatively low R-squared compared to hedonic models, being around 0.13 (Anglin et al., 2003; Dubé & Legros, 2016).

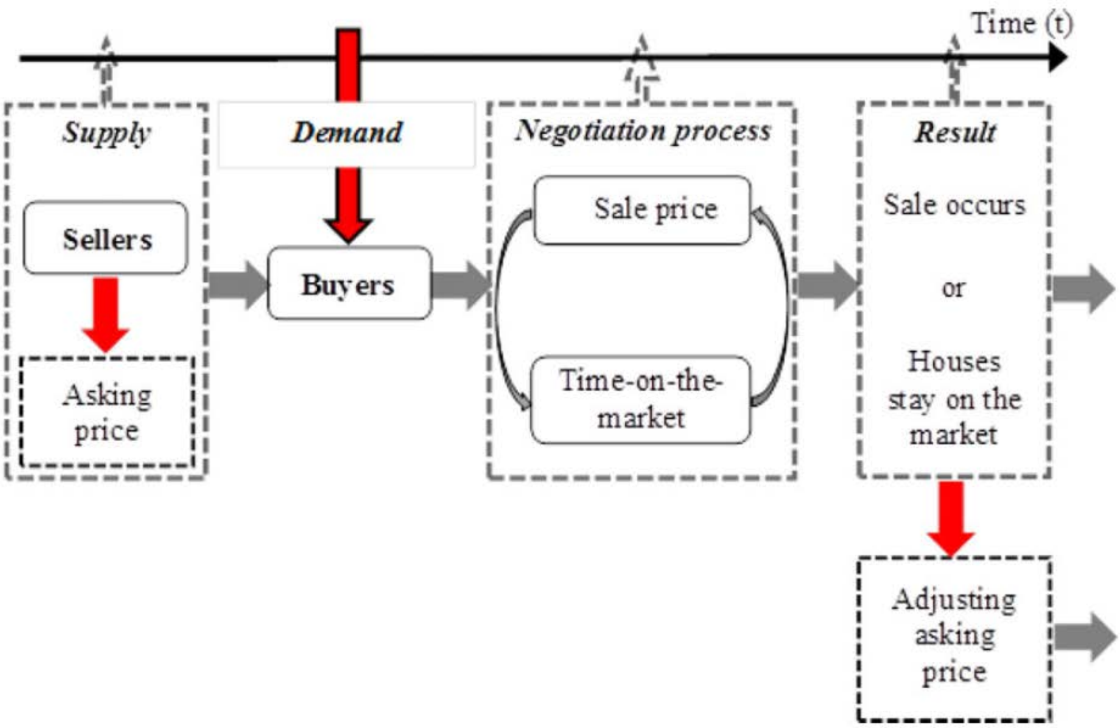


Figure 2.1: The selling process of a house. Source: Dubé and Legros (2016, p.852).

The listing price choice is the first step in the selling process; however, Figure 2.1 shows that the simultaneous optimization problem is finally solved during the bargaining or negotiation process between buyer and seller (Dubé & Legros, 2016). Both the selling price and TOM are affecting each other and are related to the motivation of buyers and sellers (Dubé & Legros, 2016, p.847):

“On the one hand, a motivated seller (buyer) can be ready to accept (propose) a lower (higher) price to quickly proceed to the transaction. On the other hand, a patient seller (buyer) can wait longer in the hope of obtaining the highest (lowest) price as possible.”

The above illustrates the simultaneity between TOM and price which causes endogeneity in a TOM model including selling price as an independent variable. Both selling price and TOM act as independent and dependent variable because they are influencing each other in the bargaining process (Dubé & Legros, 2016; Knight, 2002). Since the motivation of seller and buyer is unobserved, it is a latent variable and hidden in the error term. The selling price is affected by the willingness of sellers and buyers to negotiate on the hedonic prices of the house attributes, thereby a connection exists to the motivation of both agents. The motivation will be hidden in the error term and is related to selling price, thereby causing an endogenous problem (Brooks & Tsolacos, 2010; Dubé & Legros, 2016). The inclusion of variables related to the listing price, such as DOP, might also cause endogeneity since the listing price choice also relates to the motivation of the seller (Dubé & Legros, 2016). A common solution to deal with endogeneity is a Two-Stage Least Squares (2SLS) regression model including Instrumental Variables (IVs). This statistical technique first estimates a selling price model, using the predicted values of selling price as an independent variable in the TOM model that is estimated in the second step (Brooks & Tsolacos, 2010; Knight, 2002). Dubé and Legros, (2016) argue that a simultaneous model including both a selling price and a TOM equation is the preferred option to deal with the simultaneous optimization problem. In their Seemingly Unrelated Regression (SUR) model they find a negative relationship between price and TOM, contrary to the positive relationship resulting from the trade-off discussed above. They theorize that houses with higher prices have better amenities and, therefore, sell faster. A 2SLS model only focusing on TOM is also an option instead of a SUR. It also estimates two equations since a selling price model is estimated in the first stage (Brooks & Tsolacos, 2010).

Although Dubé and Legros (2016) mention that the motivation of the seller is often unobserved, there are studies trying to operationalize seller motivation. Springer (1996) includes a binary variable on whether the seller stated that he or she is motivated, anxious or must sell. This eagerness, however, significantly increased TOM, possibly because they have houses that are difficult in terms of marketing (Springer, 1996). Glower et al. (1998) expect that seller motivation is increased by setting a move date, having accepted a new job somewhere else, making an offer on another house or having already bought another property. A logistic survival model is used to test the effect of seller motivation on TOM, using a relatively small sample of 115 cases. Having plans to move quickly and a job change is shown to shorten TOM (Glower et al., 1998). However, the above models are not able to take into account the fact that the urgency to sell might change over time (De Kam & Mey, 2017). This thesis is not able to include a variable on seller motivation; therefore, the TOM models have to deal with the simultaneity with selling price which is done using the 2SLS technique. Besides the trade-off between price and TOM, other factors affect TOM, being the characteristics of the seller, the structural and locational attributes of the house, and market conditions. These factors are discussed in the following paragraphs.

2.3 Seller characteristics

The influence of the characteristics of the seller on the selling process is given minor attention in earlier studies. The personal situation of the seller can influence the urgency towards a sale, thereby affecting the listing price choice and the strategy during the bargaining process (De Kam & Mey, 2017; Dubé & Legros, 2016). The capabilities of the seller could play a role in the search costs experienced by the seller or the extent the seller can influence the outcome of the bargaining process (De Kam & Mey, 2017; Evans, 2004). De Kam and Mey (2017) include seller characteristics in their analysis. They assume that younger sellers are more flexible in the selling process due to a stronger focus on their future career. Furthermore, a higher educated seller is expected to have more control over the search and bargaining process. A smoother selling process is also expected for sellers that have already lived in the region since they can use their local social network. Unfortunately, De Kam and Mey (2017) did not include these factors in a comprehensive TOM model, although they do analyze correlations between TOM and their independent variables. They found the expected relationships with age, education level and originating from the region. The seller can also employ a broker to enhance their bargaining capabilities which usually leads to a higher selling price and lower TOM (Jud et al., 1996). A broker with an attitude characterized by openness about property damage due to earthquakes is shown to reduce selling difficulties around the Groningen gas field (De Kam & Mey, 2017).

The personal financial situation of the seller might also play a role. De Kam and Mey (2017) studied the effect of the financial leeway of the seller, operationalized by the ratio of annual income to housing value. A higher financial leeway is expected to decrease the urgency to sell, meaning sellers are able to sustain a longer TOM to realize a higher selling price. De Kam and

Mey (2017) did now find a significant correlation; however, financial distress is shown to play a role in TOM studies using more comprehensive modeling (e.g. Genesove & Mayer, 1997; Sirmans et al., 1995). Sirmans et al. (1995) show that sellers with high holdings costs, for example related to their mortgage, have a higher probability on a quick sale since financial distress is forcing sellers to settle for a lower selling price to facility a lower TOM. Genesove & Mayer (1997) show that sellers with a higher loan-to-value (LTV) ratio have higher listing prices, higher selling prices, and a longer TOM. Their explanation for this result is that financially constrained sellers will chose a reservation price that combines the down payment on a new house, the outstanding mortgage debt on the house, and brokerage costs. The higher reservation price leads to a higher listing price which usually causes a longer TOM and higher selling price (Genesove & Mayer, 1997). In general, the results of financial distress appear to be ambiguous. High holdings costs and a low financial leeway increase the seller's motivation to sell faster, while a high LTV causes sellers to wait longer to achieve a higher selling price. Unfortunately, the TOM models in this thesis are not able to include variables on the characteristics of the seller.

2.4 Market conditions

The housing market is characterized by booms and busts in prices. Housing market cycles are strongly related to macroeconomic developments such as business cycles, income growth, credit availability, industrial production, and the unemployment rates (Agnello & Schuknecht, 2011). A strong housing boom started in the late 1990s in many industrialized countries, including the Netherlands. However, the global financial crisis starting in 2008 caused a major downturn in the housing market. Currently, housing markets are recovering from the major bust (Agnello & Schuknecht, 2011; Immergluck, 2015).

In a period of a boom, TOM is usually lower. Economic progress lowers interest rates, enabling buyers to afford a higher mortgage which increases their reservation price. Therefore, there is a larger probability that buyers will meet the reservation price of the seller, causing houses to sell more quickly. Properties are sold at a fast rate, thereby decreasing the number of properties on the market. This shortage will cause an increase in housing prices (Evans, 2004). Contrary to in a boom, TOM is expected to be longer in a bust. In a situation of rising interest rates, buyers decrease their reservation prices and search longer for a suitable house; therefore, it takes sellers longer to find a buyer. Seller will first wait with reducing their listing price, thereby increasing the amount of properties on the market which will lower prices eventually (Evans, 2004). Therefore, TOM appears to have an inverse relationship to housing prices.

In the Netherlands, the housing market is currently characterized by a boom, recovering from the major bust due to the recent financial crisis. This is reflected in the high growth rate in housing prices of 7.6% in 2017 (Lennartz et al., 2018). Houses offered by agents affiliated to the Dutch Association of Real Estate Brokers (NVM) are sold within 56 days on average in the first quarter of 2018 which is a decrease compared to the 75 days a year earlier (NVM, 2018). The current boom is characterized by interest rates that remain low, causing further increases in housing prices (Lennartz et al., 2018). The general market trend can be included in a TOM model using the number of sales, the number of houses for sale or the interest rate. Several studies also control for seasonality differences since more houses are offered in the summer. However, the most common approach is to include time fixed effects (Anglin et al., 2003; Dubé & Legros, 2016; Haurin et al., 2010; Springer, 1996), which is done in the TOM models in this thesis using the year of sale.

2.5 Structural and locational characteristics

The influence of property attributes on TOM is the focus in the research of Haurin (1988) and later Haurin et al. (2010). This line of research states that the atypicality of a house plays an important role in explaining the variety in TOM between houses. Atypical houses have less common structural and/or locational characteristics. The difficulties in valuing such atypical properties lead to a higher variety in offers, causing the seller to increase the reservation price and wait longer for a buyer willing to pay a higher price (Haurin, 1988). Haurin et al. (2010) constructed an atypicality measure that compares the implicit price of each structural and locational characteristic of the property with the mean value of these characteristics in the surrounding neighborhood. Haurin et al. (2010) include this measure in a hazard model which shows that atypical houses have a longer TOM. De Kam and Mey (2017) also found this effect around the Groningen gas field, where (semi) detached houses take longer to sell than townhouses and apartments.

A different approach is to include structural characteristics of the property directly into the model to account for a varying TOM between different types of houses (Anglin et al., 2003). Under housing characteristics can be thought of the age and size of the property, the amount of bathrooms and bedrooms, the presence of a fireplace, a pool and a garage, the number of stories, and the price class. Most structural characteristics appear to be insignificant in explaining TOM, with sometimes exceptions for variables related to the age and size of the house (Anglin et al., 2003; Forgey et al., 1996; Knight, 2002; Springer, 1996). Older and larger houses appear to have a longer TOM (Forgey et al., 1996). The smaller importance of structural characteristics of the property in TOM models is contrary to hedonic models, where they play a considerable role in explaining the variation in housing prices (Daams et al., 2016; Dubé & Legros, 2016). The maintenance status of the house might also play a role, especially in the region around the Groningen gas field where buyers are aware of the risk at earthquake damage (Atlas voor Gemeenten, 2017). The atypicality index of Haurin et al. (2010) is beyond the scope of this thesis; however, the TOM models do include a large amount of structural characteristics such as the type of the house and the building year.

Besides the physical characteristics of the house, hedonic models often include locational characteristics such as the proximity to a highway, a park or a school. These amenities can have positive externalities on housing prices (Daams et al., 2016; Schwartz et al., 2006). The importance of accessibility dates back to the work on bid rent models of Von Thünen (1842) and later Alonso (1964), assuming that households are willing to pay more for a location closer to the Central Business District (CBD) since most jobs are located there (McCann, 2013). A higher degree of urbanization which can be measured using the address density, also increases housing prices (Daams et al., 2016). Negative externalities are also possible, for example, related to noise pollution from development projects (Schwartz et al., 2006). Spatial fixed effects are often included in a hedonic model to control for locational characteristics that are omitted from the model (Livy & Klaiber, 2016). Locational characteristics and spatial fixed effects can also be added to a model estimating TOM (Dubé & Legros, 2016), which is also done in this thesis. Positive externalities related to the proximity to certain amenities might increase the arrival rate of buyers and increase their reservation price, thereby decreasing TOM (Yavas & Yang, 1995). On the other hand, negative externalities could make buyers less willing to buy houses in a certain region which increases TOM. Earthquake risk could be such a negative externality.

2.6 Earthquakes and the housing market

The area around the Groningen gas field is characterized by the frequent occurrence of earthquakes which affects the surrounding housing market. Research has mainly focused on the negative impact on housing prices, where Koster and Van Ommeren (2015) identify three main effects:

- Damage to properties induced by earthquakes can lower housing values if not (adequately) repaired.
- Past earthquakes might indicate a high probability of future earthquakes causing damage to the property.
- The presence of earthquakes in the region might decrease the quality of life in surrounding region, for example, due to unsafety or insecurity regarding future earthquakes.

Existing compensation schemes for incurred earthquake damage to a property could ensure a less important role of the first two effects (Koster & Van Ommeren, 2015); however, these schemes require long procedures which impose considerable transaction costs to home-owners around the Groningen gas field (Van der Voort & Vanclay, 2015). Furthermore, the granted repair budgets do not always enable a structural solution to earthquake damage of the property (De Kam & Spijkerboer, 2015). Combined with a decreased quality of life (Boelhouwer et al., 2016; Postmes et al., 2017), the risk of future damage creates a negative image of the Groningen gas field region. Buyers are less willing to buy a house in an earthquake region which is reflected in lower selling prices. Therefore, sellers have to search longer for a suitable buyer and have a weaker bargaining position, being the characteristics of a buyer's market (De Kam & Mey, 2017). The difficult search period and bargaining phase, especially if sellers do not want to lower their reservation price, causes a long TOM. Many sellers are not willing to lower the listing price since their house is under water (De Kam et al., 2018; De Kam & Mey, 2017). The awareness of earthquake risk might be enhanced by a recent earthquake, meaning there is both a spatial and a temporal dimension to the earthquake effect on the housing market (Duran & Elhorst, 2017). However, selling difficulties around the Groningen gas field could also be related to population decline (Boelhouwer et al., 2016; CBS, 2017b; De Kam & Mey, 2017; Koster, 2016). Earlier studies used a variety of indicators to gain insight into the specific effect of earthquakes on housing prices and TOM.

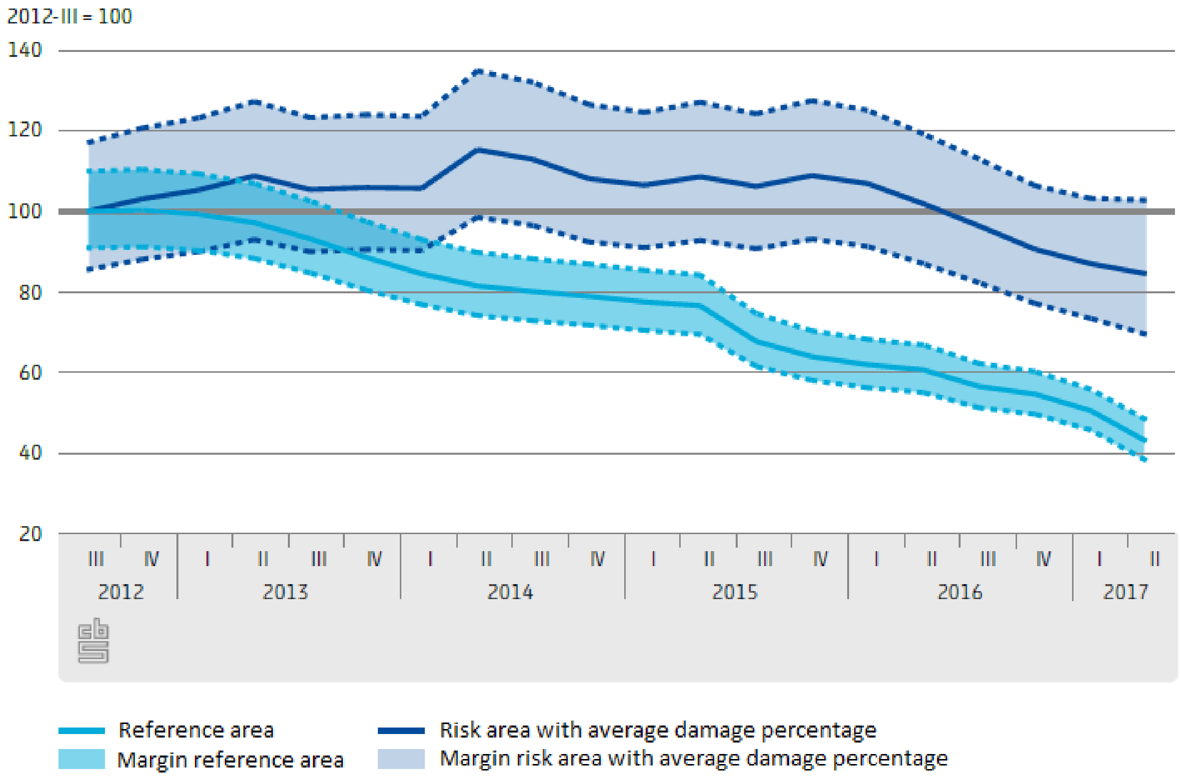


Figure 2.2: TOM development in the risk area with an average damage percentage and the reference area from the third quarter of 2012 until the second quarter of 2017. Source: CBS (2017, p.33) (edited).

One approach that can be used to analyze earthquake impact is to compare areas at risk of earthquakes, risk areas, with similar areas without earthquakes, reference areas (CBS, 2017b). This approach is related to the Difference-In-Difference technique (DID), where treatment areas are compared to control areas. It is important that these control areas have similar characteristics and are not affected by earthquakes (Schwartz et al., 2006). CBS (2017b) selected neighborhoods with damaged houses as risk area and neighborhoods around the Groningen gas field with similar socioeconomic attributes as reference area. The risk area is further divided based on the percentage of damaged houses of the total amount of houses, being low (<31%), average (31-54%), and high (>54). Regarding housing prices, CBS (2017b) uses a hedonic model to compensate for price changes in housing attributes. Areas with a high and average damage percentage are lagging behind reference areas since 2012; however, low damage areas appear to increase faster in prices (CBS, 2017b). The trends in TOM in the average damage percentage areas are highlighted in Figure 2.2. In the reference areas, recovery of the housing market started around 2013, leading to a decreasing TOM. However, in the average damage percentage area, TOM remains stable and increases even slightly. Recovery starts in 2016. The rise was sharper in high damage percentage areas and recovery started later in 2016. In low damage percentage areas, recovery already started in 2014 and did not significantly differ, using a confidence interval from 90%, from the reference areas in the second quarter of 2017. In risk areas with population decline, TOM is decreasing slower from 2016 onwards compared to those without population decline (CBS, 2017b). CBS (2017b) excluded the City of Groningen from their analyses, leaving the question unanswered whether the housing market here is also affected by earthquakes. Furthermore, reference areas are located adjacent to the risk area and might be affected by earthquake risk.

Bosker et al. (2016) created a TOM regression model, but they did not find a significant effect of being located in the risk area. Their risk area comprises of eight municipalities where properties are assumed to have been damaged by earthquakes and they analyze the time period after the Huizinge earthquake of 2012. However, the earthquake effect might be unevenly spread over the region and could have started before 2012 (De Kam, 2016). Bosker et al. (2016) use reference properties in the whole of the Netherlands instead of reference areas. These reference properties cannot be located in a buffer around the risk area; however, properties close to the Groningen gas field might still be affected by earthquake risk (De Kam, 2016). Finally, the TOM model of Bosker et al. (2016) does not include the trade-off with selling price and spatial fixed effects.

Bosker et al. (2016) also estimate a hedonic model which is improved in the study of Atlas voor Gemeenten (2017), using ZIP-code areas where more than 20% of the houses has earthquake damage as risk area. They find an average negative price effect of 2.2% of being located in the risk area compared to reference properties, although vast regional differences are highlighted. A meta-analysis of international literature by Koopmans and Rougoor (2017) confirms the general negative price effect of earthquake risk, for example, evidence is found in the Tokyo Metropolitan Area (Nakagawa et al., 2007, 2009). Atlas voor Gemeenten (2017) only analyzes the period after the 2012 Huizinge earthquake because the hedonic model of Bosker et al. (2016) did not show an effect before 2012. However, the latter study did not include housing transaction before 2011, while the price effect might have started earlier (De Kam, 2016). Murdoch et al. (1993) and Beron et al. (1997) find a significant decline in housing prices in the San Francisco Bay Area after the Low Prieta Earthquake of 1989. Buyers might have underestimated earthquake risk before the earthquake; however, Koopmans and Rougoor (2017) state that the occurrence of a recent earthquake does not affect the impact of earthquake risk on housing prices. Inhabitants around the Groningen gas field already showed awareness of earthquake risk before the Huizinge earthquake of 2012 (De Kam & Raemakers, 2014). Nevertheless, the Huizinge earthquake did trigger a considerable amount of media attention which can influence housing prices (Bosker et al., 2016; Koopmans & Rougoor, 2017).

Atlas voor Gemeenten (2017) also shows that houses having received compensation for damage sell for higher prices. The maintenance status might be of great importance in a region characterized by earthquake damage. Francke and Lee (2014) investigate the physical damage to individual properties and find that the average TOM for houses with damage is higher compared to those without damage. Their hedonic model does not find a significant negative effect on housing prices; however, they only use a limited number of transactions. De Kam and Mey (2017) show that many buyers state that they would offer more for a house without earthquake damage compared to an identical one with damage.

Besides using risk and reference areas and damage to properties to investigate the effect of earthquakes on the housing market, the impact of earthquakes can also be measured using the earthquake intensity at the location of the house. Koster and Van Ommeren (2015) and later Koster (2016) use the Peak Ground Velocity (PGV) to measure the intensity of an earthquake at a specific location. The PGV is based on the magnitude, depth, and the distance to the epicenter. An earthquake with a PGV above 0.5 cm/s is noticeable and the number of these noticeable earthquakes can be included in a hedonic model. Koster and Van Ommeren (2015) show that a noticeable earthquake has a negative effect on housing prices of 1.9%. Duran and Elhorst (2017) argue that counting the number of noticeable earthquakes might not be the best approach. It is difficult information to retrieve for a buyer and houses close to one another can have different counts while the risk perception might be comparable. Koster (2016) adds a variable on the percentage of damaged houses in the surrounding ZIP-code area, showing that a 1% increase in the damage percentage lowers housing prices with 0.2%. Koster (2016) controls for the effect of population decline using spatial fixed effects at a ZIP-code 6 level. The approach of Duran and Elhorst (2017) is to accumulate the PGV of all previous earthquakes at the house location. The preliminary model also discounted the PGV by time and calculated it both at the house location and at the neighborhood level, showing a price effect of -0.3% when the PGV that hits the neighborhood doubles. Currently, the PGV model of Duran and Elhorst (2017) is under further construction, removing the time discounting effect. It acted as a memory effect since a recent earthquake might have a larger impact on risk perception;

however, Koopmans and Rougoor (2017) show that the occurrence of a recent earthquake does not have a significant impact on the effect of earthquakes on housing prices.

In conclusion, earlier studies on the effect of earthquakes on the housing market employed a variety of indicators to capture earthquake risk. The precise impact differs per method, although they all studies indicate a negative price effect. CBS (2017b) shows an increasing effect of earthquake risk on TOM, while Bosker et al. (2016) do not find a significant effect. Results are influenced by the chosen region assumed to be at risk of earthquakes and the time period where earthquakes are expected to influence the housing market. This thesis employs a DID and a PGV approach to measure earthquake risk. Unfortunately, it is not able to include the physical damage to properties or the damage percentage variable of Koster (2016). The DID model compares a risk and a reference area before and after the 2012 Huizinge earthquake or the 2006 Middelstum earthquake. It also employs the Middelstum earthquake since the impact could have started before 2012. The DID model gives insight into TOM changes in areas with and without earthquake risk and the effect of a major earthquake. However, it depends on predetermining a risk area based on damage percentages and does not take into account that many earthquakes take place around the Groningen gas field. Therefore, the accumulated PGV at the house location is also used to understand the complex spatial and temporal dimension of earthquake risk and the impact on TOM.

2.7 Conceptual model and hypotheses

The discussed theories and empirical work have given insight into the first two sub questions on the factors influencing TOM and the impact of earthquakes on the housing market in Groningen. The relationships shown by earlier work are captured in the conceptual model in Figure 2.3 which creates a foundation for the empirical study.

The conceptual model captures the trade-off between selling price and TOM, assuming that seller motivation is unobserved. The listing price choice is not included since the simultaneous optimization problem is finally solved in the bargaining process between buyer and seller. It also acknowledges other factors influencing TOM such as seller characteristics, structural and locational attributes of the house, and market conditions. Furthermore, it visualizes the effect of earthquake risk. These factors also influence the selling price; however, the main focus is on TOM which is visualized with the thick arrow.

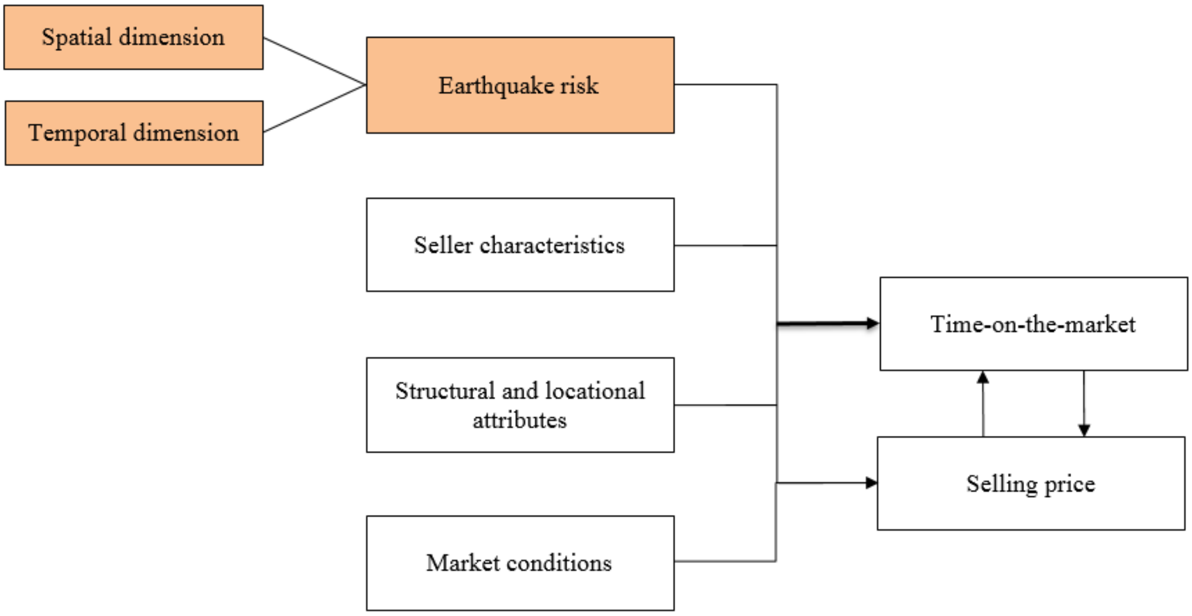


Figure 2.3: The conceptual model for this research, based on the theoretical framework and the empirical research questions.

The conceptual model forms the basis for the empirical modeling since it shows which aspects to include in the regression model. Unfortunately, the dataset does not allow the inclusion of seller characteristics. The conceptual model also shows the suitability of a 2SLS model since it can take into account the simultaneity between selling price and TOM. The first stage can regress price based on the factors on the left side of the model, the second stage then estimates TOM. The main interest of this thesis is in the impact of earthquakes; the spatial and temporal dimension of earthquake risk are highlighted in the conceptual model, thereby the model also covers sub question 3 and 4. Based on the discussed theories and empirical studies, two hypotheses are formulated regarding these sub questions:

- Being located in an area at risk of earthquakes increases the time-on-the-market of houses around the Groningen gas field.
- The event of a major earthquake increases the time-on-the-market of houses around the Groningen gas field.

Chapter 3 discusses the data and methods used to study the relationship between earthquakes and TOM covered by sub questions 3 and 4. The final statistical model presented in chapter 4 indicates whether to accept or reject the above hypotheses.

3 Methodology

3.1 Quantitative approach

This Master's thesis uses a quantitative approach to gain insight into the variety of factors influencing the time-on-the-market (TOM) of houses around the Groningen gas field. The strength of such an approach is the ability to find patterns representative for the housing market in Groningen. Controlling for a broad range of aspects, statistical analyses are able to highlight the specific effect of earthquakes on TOM. It allows a certain degree of generalization, also to other regions characterized by earthquakes (Babbie, 2013; Brooks & Tsolacos, 2010). A qualitative approach might be more suitable to gain an in-depth insight into the selling process. Questionnaires and interviews can show the experiences and motivations of buyers and sellers; however, the generalization of these results is limited (Babbie, 2013; De Kam & Mey, 2017). A mixed methods approach is the best choice to research the full dynamics of the housing market. This thesis employs the results from qualitative studies to better identify factors affecting TOM. A quantitative approach using statistical modeling then enables this study to find general patterns regarding TOM and to highlight the specific effect of earthquake risk, thereby contributing to a better understanding of the selling process of houses around the Groningen gas field.

3.2 Regression models

The main challenge in modeling the effect of earthquakes is including both the spatial and temporal dimension. Earthquake risk is perceived to be present in the region around the Groningen gas field where earthquakes take place; however, the awareness of this risk might be triggered by a major earthquake. The Huizinge earthquake of 3.6 on the Richter scale which took place on August 16th, 2012 is often seen as a turning point increasing the attention given to earthquake damage to houses and necessary compensation schemes (Bosker et al., 2016). However, the connection between gas extraction and earthquakes was already known in the 1990s and major earthquakes were already present before the 2012 Huizinge earthquake. An earthquake of 3.5 of the Richter scale took place in 2006 near Middelstum (Bosker et al., 2016). The risk awareness could also have been influenced by earthquakes after Huizinge such as the recent Zeerijp earthquake in January 2018 with a magnitude of 3.4. Two different approaches to measuring earthquake risk are applied in the analysis. The first is based on a Difference-In-Difference (DID) technique inspired by Schwartz et al. (2006) and Van Duijn et al. (2016), while the second approach employs a variable showing the accumulated Peak Ground Velocity (PGV) of previous earthquakes at the house location.

3.2.1 Difference-In-Difference approach

The first two models apply a DID technique developed by Schwartz et al. (2006) and later improved by Van Duijn et al. (2016) to measure to external effects of redevelopment projects on housing prices. This thesis adapts the method to measure the effects of earthquakes on TOM. This DID approach compares the differences between a treatment area and a control area before and after a certain event. It is of crucial importance that these areas have similar characteristics (Schwartz et al., 2006). This approach can also be applied to the earthquake risk which is assumed to have an increasing effect on TOM. The Huizinge earthquake of 2012 could be a turning point in the awareness of earthquake risk, although it has to be taken into account that heavy earthquakes already took place before 2012 such as the Middelstum earthquake of 2006 (Bosker et al., 2016). The treatment area consists of the regions around the Groningen gas field at risk of earthquake damage to houses. This DID approach is applied in both an OLS and a 2SLS model that control for other factors influencing TOM. The first model estimates the following equation:

$$\log TOM_{ijt} = \alpha + \beta E_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varepsilon_{it} \tag{3.1}$$

The dependent variable is the natural logarithm of TOM in days of house *i* in region *j* that is sold at time *t*. The vector *X* captures the structural and locational characteristics of the property. The included temporal and spatial fixed effects are represented respectively by *T* and *S*. Unfortunately, the available variables in the dataset did not allow seller characteristics to be included in the model. Furthermore, the equation includes a constant and an error term. The earthquake variables are captured by *E* and consist of two dummy variables. The first takes the value of 1 if the house is located within the area at risk of earthquakes and 0 for the reference area, this variable acts as a baseline for the differences between the areas and is called ‘before’. The second dummy takes the value of 1 if the house is located in the earthquake area and the transaction took place after 2012, the year of the Huizinge earthquake. This variable is called ‘after’ and captures the effect on prices after the major earthquake in the area at risk of earthquakes. An alternative specification tests the differences before and after the Middelstum earthquake of 2006. The selection of risk and reference areas is based on the approach of CBS

(2017b) using the percentage of damaged houses in the neighborhood (see paragraph 3.3). The above equation is estimated using an OLS model; however, it is unable to capture the effect of selling price. Including it as an independent variable would cause endogeneity since selling price and TOM are simultaneously determined (Brooks & Tsolacos, 2010; Dubé & Legros, 2016). Therefore, a 2SLS model is constructed that estimates the natural logarithm of the selling price (P) in the first stage:

$$\log P_{ijt} = \alpha + \beta E_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varphi I_{it} + \varepsilon_{it} \tag{3.2}$$

The first stage regresses the housing price on the same variables as the first TOM model, being E, X, T and S, and two instrumental variables denoted by I. The employed instrumental variables are the number of disability benefits per 1,000 inhabitants in the surrounding neighborhood and a dummy on the distance to the nearest train station; they are discussed in more detail in paragraph 3.3. The earthquake indicators are also included in the first stage selling price regression since earlier research has shown the negative influence of earthquakes on housing values (Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster, 2016). The first stage coefficients are used to estimate predicted values for the log of selling price ($\log \hat{P}$); these predicted values are then included as an independent variable in the TOM model estimated in the second stage:

$$\log TOM_{ijt} = \alpha + \beta E_{it} + \omega \log \hat{P}_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varepsilon_{it} \tag{3.3}$$

3.2.2 PGV approach

The downside of the DID approach is that it assumes that the impact of earthquakes mainly started after the Huizinge earthquake of 2012 or alternatively after the Middelstum earthquake of 2006. The effect might be more complex since a large amount of earthquakes can be felt every year (Duran & Elhorst, 2017). Furthermore, the use of a treatment area disregards differences within the areas at risk of earthquakes (Koster & Van Ommeren, 2015). Earlier research has used different municipalities or neighborhoods as risk area. The selection of reference areas is also disputed since some studies use control areas that are risk areas in the models of other studies (Atlas voor Gemeenten, 2017; Boelhouwer et al., 2016; CBS, 2017b). The DID approach might ignore the spatial and the temporal complexity of earthquake risk; therefore, it might be more useful to include a variable that measures the impact received by earthquakes at the house location in the period of time before the sale (Duran & Elhorst, 2017; Koster & Van Ommeren, 2015). The PGV approach employs a variable that shows the Peak Ground Velocity (PGV) of all earthquakes received at the house location. The first PGV model estimates the following equation:

$$\log TOM_{ijt} = \alpha + \beta \log PGV_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varepsilon_{it} \tag{3.4}$$

Similar to the DID modeling, the first model is an OLS regression that does not include selling price. The PGV variable is transformed into a natural logarithm and is denoted as logPGV. It is not included in its raw form. First, it is divided into 50 equal groups which are included in the regression as 49 dummies to analyze when the PGV starts to have a significant impact on TOM. The lowest groups are then set to zero in the logPGV variable that is included in the final regression model. The equation again includes the structural and locational attributes of the house, and spatial and time fixed effects. The PGV model does not rely on treatment and control areas; consequently, it can be applied to a larger area around the Groningen gas field. The dataset and the study area are discussed in more detail in paragraph 3.3. The next step is to estimate a 2SLS model including selling price as an endogenous explanatory variable. The goal of the first stage is to estimate selling price, while the second stage focuses on TOM and the effect of earthquakes. The 2SLS model is based on the following two equations:

$$\log P_{ijt} = \alpha + \beta \log PGV_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varphi I_{it} + \varepsilon_{it} \tag{3.5}$$

$$\log TOM_{ijt} = \alpha + \beta \log PGV_{it} + \omega \log \hat{P}_{it} + \delta X_{it} + \gamma_t T_t + \theta_j S_j + \varepsilon_{it} \tag{3.6}$$

3.3 Data and variables

3.3.1 Dataset

The dataset employed to run the above models is acquired from the Dutch Association of Real Estate Brokers (NVM). It contains data on 216,126 housing transactions between 1994 and 2014 in the three Northern provinces in the Netherlands, being Friesland, Groningen and Drenthe. The analysis focuses on the on the time period from 2003 until 2014, thereby leaving 130,062 observations. It is often assumed that the impact of earthquakes on the housing market started after the 2012 Huizinge earthquake (Atlas voor Gemeenten, 2017); however, the larger time period also allows to investigate if the effect did not start earlier, for example, after the Middelstum earthquake of 2006. The data also includes the years before the recent economic crisis which are characterized by a housing market boom. The full dataset can be used for the PGV models, including 122,908 cases after various variable transformations. The DID models, however, rely on the selection of treatment and control areas (see paragraph 3.3.3), meaning several regions are excluded from the dataset, leaving 53,315 housing transactions.

The NVM covers about 75% of the real estate transactions in the Netherlands (NVM, 2018). The coverage is lower in earlier years, being around 50% between 2000 and 2010, while it is around 90% after 2010 (Boelhouwer et al., 2016). The average price in the large PGV dataset is 176,034 euros, after keeping the central 99% to reduce the effect of outliers. The minimum price is €54,375 and the maximum €449,291. The price distribution can be seen in appendix II. The characteristics of the smaller DID dataset are discussed further below. The average price fluctuates over the years (see Figure 3.1), rising before the bust due to the economic crisis and slightly recovering since 2013. It is hard to compare these averages to CBS data on all transactions since CBS takes into account composition differences in calculating the average.

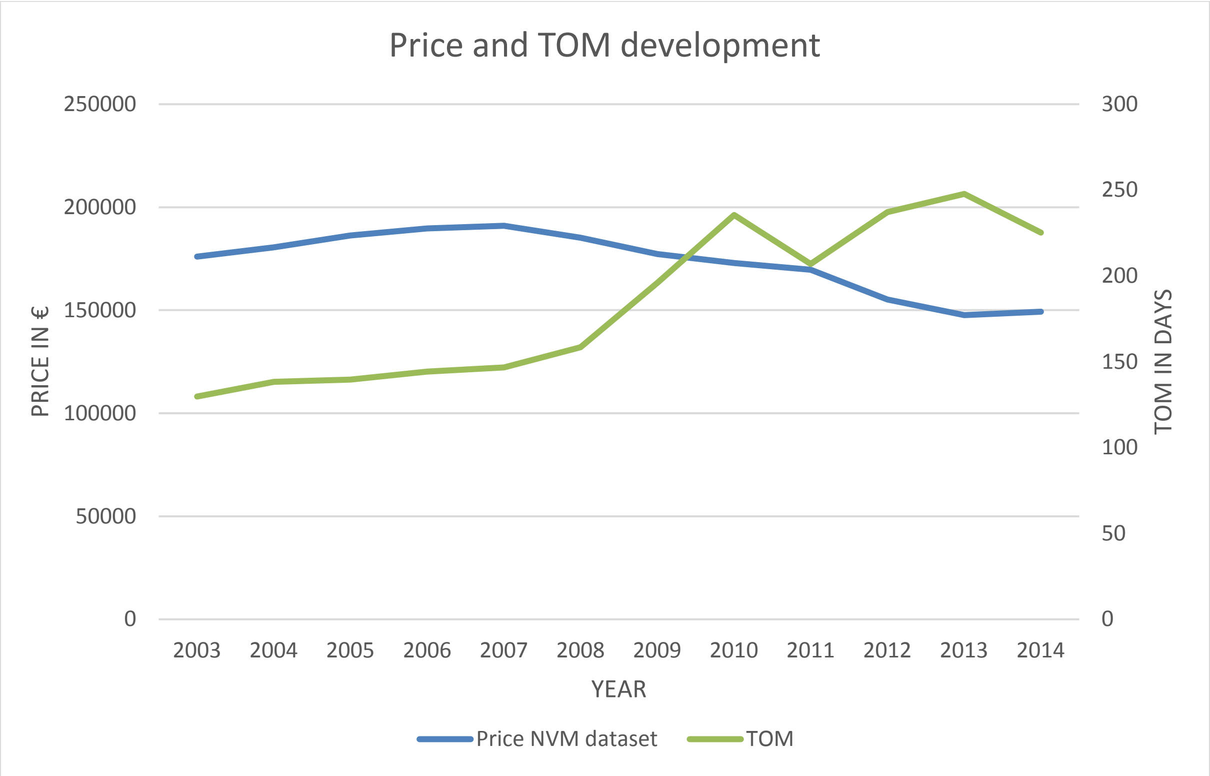


Figure 3.1: The development of price and TOM between 2003 and 2014 in the PGV dataset.

The inverse relationship between housing prices and TOM can also be seen in Figure 3.1. More housing transactions take place in years characterized by a housing market boom compared to a bust (see appendix II). CBS (2017a) shows the average TOM in the Northern Netherlands for each quarter in 2014, ranging from 13 to 16 months. Quarterly averages cannot be calculated for the NVM dataset, but the CBS averages are considerably higher than the average for 2014 in the NVM dataset of 225 days. However, both institutions measure the transaction date differently: NVM uses the signing of the selling contract and CBS employs the finalization at the notary. Furthermore, TOM counting might be influenced by sellers who withdraw their house from the market and then reoffer it after a certain time. De Kam and Mey (2017) find higher TOMs among seller around the Groningen gas field than reported by NVM. The NVM states, however, that they compensate for reentering sellers by continuing the TOM if the retracted house is reoffered within two weeks (NVM, 2015). The different types of houses in the dataset are shown in Table 3.1, apartments are excluded from the dataset due to the small number. The majority of the houses is built between 1960 and 2000.

Type of house	**Amount**
Townhouse	35,311
Corner house	18,129
Half double	32,275
Detached	37,193
Total	122,908

Table 3.1: The distribution between the different types of houses in the PGV dataset.

Finally, the geographical coverage of the housing transactions in the dataset over the municipalities in the Northern Netherlands is shown in Figure 3.2. It shows a higher amount of transactions in urbanized areas such as the Municipalities of Groningen, Leeuwarden, Assen, and Emmen. The area around the Groningen gas field is characterized by a low number of transactions. The low amount of cases in the risk area could cause difficulties in finding the effect of earthquakes in a dataset covering the whole of the Northern Netherlands.

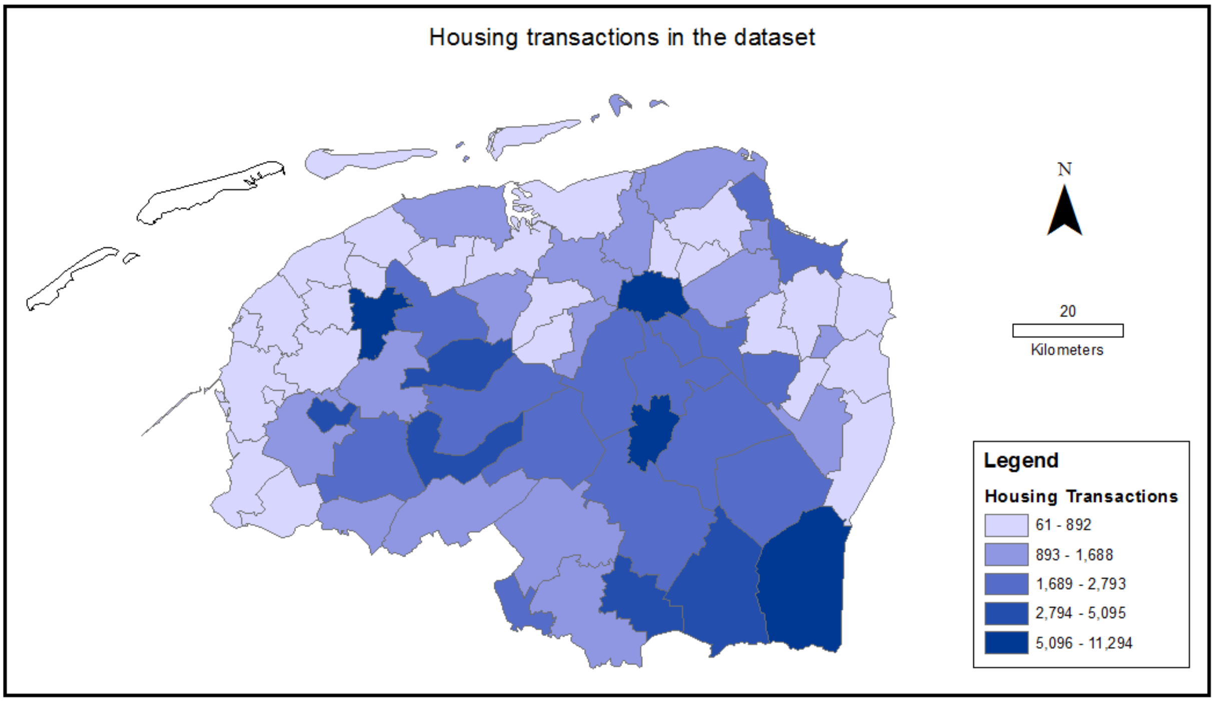


Figure 3.2: The geographical distribution of the housing transactions over the 2009 municipalities.

3.3.2 Earthquake indicators

The DID approach is based on the comparison between treatment and control areas which are supposed to have similar characteristics (Schwartz et al., 2006). The treatment areas are the regions at risk of earthquakes that can cause damage to houses or decrease the quality of life (Koster & Van Ommeren, 2015). Different risk areas are used in earlier studies, ranging from certain municipalities to ZIP-code areas with a certain percentage of damaged houses. A variety of reference areas is also used: municipalities surrounding the risk area or reference properties in other regions in the Netherlands (Atlas voor Gemeenten, 2017; Boelhouwer et al., 2016). CBS (2017b) used the percentage of damaged houses in a neighborhood to identify risk areas and the result can be seen in Figure 3.3. The surrounding neighborhoods are used as reference areas, excluding certain neighborhoods based on population structure, median income, employment percentage, housing value, percentage of owner-occupied houses, and address density.

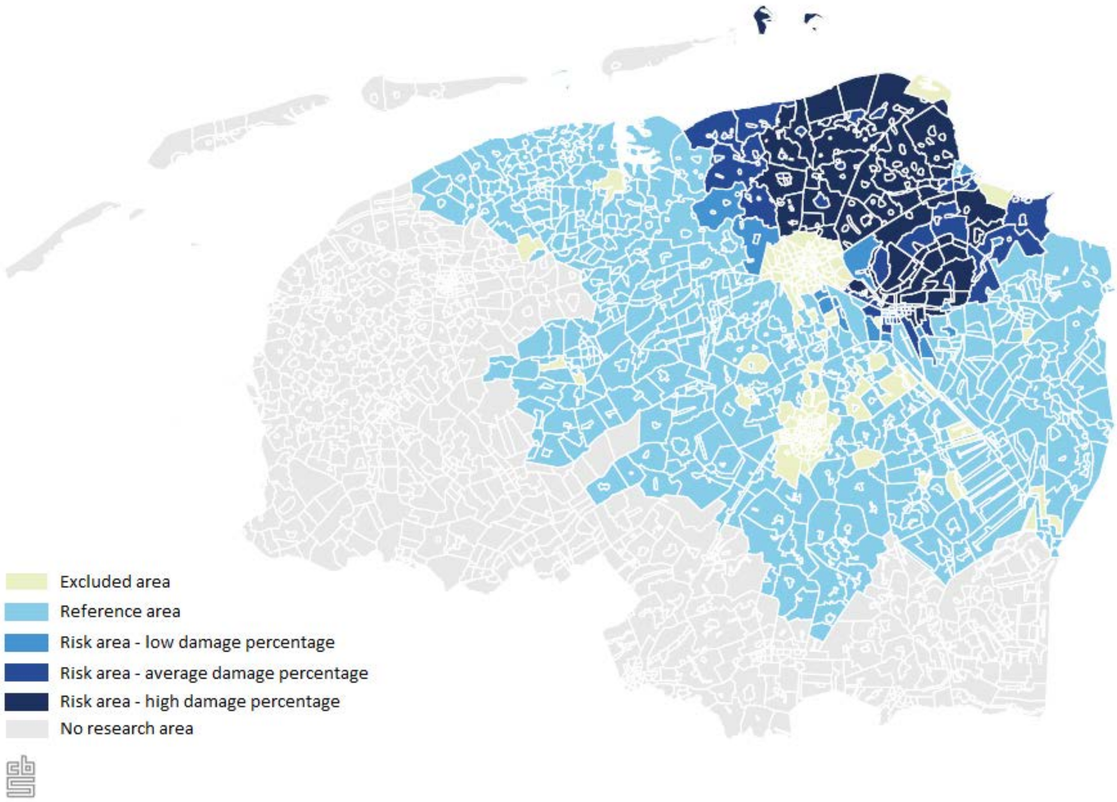


Figure 3.3: The risk and reference areas. Source: CBS (2017, p.13) (edited).

The DID models employ the risk and reference areas of CBS (2017b) since they are detailed at a neighborhood level. However, ambiguity exists concerning the size of the risk area and appropriate reference areas. Comparing the house type and building year of the housing transactions in the dataset between the risk and reference areas does indicate significant differences regarding the characteristics of the transacted houses (see Table 3.2). Furthermore, several areas close to the gas field such as the Municipality of Groningen are excluded, while a negative image might exist for the whole Province of Groningen (Atlas voor Gemeenten, 2017). Reference areas close to the Groningen gas field could be associated with earthquake risk. Unfortunately, this thesis was not able to use the whole Province of Groningen as a risk area due to multicollinearity issues that are elaborated upon in Chapter 4. The dataset did not allow the use of the sustained damage to a house or damage percentages at a ZIP-code level that are included in a hedonic model in Koster (2016). The DID approach does not analyze differences within the risk area and only compares TOM before and after one major earthquake. The time effect might be more complex considering many earthquakes have been taking place since the 1990s and also after 2012. DID modeling is useful to give a first impression of the effect of earthquake risk on TOM, but an indicator including a more detailed representation of the spatial and temporal dimension of earthquake risk is necessary to fully understand the impact of earthquakes.

	Risk area		Reference area		Ref – Risk
	Mean	**SE**	**Mean**	**SE**	**Mean difference**
Type of house					
- Townhouse	0.195	0.396	0.187	0.390	-0.007*
- Corner house	0.120	0.325	0.107	0.309	-0.013***
- Half double	0.298	0.457	0.290	0.454	-0.008*
- Detached	0.388	0.487	0.416	0.493	0.029***
Building year					
- <1945	0.322	0.467	0.201	0.401	-0.120***
- 1945-1980	0.456	0.498	0.514	0.500	0.057***
- >1980	0.222	0.416	0.285	0.451	0.063***
N	13,929		39,386		53,315

*Table 3.2: The characteristics of the risk and references areas. A mean comparison t-test is used to calculate the differences between the means with *, ** and *** indicating significance levels of respectively 10%, 5% and 1%.*

The accumulated PGV at the house location could be suitable to capture both the spatial and temporal effect of earthquakes. The Peak Ground Velocity (PGV) can be used to measure the intensity of earthquake *e* in centimeters per second at the location of the house *i* and is calculated using the following formula:

$$\log_{10} PGV_{ei} = -1.53 + 0.74M_e - 1.33 \log_{10} R_{ie} - 0.00139R_{ie} \quad (3.7)$$

The magnitude at the epicenter of the earthquake is represented by *M* and *R* is the hypocentral distance between the house location and the epicenter. This distance is calculated using the distance in kilometers and the depth of the earthquake (Duran & Elhorst, 2017; Koster & Van Ommeren, 2015). To take into account the temporal dimension, Koster and Van Ommeren (2015) count every variable with a PGV above 0.5 cm/s at the house location before the sale of the house, these are noticeable earthquakes. This thesis employs a variable that calculates the PGV for every earthquake at the house location and accumulates them. It takes into account the magnitude of different earthquakes and can be used to assess both the spatial and temporal dimension of earthquake risk.

3.3.3 Variables

The dependent variable of the estimated regression models is the time-on-the-market measured in days. The variable has been transformed into a natural logarithm to achieve a more normal distribution (see Appendix II). In the DID and PGV model, the logTOM variable is regressed on a number of independent variables that can be seen in Table 3.3 and Table 3.4, where the main interest is in the earthquake indicators. The before and after dummy variables are used in the DID approach. The before dummy acts as a baseline for the difference between the risk and references areas of CBS (2017b), while the after dummy shows the effect of the Huizinge earthquake or the Middelstum earthquake. The year of sale is used to identify whether a house sale can be considered to be after an earthquake, for example, house sales are positioned after the 2012 Huizinge earthquake from 2013 onwards. The accumulated PGV is the other variable used to measure earthquake risk. The accumulation is done on a logarithmic scale with base 10 but it has been transformed into a natural logarithm to ease interpretation. The logPGV variable is not included in its original form in the regression models. It can be assumed that below a certain intensity the effect of earthquake risk on the housing market is negligible (Koster & Van Ommeren, 2015). LogPGV is divided into 50 equal groups to analyze when earthquakes start to have a significant impact on TOM, the lowest groups are then set to zero in the logPGV variable that is included in the regression model.

Variable name	Description
Dependent	
logTOM	Natural logarithm of the time-on-the-market in days
Earthquake risk	
Before	Dummy whether the house is located in the risk (1) or reference (0) area
AfterHuizinge	Dummy whether the house is located in the risk area and the sale took place after the Huizinge earthquake (1) or not (0)
AfterMiddelstum	Dummy whether the house is located in the risk area and the sale took place after the Middelstum earthquake (1) or not (0)
logPGVoriginal	The natural logarithm of the accumulated PGV at the house location
logPGVgroups	logPGVoriginal divided into 50 equal groups.
logPGV	The natural logarithm of the accumulated PGV at the house location, where insignificant groups in logPGVgroups have been set to zero
Selling price	
logPrice	Natural logarithm of the transaction price in euros

Structural	
LogArea	Natural logarithm of the square meters of the house
LogLot	Natural logarithm of the lot size.
Rooms	Dummy for number of rooms (1 = five or more)
Stories	Dummy for the number of stories (1 = three or more)
Type	Categorical variable for the type of house
Built	Categorical variable for the building period
Balcony	Dummy for balcony (1 = Yes)
Garage	Dummy for garage (1 = Yes)
Attic	Dummy for attic (1 = Yes)
Loft	Dummy for loft (1 = Yes)
Barn	Dummy for barn (1 = Yes)
Fire	Dummy for fireplace (1 = Yes)
Maintenance	Dummy for inside Maintenanceenance status (1 = Good)
Locational	
Density	Categorical variable for the address density
Elderly	Dummy for the percentage of people with 65 or more years of age (1 = higher than the median)
Fixed effects	
Year	Categorical variable for year of sale
Municipality	Categorical variable for municipality
Instruments	
Unemployment	Unemployment benefits granted per 1000 inhabitants in the neighborhood
Disability	Disability benefits granted per 1000 inhabitants in the neighborhood
Train	Dummy for the distance to the nearest train station (1 = higher than the median)

Table 3.3: An overview of the dependent and control variables.

The models also include variables to control for the TOM and price trade-off, structural and locational characteristics, and market conditions. The selling price is also taken as a natural log to approach a more normal distribution. The logPrice variable is not included directly into the TOM model due to a risk of endogeneity since the selling price and TOM are simultaneously determined (Dubé & Legros, 2016). Therefore, a 2SLS model is used that estimates selling price in the first stage and uses the predicted values for selling price as an independent variable in the TOM model estimated in the second stage (Brooks & Tsolacos, 2010). The structural characteristics can be divided in those related to the size of the house, being the square meters of the house and the lot (both as a natural logarithm), and the number of rooms. Furthermore, the type of the house is included, complemented with other physical characteristics, to account for the effect of the atypicality of a house on TOM (Haurin et al., 2010).

	PGV dataset				DID dataset			
Variable	**Mean**	**SE**	**Min**	**Max**	**Mean**	**SE**	**Min**	**Max**
logTOM	4.560	1.189	0.693	6.968	4.633	1.179	0.693	6.968
Before	-	-	-	-	0.261	0.439	0	1
AfterHuizinge	-	-	-	-	0.032	0.177	0	1
After Middelstum	-	-	-	-	0.159	0.366	0	1
logPGV	0.829	0.739	-0.345	3.439	-	-	-	-
logPrice	12.005	0.381	10.904	13.015	11.984	0.402	10.674	13.046
logArea	4.764	0.261	3.401	5.517	4.776	0.265	3.401	5.517
logLot	5.698	0.777	2.398	8.516	5.952	0.751	2.773	8.516
Rooms (1 = >4)	0.574	0.495	0	1	0.575	0.494	0	1
Type townhouse	0.287	0.453	0	1	0.189	0.392	0	1
Type corner house	0.148	0.355	0	1	0.110	0.313	0	1
Type half double	0.263	0.440	0	1	0.292	0.455	0	1
Type detached	0.303	0.459	0	1	0.409	0.492	0	1
Built 1500-1905	0.046	0.209	0	1	0.044	0.205	0	1
Built 1906-1930	0.112	0.315	0	1	0.119	0.324	0	1
Built 1931-1944	0.074	0.262	0	1	0.070	0.254	0	1
Built 1945-1959	0.064	0.245	0	1	0.080	0.272	0	1
Built 1960-1970	0.151	0.358	0	1	0.164	0.371	0	1
Built 1971-1980	0.225	0.418	0	1	0.254	0.435	0	1
Built 1981-1990	0.130	0.336	0	1	0.113	0.317	0	1
Built 1991-2001	0.145	0.352	0	1	0.114	0.318	0	1
Built >2001	0.053	0.225	0	1	0.041	0.199	0	1
Balcony (1 = Yes)	0.067	0.250	0	1	0.059	0.236	0	1

Garage (1 = Yes)	0.470	0.499	0	1	0.569	0.495	0	1
Attic (1 = Yes)	0.329	0.470	0	1	0.307	0.461	0	1
Loft (1 = Yes)	0.124	0.329	0	1	0.130	0.336	0	1
Barn (1 = Yes)	0.657	0.475	0	1	0.626	0.484	0	1
Fire (1 = Yes)	0.059	0.235	0	1	0.061	0.240	0	1
Main (1 = Good)	0.864	0.343	0	1	0.851	0.356	0	1
Density <500	0.374	0.484	0	1	0.512	0.500	0	1
Density 500-1000	0.261	0.439	0	1	0.304	0.460	0	1
Density 1000-1500	0.201	0.401	0	1	0.141	0.348	0	1
Density 1500-2500	0.107	0.310	0	1	0.042	0.201	0	1
Density >2500	0.057	0.231	0	1				
Elderly	0.482	0.500	0	1	0.495	0.500	0	1
Unemployment	36.805	21.279	0	300.000	-	-	-	-
Disability	-	-	-	-	91.390	35.047	5.000	648.000
Train	0.494	0.500	0	1	0.499	0.500	0	1
N			122,908				53,315	

Table 3.4: Descriptive statistics of the employed variables in both datasets.

The locational characteristics include the address density to account for the effect of more vibrant urban markets. Several regions in Groningen are characterized by population decline and are economically deprived which is partly absorbed by the share of the elderly in the surrounding neighborhood (Koster & Van Ommeren, 2015). A proxy for population decline is not included in the model since it is expected that this effect is accounted for by the spatial fixed effects. Municipalities from 2009 are used for these fixed effects and they account for unobserved locational attributes. Municipalities are favored above neighborhood fixed effects since the small size and substantial amount of the latter might absorb the explanatory effect of the other variables (Livy & Klaiber, 2016). Besides spatial fixed effects, the model also includes year fixed effects to control for booms and busts in the housing market that affect TOM. The correlation matrix including the independent variables did not show correlations large enough to become problematic by causing multicollinearity in the regression model. However, the before variable used for the DID model shows considerable overlap with the municipalities. In the PGV dataset, the logPGV variable accumulates earthquakes over time and might be correlated with the year fixed effects. Both cases might cause multicollinearity issues.

The 2SLS model in both the DID and PGV approach requires two instrumental variables to run the first stage estimating logPrice. Valid instruments are exogenous and relevant: they are uncorrelated with the error term and are highly correlated with the endogenous independent variable. This means that the instrument has to be unrelated to TOM but has a strong relationship to selling price (Brooks & Tsolacos, 2010; Dubé & Legros). The search for instrumental variables was inspired by the fact that locational variables are often excluded from TOM model, while they generally have a strong explanatory power in hedonic models. Furthermore, Dubé & Legros (2016) show that locational characteristics play a larger role in explaining selling price than TOM. The variables on the unemployment benefits and disability benefits per 1000 inhabitants in the surrounding neighborhood might be suitable instruments with a Pearson correlation coefficient of below 0.1 with the dependent variable logTOM. The correlation with the endogenous independent variable logPrice is stronger, though still considered weak since they are between 0.2 and 0.3. Unemployment and disability benefits could be related to regions with lower incomes, where housing prices might be lower (Koster & Van Ommeren, 2015). This could not necessarily affect TOM since houses in lower price classes still have a substantial arrival rate of buyers (Knight, 2002). Dubé and Legros (2016) indicate that the proximity to certain amenities such as a school, highway, major boulevard or park have substantial higher t-values in the selling price model compared to the TOM model.

The dataset includes a variable on the distance to the nearest train station that is transformed into a dummy since the distributions deviates considerable from a normal one. Although the train dummy has similar weak correlations with logTOM and logPrice, the variable is not significant in the OLS model estimating TOM and highly significant if selling price is used as a dependent variable. The diagnostic tests after the 2SLS model show that disability benefits and the train dummy are valid instruments in the DID model and unemployment benefits and the train dummy in the PGV model (see paragraphs 4.2.1 and 4.3.1). The differences in the combination of two instrument might be caused by a change in context: the PGV model employs a larger sample covering the whole of the northern Netherlands including more urbanized areas than the DID model. The 2SLS models and also the OLS specifications are discussed in the next chapter; however, the ethical issues surrounding this study are given attention first.

3.4 Ethical considerations

The main ethical concern is to avoid harm to people (Babbie, 2013). In this research, the housing transactions are information that can be connected to persons. Furthermore, the dataset might prove valuable for commercial or other purposes. Therefore, I signed an agreement to not share the data and only use it for this Master's thesis. The selection of the data available for this thesis did not include the exact house location or the specific transaction date. The report only shows general trends found using the data and does not discuss specific cases, thereby ensuring anonymity. The confidentiality of the data is guaranteed by saving the data on a university account protected by a password. The thesis is transparent about the process leading to the presented findings to make sure no unsubstantiated positive or negative view is given about the housing market in the earthquake region around the Groningen gas field.

4 Results

4.1 Exploratory analysis

A first insight in the effect of earthquake risk on TOM can be seen in Figure 4.1 which plots the average TOM in the risk and reference areas over the years 2003 to 2014. The classification of risk and reference areas based on the percentage of houses in the neighborhood sustaining earthquake damage of CBS (2017b) is employed; the average TOM is calculated using the housing transactions in the Difference-In-Difference (DID) dataset. The graph shows a similar TOM trend in risk and reference areas between 2003 and 2008, being low due to the boom in the housing market. The financial crisis causing a major bust increases the TOM from 2008 onwards, with recovery starting around 2013 (Agnello & Schuknecht, 2011).

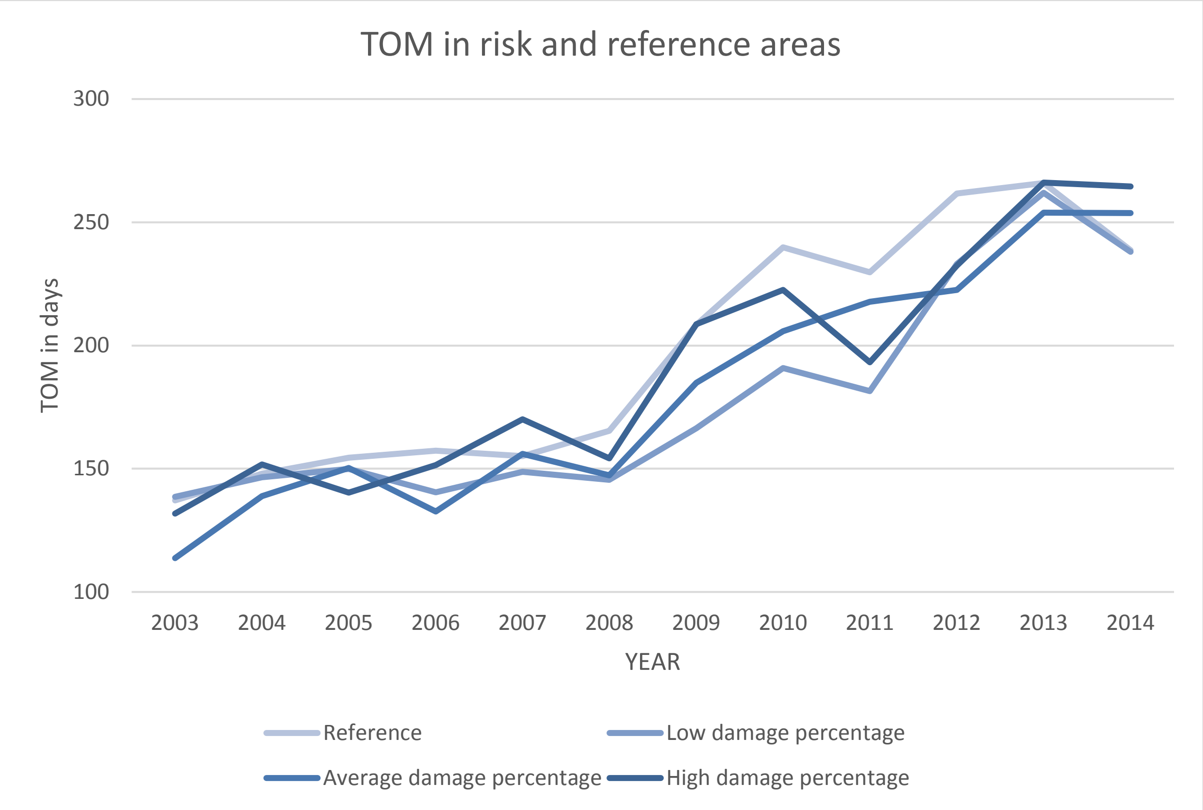


Figure 4.1: The development of TOM between 2003 and 2014 in the risk areas divided based on damage percentages and the reference areas of CBS (2017b), N = 53,315.

The risk areas outperform the reference area in the crisis. However, a sharp decrease can be seen in the TOM after 2013 in the reference areas and risk area with a low percentage of damaged houses, while the TOM line in the risk areas with average and high damage percentages remains flat. This could indicate the effect of the Huizinge earthquake that took place in August 2012. It could have increased the awareness of earthquake risk in the area around the Groningen gas field (Bosker et al., 2017), thereby highlighting the temporal dimension of the earthquake impact on the housing market. The time effect might be more complex since (heavy) earthquakes also occurred before and after the Huizinge earthquake. Therefore, it might be suitable to attempt a time series analysis of the occurrence of earthquakes and the development of TOM which might also enable forecasting the effect of an occurring earthquake on TOM (Brooks & Tsolacos, 2010). However, Figure 4.1 also shows the differentiated spatial effect of earthquakes: the impact appears to be stronger in areas with high damage and average damage percentages. The outcome is similar to the research of CBS (2017b) that includes data until 2017, showing that reference areas and areas with low damage percentages recover faster from the housing market bust. The complex spatial and temporal dimension provides an incentive to create a model employing panel data since it is able to comprise both spatial and time variables (Brooks & Tsolacos, 2010).

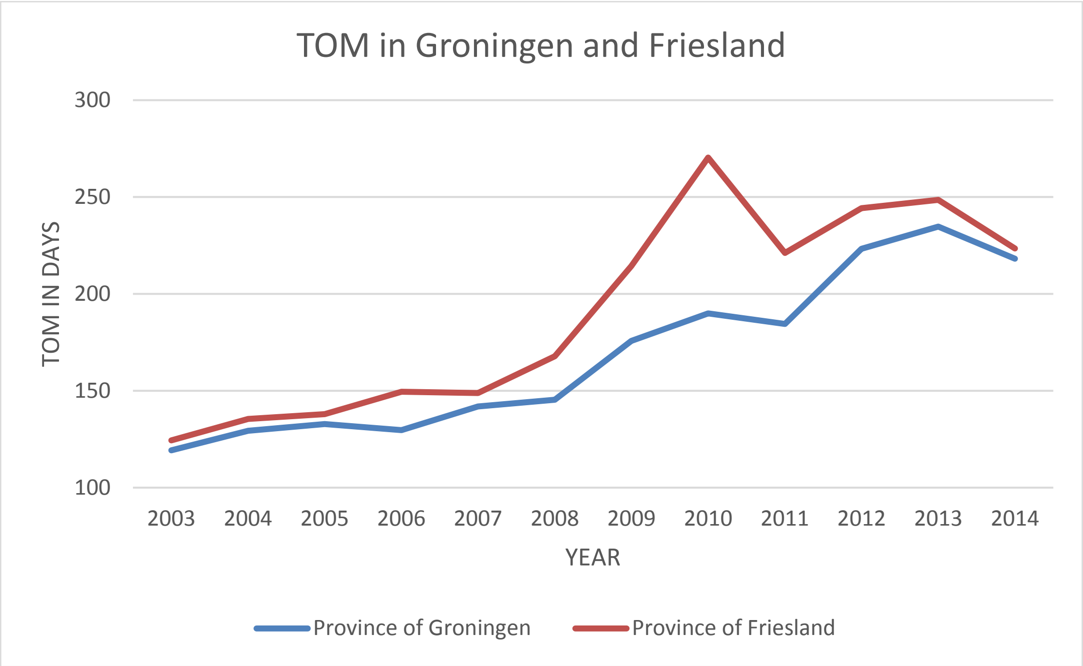


Figure 4.2: The development of TOM between 2003 and 2014 in the Province of Groningen (risk area) and the Province of Friesland reference area, N = 82,327.

The CBS (2017b) classification can be questioned since it excludes areas close to the earthquake region such as the City of Groningen. Furthermore, several reference areas close to the gas field might be influenced by earthquake risk or the negative image that could exist for the whole Province of Groningen (Atlas voor Gemeenten, 2017). Therefore, it could be a suitable option to employ the Province of Groningen as a risk area. The Province of Friesland could be employed as a reference area. The urban hierarchy in both provinces is dominated by one major city, being Leeuwarden and Groningen (the city). Furthermore, both provinces have municipalities that are characterized by population decline. Figure 4.2 shows the TOM development in both provinces using the large PGV dataset. The Province of Drenthe has been excluded since it is unclear if parts in the north close to the gas field are affected by the earthquakes. Both provinces have a similar TOM trend, though the average is consistently lower in the Province of Groningen. Both provinces appear to recover between 2013 and 2014 after the major bust starting in 2008. Although the Province of Friesland recovers slightly faster, a clear earthquake effect cannot be distinguished in the graph.

The trend in the Province of Groningen might be dominated by the City of Groningen, covering the TOM development in the more rural areas around the gas field. Therefore, the DID regression models use the risk area classification of CBS (2017b) since it identifies risk areas at a low spatial level around the gas field, excluding the effect of the large housing market in the City of Groningen. Furthermore, employing the Province of Groningen as a risk area causes multicollinearity issues in the DID regressions if municipalities are included as fixed effects, solving these issues is beyond the scope of this thesis. The CBS (2017b) classification does have its shortcomings and so does the DID approach that simplifies the effect by using only one risk area and comparing it to a reference area before and after one point in time. The logPGV variable indicating the intensity of occurred earthquakes at a specific location could be a more suitable option since it accounts for both the complex spatial and temporal effect of earthquake risk. Furthermore, the regression model can include the whole of the Northern Netherlands, thereby not predetermining the risk area. The results of the DID and the PGV regression models are discussed in the following paragraphs.

4.2 Difference-In-Difference regression models

4.2.1 Model performance

The DID models are based on comparing the area at risk of earthquakes, being neighborhoods with damaged houses, to a reference area with similar locational characteristics before and after the Huizinge earthquake of 2012 or the Middelstum earthquake of 2006. This is done using two dummy variables. The before variable has a value of one if the house is located in the risk area and acts as a baseline. It shows the differences between the areas before the earthquake

(Schwartz et al., 2006; Van Duijn et al., 2016). The after variable then gets a value of 1 if the house is located in the risk area after the earthquake. The results are shown in Table 4.1 which includes three specifications. First, the Huizinge earthquake is used as a turning point. The first model is an OLS model, containing the earthquake indicators, structural and locational attributes, spatial fixed effects to control for unobserved regional characteristics and temporal fixed effects to take into account market conditions. The 2SLS model then adds logPrice as an endogenous explanatory variable. The second 2SLS model employs the Middelstum earthquake as the before and after event.

The OLS model has an R-squared of 0.1015 which is lower than comparable OLS models estimating TOM, having an R-squared ranging from 0.13 to 0.15 (e.g. Anglin et al., 2003; Dubé & Legros, 2016). An explanation for this difference could be that Anglin et al. (2003) use more variables to control for market conditions, while Dubé and Legros (2016) include a larger amount of structural and locational attributes. The estimator in the OLS model can be considered a Best Linear Unbiased Estimator (BLUE) if the four assumptions of the classical linear regression model hold (Brooks & Tsolacos, 2010):

1. The errors have zero mean.
2. The errors have a constant variance.
3. The errors are independent from one another.
4. There is no relationship between the error and corresponding independent variable.

The first assumption always holds if a constant is included in the model. The second assumption implies that the residuals are homoscedastic; however, heteroscedasticity has been detected in the second OLS model by the Breusch-Pagan/Cook-Weisberg test. Robust standard errors are used to take this into account, although they make it harder to achieve a significant outcome. The third assumption might be violated due to autocorrelation between errors over time or over space which is often the case in the housing market (Brooks & Tsolacos, 2010). However, it is hard to test this since the used dataset does not include specific time or location data. The fourth assumption could have been violated if the selling price was added directly to the model due to its endogeneity with the dependent variable logTOM (Dubé & Legros, 2016). A fifth assumption is needed for making inferences about population parameters based on the sample parameters estimated on a finite amount of data, namely that the error term is normally distributed (Brooks & Tsolacos, 2010). The Q-Q and P-P plot of the residuals of the second OLS model show slight deviations from normality; the significance of the normality test confirms this. However, the consequences of violating a normality assumption are negligible when using a large sample due to the Central Limit Theorem (Brooks & Tsolacos, 2010). Finally, the model is characterized by multicollinearity issues due to a high overlap between the risk and reference area and the municipalities used for the spatial fixed effects. Near multicollinearity does not affect the BLUE properties of the OLS model; however, it can cause difficulties with obtaining small standard errors (Brooks & Tsolacos, 2010). This might be the case for the before variable that has a VIF score around 10.

	OLS model		**2SLS model (1)**		**2SLS model (2)**	
	Coefficient	Robust SE	Coefficient	Robust SE	Coefficient	Robust SE
Before	-0.0707**	0.0358	-0.0915**	0.0368	-0.0211	0.0389
AfterHuizinge	0.0675*	0.0348	0.0599*	0.0353		
AfterMiddelstum					-0.1026***	0.0225
logPrice			-1.3787***	0.2936	-1.3822***	0.2935
logArea	0.3037***	0.0243	1.1223***	0.1757	1.1226***	0.1757
logLot	-0.0886***	0.0104	0.1137**	0.0445	0.1154***	0.0445
Rooms	-0.0187*	0.0111	0.0217	0.0141	0.0216	0.0141
Type corner	-0.0255	0.0184	-0.0021	0.0191	-0.0013	0.0191
Type half double	0.1267***	0.0174	0.2299***	0.0281	0.2314***	0.0281
Type detached	0.4885***	0.0209	0.7162***	0.0528	0.7180***	0.0528
Balcony	0.1054***	0.0209	0.1474***	0.0228	0.1489***	0.0228
Garage	0.0485***	0.0125	0.1595***	0.0268	0.1594***	0.0268
Attic	-0.0457***	0.0115	-0.0179	0.0130	-0.0189	0.0130
Loft	0.0373**	0.0155	0.0726***	0.0172	0.0729***	0.0172
Barn	-0.0087	0.0110	0.0166	0.0122	0.0164	0.0122
Fire	0.0411**	0.0211	0.1433***	0.0302	0.1434***	0.0303
Maintenance	0.2583***	0.0146	0.5015***	0.0538	0.5024***	0.0538
Density 500-1000	-0.0706***	0.0139	0.0183	0.0235	0.0204	0.0235
Density 1000-1500	-0.0873***	0.0207	0.0362	0.0337	0.0394	0.0337
Density 1500-2500	0.0280	0.0307	0.1061***	0.0350	0.1066***	0.0350
Elderly	-0.0395***	0.0107	-0.0035	0.0133	-0.0028	0.0133
Built dummies	YES		YES		YES	
Year FE	YES		YES		YES	
Municipality FE	YES		YES		YES	
N	53,315		53,315		53,315	
R^2	0.1015		0.0877		0.0878	
Adjusted R^2	0.1002					
Joint sign. F-test	129.71***					
Wald Chi^2			6188.15***		6202.96***	

*Table 4.1: The coefficients and the robust standard errors within brackets of the DID models, with *, ** and *** indicating significance levels of respectively 10%, 5% and 1%.*

To include the effect of the selling price on TOM, a 2SLS model has been estimated on the effect of the Huizinge earthquake using the disability benefits per 1000 inhabitants in the surrounding neighborhood and a dummy on the distance to the nearest train station as instruments. Since two instruments are used on only one endogenous variable, the Wooldridge's robust score test of over-identifying restrictions can be performed to check whether the instruments are uncorrelated to the error term. Wooldridge's test is used because robust standard errors where used to compensate for heteroscedasticity. The over-identification test appeared to be insignificant ($p = 0.3414$) and shows that the employed instruments are exogenous. The relevance of the instruments can be analyzed with an F-test on the value of the additional instruments in the first stage regression which is highly significant (0.000) and has an F-statistic far above 10 (163.36). These tests indicate that unemployment and train are valid instruments. The Wooldridge's score test and a regression-based F-test are used to test whether logPrice has to be treated as endogeneous, meaning an OLS model including selling price as an independent variable is not efficient. Both test are significant at a 5% level, meaning selling price cannot be treated as exogenous from TOM. This was expected based on the optimization problem that sellers face between maximizing the price and minimizing TOM (Dubé & Legros, 2016; Knight, 2002; Yavas & Yang, 1995). It is not out of the ordinary that the 2SLS model has a lower R-squared than the OLS model. The above tests have similar results for the 2SLS model on the Middelstum earthquake, showing that the instruments are valid and that selling price is endogenous. Therefore, the 2SLS models are preferred of the OLS specifications since those coefficients might be biased.

4.2.2 Interpretation

The main interest of this research is in the coefficients of the before and after variables in the 2SLS models. The 2SLS model is more comprehensive by including the effect of selling price, the OLS estimators are at risk of omitted variable bias (Brooks & Tsolacos, 2010). The before dummy in the first 2SLS model indicates the difference between the risk and reference area before the 2012 Huizinge earthquake. The coefficient is significantly different from zero at the 5% level and indicates that TOM in the risk area is 8.7% ($100 * (e^{-0.0915} – 1)$) lower compared to the reference area. It confirms the trend shown in Figure 4.1 in the exploratory analysis, where the annual average TOM is consistently higher in the reference areas until 2013. It does not indicate any anticipation effects of buyers at a major earthquake. The after dummy then shows the external effect of the event of a major earthquake, being the Huizinge earthquake in

2012. In the 2SLS model, TOM is 6.2% (100 * ($e^{0.0599}$ – 1)) higher in the risk area compared to the reference area after 2012. However, the coefficient is only significant at the 10% level.

The first 2SLS model assumed that the Huizinge earthquake acted as a major turning point increasing the awareness of earthquake risk around the Groningen gas field (Atlas voor Gemeenten, 2017; Bosker et al., 2016; Koster & Van Ommeren, 2015). An alternative assumption might be that buyers and sellers were already aware of the risk of earthquakes before 2012 (De Kam, 2016; De Kam & Raemakers, 2014), for example, after the Middelstum earthquake of 2006 with a similar magnitude as the Huizinge earthquake. The after variable can be changed into a dummy having the value of 1 if the house is located in the risk area after 2006. Table 4.1 shows that the before variable loses its significance, while the after variable becomes negative and highly significant. This does not indicate selling difficulties, although sellers in the earthquake region might chose to sell faster by lowering the listing price and prevent a further lowering of housing values (De Kam & Mey, 2017). However, it is also likely that the before and after dummies are picking up other effects. The exploratory analysis showed that the risk areas had lower TOM averages before 2013 than the reference areas, a pattern confirmed by the before dummy in the first 2SLS model. Therefore, the Huizinge earthquake might have been a turning point in the effect of earthquake on the housing market.

Earlier studies have already shown the lowering of housing prices in the area around the Groningen gas field (Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster, 2016), indicating that sellers have to lower prices to find a buyer and that the risk of earthquakes has a negative external effect on housing value. The DID models give more insight into the situation of seller by showing that they also face a lengthier selling process after the 2012 Huizinge earthquake compared to sellers in similar regions. The research of CBS (2017b) showed a similar pattern, where the average TOM in risk areas takes longer to recover from the housing market bust compared to reference areas. The DID models provide evidence for the importance of the Huizinge earthquake as a turning point increasing the awareness of earthquake risk around the Groningen gas field, thereby showing the temporal dimension of earthquake risk (Atlas voor Gemeenten, 2017; Duran & Elhorst, 2017). Buyers might be less willing to buy houses due to the risk of earthquake damage and a lower quality of life due to feelings of unsafety (Boelhouwer et al., 2016; Koster & Van Ommeren, 2015; Postmes et al., 2017). Compensation schemes might not be able to counteract earthquake effects since they impose high transactions costs and not always enable adequate repairs (De Kam & Spijkerboer, 2015; Van der Voort & Vanclay, 2015).

Contrary to the TOM regression model of Bosker et al. (2016), the DID models show that being located in an area at risk of earthquakes increases TOM, indicating the spatial dimension of earthquake risk. Bosker et al. (2016) only included one year before the Huizinge earthquake and use a different risk area which is compared to reference properties in the Netherlands. De Kam (2016) notes that the maintenance status of houses around the Groningen gas field has declined since 2012 and it might influence the results if the reference properties do not follow that pattern. The decline in maintenance status can be related to earthquake damage and lacking repair budgets (De Kam, 2016). Bosker et al. (2016) do include the received repair budgets in their TOM model. However, the model does not take into account the simultaneity with selling price and spatial fixed effects.

Besides the earthquake indicators, the 2SLS models included a large amount of variables controlling for the effect of selling price, structural and locational attributes, and market conditions on TOM. The model on the Huizinge earthquake is interpreted here since it appears to give the best coverage of the earthquake effect. The 2SLS model adds the effect of selling price as an endogenous explanatory variable, resulting in a negative coefficient for logPrice. The assumed trade-off between selling price and TOM implied a positive relationship since patient sellers are expected to wait longer for a buyer with a higher reservation price or to take more time in the bargaining process to achieve a higher transaction price (Dubé & Legros, 2016; Knight, 2002; Yavas & Yang, 1995). Dubé & Legros, (2016) also find a negative relationship between selling price and TOM, explaining the result by stating that houses with better amenities sell faster. Higher prices could be a signal of the high quality of the house due to repairs and improvements which could be of great importance in a region characterized by earthquake risk. In general, the 2SLS model shows the strong relationship between the selling price and TOM that are simultaneously determined in the selling process (Dubé & Legros, 2016).

Several structural and locational attributes lose their significance after adding logPrice to the 2SLS model. The different coefficients and significance levels in the OLS model can be caused by the fact that logPrice acts as an omitted variable (Brooks & Tsolacos, 2010). Contrary to models in earlier empirical studies (e.g. Anglin et al., 2003; Dubé & Legros, 2016; Springer, 1996), the majority of the structural attributes of the house have significant coefficients (). Detached and half double houses have a longer TOM compared to the reference category of townhouses. The size of the house and the lot increase the time it takes to find a seller since the market for these houses might be smaller. The building year dummies indicate that newer houses take longer to sell (see Appendix III); this is in line with the effect found by Anglin et al. (2003). A better maintenance status causes TOM to increase, showing that these seller might be more confident and is waiting for a buyer willing to pay a higher price. The density dummies and elderly percentage play a limited role in explaining TOM, although the spatial fixed effects might absorb their impact. The fixed effects are expected to control for the local economic situation and the effect of population decline (Koster, 2016). Finally, the effect of economic cycles is represented by temporal fixed effects. The coefficients of the time dummies increase during the housing market bust starting in 2008 and then start to decline after 2012 (see Appendix III).

4.2.3 Robustness

The results depend on the chosen demarcation the risk and reference areas which is based on the percentage of damaged houses in a neighborhood. The reference areas are selected on similar socioeconomic attributes (CBS, 2017b). The risk areas contain fewer housing transactions than the reference areas; however, this did not limit the analysis to highlight the effect of earthquakes on TOM. The risk area selection of CBS (2017b) can be criticized regarding excluding the City of Groningen which is located close to the Groningen gas field and using reference areas that might be affected by earthquake risk. A viable alternative might be to assume that a negative image exists for the whole Province of Groningen and use that as a risk area. The Province of Friesland could be used as a reference area. The two provinces are already compared in the exploratory analysis in paragraph 4.1, where no particular pattern regarding earthquakes could be discovered. This risk-reference specification overlaps perfectly with the municipalities used for the spatial fixed effect causing perfect multicollinearity, requiring modeling that is beyond the scope of this thesis.

The presented DID models only employ one risk area; however, spatial differentiation in the earthquake impact exists within this region (Koster & Van Ommeren, 2015). The damage percentages can be used to create two categories within the risk area: neighborhoods with a low damage percentage (below 31%) and those with an average and high damage percentage (above

31%). Before and after dummies can then be created for both categories (Van Duijn et al., 2016). The Huizinge earthquake of 2012 is used as the turning point and the results are shown in Table 4.2. The coefficient for the before dummy for both the low and high damage percentage areas is significant and negative, confirming the pattern that risk areas had a lower TOM before 2013. The pattern switches after 2013, the coefficients of the after dummies are positive; however, they are not significantly different from zero. Even by further specifying the risk area, the DID approach might not be able to capture to complex spatial dimension of earthquake risk. Furthermore, the DID technique simplifies the temporal dimension since many (weaker) earthquake have been taking place over the years besides the Huizinge and Middelstum earthquakes. The PGV approach calculates the accumulated earthquake impact at the exact location of the house and might be able to give more insight into the effect of earthquake risk on the housing market.

	2SLS model (1)		**2SLS model (3)**	
	Coefficient	Robust SE	Coefficient	Robust SE
Before	-0.0915**	0.0368		
AfterHuizinge	0.0599*	0.0353		
BeforeLow			-0.0718*	0.0387
Before High			-0.1373***	0.0431
AfterLow			0.0621	0.0563
AfterHigh			0.0601	0.0416
Selling price	YES		YES	
Structural and locational	YES		YES	
Year FE	YES		YES	
Municipality FE	YES		YES	
N	53,315		53,315	
R^2	0.0877		0.0891	
Wald Chi^2	6188.15***		6209.16***	

*Table 4.2: The coefficients and robust standard errors within brackets of the final DID model and a DID model with a more detailed specification of the risk area, with *, ** and *** indicating significance levels of respectively 10%, 5% and 1%.*

4.3 PGV regression models

4.3.1 Model performance

The PGV models measure the risk of earthquakes by means of the accumulated Peak Ground Velocity (PGV) at the location of the house. The logPGV variable is not bound to identifying control areas that are similar to the treatment area; therefore, the model can be applied to the whole of the Northern Netherlands. The original logPGV variable is first divided into fifty equal groups that are included in an OLS and a 2SLS model as 49 dummies, excluding the first dummy as a reference group. The models also contain structural and locational attributes, spatial fixed effects, and temporal fixed effects. The 2SLS model adds selling price as an endogenous explanatory variable. The dummies for group 2 to 4 are insignificant in both the OLS and 2SLS model (see Figure 4.3 and Appendix II). Therefore, it is assumed that these low PGV values have a negligible effect on the housing market and they are set to zero, operating to some extent as a reference area. Finally, the transformed logPGV variable is then included in an OLS and a 2SLS model, the estimates of these TOM models are shown in Table 4.3.

The OLS model has a higher (adjusted) R-squared than the OLS models in the DID approach, although 0.1120 is still below the R-squared found in other TOM studies (e.g. Anglin et al., 2003; Dubé & Legros, 2016). Similar to the second OLS model in the DID approach, the OLS model with logPGV is also characterized by heteroscedasticity, thereby violating the second assumption of the classical linear regression model. Therefore, robust standard errors are again used to account for this. To avoid a correlation between the error term and an independent variable, the fourth assumption, selling price is not added directly to the OLS model as an independent variable (Dubé & Legros, 2016). The error term of the PGV OLS model also shows deviations from normality; however, this does not appear to be problematic due to the use of a large sample (Brooks & Tsolacos, 2010). Multicollinearity issues also play a role in the PGV approach since the PGV may show overlap with certain municipalities but also with the year dummies since the logPGV variable also has a temporal dimension. Near multicollinearity does not affect the BLUE properties of the OLS model; however, it can have an increasing effect on the standard errors (Brooks & Tsolacos, 2010).

The effect of selling price on TOM is accounted for by estimating a 2SLS model employing unemployment benefits per 1000 inhabitants in the surrounding neighborhood and a dummy on the distance to the nearest train station as instruments. The over-identification test and first stage F-statistic indicate that these instrumental variables are exogenous and relevant, meaning both can be considered to be a valid instrument. The endogenous tests again indicate that logPrice has to be treated as an endogenous variable, again confirming the expected simultaneity between selling price and TOM (Dubé & Legros, 2016; Knight, 2002; Yavas & Yang, 1995). This indicates that the OLS specification including logPrice as an independent variable is not efficient. Contrary to the DID approach, the 2SLS model in the PGV approach has a higher R-squared than the OLS specification. The structural and locational attributes keep their significant effect on logTOM and are not absorbed by the effect of selling price. The R-squared of 0.1156 is lower than the 0.1319 in the Seemingly Unrelated Regression (SUR) of Dubé and Legros (2016) that also included the log of selling price using a two-stage approach with instrumental variables. They claim that this technique is better able to take into account the simultaneity between TOM and selling price; furthermore, they include a wider variety of structural and locational characteristics in their model.

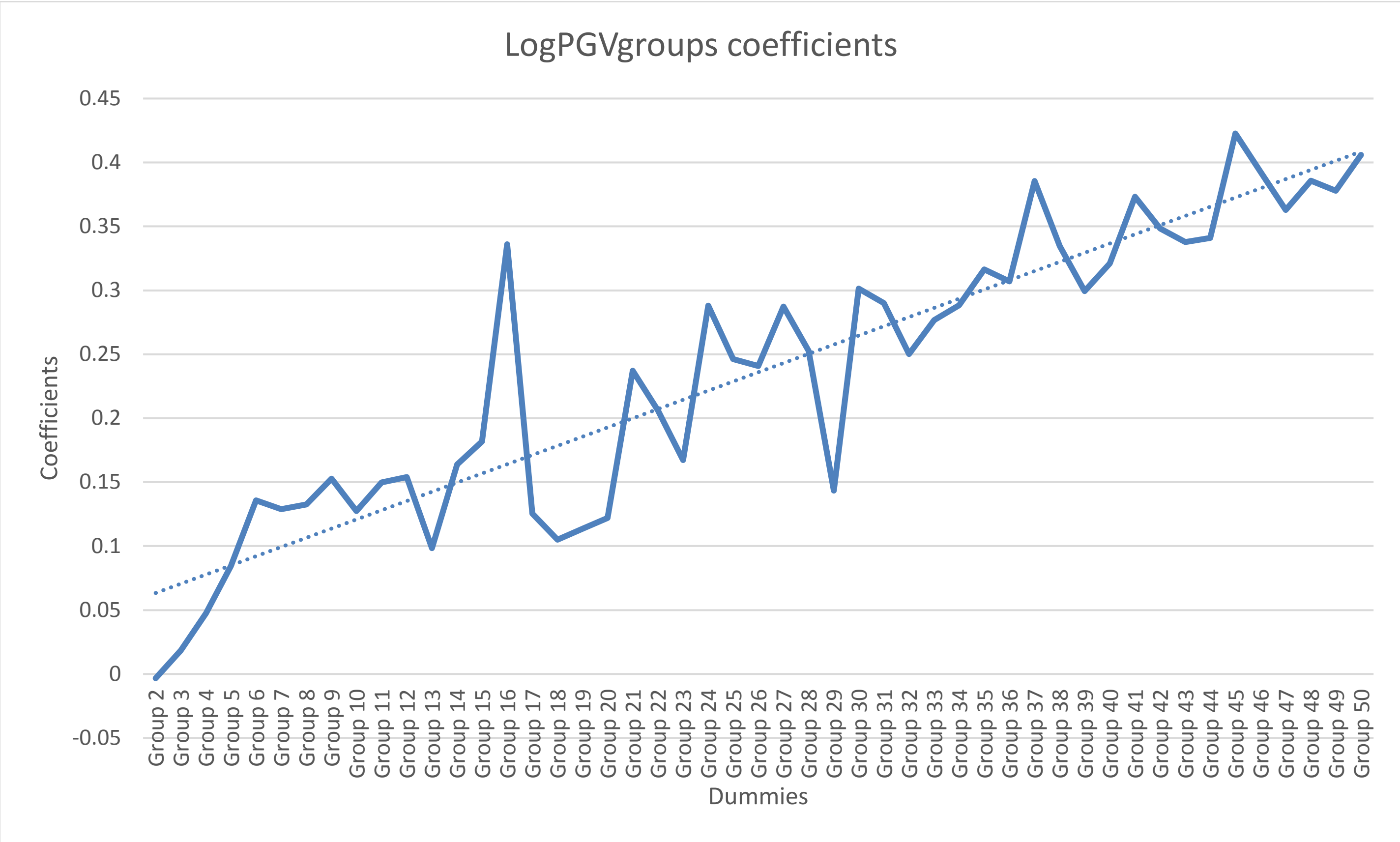


Figure 4.3: The coefficients of the logPGVgroup dummies in the 2SLS model given by the solid blue line, the dashed line is a linear trend line.

4.3.2 Interpretation

The main interest of this research is in the coefficients of the logPGV dummies and the final logPGV variable. The 2SLS model is more comprehensive by including the effect of selling price, while the OLS estimates might be biased by excluding logPrice. The 2SLS model also has a higher explained variation, contrary to the 2SLS model in the DID approach. The logPGV variable represents earthquake risk by accumulating the PGV of past earthquakes at the house location. The coefficients from the logPGV dummies are plotted in Figure 4.3, showing a trend where an increasing PGV correspondents with a rising TOM. The dummies for group 2 to 4 are insignificant, indicating that earthquake risk starts to play a role from logPGV group 5. Taking the exponent of the lowest log value of group 5 shows a minimum PGV of 0.7 m/s. The lowest impact groups have then been set to zero in the logPGV variable that is than included in the 2SLS model (see Table 4.3). The model indicates that a 10% increase in PGV impact increases TOM with 0.5%.

	OLS model		**2SLS model**	
	Coefficient	Robust SE	Coefficient	Robust SE
logPGV	0.0468***	0.0152	0.0482***	0.0152
logPrice			-0.2188**	0.1082
logArea	0.2621***	0.0168	0.3914***	0.0660
logLot	-0.0893***	0.0074	-0.0589***	0.0168
Rooms	-0.0289***	0.0074	-0.0239***	0.0078
Type corner	0.0321***	0.0104	0.0332***	0.0104
Type half double	0.1581***	0.0112	0.1729***	0.0133
Type detached	0.5603***	0.0139	0.5950***	0.0220
Balcony	0.0642***	0.0133	0.0765***	0.0146
Garage	0.0479***	0.0087	0.0661***	0.0126
Attic	-0.0697***	0.0074	-0.0673***	0.0075
Loft	0.0266***	0.0104	0.0312***	0.0106
Barn	-0.0347***	0.0076	-0.0331***	0.0076
Fire	0.0310**	0.0145	0.0471***	0.0165
Maintenance	0.2448***	0.0100	0.2802***	0.0202
Density 500-1000	-0.0860***	0.0101	-0.0719***	0.0122
Density 1000-1500	-0.0983***	0.0125	-0.0835***	0.0144
Density 1500-2500	-0.0469***	0.0157	-0.0284	0.0181
Density >2500	-0.0249	0.0215	0.0256	0.0329
Elderly	-0.0446***	0.0070	-0.0386***	0.0076
Built dummies	YES		YES	
Year FE	YES		YES	
Municipality FE	YES		YES	
N	122,908		122,908	
R^2	0.1120		0.1156	
Adjusted R^2	0.1113			
Joint sign. F-test	288,24***			
Wald Chi^2			16,439.28***	

*Table 4.3: The coefficients and the robust standard errors within brackets of the PGV models, with *, ** and *** indicating significance levels of respectively 10%, 5% and 1%.*

The logPGV variable incorporates both the spatial and temporal dimension of earthquake risk. The accumulated PGV will be higher at a location with a large frequency of (heavy) earthquakes over a long period of time. Buyers are expected to be less willing to buy houses at a location where many earthquakes occurred and are aware of this risk after the appearance of a one major earthquake or several smaller ones. Houses with a high logPGV value can be associated with risk of future earthquake damage or a lower quality of life in the surrounding region, for example, due to feelings of unsafety (Boelhouwer et al., 2016; Duran & Elhorst, 2017; Koster & Van Ommeren, 2015; Postmes et al., 2017). Duran and Elhorst (2017) have already hinted at a negative relationship between PGV and housing prices. The PGV models in this research show that earthquakes also affect TOM, giving more insight into the external effects of earthquakes in the region around the Groningen gas field.

The model also includes a variety of variables accounting for the influence of selling price, structural and locational attributes, and market conditions on TOM. The coefficient of logPrice implies that if the housing price increases with 10%, then TOM decreases with 2.1%. This might indicate that houses with better amenities sell faster (Dubé & Legros, 2016). The variable on the area of the house has a strong positive effect on TOM, a 10% increase in square meters increases TOM with 3.9%. It might have been interesting to include the price per square meter in the model to account for the effect of both selling price and house area. It is expected to have a negative relationship with TOM. A higher price per m^2 indicates the presence of amenities buyers are willing to pay for and accept a lower house area. The price per m^2 is often higher in highly urbanized areas, where a high arrival rate of buyers is expected to shorten TOM. This model chose to include selling price directly instead of the price per m^2 since it better captures the trade-off between selling price and TOM (Dubé & Legros, 2016). Contrary to the DID model, the coefficients of the structural and locational characteristics remain significant in the 2SLS model. This might be caused by the fact that a larger sample is used and the model is better able to estimate the coefficients and standard errors (Brooks & Tsolacos, 2010). However, the spatial context also changed by adding more parts of the Northern Netherlands. Above is noted that the size of the house increases TOM; however, the addition of amenities such as more rooms or a larger lot size might decline TOM. Detached houses take longer to sell compared to townhouses, a pattern also found by De Kam and Mey (2017). Unfortunately, the model was not able to include the atypicality index of Haurin et al. (2010), they showed that sellers of an atypical houses require more time to find a suitable buyer. The maintenance status still has an increasing effect on TOM. The elderly percentage represents the economic situation of a region and partly population decline (Koster & Van Ommeren, 2015). The 2SLS model indicates that a higher percentage decreases TOM, while it could be expected that it is more difficult to find a seller in a more deprived area. However, housing prices might be lower in these regions which increases the amount of buyers or sellers are motivated to leave this area and chose lower listing prices (Yavas & Yang, 1995). Furthermore, the spatial fixed effects could already have absorbed the effect of the local economic situation and population decline. The density dummies show a pattern where an increased degree of urbanization lowers TOM, although the effect decreases when moving up the density ladder and loses its significance. An extra density dummy is added compared to the DID model since the sample now includes more urbanized areas. The result might indicate the effect of a more vibrant urban market, an increased density is also been shown to increase housing prices (Daams et al., 2016).

4.3.3 Robustness

The logPGV variable measures the earthquake intensity at the location of the house; however, it can be questioned whether buyers analyze earthquake risk at this scale. Buyers might focus more on the surrounding neighborhood and attach a negative image to certain regions (Atlas voor Gemeenten, 2017; Duran & Elhorst, 2017). Duran and Elhorst (2017) also apply a logPGV variable at the neighborhood level, it might have been interesting to also include this variable in the 2SLS models. However, the DID approach was used to analyze the effect on TOM of being located in a risk area. Instead of accumulating PGV, Koster and Van Ommeren (2015) count the number of noticeable earthquakes, having a PGV above 0.5 m/s, at the house location. However, the logPGV uses a more advanced measurement of earthquake risk by also including the possible accumulative impact of a large amount of small earthquakes occurring at the house location. The fact that the area around the Groningen gas field contained a relative small amount of housing transactions compared to other regions in the Northern Netherlands did not limit the regression model in finding a strong effect of logPGV on TOM.

The logPGV dummies are used to investigate at which PGV value the earthquake intensity starts to affect TOM, being around 0.7 m/s. Following Koster and Van Ommeren (2015), an alternative approach to choose a cut-off value for the logPGV variable is to assume that at least one noticeable earthquake at the house location is necessary to affect TOM. The alternative cut-off point for the logPGV variable is then 0.5 m/s. Since logPGV is a natural logarithm, the natural log is taken of 0.5 m/s which is then used as a cut-off value. The cut-off value of 0.7

m/s was also originally a log value and was obtained by taking the exponent. Table 4.4 shows that the alternative cut-off value of 0.5 m/s increases the impact of logPGV on TOM, The first 2SLS model using 0.7 m/s as a cut-off value sets more cases to zero than the second 2SLS model employing 0.5 m/s which causes the considerable difference between the logPGV coefficients. The 2SLS model using a logPGV variable appears to be highly dependent on the employed cut-off value, an alternative might be to focus on the coefficients of the 49 dummies since they do not depend on a cut-off value. However, the logPGV variable is chosen since it gives a valuable insight into earthquake impact on TOM using one coefficient. The cut-off value of 0.7 m/s employed in the final 2SLS model can be substantiated by the analysis of the logPGV dummies.

	2SLS model (1)		**2SLS model (2)**	
	Coefficient	Robust SE	Coefficient	Robust SE
logPGV	0.0482***	0.0152		
logPGValt			0.0835***	0.0160
Selling price	YES		YES	
Structural and locational	YES		YES	
Year FE	YES		YES	
Municipality FE	YES		YES	
N	122,908		122,908	
R^2	0.1156		0.1155	
Wald Chi^2	16,447.89***		16,447.57***	

*Table 4.4: The coefficients and robust standard errors within brackets of the final PGV model and a PGV model with an alternative specification of the PGV variable, with *, ** and *** indicating significance levels of respectively 10%, 5% and 1%.*

5 Conclusions

The region around the Groningen gas field is characterized by earthquakes, causing damage to properties and lowering the quality of life (Boelhouwer et al., 2016; Koster & Van Ommeren, 2015; Postmes et al., 2017). Earthquake risk creates difficulties for sellers to find a suitable buyer; they often face a lengthy selling process (De Kam & Mey, 2017). This Master's thesis gives insight into the relationship between earthquakes and time-on-the-market (TOM) using regression models that are able to control for other factors playing a role in the selling process. The research focuses on answering the following research question:

To what extent do earthquakes influence the time-on-the-market of house sales around the Groningen gas field?

The first step of this thesis consisted of exploring a variety of factors affecting TOM using theoretical work and earlier empirical studies. Literature regarding the relationship between earthquakes and the housing market was also discussed. The theoretical framework creates the foundation for the statistical modeling of the effect of earthquake risk on TOM. The regression models give insight into both the spatial and the temporal dimension of the earthquake impact. The findings of the empirical analysis are discussed further below, but first the results from the theoretical exploration are given attention.

5.1 Factors affecting selling process

The theoretical framework covered sub questions 1 and 2, the first one being: which factors influence the selling process of houses? The literature review showed a variety of aspects playing a role:

- Selling price. TOM is simultaneously determined with selling price in the bargaining process between buyer and seller. Both are related to the motivation of buyer and seller that is often unobserved. The seller faces a trade-off between maximizing selling price and minimizing TOM (Dubé & Legros, 2016; Knight, 2002; Yavas & Yang, 1995).
- Seller characteristics. The capabilities of the seller can influence the extent to which he or she can influence to search and bargaining process, while the personal (financial) situation might affect the urgency towards a sale (De Kam & Mey, 2017).

- Structural and locational characteristics. Structural attributes such as age and size impact TOM in some cases (Anglin et al., 2003; Forgey et al., 1996; Knight, 2002; Springer, 1996). Atypical houses appear to have a longer TOM than houses with a more common combination of attributes (Haurin, 1988; Haurin et al., 2010). Hedonic models often include locational characteristics related to urban density or the distance to certain amenities (Daams et al., 2016; Schwartz et al., 2006). Dubé and Legros (2016) show that they also play a significant role in a TOM model.
- Market conditions. Housing booms are decreasing TOM, while housing busts have an increasing effect (Evans, 2004).

The above factors are affecting the TOM of a house; therefore, they have been included to a large extent in the regression models estimating TOM to be able to filter out the specific effect of earthquakes on TOM around the Groningen gas field.

The theoretical framework also elaborated on sub question 2: how are earthquakes affecting the housing market? Buyers appear to be less willing to buy houses in a region at risk of earthquakes. The house might sustain earthquake damage, while compensation schemes impose considerable transactions costs or do not grant adequate repair budgets (De Kam & Spijkerboer, 2015; Van der Voort & Vanclay, 2015). Furthermore, feelings of unsafety might have decreased the quality of life (Boelhouwer et al., 2016; Postmes et al., 2017). Earlier studies have shown that being located in an area at risk of earthquakes decreases property values and increases TOM (Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster, 2016). Besides this spatial dimension, earthquake risk also has a temporal aspect since multiple earthquakes take place around the Groningen gas field and a major earthquake such as the Huizinge earthquake of 2012 might have increased the awareness of the risk of earthquakes (Bosker et al., 2016; Duran & Elhorst, 2017).

The theories and earlier empirical work discussed in the theoretical framework are used to create a conceptual model on the factors affecting TOM. It includes the characteristics of the seller, structural and locational attributes, and market conditions. Furthermore, it shows the simultaneous relationship between selling price and TOM. Finally, the spatial and temporal dimension of earthquake risk are added to the model. The conceptual model shows the factors the regression model estimating TOM has to take into account to uncover the relationship between earthquake risk and TOM.

5.2 Earthquakes and time-on-the-market

The empirical analysis focuses on highlighting the specific effect of earthquakes on TOM and is based on sub questions 3 and 4:

- To what extent does the location in an area at risk of earthquakes affect the time-on-the-market of houses around the Groningen gas field?
- How does a major earthquake influence the time-on-the-market of houses around the Groningen gas field?

Two approaches are used to include the impact of earthquakes in a regression model. The first is a Difference-In-Differences (DID) technique comparing risk and reference areas before and after the Huizinge earthquake of 2012. The risk areas are neighborhoods with houses that have sustained earthquake damage (CBS, 2017b). The second approach uses a variable measuring the Peak Ground Velocity (PGV) of each previous earthquake at the house location. The models employ a NVM dataset including housing transactions from 2003 until 2014 in the Provinces of Friesland, Groningen and Drenthe. The PGV approach uses a larger sample since it is able to cover the whole of the Northern Netherlands, while the DID approach is bound to include only similar neighborhoods as control areas. The final result of both approaches is a 2SLS model including the earthquake indicator, the selling price, structural and locational attributes, spatial fixed effects and time fixed effects. The employed NVM dataset did not allow the inclusion of seller characteristics.

The 2SLS model in the DID approach showed that TOM is 6.2% higher in risk areas compared to reference areas after the Huizinge earthquake of 2012, while TOM used to be 8.7% lower before the earthquake. This result indicates the increasing effect on TOM of being located in a region associated with earthquake risk. Furthermore, it indicates the influence of the Huizinge earthquake on the awareness of buyers of the earthquake risk in the area. The DID model did not show that the Middelstum earthquake of 2006 acted as a turning point. The 2SLS model in the PGV approach show that a 10% increase in the PGV at the house location increases TOM with 0.5%. Earthquakes appear to start influence TOM above a PGV of 0.7 m/s. The logPGV variable combines the spatial and temporal dimension of earthquake risk. Houses with a large PGV are located in an area characterized by earthquakes and PGV will be higher after the occurrence of an (heavy) earthquake. The results of both models confirm the two hypotheses:

- Being located in an area at risk of earthquakes increases the time-on-the-market of houses around the Groningen gas field.

- The event of a major earthquake increases the time-on-the-market of houses around the Groningen gas field.

The findings confirm the pattern found by CBS (2017b), where the TOM in neighborhoods with a high percentage of damaged houses is higher and recovers slower from the housing bust compared to areas with similar socioeconomic attributes. The results are opposite to the TOM regression model of Bosker et al. (2016) who did not find an effect of being located in the area around the Groningen gas field. However, they employ a different risk area that is compared to reference properties in The Netherlands after the 2012 Huizinge earthquake. This thesis presents a more comprehensive TOM regression model by adding selling price as an endogenous explanatory variable, spatial fixed effects and a PGV variable to measure earthquake risk. TOM studies complement earlier research showing that housing prices are declining due to the risk of earthquakes (Atlas voor Gemeenten, 2017; CBS, 2017b; Duran & Elhorst, 2017; Koster & Van Ommeren, 2015). This thesis is part of the growing body of research on the effects of earthquake risk on the housing market around the Groningen gas field which increases the insight into housing market dynamics in regions characterized by earthquakes. The TOM models presented in this thesis take into account a variety of factors affecting TOM and investigate both the spatial and temporal dimension of earthquake risk. The results lead to the following answer to the main research question: the occurrence of earthquakes is increasing the TOM of houses around the Groningen gas field.

5.3 Policy recommendations

The regression models in this research gave more insight into difficulties faced by sellers of houses around the Groningen gas field. They have a lengthier selling process due to the risk of earthquakes. Buyers might be less willing to buy a house in the earthquake region to risk of damage to the property and a declined quality of life in the region (Boelhouwer et al., 2016; Koster & Van Ommeren, 2015). An important role can be played by creating compensation schemes that do not impose high transactions costs to households. This might cause earthquake risk to be less of a constraint to buying a house around the Groningen gas field, thereby increasing the arrival rate of buyers which might shorten TOM. A compensation scheme that also covers the loss in housing value ensures that sellers are willing to accept a lower selling price. A lower listing or selling price also increases the arrival rate of buyers, making it easier to find a suitable buyer which might shorten the selling process. Another option is a policy strategy focusing on improving the economic structure and spatial quality of the region which could stimulate the attractiveness of the area around the Groningen gas field to buyers and compensate for the negative effect of the appearance of earthquakes. In conclusion, the findings from this Master's thesis show an increased TOM for houses around the Groningen gas field due to earthquake risk which calls for a policy strategy focusing on a solid compensation scheme for earthquake damage and declining housing values and improving the quality of life in the region.

5.4 Limitations and further research

The Master's thesis presented a comprehensive model estimating TOM, including earthquake indicators, selling price, structural and locational characteristics, and market conditions. It gained insight into the effect of earthquake on the housing market around the Groningen gas field. However, improvements are still possible which also provide options for further research:

- The DID approach based the risk area on the percentage of damaged houses in neighborhoods around the gas field. However, ambiguity exists regarding the size of the risk area, it can also be argued that a negative image is present for the whole Province of Groningen (Atlas voor Gemeenten, 2017). The overlap with the municipalities used for the spatial fixed effects cause multicollinearity issues that require beyond the scope of this thesis. The DID model simplifies the spatial dimension of earthquake risk by using one risk area and the temporal dimension by using one time point. Furthermore, the year of sale is used to select houses after a major earthquake, it would be more precise to base the selection on the exact sale date.
- The logPGV variable better takes into account spatial and temporal complexities; however, the variable can also be applied on a neighborhood level since this might be the scale buyers consider when they estimate earthquake risk (Duran & Elhorst, 2017). It could have been interesting to compare the findings of the PGV model to the variable of Koster (2016) on the percentage of damaged houses in the surrounding region.
- It has to be noted that the logPGV variable is correlated with municipality and year fixed effects, causing multicollinearity issues. Future research might focus on identifying an appropriate alternative to the inclusion of spatial and temporal fixed effects in a regression model. The coefficient of the logPGV variable also depends on the chosen cut-off value, future studies could focus on improving the methods to identify this value.

- A 2SLS model was used to account for the simultaneity between selling price and TOM; however, Dubé and Legros (2016) shows the potential of a Seemingly Unrelated Regression (SUR) estimating both a TOM and selling price equation simultaneously.
- The temporal effect of earthquake risk on TOM can be further explored using a time series analysis also with the goal of forecasting the effect of an occurring earthquake (Brooks & Tsolacos, 2010).
- Instead of including selling price directly, another approach is to include the sale price premium (difference selling price and listing price) since it better accounts for the relative height of the selling price and the motivation of the seller (Knight, 2002). Furthermore, the relative height of the listing price can also influence TOM (Anglin et al., 2003).
- The TOM model can be complemented with variables related to seller motivation and seller characteristics (Glower et al., 1998; Springer, 1996). Furthermore, it might be suitable to add to the atypicality index of Haurin et al. (2010) instead of separate structural and locational attributes.
- The accuracy of the regression models also depend on the quality of the data. De Kam and Mey (2017) mention that NVM data might underestimate TOM. Many variables where also transformed into dummy variables. This simplification also affects the extent to which to model is able to represent housing market dynamics. A higher R-squared might be reached by including more structural and locational attributes.

In conclusion, further research can continue the search for a suitable earthquake indicators and try to create a regression model able to capture a variety of factors affecting TOM. It can also be useful to apply TOM models to other earthquake regions besides the Groningen gas field to enable better generalization of results. Finally, quantitative studies can be combined with qualitative research able to give insight into the experiences of buyers and sellers. Further studies can complement the research presented in this thesis by analyzing the housing market effect of recent developments regarding the extraction of the Groningen gas field and compensation schemes, thereby increasing the understanding of the selling process of houses around the Groningen gas field.

References

Agnello, L. & Schuknecht, L. (2011). Booms and busts in housing markets: Determinants and applications. *Journal of Housing Economics*, 20(3), 171-190.

Alonso, W. (1964). Location and Land Use. Cambridge, MA: Harvard University Press.

Anglin, P.M., Rutherford, R. & Springer, T.M. (2003). The trade-off between the selling price of residential properties and time-on-the-market: The impact of price setting. *Journal of Real Estate Finance and Economics*, 26(1), 95-111.

Atlas voor Gemeenten (2017). *Vijf jaar na Huizinge: Het effect van aardbevingen op de huizenprijzen in Groningen.* Utrecht: Atlas voor Gemeenten.

Babbie, E. (2013). *The Practice of Social Research.* 13th edition. Belmont: Wadsworth Cangage Learning.

Beron, K.J., Murdoch, J.C., Thayer, M.A. & Vijverberg, W.P.M. (1997). An analysis of the housing market before and after the 1989 Loma Prieta Earthquake. *Land Economics*, 73(1), 101-113.

Boelhouwer, P., Boumeester, H., Groetelaers, D., Hoekstra, J., Van der Heijden, H., Jansen, S. … Ringersma, R. (2016). *Woningmarkt- en leefbaarheidsonderzoek aardbevingsgebied Groningen.* Delft: OTB – Research for the Built Environment, Faculty of Architecture and the Built Environment, Delft University of Technology.

Boeri, T. & Van Ours, J. (2013). *The Economics of Imperfect Labor Markets.* 2nd edition. Princeton: Princeton University Press.

Bosker, M., Garretsen, H., Marlet, G., Ponds, R., Poort, J., Van Dooren, R. & Van Woerkens, C. (2016). *Met angst en beven: Verklaringen voor dalende huizenprijzen in het Groningse aardbevingsgebied.* Utrecht: Atlas voor Gemeenten.

Brooks, C. & Tsolacos, S. (2010). *Real Estate Modelling and Forecasting.* Cambridge: Cambridge University Press.

CBS (2017a). *Bestaande woningen in verkoop; vraagprijs, aanbodtijd, oppervlakte, regio.* Retrieved on 13 March 2018 from http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=82534ned&D1=15-21,31&D2=0&D3=0-1&D4=0-7&HDR=G1,G3&STB=T,G2&VW=T.

CBS (2017b). *Woningmarktontwikkelingen rondom het Groningerveld: 1e kwartaal 1995 tot en met 2e kwartaal 2017.* Den Haag: Centraal Bureau voor de Statistiek (CBS).

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

De Kam, G. (2016). *Waardedaling door aardbevingen nader beschouwd: Een reactie op ESB en het onderzoek met angst en beven.* URSI Research Report 354. Groningen: Faculty of Spatial Sciences, University of Groningen.

De Kam, G. & Mey, E. (2017). *Ervaringen bij verkoop van woningen in aardbevingsgebied.* URSI Research Report 361. Groningen: Faculty of Spatial Sciences, University of Groningen.

De Kam, G. & Raemakers, J. (2014*). Opvattingen van bewoners over de effecten van aardbevingen op het woongenot en de woningwaarde in Groningen: Een vergelijkend onderzoek in drie woonbuurten in Middelstum, Loppersum en Slochteren in 2009 en 2013.* URSI Research Report 346. Groningen: Faculty of Spatial Sciences, University of Groningen.

De Kam, G. & Spijkerboer, R. (2015). *Wat willen de Groningers zelf?* URSI Research Report 351. Groningen: Faculty of Spatial Sciences, University of Groningen.

De Kam, G., Van den Broek, J. & Van Maanen, J. (2018). *Gegijzeld in eigen huis: Zorgen en problemen rond hypotheken in het Groningse aardbevingsgebied.* URSI Research Report 363. Groningen: Faculty of Spatial Sciences, University of Groningen.

Dubé, J. & Legros, D. (2016). A spatiotemporal solution for the simultaneous sale price and time-on-the-market problem. *Real Estate Economics*, 44(4), 846-877.

Duran, N. & Elhorst, J.P. (2017). *The willingness to accept moderate seismic risk: evidence from the North of the Netherlands.* Paper prepared for ERSA 2017 Conference, Groningen. Preliminary version

Evans, A.W. (2004). *Economics, Real Estate and the Supply of Land.* Oxford: Blackwell Publishing.

Forgey, F.A., Rutherford, R.C. & Springer, T.M. (1996). Search and liquidity in single-family housing. *Real Estate Economics*, 24(3), 273-292.

Francke, M.K. & Lee, K.M. (2014). *De invloed van fysieke schade op verkopen van woningen rond het Groningenveld.* Rotterdam: Ortec Finance Research Center.

Genesove, D. & Mayer, C.J. (1997). Equity and time to sale in the real estate market. *The American Economic Review*, 87(3), 255-269.

Glower, M., Haurin, D.R. & Hendershott, P.H. (1998). Selling time and selling price: The influence of seller motivation. *Real Estate Economics*, 26(4), 719-740.

Haurin, D. (1988). The duration of marketing time of residential housing. *Real Estate Economics*, 16(4), 396-410.

Haurin, D.R., Haurin, J.L., Nadauld, T. & Sanders, A. (2010). List prices, sale prices and marketing time: An application to U.S. housing markets. *Real Estate Economics*, 38(4), 659-685.

Immergluck, D. (2015). A look back: What we now know about the causes of the US mortgage crisis. *International Journal of Urban Sciences*, 19(3), 269-285.

Jud, G.D., Seaks, T.G. & Winkler, D.T. (1996). Time on the market: The impact of residential brokerage. *Journal of Real Estate Research*, 12(3), 447-458.

Knight, J.R. (2002). Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2), 213-237.

Koopmans, C. & Rougoor, W. (2017). *Earthquakes and house prices: a meta-analysis.* Paper prepared for ERSA 2017 Conference, Groningen. Preliminary version

Koster, H.R.A. (2016). Gaswinning, aardbevingen en huizenprijzen. Amsterdam: Vrije Universiteit Amsterdam.

Koster, H.R.A. & Van Ommeren, J. (2015). A shaky business: natural gas extraction, earthquakes and house prices. *European Economic Review*, 80(1), 120-139.

Lennartz, C., Vrieselaar, N. & Hoving, L. (2018). *Dutch housing market rumbles on: even higher price hike expected in 2018.* Dutch Housing Market Quarterly. Retrieved on 5 March 2018 from https://economics.rabobank.com/publications/2018/february/dutch-housing-market-even-higher-price-hike-expected-in-2018/. Utrecht: Rabobank Research.

Livy, R.L. & Klaiber, H.A. (2016). Maintenanceaining public goods: The capitalized value of local park renovations. *Land Economics*, 92(1), 96-116.

McCann, P. (2013). *Modern Urban and Regional Economics.* 2nd edition. Oxford: Oxford University Press.

Murdoch, J.C., Singh, H. & Thayer, M. (1993). The impact of natural hazards on housing values: The Loma Prieta Earthquake. *Real Estate Economics*, 21(2), 167-184.

Nakagawa, M., Saiko, M. & Yamaga, H. (2007). Earthquake risk and housing rents: Evidence from the Tokyo Metropolitan Area. *Regional Science and Urban Economics*, 37(1), 87-99.

Nakagawa, M., Saito, M. & Yamaga, H. (2009). Earthquake risks and land prices: Evidence from the Tokyo Metropolitan Area. *Japanese Economic Review*, 60(2), 208-222.

NVM (2015). *Analyse woningmarkt van de bestaande koopwoningen: 2e kwartaal 2015.* Nieuwegein: De Nederlandse Vereniging van Makelaars en Taxateurs (NVM).

NVM (2018). *Analyse woningmarkt: 1e kwartaal 2018.* Nieuwegein: De Nederlandse Vereniging van Makelaars en Taxateurs (NVM).

Postmes, T., Stroebe, K., Richardson, J., LeKander, B., Oldersma, F., Broer, J. & Greven, F. (2017). Veiligheidsbeleving, gezondheid en toekomstperspectief van Groningers. Groningen: Rijksuniversiteit Groningen.

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

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

Sirmans, C.F., Turnbull, G.K. & Dombrow, J. (1995). Quick house sales: Seller mistake or luck? *Journal of Housing Economics*, 4(3), 230-243.

Springer, T.M. (1996). Single-family housing transactions: Seller motivations, price, and marketing time. *Journal of Real Estate Finance and Economics*, 13(3), 237-254.

Taylor, C.R. (1999). Time-on-the-market as a sign of quality. *Review of Economic Studies*, 66(3), 555-578.

Van der Voort, N. & Vanclay, F. (2015). Social impacts of earthquakes caused by gas extraction in the Province of Groningen, The Netherlands. *Environmental Impact Assessment Review*, 50(4), 1-15.

Van Duijn, M., Rouwendal, J. & Boersema, R. (2016). Redevelopment of industrial heritage: Insights into external effects on housing prices. *Regional Science and Urban Economics*, 57, 91-107.

Van Hoven, B. & Meijering, L. (2011). On the ground. In V.J. Del Casino, M.E. Thomas, P. Cloke & R. Panelli (Eds.), *A companion to social geography* (pp.161-180). Malden: Blackwell Publishing Ltd.

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

Whaley, J. (2009). *The Groningen Gas Field.* Retrieved on January 10, 2018 from https://www.geoexpro.com/articles/2009/04/the-groningen-gas-field.

Yavas, A. & Yang, S. (1995). The strategic role of listing price in marketing real estate: Theory and evidence. *Real Estate Economics*, 23(3), 347-368.

Appendix I: Syntax

DID syntax

```
*CREATE LOG FILE
clear
cap log close
log using "X:\My Desktop\Master thesis\Empirical analysis\Output\logDIDmodels.log", replace

*OPEN DATASET
set excelxlsxlargefile on
import excel "X:\My Desktop\Master thesis\Empirical analysis\NVM
data\model_data_Paul_students_projects_update.xlsx", sheet("model_data") firstrow clear

*SAMPLE SELECTION
drop if t<=2002

*VARIABLE SELECTION AND TRANSFORMATION
*Y: TOM
drop if tom<=1
gen logtom = ln(tom)

*X: Earthquake impact
clonevar municipality2013 = municipality
recode municipality2013 (39 = 1895) (52 = 1895) (1661 = 1895)
recode municipality2013 (64 = 1900) (91 = 1900) (104 = 1900) (683 = 1900) (710 = 1900)
recode municipality2013 (83 = 1908)

gen damage1 = 0
replace damage1 = 1 if
inlist(nbh,100000,100002,100202,170000,170001,170005,170101,170108,180101,180103,180105,180107,1802
01,180205,180207,180210,180501,180601,180604,180606,180701,181101,400309,470100,530100,560200,560
209,560300,560302,560309,16630100,16630101,16630102,16630106,16630300,16630302,19870000,19870300
,19870304)
replace damage1 = 1 if
inlist(nbh,30000,30002,50000,100001,100003,100005,100200,100201,100203,100209,180102,180202,180208,1
80302,180304,180305,180602,180605,180702,180803,180804,180901,181103,181104,240200,400109,400209,
400308,530001,530002,530101,530102,530109,530200,530202,530301,560203,16510001,16510100,16510106,
16510400,16630103,16630105,16630109,16630200,16630201,16630202,16630203,16630208,16630209,16630
304,16630309,390200,390201,390202,390209,390400,390403,19870002,19870005,19870100,19870302,19870
303)
```

```
replace damage1 = 1 if
inlist(nbh,30001,30007,30008,30009,50001,50002,50003,50004,50005,50006,50007,90000,90001,90002,90003,
90005,90009,90100,90101,90102,90104,90106,90109,100004,100103,100104,100107,100205,100207,100300,1
00301,100302,100303,100304,100305,100309,180104,180106,180303,180801,181102,181105,181106,181107,
181108,181109,240000,240001,240002,240003,240008,240009,240100,240101,240102,240109,240201,240202
,240203,240209,240300,240301,240302,240303,240304,240305,240309,400000,400001,400002,400003,40000
4,400006,400007,400008,400009,400100,400101,400102,400108,400200,400201,400203,400204,400208,4003
00,400301,400302,400304,400400,400409,530000,530003,530004,530009,530201,530203,530204,530209,530
300,530302,530309,16510002,16510008,16510009,16510101,16510102,16510103,16510104,16510105,165101
07,16510109,16510200,16510201,16510202,16510203,16510204,16510205,16510209,16510300,16510308,165
10309,16510409,16630104,16630303,19870001,19870003,19870009,19870101,19870109)
replace damage1 = 2 if
inlist(municipality2013,14,51,55,60,63,70,72,74,80,81,82,83,88,98,106,109,114,118,119,140,653,737,1690,170
1,1900)
replace damage1 = 2 if
inlist(nbh,100108,170002,170003,170102,180203,180301,180603,220204,220107,240005,370100,370101,4804
09,580209,900124,900126,16510108,16630006,16800009,16800309,16800509,16800700,16800809,16800900,
16800909,16801100,16801209,16801609,16801908,16810209,16810809,16990410,16990440,17300009,17300
402,17300609,17300701,17300809,17300909,18910209,520004)
replace damage1 = . if missing(nbh)
drop if damage1==2

gen damage2 = 0
replace damage2 = 1 if
inlist(nbh,100000,100002,100202,170000,170001,170005,170101,170108,180101,180103,180105,180107,1802
01,180205,180207,180210,180501,180601,180604,180606,180701,181101,400309,470100,530100,560200,560
209,560300,560302,560309,16630100,16630101,16630102,16630106,16630300,16630302,19870000,19870300
,19870304)
replace damage2 = 2 if
inlist(nbh,30000,30002,50000,100001,100003,100005,100200,100201,100203,100209,180102,180202,180208,1
80302,180304,180305,180602,180605,180702,180803,180804,180901,181103,181104,240200,400109,400209,
400308,530001,530002,530101,530102,530109,530200,530202,530301,560203,16510001,16510100,16510106,
16510400,16630103,16630105,16630109,16630200,16630201,16630202,16630203,16630208,16630209,16630
304,16630309,390200,390201,390202,390209,390400,390403,19870002,19870005,19870100,19870302,19870
303)
replace damage2 = 2 if
inlist(nbh,30001,30007,30008,30009,50001,50002,50003,50004,50005,50006,50007,90000,90001,90002,90003,
90005,90009,90100,90101,90102,90104,90106,90109,100004,100103,100104,100107,100205,100207,100300,1
00301,100302,100303,100304,100305,100309,180104,180106,180303,180801,181102,181105,181106,181107,
181108,181109,240000,240001,240002,240003,240008,240009,240100,240101,240102,240109,240201,240202
```

```
,240203,240209,240300,240301,240302,240303,240304,240305,240309,400000,400001,400002,400003,40000
4,400006,400007,400008,400009,400100,400101,400102,400108,400200,400201,400203,400204,400208,4003
00,400301,400302,400304,400400,400409,530000,530003,530004,530009,530201,530203,530204,530209,530
300,530302,530309,16510002,16510008,16510009,16510101,16510102,16510103,16510104,16510105,165101
07,16510109,16510200,16510201,16510202,16510203,16510204,16510205,16510209,16510300,16510308,165
10309,16510409,16630104,16630303,19870001,19870003,19870009,19870101,19870109)
replace damage2 = . if missing(nbh)

gen before = 0
replace before = 1 if damage1==1
gen afterHuizinge = 0
replace afterHuizinge = 1 if t>=2013 & damage1==1
gen afterMiddelstum = 0
replace afterMiddelstum = 1 if t>=2007 & damage1==1

gen before_2cat1 = 0
replace before_2cat1 = 1 if damage2==1
gen after_2cat1 = 0
replace after_2cat1 = 1 if t>=2013 & damage2==1

gen before_2cat2 = 0
replace before_2cat2 = 1 if damage2==2
gen after_2cat2 = 0
replace after_2cat2 = 1 if t>=2013 & damage2==2

*Z: Price
drop if y<=10.67113
drop if y>=13.04589

*Z: Housing characteristics
drop if m2<=25
drop if m2>=250
gen logm2 = ln(m2)

drop if lot<=10
drop if lot>=5000
gen loglot = ln(lot)

drop if rooms==0
recode rooms (1/4 = 0)
```

```
recode rooms (5/8 = 1)

drop if nvm_typ==1
drop if nvm_typ>=8
recode nvm_typ (3 = 2)

drop if per==0

recode balc (1/3 = 1)

recode gar (1/5 = 1)

recode barn (1/6 = 1)

recode fire (1/2 = 1)

recode ins_m (1/6 = 0)
recode ins_m (7/9 = 1)

*Z: Locational characteristics
gen urb_dens_gr = 0
replace urb_dens_gr = 1 if add_km2>=500 & add_km2<1000
replace urb_dens_gr = 2 if add_km2>=1000 & add_km2<1500
replace urb_dens_gr = 3 if add_km2>=1500 & add_km2<2500
replace urb_dens_gr = 4 if add_km2>=2500
replace urb_dens_gr = . if missing(add_km2)

gen elderly_dummy = 0
replace elderly_dummy = 1 if elderly>16

*I: Instruments
gen train_dummy = 0
replace train_dummy = 1 if train_d>8.9

*SUMMARY STATISTICS, CORRELATION MATRIX
xi:sum logtom before afterHuizinge afterMiddelstum y logm2 loglot rooms 2.nvm_typ i.nvm_typ 1.per i.per balc gar attic loft barn fire ins_m 0.urb_dens_gr i.urb_dens_gr elderly_dummy ds__000 train_dummy
xi:pwcorr logtom y before afterHuizinge afterMiddelstum logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire i.urb_dens_gr elderly_dummy
```

```
*COMPARISON TREATMENT AND CONTROL
gen per_old = 0
replace per_old = 1 if inlist(per,1,2,3)
gen per_middle = 0
replace per_middle = 1 if inlist(per,4,5,6)
gen per_new = 0
replace per_new = 1 if inlist(per,7,8,9)
gen type_town = 0
replace type_town = 1 if nvm_typ==2
gen type_corner = 0
replace type_corner = 1 if nvm_typ==4
gen type_half = 0
replace type_half = 1 if nvm_typ==5
gen type_detached = 0
replace type_detached = 1 if nvm_typ==6
sum type_town type_corner type_half type_detached per_old per_middle per_new if damage1==1
sum type_town type_corner type_half type_detached per_old per_middle per_new if damage1==0
ttest per_old, by(damage1)
ttest per_middle, by(damage1)
ttest per_new, by(damage1)
ttest type_town, by(damage1)
ttest type_corner, by(damage1)
ttest type_half, by(damage1)
ttest type_detached, by(damage1)

*OLS REGRESSION
xi:areg logtom before afterMiddelstum logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t, robust absorb(municipality)
xi:regress logtom before afterMiddelstum logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality, robust
xi:areg logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t, robust absorb(municipality)
xi:regress logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality, robust

*DIAGNOSTICS OLS REGRESSION
quietly xi:regress logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire
ins_m i.urb_dens_gr elderly_dummy i.t i.municipality, robust
predict OlSResiduals,residuals
estat vif
```

```
quietly xi:regress logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality
rvfplot, yline(0)
estat hettest

pnorm OlSResiduals
qnorm OlSResiduals
sktest OlSResiduals

*2SLS REGRESSION
xi:ivregress 2sls logtom before afterMiddelstum logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality (y = ds__000 train_dummy), robust
xi:ivregress 2sls logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality (y = ds__000 train_dummy), robust

*DIAGNOSTICS 2SLS REGRESSION
xi: ivregress 2sls logtom before afterHuizinge logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality (y = ds__000 train_dummy), robust first
estat overid
estat firststage
estat endogenous y

predict SLSResiduals, residuals
pnorm SLSResiduals
qnorm SLSResiduals
sktest SLSResiduals

*ROBUSTNESS
xi:ivregress 2sls logtom before_2cat1 after_2cat1 before_2cat2 after_2cat2 logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality (y = ds__000 train_dummy), robust

*SAVE AND EXIT
log close
exit
```

PGV syntax

```
*CREATE LOG FILE
clear
```

```
cap log close
log using "X:\My Desktop\Master thesis\Empirical analysis\Output\logPGVmodels.log", replace

*OPEN DATASET
set excelxlsxlargefile on
import excel "X:\My Desktop\Master thesis\Empirical analysis\NVM data\model_data_Paul_students_projects_update.xlsx", sheet("model_data") firstrow clear

*SAMPLE SELECTION
drop if t<=2002

*VARIABLE SELECTION AND TRANSFORMATION
*Y: TOM
drop if tom<=1
gen logtom = ln(tom)

*Z: Price
drop if y<=10.90366
drop if y>=13.01543

*Z: Housing characteristics
drop if stor>=6
recode stor (1/2 = 0)
recode stor (3/4 = 1)

drop if m2<=25
drop if m2>=250
gen logm2 = ln(m2)

drop if lot<=10
drop if lot>=5000
gen loglot = ln(lot)

drop if rooms==0
recode rooms (1/4 = 0)
recode rooms (5/8 = 1)

drop if nvm_typ==1
drop if nvm_typ>=8
recode nvm_typ (3 = 2)
```

```
drop if per==0

recode balc (1/3 = 1)

recode gar (1/5 = 1)

recode barn (1/6 = 1)

recode fire (1/2 = 1)

recode ins_m (1/6 = 0)
recode ins_m (7/9 = 1)

*Z: Locational characteristics
gen urb_dens_gr = 0
replace urb_dens_gr = 1 if add_km2>=500 & add_km2<1000
replace urb_dens_gr = 2 if add_km2>=1000 & add_km2<1500
replace urb_dens_gr = 3 if add_km2>=1500 & add_km2<2500
replace urb_dens_gr = 4 if add_km2>=2500
replace urb_dens_gr = . if missing(add_km2)

gen elderly_dummy = 0
replace elderly_dummy = 1 if elderly>17

*X: Earthquake impact
egen logPGVgroups = cut(pgv), group(50)
clonevar pgv_new = pgv
replace pgv_new = 0 if logPGVgroups<=4
clonevar pgv_alt = pgv
replace pgv_alt = 0 if pgv<-0.69314718

*I: Instruments
gen train_dummy = 0
replace train_dummy = 1 if train_d>4.4

*SUMMARY STATISTICS, CORRELATION MATRIX
xi:sum logtom pgv_new y logm2 loglot rooms 2.nvm_typ i.nvm_typ 1.per i.per balc gar attic loft barn fire ins_m
0.urb_dens_gr i.urb_dens_gr elderly_dummy u_b_000 train_dummy
```

```
xi:pwcorr logtom y pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr
elderly_dummy

*OLS REGRESSION
xi:areg logtom i.logPGVgroups logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t, robust absorb(municipality)
xi:regress logtom i.logPGVgroups logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality, robust
xi:areg logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr
elderly_dummy i.t, robust absorb(municipality)
xi:regress logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr
elderly_dummy i.t i.municipality, robust

*DIAGNOSTICS OLS REGRESSION
quietly xi:regress logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality, robust
predict OlSResiduals,residuals
estat vif

quietly xi:regress logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality
rvfplot, yline(0)
estat hettest

pnorm OlSResiduals
qnorm OlSResiduals
sktest OlSResiduals

*2SLS REGRESSION
xi:ivregress 2sls logtom i.logPGVgroups logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality (y = u_b_000 train_dummy), robust
xi:ivregress 2sls logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality (y = u_b_000 train_dummy), robust

*DIAGNOSTICS 2SLS REGRESSION
xi: ivregress 2sls logtom pgv_new logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m
i.urb_dens_gr elderly_dummy i.t i.municipality (y = u_b_000 train_dummy), robust first
estat overid
estat firststage
estat endogenous y
```

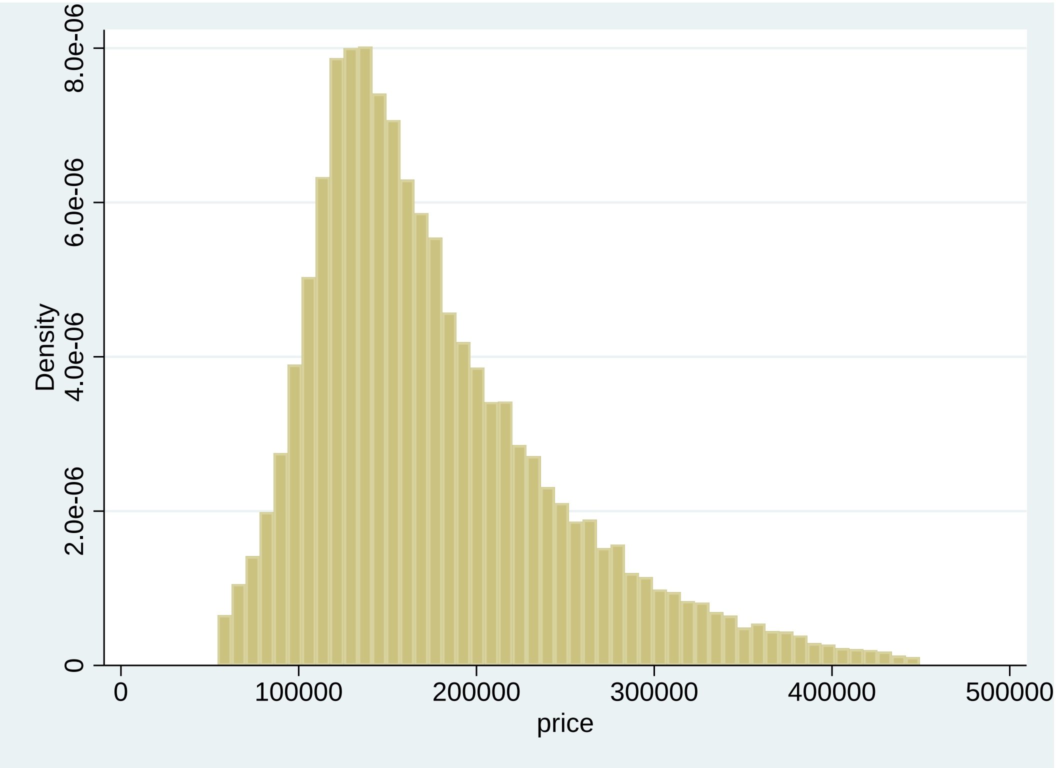
```
predict SLSResiduals, residuals
pnorm SLSResiduals
qnorm SLSResiduals
sktest SLSResiduals

*ROBUSTNESS
xi:ivregress 2sls logtom pgv_alt logm2 loglot rooms i.nvm_typ i.per balc gar attic loft barn fire ins_m i.urb_dens_gr elderly_dummy i.t i.municipality (y = u_b_000 train_dummy), robust

*SAVE AND EXIT
log close
exit
```

Appendix II: Other figures and tables

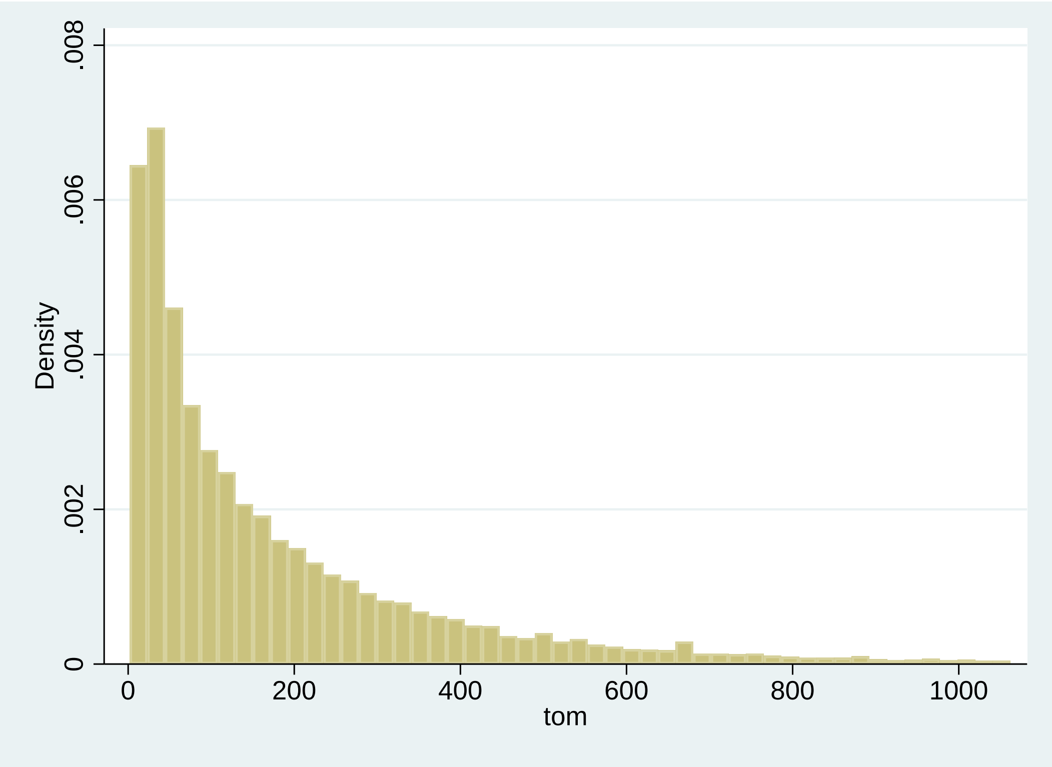
Distribution Price in PGV dataset

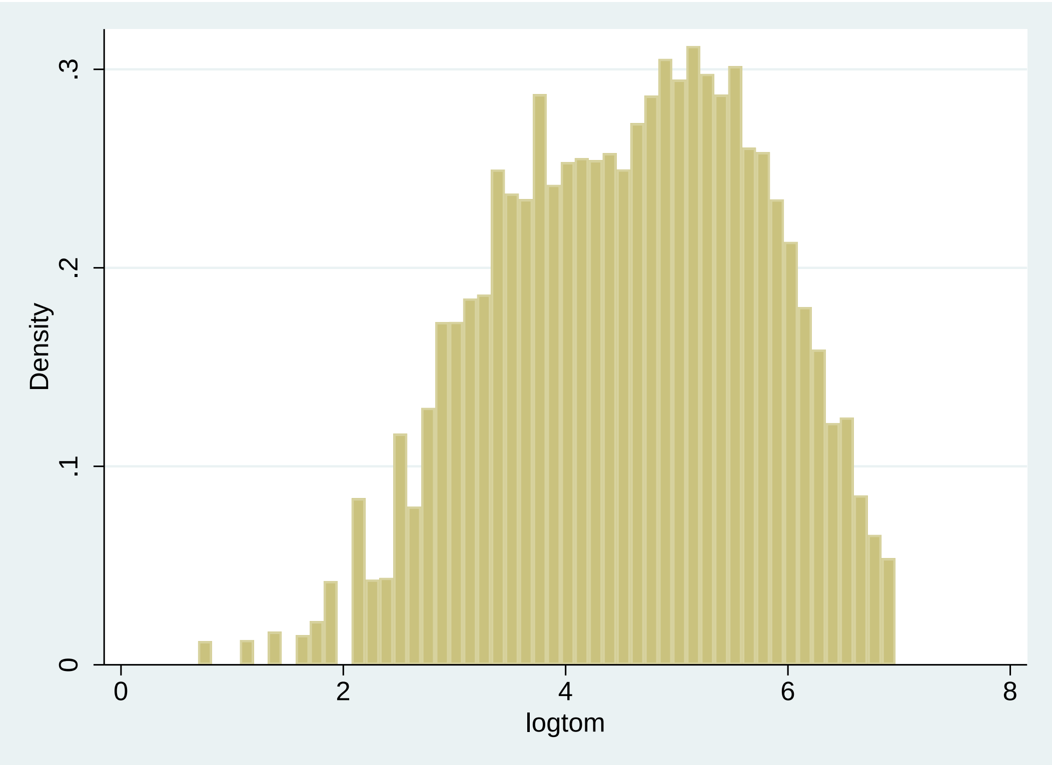


Number of properties per year in PGV dataset

Year	Freq.	Percent
2003	11,305	9.2
2004	11,480	9.3
2005	12,800	10.4
2006	13,498	11.0
2007	13,330	10.9
2008	11,245	9.2
2009	8,615	7.0
2010	9,340	7.6
2011	7,471	6.1
2012	7,553	6.2
2013	7,008	5.7
2014	9,263	7.5
Total	122,908	100

Distribution TOM and logTOM in PGV dataset





LogPGVgroups estimates in OLS model

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
logPGVgroups2	-0.0026170	0.0319897	-0.08	0.935	-0.0653163	0.0600824
logPGVgroups3	0.0219414	0.0329223	0.67	0.505	-0.0425857	0.0864686
logPGVgroups4	0.0477024	0.0336369	1.42	0.156	-0.0182254	0.1136303
logPGVgroups5	0.0854110	0.0345909	2.47	0.014	0.0176134	0.1532086
logPGVgroups6	0.1379436	0.0352374	3.91	0.000	0.0688788	0.2070084
logPGVgroups7	0.1312562	0.0365505	3.59	0.000	0.0596178	0.2028947
logPGVgroups8	0.1356882	0.0364925	3.72	0.000	0.0641636	0.2072128
logPGVgroups9	0.1576680	0.0366412	4.30	0.000	0.0858518	0.2294842
logPGVgroups10	0.1319708	0.0373662	3.53	0.000	0.0587337	0.2052078
logPGVgroups11	0.1560484	0.0377930	4.13	0.000	0.0819747	0.2301221
logPGVgroups12	0.1587540	0.0385857	4.11	0.000	0.0831266	0.2343813
logPGVgroups13	0.1049482	0.0392416	2.67	0.007	0.0280354	0.1818610
logPGVgroups14	0.1714537	0.0400654	4.28	0.000	0.0929262	0.2499812
logPGVgroups15	0.1886908	0.0408590	4.62	0.000	0.1086079	0.2687737
logPGVgroups16	0.3463003	0.0418611	8.27	0.000	0.2642532	0.4283473
logPGVgroups17	0.1340784	0.0424430	3.16	0.002	0.0508908	0.2172660
logPGVgroups18	0.1139060	0.0432609	2.63	0.008	0.0291153	0.1986967
logPGVgroups19	0.1204131	0.0441818	2.73	0.006	0.0338175	0.2070086
logPGVgroups20	0.1302812	0.0448541	2.90	0.004	0.0423679	0.2181945
logPGVgroups21	0.2463800	0.0458272	5.38	0.000	0.1565594	0.3362006
logPGVgroups22	0.2139368	0.0469567	4.56	0.000	0.1219025	0.3059711
logPGVgroups23	0.1767975	0.0480637	3.68	0.000	0.0825934	0.2710016
logPGVgroups24	0.3008088	0.0494519	6.08	0.000	0.2038839	0.3977338
logPGVgroups25	0.2565383	0.0493431	5.20	0.000	0.1598267	0.3532499
logPGVgroups26	0.2497997	0.0497306	5.02	0.000	0.1523287	0.3472708
logPGVgroups27	0.2963499	0.0509877	5.81	0.000	0.1964148	0.3962851
logPGVgroups28	0.2610276	0.0515280	5.07	0.000	0.1600336	0.3620217
logPGVgroups29	0.1555851	0.0523393	2.97	0.003	0.0530009	0.2581693
logPGVgroups30	0.3120553	0.0524871	5.95	0.000	0.2091814	0.4149292
logPGVgroups31	0.3071589	0.0532371	5.77	0.000	0.2028150	0.4115028
logPGVgroups32	0.2635425	0.0539289	4.89	0.000	0.1578427	0.3692423
logPGVgroups33	0.2889551	0.0546205	5.29	0.000	0.1818998	0.3960104
logPGVgroups34	0.3052008	0.0551630	5.53	0.000	0.1970822	0.4133194
logPGVgroups35	0.3332003	0.0554829	6.01	0.000	0.2244547	0.4419459
logPGVgroups36	0.3161513	0.0561441	5.63	0.000	0.2061098	0.4261929
logPGVgroups37	0.3969547	0.0567337	7.00	0.000	0.2857575	0.5081518
logPGVgroups38	0.3414841	0.0569615	5.99	0.000	0.2298405	0.4531277
logPGVgroups39	0.3085403	0.0573336	5.38	0.000	0.1961675	0.4209132
logPGVgroups40	0.3319205	0.0583940	5.68	0.000	0.2174692	0.4463718
logPGVgroups41	0.3854169	0.0585045	6.59	0.000	0.2707491	0.5000848
logPGVgroups42	0.3590032	0.0592540	6.06	0.000	0.2428664	0.4751400
logPGVgroups43	0.3506031	0.0602903	5.82	0.000	0.2324350	0.4687711

logPGVgroups44	0.3511490	0.0609974	5.76	0.000	0.2315951	0.4707029
logPGVgroups45	0.4335936	0.0620012	6.99	0.000	0.3120723	0.5551150
logPGVgroups46	0.4004498	0.0629432	6.36	0.000	0.2770822	0.5238175
logPGVgroups47	0.3691020	0.0636984	5.79	0.000	0.2442543	0.4939497
logPGVgroups48	0.3910398	0.0645285	6.06	0.000	0.2645649	0.5175146
logPGVgroups49	0.3865768	0.0658198	5.87	0.000	0.2575712	0.5155825
logPGVgroups50	0.4215375	0.0690300	6.11	0.000	0.2862397	0.5568352

LogPGVgroups estimates in 2SLS model

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
logPGVgroups2	-0.0033124	0.0319340	-0.10	0.917	-0.0659018	0.0592770
logPGVgroups3	0.0183877	0.0329646	0.56	0.577	-0.0462218	0.0829972
logPGVgroups4	0.0474911	0.0335960	1.41	0.157	-0.0183559	0.1133381
logPGVgroups5	0.0842573	0.0345530	2.44	0.015	0.0165347	0.1519799
logPGVgroups6	0.1358321	0.0352329	3.86	0.000	0.0667770	0.2048873
logPGVgroups7	0.1287819	0.0365333	3.53	0.000	0.0571779	0.2003859
logPGVgroups8	0.1325662	0.0364978	3.63	0.000	0.0610319	0.2041005
logPGVgroups9	0.1528428	0.0367495	4.16	0.000	0.0808152	0.2248704
logPGVgroups10	0.1273297	0.0374577	3.40	0.001	0.0539139	0.2007456
logPGVgroups11	0.1499166	0.0379548	3.95	0.000	0.0755265	0.2243067
logPGVgroups12	0.1538831	0.0386787	3.98	0.000	0.0780743	0.2296920
logPGVgroups13	0.0984177	0.0394031	2.50	0.012	0.0211889	0.1756464
logPGVgroups14	0.1638187	0.0403270	4.06	0.000	0.0847792	0.2428582
logPGVgroups15	0.1817630	0.0410410	4.43	0.000	0.1013240	0.2622020
logPGVgroups16	0.3361658	0.0423004	7.95	0.000	0.2532585	0.4190731
logPGVgroups17	0.1254886	0.0427733	2.93	0.003	0.0416546	0.2093227
logPGVgroups18	0.1049249	0.0435751	2.41	0.016	0.0195192	0.1903306
logPGVgroups19	0.1136378	0.0443533	2.56	0.010	0.0267070	0.2005686
logPGVgroups20	0.1218952	0.0451077	2.70	0.007	0.0334856	0.2103047
logPGVgroups21	0.2370805	0.0461337	5.14	0.000	0.1466601	0.3275008
logPGVgroups22	0.2061139	0.0471525	4.37	0.000	0.1136968	0.2985311
logPGVgroups23	0.1671817	0.0484002	3.45	0.001	0.0723190	0.2620444
logPGVgroups24	0.2881228	0.0499970	5.76	0.000	0.1901305	0.3861152
logPGVgroups25	0.2461649	0.0496711	4.96	0.000	0.1488113	0.3435184
logPGVgroups26	0.2409266	0.0499616	4.82	0.000	0.1430037	0.3388496
logPGVgroups27	0.2872291	0.0512290	5.61	0.000	0.1868221	0.3876361
logPGVgroups28	0.2525413	0.0517285	4.88	0.000	0.1511553	0.3539273
logPGVgroups29	0.1433306	0.0528950	2.71	0.007	0.0396583	0.2470028
logPGVgroups30	0.3014309	0.0528118	5.71	0.000	0.1979217	0.4049401
logPGVgroups31	0.2900851	0.0542015	5.35	0.000	0.1838521	0.3963181
logPGVgroups32	0.2500935	0.0544807	4.59	0.000	0.1433133	0.3568738
logPGVgroups33	0.2767607	0.0550528	5.03	0.000	0.1688592	0.3846623
logPGVgroups34	0.2883557	0.0560482	5.14	0.000	0.1785033	0.3982082
logPGVgroups35	0.3162877	0.0563625	5.61	0.000	0.2058192	0.4267562

logPGVgroups36	0.3069073	0.0563511	5.45	0.000	0.1964612	0.4173533
logPGVgroups37	0.3854819	0.0570864	6.75	0.000	0.2735947	0.4973691
logPGVgroups38	0.3346350	0.0570568	5.86	0.000	0.2228057	0.4464642
logPGVgroups39	0.2994787	0.0575387	5.20	0.000	0.1867050	0.4122524
logPGVgroups40	0.3209538	0.0587038	5.47	0.000	0.2058964	0.4360112
logPGVgroups41	0.3731732	0.0589220	6.33	0.000	0.2576882	0.4886582
logPGVgroups42	0.3484402	0.0595554	5.85	0.000	0.2317138	0.4651665
logPGVgroups43	0.3378296	0.0607396	5.56	0.000	0.2187822	0.4568770
logPGVgroups44	0.3410743	0.0612218	5.57	0.000	0.2210817	0.4610668
logPGVgroups45	0.4226868	0.0622822	6.79	0.000	0.3006160	0.5447577
logPGVgroups46	0.3925642	0.0630358	6.23	0.000	0.2690163	0.5161121
logPGVgroups47	0.3628545	0.0637258	5.69	0.000	0.2379543	0.4877548
logPGVgroups48	0.3856666	0.0645103	5.98	0.000	0.2592287	0.5121046
logPGVgroups49	0.3778186	0.0659499	5.73	0.000	0.2485592	0.5070781
logPGVgroups50	0.4060007	0.0696301	5.83	0.000	0.2695283	0.5424731

Appendix III: Regression models

DID OLS model

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
Before	-0.0707177	0.0358024	-1.98	0.048	-0.1408907	-0.0005447
AfterHuizinge	0.0675304	0.0348431	1.94	0.053	-0.0007624	0.1358232
logArea	0.3036799	0.0242941	12.50	0.000	0.2560634	0.3512965
logLot	-0.0886217	0.0104354	-8.49	0.000	-0.1090750	-0.0681683
Rooms	-0.0187277	0.0110530	-1.69	0.090	-0.0403917	0.0029362
Type corner	-0.0254976	0.0183916	-1.39	0.166	-0.0615453	0.0105501
Type half double	0.1266561	0.0174493	7.26	0.000	0.0924553	0.1608569
Type detached	0.4885081	0.0209050	23.37	0.000	0.4475342	0.5294820
Built 1906-1930	-0.0059903	0.0287155	-0.21	0.835	-0.0622730	0.0502923
Built 1931-1944	0.0266406	0.0317096	0.84	0.401	-0.0355104	0.0887917
Built 1945-1959	-0.0847397	0.0315508	-2.69	0.007	-0.1465795	-0.0228998
Built 1960-1970	0.0435640	0.0292140	1.49	0.136	-0.0136957	0.1008237
Built 1971-1980	0.0503020	0.0282486	1.78	0.075	-0.0050655	0.1056695
Built 1981-1990	0.0405966	0.0302956	1.34	0.180	-0.0187830	0.0999762
Built 1991-2001	0.0467755	0.0299768	1.56	0.119	-0.0119794	0.1055303
Built >2001	0.0025975	0.0359103	0.07	0.942	-0.0677870	0.0729819
Balcony	0.1054267	0.0209266	5.04	0.000	0.0644103	0.1464431
Garage	0.0485304	0.0125250	3.87	0.000	0.0239813	0.0730794
Attic	-0.0457350	0.0114539	-3.99	0.000	-0.0681847	-0.0232853
Loft	0.0373063	0.0154524	2.41	0.016	0.0070195	0.0675930
Barn	-0.0087255	0.0109888	-0.79	0.427	-0.0302636	0.0128126
Fire	0.0410839	0.0210519	1.95	0.051	-0.0001780	0.0823458
Maintenance	0.2582598	0.0145600	17.74	0.000	0.2297220	0.2867976
Density 500-1000	-0.0705719	0.0139011	-5.08	0.000	-0.0978182	-0.0433255
Density 1000-1500	-0.0873062	0.0207133	-4.21	0.000	-0.1279045	-0.0467079
Density 1500-2500	0.0280440	0.0307242	0.91	0.361	-0.0321757	0.0882638
Elderly	-0.0394989	0.0106512	-3.71	0.000	-0.0603753	-0.0186224
Year 2004	0.0964561	0.0226170	4.26	0.000	0.0521267	0.1407855
Year 2005	0.1733365	0.0215159	8.06	0.000	0.1311652	0.2155078
Year 2006	0.1237431	0.0216686	5.71	0.000	0.0812724	0.1662137
Year 2007	0.1684243	0.0220486	7.64	0.000	0.1252089	0.2116396
Year 2008	0.2103555	0.0227461	9.25	0.000	0.1657729	0.2549380
Year 2009	0.4823500	0.0249768	19.31	0.000	0.4333952	0.5313048
Year 2010	0.5777878	0.0247830	23.31	0.000	0.5292129	0.6263628
Year 2011	0.5256104	0.0263091	19.98	0.000	0.4740443	0.5771766
Year 2012	0.6900205	0.0262127	26.32	0.000	0.6386433	0.7413977
Year 2013	0.7273664	0.0281894	25.80	0.000	0.6721149	0.7826180
Year 2014	0.6108702	0.0260138	23.48	0.000	0.5598829	0.6618575
Municipality 5	-0.2330266	0.0582317	-4.00	0.000	-0.3471613	-0.1188919
Municipality 7	0.2328740	0.0737238	3.16	0.002	0.0883748	0.3773733
Municipality 9	-0.0970440	0.0581739	-1.67	0.095	-0.2110654	0.0169774

Municipality 10	0.1298820	0.0451283	2.88	0.004	0.0414301	0.2183338
Municipality 15	-0.3720020	0.0682377	-5.45	0.000	-0.5057485	-0.2382554
Municipality 17	-0.0835522	0.0457402	-1.83	0.068	-0.1732034	0.0060991
Municipality 18	-0.0180411	0.0433564	-0.42	0.677	-0.1030200	0.0669378
Municipality 22	-0.2710436	0.0593734	-4.57	0.000	-0.3874159	-0.1546713
Municipality 24	-0.1211237	0.0575757	-2.10	0.035	-0.2339725	-0.0082749
Municipality 25	-0.2774890	0.0746199	-3.72	0.000	-0.4237445	-0.1312334
Municipality 37	0.0253779	0.0586043	0.43	0.665	-0.0894871	0.1402429
Municipality 39	0.2285103	0.0657563	3.48	0.001	0.0996274	0.3573931
Municipality 40	-0.1904346	0.0523285	-3.64	0.000	-0.2929989	-0.0878703
Municipality 47	0.2032797	0.0554774	3.66	0.000	0.0945436	0.3120158
Municipality 48	-0.1384631	0.0656508	-2.11	0.035	-0.2671392	-0.0097870
Municipality 52	0.2539710	0.0582488	4.36	0.000	0.1398029	0.3681392
Municipality 53	-0.1369065	0.0494736	-2.77	0.006	-0.2338751	-0.0399379
Municipality 56	-0.3182193	0.0555390	-5.73	0.000	-0.4270762	-0.2093625
Municipality 58	0.0634868	0.0634781	1.00	0.317	-0.0609309	0.1879045
Municipality 59	0.0002438	0.0614955	0.00	0.997	-0.1202879	0.1207754
Municipality 79	0.0683893	0.0713572	0.96	0.338	-0.0714714	0.2082501
Municipality 83	-0.0430065	0.0704580	-0.61	0.542	-0.1811048	0.0950918
Municipality 85	-0.1077173	0.0580253	-1.86	0.063	-0.2214473	0.0060127
Municipality 86	-0.2936396	0.0575407	-5.10	0.000	-0.4064200	-0.1808593
Municipality 90	-0.2672513	0.0541090	-4.94	0.000	-0.3733054	-0.1611972
Municipality 765	0.2476578	0.0648639	3.82	0.000	0.1205241	0.3747916
Municipality 1651	-0.1001048	0.0540183	-1.85	0.064	-0.2059812	0.0057716
Municipality 1661	0.2320869	0.0792572	2.93	0.003	0.0767421	0.3874317
Municipality 1663	0.0475966	0.0602884	0.79	0.430	-0.0705692	0.1657623
Municipality 1680	0.0650908	0.0591867	1.10	0.271	-0.0509157	0.1810973
Municipality 1681	0.0791496	0.0580621	1.36	0.173	-0.0346526	0.1929518
Municipality 1699	-0.1924623	0.0568203	-3.39	0.001	-0.3038306	-0.0810940
Municipality 1722	0.1529609	0.0798100	1.92	0.055	-0.0034674	0.3093892
Municipality 1730	-0.2364611	0.0570956	-4.14	0.000	-0.3483690	-0.1245531
Municipality 1731	-0.0169915	0.0568451	-0.30	0.765	-0.1284084	0.0944254
Municipality 1891	-0.1398430	0.0692745	-2.02	0.044	-0.2756217	-0.0040644
Municipality 1987	0.0771500	0.0544018	1.42	0.156	-0.0294781	0.1837780
Constant	3.0406720	0.1288035	23.61	0.000	2.7882160	3.2931280
N	53,315					
R^2	0.1015					
Adjusted R^2	0.1002					
Joint sign. F-test	129.71***					

DID 2SLS model (first stage)

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
Before	-0.0121753	0.0057515	-2.12	0.034	-0.0234481	-0.0009024
AfterHuizinge	-0.0062558	0.0059170	-1.06	0.290	-0.0178532	0.0053417

logArea	0.5881815	0.0054751	107.43	0.000	0.5774504	0.5989127
logLot	0.1464741	0.0020533	71.33	0.000	0.1424495	0.1504987
Rooms	0.0294769	0.0017205	17.13	0.000	0.0261048	0.0328491
Type corner	0.0173919	0.0025767	6.75	0.000	0.0123416	0.0224422
Type half double	0.0757649	0.0028006	27.05	0.000	0.0702757	0.0812540
Type detached	0.1685095	0.0040426	41.68	0.000	0.1605860	0.1764331
Built 1906-1930	-0.0036959	0.0060818	-0.61	0.543	-0.0156162	0.0082245
Built 1931-1944	0.0443734	0.0062728	7.07	0.000	0.0320786	0.0566682
Built 1945-1959	0.0271724	0.0061682	4.41	0.000	0.0150826	0.0392623
Built 1960-1970	0.0325854	0.0059037	5.52	0.000	0.0210141	0.0441566
Built 1971-1980	0.0366332	0.0059836	6.12	0.000	0.0249053	0.0483611
Built 1981-1990	0.1137284	0.0059328	19.17	0.000	0.1021001	0.1253568
Built 1991-2001	0.2011373	0.0059237	33.95	0.000	0.1895269	0.2127477
Built >2001	0.2255628	0.0074254	30.38	0.000	0.2110089	0.2401167
Balcony	0.0305861	0.0031780	9.62	0.000	0.0243571	0.0368151
Garage	0.0795098	0.0023572	33.73	0.000	0.0748896	0.0841299
Attic	0.0202226	0.0017317	11.68	0.000	0.0168283	0.0236168
Loft	0.0251655	0.0023033	10.93	0.000	0.0206510	0.0296800
Barn	0.0185017	0.0019217	9.63	0.000	0.0147352	0.0222683
Fire	0.0733968	0.0033335	22.02	0.000	0.0668632	0.0799305
Maintenance	0.1768383	0.0026619	66.43	0.000	0.1716209	0.1820558
Density 500-1000	0.0596903	0.0021299	28.03	0.000	0.0555158	0.0638648
Density 1000-1500	0.0851811	0.0031535	27.01	0.000	0.0790002	0.0913619
Density 1500-2500	0.0590073	0.0045110	13.08	0.000	0.0501658	0.0678489
Elderly	0.0268745	0.0018191	14.77	0.000	0.0233092	0.0304399
Year 2004	0.0285729	0.0037153	7.69	0.000	0.0212910	0.0358549
Year 2005	0.0676190	0.0035090	19.27	0.000	0.0607413	0.0744968
Year 2006	0.0595770	0.0037159	16.03	0.000	0.0522938	0.0668602
Year 2007	0.0924585	0.0034543	26.77	0.000	0.0856880	0.0992291
Year 2008	0.0866566	0.0035242	24.59	0.000	0.0797492	0.0935639
Year 2009	0.0332663	0.0038621	8.61	0.000	0.0256966	0.0408360
Year 2010	0.0074578	0.0038169	1.95	0.051	-0.0000234	0.0149389
Year 2011	-0.0231445	0.0040126	-5.77	0.000	-0.0310092	-0.0152798
Year 2012	-0.1292230	0.0042182	-30.63	0.000	-0.1374907	-0.1209552
Year 2013	-0.1970066	0.0046397	-42.46	0.000	-0.2061005	-0.1879127
Year 2014	-0.1885822	0.0041629	-45.30	0.000	-0.1967416	-0.1804228
Municipality 5	0.1511606	0.0086406	17.49	0.000	0.1342250	0.1680962
Municipality 7	-0.0938652	0.0137120	-6.85	0.000	-0.1207407	-0.0669896
Municipality 9	0.0908392	0.0092086	9.86	0.000	0.0727903	0.1088881
Municipality 10	-0.0909337	0.0072938	-12.47	0.000	-0.1052296	-0.0766377
Municipality 15	0.0880857	0.0100794	8.74	0.000	0.0683299	0.1078414
Municipality 17	0.3833070	0.0073731	51.99	0.000	0.3688557	0.3977584
Municipality 18	-0.0398364	0.0067254	-5.92	0.000	-0.0530182	-0.0266546
Municipality 22	0.1806397	0.0092972	19.43	0.000	0.1624172	0.1988623
Municipality 24	0.0099450	0.0100331	0.99	0.322	-0.0097199	0.0296099
Municipality 25	0.1750069	0.0109991	15.91	0.000	0.1534485	0.1965653

Municipality 37	-0.0263998	0.0099095	-2.66	0.008	-0.0458225	-0.0069770
Municipality 39	-0.1082490	0.0112956	-9.58	0.000	-0.1303885	-0.0861095
Municipality 40	0.0556920	0.0083510	6.67	0.000	0.0393240	0.0720600
Municipality 47	-0.0798366	0.0089799	-8.89	0.000	-0.0974372	-0.0622360
Municipality 48	-0.0050490	0.0108296	-0.47	0.641	-0.0262752	0.0161772
Municipality 52	-0.0548937	0.0093782	-5.85	0.000	-0.0732751	-0.0365124
Municipality 53	0.1110177	0.0076366	14.54	0.000	0.0960500	0.1259855
Municipality 56	0.0854129	0.0086746	9.85	0.000	0.0684106	0.1024153
Municipality 58	0.0925442	0.0114811	8.06	0.000	0.0700410	0.1150473
Municipality 59	0.1328310	0.0094973	13.99	0.000	0.1142162	0.1514458
Municipality 79	0.0542686	0.0109521	4.96	0.000	0.0328024	0.0757349
Municipality 83	0.0725508	0.0128551	5.64	0.000	0.0473547	0.0977469
Municipality 85	0.1315347	0.0094106	13.98	0.000	0.1130899	0.1499795
Municipality 86	0.2141212	0.0093153	22.99	0.000	0.1958632	0.2323793
Municipality 90	0.1362351	0.0090405	15.07	0.000	0.1185157	0.1539545
Municipality 765	-0.1453447	0.0108358	-13.41	0.000	-0.1665830	-0.1241063
Municipality 1651	-0.0401183	0.0085610	-4.69	0.000	-0.0568979	-0.0233386
Municipality 1661	-0.2697266	0.0146339	-18.43	0.000	-0.2984090	-0.2410441
Municipality 1663	-0.0172742	0.0101074	-1.71	0.087	-0.0370848	0.0025365
Municipality 1680	0.1815570	0.0098417	18.45	0.000	0.1622671	0.2008468
Municipality 1681	-0.0030488	0.0106622	-0.29	0.775	-0.0239468	0.0178493
Municipality 1699	0.2196260	0.0093624	23.46	0.000	0.2012757	0.2379763
Municipality 1722	0.0275149	0.0131290	2.10	0.036	0.0017818	0.0532479
Municipality 1730	0.2546304	0.0093662	27.19	0.000	0.2362725	0.2729883
Municipality 1731	0.1429659	0.0089096	16.05	0.000	0.1255030	0.1604287
Municipality 1891	0.0930398	0.0103777	8.97	0.000	0.0726994	0.1133802
Municipality 1987	-0.0451413	0.0086424	-5.22	0.000	-0.0620806	-0.0282020
Disability	-0.0004347	0.0000363	-11.98	0.000	-0.0005058	-0.0003636
Train	-0.0369897	0.0030642	-12.07	0.000	-0.0429956	-0.0309839
Constant	7.8352420	0.0264136	296.64	0.000	7.7834720	7.8870130
N	53,315					
R^2	0.7912					
Adjusted R^2	0.7909					
Joint sign. F-test	2993.28***					

DID 2SLS model (second stage)

	Coefficient	Robust SE	z	Sign.	95% Conf. Interval	
logPrice	-1.3786960	0.2935824	-4.70	0.000	-1.9541070	-0.8032850
Before	-0.0921288	0.0368565	-2.50	0.012	-0.1643661	-0.0198914
AfterHuizinge	0.0603832	0.0353449	1.71	0.088	-0.0088916	0.1296579
logArea	1.1223110	0.1756668	6.39	0.000	0.7780109	1.4666120
logLot	0.1137424	0.0445460	2.55	0.011	0.0264338	0.2010509
Rooms	0.0217054	0.0140816	1.54	0.123	-0.0058940	0.0493048
Type corner	-0.0020749	0.0191382	-0.11	0.914	-0.0395851	0.0354352

Type half double	0.2299393	0.0281488	8.17	0.000	0.1747686	0.2851100
Type detached	0.7162034	0.0528168	13.56	0.000	0.6126844	0.8197224
Built 1906-1930	-0.0115825	0.0297534	-0.39	0.697	-0.0698981	0.0467330
Built 1931-1944	0.0878224	0.0349908	2.51	0.012	0.0192416	0.1564031
Built 1945-1959	-0.0493796	0.0332428	-1.49	0.137	-0.1145343	0.0157751
Built 1960-1970	0.0857818	0.0312959	2.74	0.006	0.0244430	0.1471206
Built 1971-1980	0.1012556	0.0310344	3.26	0.001	0.0404293	0.1620819
Built 1981-1990	0.1972067	0.0456894	4.32	0.000	0.1076570	0.2867563
Built 1991-2001	0.3246281	0.0667387	4.86	0.000	0.1938227	0.4554335
Built >2001	0.3133077	0.0769182	4.07	0.000	0.1625508	0.4640646
Balcony	0.1474324	0.0227733	6.47	0.000	0.1027975	0.1920672
Garage	0.1594536	0.0267772	5.95	0.000	0.1069713	0.2119359
Attic	-0.0178619	0.0129566	-1.38	0.168	-0.0432565	0.0075326
Loft	0.0726203	0.0172009	4.22	0.000	0.0389072	0.1063334
Barn	0.0165677	0.0122154	1.36	0.175	-0.0073740	0.0405094
Fire	0.1432890	0.0302443	4.74	0.000	0.0840113	0.2025668
Maintenance	0.5014973	0.0538434	9.31	0.000	0.3959661	0.6070284
Density 500-1000	0.0182645	0.0235389	0.78	0.438	-0.0278708	0.0643998
Density 1000-1500	0.0362029	0.0337233	1.07	0.283	-0.0298936	0.1022994
Density 1500-2500	0.1060702	0.0349866	3.03	0.002	0.0374978	0.1746426
Elderly	-0.0035429	0.0132555	-0.27	0.789	-0.0295231	0.0224373
Year 2004	0.1347727	0.0245897	5.48	0.000	0.0865778	0.1829676
Year 2005	0.2667924	0.0295042	9.04	0.000	0.2089651	0.3246196
Year 2006	0.2104496	0.0286072	7.36	0.000	0.1543805	0.2665187
Year 2007	0.2943595	0.0347335	8.47	0.000	0.2262831	0.3624360
Year 2008	0.3285920	0.0341299	9.63	0.000	0.2616986	0.3954855
Year 2009	0.5379904	0.0280109	19.21	0.000	0.4830901	0.5928907
Year 2010	0.5944084	0.0251619	23.62	0.000	0.5450921	0.6437247
Year 2011	0.4933909	0.0274861	17.95	0.000	0.4395192	0.5472626
Year 2012	0.5212744	0.0450230	11.58	0.000	0.4330310	0.6095179
Year 2013	0.4553407	0.0646593	7.04	0.000	0.3286108	0.5820706
Year 2014	0.3498136	0.0615486	5.68	0.000	0.2291805	0.4704466
Municipality 5	-0.0186010	0.0741788	-0.25	0.802	-0.1639889	0.1267868
Municipality 7	0.0549934	0.0849123	0.65	0.517	-0.1114317	0.2214185
Municipality 9	0.0570929	0.0676188	0.84	0.398	-0.0754374	0.1896232
Municipality 10	0.0144213	0.0518151	0.28	0.781	-0.0871343	0.1159770
Municipality 15	-0.2502539	0.0736663	-3.40	0.001	-0.3946372	-0.1058706
Municipality 17	0.4656890	0.1256163	3.71	0.000	0.2194855	0.7118924
Municipality 18	-0.0795853	0.0457697	-1.74	0.082	-0.1692923	0.0101217
Municipality 22	-0.0669813	0.0742309	-0.90	0.367	-0.2124713	0.0785087
Municipality 24	-0.0863498	0.0589099	-1.47	0.143	-0.2018111	0.0291114
Municipality 25	-0.0731923	0.0868472	-0.84	0.399	-0.2434097	0.0970250
Municipality 37	-0.0734272	0.0631825	-1.16	0.245	-0.1972626	0.0504083
Municipality 39	0.0755167	0.0748578	1.01	0.313	-0.0712018	0.2222352
Municipality 40	-0.1097729	0.0556568	-1.97	0.049	-0.2188583	-0.0006875
Municipality 47	0.0739955	0.0626368	1.18	0.237	-0.0487704	0.1967614

Municipality 48	-0.2156689	0.0685578	-3.15	0.002	-0.3500397	-0.0812981
Municipality 52	0.1645737	0.0619937	2.65	0.008	0.0430684	0.2860791
Municipality 53	0.0355789	0.0619820	0.57	0.566	-0.0859036	0.1570615
Municipality 56	-0.1804417	0.0636351	-2.84	0.005	-0.3051642	-0.0557193
Municipality 58	0.1455533	0.0671724	2.17	0.030	0.0138977	0.2772088
Municipality 59	0.1624516	0.0713640	2.28	0.023	0.0225809	0.3023224
Municipality 79	0.1447071	0.0743605	1.95	0.052	-0.0010368	0.2904510
Municipality 83	0.0794948	0.0753719	1.05	0.292	-0.0682313	0.2272209
Municipality 85	0.0189097	0.0648687	0.29	0.771	-0.1082305	0.1460500
Municipality 86	-0.0373773	0.0801990	-0.47	0.641	-0.1945644	0.1198099
Municipality 90	-0.1189588	0.0632960	-1.88	0.060	-0.2430167	0.0050990
Municipality 765	0.0049197	0.0835753	0.06	0.953	-0.1588849	0.1687243
Municipality 1651	-0.1458983	0.0555023	-2.63	0.009	-0.2546807	-0.0371159
Municipality 1661	-0.1616897	0.1175107	-1.38	0.169	-0.3920064	0.0686269
Municipality 1663	0.0010547	0.0616184	0.02	0.986	-0.1197152	0.1218246
Municipality 1680	0.2922211	0.0767773	3.81	0.000	0.1417403	0.4427018
Municipality 1681	0.0376911	0.0590692	0.64	0.523	-0.0780825	0.1534646
Municipality 1699	0.0729306	0.0807562	0.90	0.366	-0.0853488	0.2312099
Municipality 1722	0.1581431	0.0810477	1.95	0.051	-0.0007076	0.3169938
Municipality 1730	0.1047929	0.0938682	1.12	0.264	-0.0791854	0.2887713
Municipality 1731	0.1760212	0.0707914	2.49	0.013	0.0372727	0.3147697
Municipality 1891	-0.0072544	0.0756243	-0.10	0.924	-0.1554754	0.1409666
Municipality 1987	0.0185841	0.0564111	0.33	0.742	-0.0919796	0.1291478
Constant	13.7416000	2.2833500	6.02	0.000	9.2663160	18.2168800
N	53,315					
R^2	0.0877					
Wald Chi^2	6188.15***					

PGV OLS model

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
logPGV	0.0467944	0.0152100	3.08	0.002	0.0169830	0.0766058
logArea	0.2620694	0.0167564	15.64	0.000	0.2292271	0.2949118
logLot	-0.0892862	0.0073770	-12.10	0.000	-0.1037451	-0.0748273
Rooms	-0.0288764	0.0073556	-3.93	0.000	-0.0432933	-0.0144595
Type corner	0.0320945	0.0103961	3.09	0.002	0.0117182	0.0524707
Type half double	0.1580801	0.0111973	14.12	0.000	0.1361337	0.1800265
Type detached	0.5603265	0.0139153	40.27	0.000	0.5330529	0.5876002
Built 1906-1930	-0.0045297	0.0190089	-0.24	0.812	-0.0417867	0.0327274
Built 1931-1944	0.0278210	0.0207220	1.34	0.179	-0.0127938	0.0684358
Built 1945-1959	-0.0543182	0.0216526	-2.51	0.012	-0.0967569	-0.0118795
Built 1960-1970	0.0462486	0.0191347	2.42	0.016	0.0087449	0.0837523
Built 1971-1980	0.0394803	0.0186605	2.12	0.034	0.0029061	0.0760546
Built 1981-1990	0.0394696	0.0195097	2.02	0.043	0.0012309	0.0777084
Built 1991-2001	0.0859727	0.0192034	4.48	0.000	0.0483344	0.1236109

Built >2001	0.0772776	0.0224952	3.44	0.001	0.0331875	0.1213678
Balcony	0.0641634	0.0132880	4.83	0.000	0.0381191	0.0902078
Garage	0.0479226	0.0087478	5.48	0.000	0.0307770	0.0650682
Attic	-0.0697379	0.0074136	-9.41	0.000	-0.0842684	-0.0552075
Loft	0.0265541	0.0103721	2.56	0.010	0.0062250	0.0468833
Barn	-0.0346922	0.0075804	-4.58	0.000	-0.0495497	-0.0198347
Fire	0.0309982	0.0145138	2.14	0.033	0.0025514	0.0594450
Maintenance	0.2447501	0.0100271	24.41	0.000	0.2250971	0.2644030
Density 500-1000	-0.0860196	0.0101218	-8.50	0.000	-0.1058582	-0.0661810
Density 1000-1500	-0.0983080	0.0125045	-7.86	0.000	-0.1228166	-0.0737995
Density 1500-2500	-0.0469373	0.0157212	-2.99	0.003	-0.0777507	-0.0161240
Elderly	-0.0249131	0.0215056	-1.16	0.247	-0.0670638	0.0172376
Year 2004	-0.0446314	0.0070341	-6.35	0.000	-0.0584181	-0.0308448
Year 2005	0.1003300	0.0146524	6.85	0.000	0.0716115	0.1290485
Year 2006	0.1443770	0.0139451	10.35	0.000	0.1170448	0.1717092
Year 2007	0.1102951	0.0144796	7.62	0.000	0.0819153	0.1386749
Year 2008	0.1500899	0.0149819	10.02	0.000	0.1207257	0.1794541
Year 2009	0.2342506	0.0158312	14.80	0.000	0.2032217	0.2652795
Year 2010	0.4949724	0.0174337	28.39	0.000	0.4608026	0.5291422
Year 2011	0.5696729	0.0179324	31.77	0.000	0.5345257	0.6048202
Year 2012	0.4924371	0.0192449	25.59	0.000	0.4547174	0.5301567
Year 2013	0.6532695	0.0199364	32.77	0.000	0.6141945	0.6923445
Year 2014	0.6536543	0.0212612	30.74	0.000	0.6119827	0.6953259
Municipality 5	0.5241607	0.0208061	25.19	0.000	0.4833811	0.5649402
Municipality 7	-0.2135745	0.0572654	-3.73	0.000	-0.3258137	-0.1013354
Municipality 9	0.3404887	0.0658852	5.17	0.000	0.2113549	0.4696226
Municipality 10	-0.1021790	0.0571492	-1.79	0.074	-0.2141905	0.0098325
Municipality 14	0.1470324	0.0448952	3.28	0.001	0.0590386	0.2350261
Municipality 15	-0.4184100	0.0400397	-10.45	0.000	-0.4968872	-0.3399328
Municipality 17	-0.2618133	0.0608906	-4.30	0.000	-0.3811578	-0.1424688
Municipality 18	0.0029343	0.0460054	0.06	0.949	-0.0872356	0.0931042
Municipality 22	0.0377868	0.0433988	0.87	0.384	-0.0472741	0.1228476
Municipality 24	-0.1418990	0.0503838	-2.82	0.005	-0.2406504	-0.0431476
Municipality 25	-0.1700925	0.0567946	-2.99	0.003	-0.2814089	-0.0587762
Municipality 37	-0.1684975	0.0680318	-2.48	0.013	-0.3018387	-0.0351563
Municipality 39	0.1388542	0.0482054	2.88	0.004	0.0443723	0.2333360
Municipality 40	0.3020223	0.0593437	5.09	0.000	0.1857096	0.4183350
Municipality 47	-0.2067712	0.0510091	-4.05	0.000	-0.3067482	-0.1067941
Municipality 48	0.3210845	0.0446274	7.19	0.000	0.2336156	0.4085534
Municipality 51	-0.0568818	0.0549844	-1.03	0.301	-0.1646503	0.0508867
Municipality 52	-0.0561308	0.0552726	-1.02	0.310	-0.1644641	0.0522025
Municipality 53	0.4004802	0.0483396	8.28	0.000	0.3057354	0.4952250
Municipality 55	-0.1023015	0.0494212	-2.07	0.038	-0.1991661	-0.0054368
Municipality 56	0.1896505	0.0564542	3.36	0.001	0.0790011	0.3002998
Municipality 58	-0.2126985	0.0493674	-4.31	0.000	-0.3094578	-0.1159393
Municipality 59	0.1928851	0.0592702	3.25	0.001	0.0767165	0.3090537

Municipality 60	0.1386159	0.0551749	2.51	0.012	0.0304740	0.2467579
Municipality 63	0.5361196	0.1062718	5.04	0.000	0.3278286	0.7444106
Municipality 64	0.2457475	0.0876310	2.80	0.005	0.0739922	0.4175029
Municipality 70	0.1070731	0.0715656	1.50	0.135	-0.0331944	0.2473406
Municipality 72	0.2294945	0.0634316	3.62	0.000	0.1051696	0.3538195
Municipality 74	0.2021704	0.0620583	3.26	0.001	0.0805371	0.3238037
Municipality 79	-0.0643432	0.0518287	-1.24	0.214	-0.1659266	0.0372403
Municipality 80	0.2013036	0.0659990	3.05	0.002	0.0719466	0.3306606
Municipality 81	0.0893706	0.0496647	1.80	0.072	-0.0079714	0.1867127
Municipality 82	0.0475448	0.0700305	0.68	0.497	-0.0897139	0.1848034
Municipality 83	0.3708756	0.0604552	6.13	0.000	0.2523844	0.4893668
Municipality 85	0.1088413	0.0676271	1.61	0.108	-0.0237067	0.2413893
Municipality 86	0.0159968	0.0493831	0.32	0.746	-0.0807931	0.1127868
Municipality 88	-0.1587008	0.0512534	-3.10	0.002	-0.2591567	-0.0582449
Municipality 90	0.2659570	0.1581699	1.68	0.093	-0.0440534	0.5759673
Municipality 91	-0.0752994	0.0478918	-1.57	0.116	-0.1691666	0.0185678
Municipality 98	0.1673521	0.0523442	3.20	0.001	0.0647583	0.2699458
Municipality 104	0.0988463	0.0594928	1.66	0.097	-0.0177586	0.2154512
Municipality 106	0.2135807	0.0672156	3.18	0.001	0.0818392	0.3453222
Municipality 109	-0.0341369	0.0394865	-0.86	0.387	-0.1115297	0.0432559
Municipality 114	0.1416973	0.0468824	3.02	0.003	0.0498087	0.2335860
Municipality 118	0.1406887	0.0399220	3.52	0.000	0.0624423	0.2189351
Municipality 119	0.0608087	0.0488206	1.25	0.213	-0.0348788	0.1564961
Municipality 140	0.1794969	0.0529715	3.39	0.001	0.0756736	0.2833202
Municipality 653	0.0745710	0.0691508	1.08	0.281	-0.0609634	0.2101053
Municipality 683	0.1168598	0.0673795	1.73	0.083	-0.0152029	0.2489225
Municipality 710	0.1031832	0.0601523	1.72	0.086	-0.0147142	0.2210807
Municipality 737	0.2004581	0.0696061	2.88	0.004	0.0640313	0.3368848
Municipality 765	0.2131780	0.0522733	4.08	0.000	0.1107233	0.3156327
Municipality 1651	0.3622492	0.0559696	6.47	0.000	0.2525496	0.4719488
Municipality 1661	-0.1152128	0.0530258	-2.17	0.030	-0.2191425	-0.0112832
Municipality 1663	0.3518851	0.0727679	4.84	0.000	0.2092611	0.4945090
Municipality 1680	0.1019914	0.0606849	1.68	0.093	-0.0169501	0.2209329
Municipality 1681	0.1469118	0.0468463	3.14	0.002	0.0550939	0.2387297
Municipality 1690	0.1716209	0.0461194	3.72	0.000	0.0812276	0.2620141
Municipality 1699	-0.0333155	0.0584548	-0.57	0.569	-0.1478859	0.0812550
Municipality 1701	-0.0695819	0.0460719	-1.51	0.131	-0.1598821	0.0207183
Municipality 1722	0.1299254	0.0544246	2.39	0.017	0.0232541	0.2365968
Municipality 1730	0.2854062	0.0771300	3.70	0.000	0.1342326	0.4365797
Municipality 1731	-0.1237628	0.0448421	-2.76	0.006	-0.2116525	-0.0358731
Municipality 1891	0.0873820	0.0456007	1.92	0.055	-0.0019946	0.1767585
Municipality 1987	-0.0037129	0.0650864	-0.06	0.955	-0.1312812	0.1238553
Constant	0.1053448	0.0538673	1.96	0.051	-0.0002342	0.2109238
N	122,908					
R^2	0.1120					
Adjusted R^2	0.1113					

Joint sign. F-test	288.24***

PGV 2SLS model (first stage)

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
logPGV	0.0144521	0.0024603	5.87	0.000	0.0096301	0.0192742
logArea	0.5864384	0.0034468	170.14	0.000	0.5796827	0.5931942
logLot	0.1385587	0.0014604	94.88	0.000	0.1356964	0.1414209
Rooms	0.0230985	0.0011056	20.89	0.000	0.0209315	0.0252655
Type corner	0.0056630	0.0014801	3.83	0.000	0.0027621	0.0085639
Type half double	0.0666078	0.0017290	38.52	0.000	0.0632189	0.0699966
Type detached	0.1578874	0.0025920	60.91	0.000	0.1528071	0.1629678
Built 1906-1930	-0.0425341	0.0037684	-11.29	0.000	-0.0499200	-0.0351482
Built 1931-1944	-0.0245372	0.0038965	-6.3	0.000	-0.0321744	-0.0169001
Built 1945-1959	-0.0376659	0.0039135	-9.62	0.000	-0.0453363	-0.0299955
Built 1960-1970	-0.0556356	0.0036786	-15.12	0.000	-0.0628455	-0.0484256
Built 1971-1980	-0.0272689	0.0037010	-7.37	0.000	-0.0345228	-0.0200150
Built 1981-1990	0.0256819	0.0036717	6.99	0.000	0.0184855	0.0328784
Built 1991-2001	0.1103206	0.0036511	30.22	0.000	0.1031645	0.1174767
Built >2001	0.1603924	0.0042865	37.42	0.000	0.1519910	0.1687937
Balcony	0.0543073	0.0020667	26.28	0.000	0.0502566	0.0583579
Garage	0.0818236	0.0015550	52.62	0.000	0.0787757	0.0848714
Attic	0.0108474	0.0010729	10.11	0.000	0.0087444	0.0129503
Loft	0.0215736	0.0015323	14.08	0.000	0.0185703	0.0245769
Barn	0.0084041	0.0012661	6.64	0.000	0.0059225	0.0108856
Fire	0.0707576	0.0022390	31.6	0.000	0.0663692	0.0751460
Maintenance	0.1609822	0.0017369	92.68	0.000	0.1575779	0.1643865
Density 500-1000	0.0550692	0.0015386	35.79	0.000	0.0520535	0.0580849
Density 1000-1500	0.0601667	0.0020049	30.01	0.000	0.0562371	0.0640962
Density 1500-2500	0.0821970	0.0025186	32.64	0.000	0.0772606	0.0871333
Elderly	0.2253449	0.0037927	59.42	0.000	0.2179113	0.2327785
Year 2004	0.0201487	0.0010997	18.32	0.000	0.0179933	0.0223041
Year 2005	0.0350637	0.0023469	14.94	0.000	0.0304638	0.0396636
Year 2006	0.0674239	0.0022257	30.29	0.000	0.0630616	0.0717862
Year 2007	0.0869673	0.0023440	37.1	0.000	0.0823732	0.0915615
Year 2008	0.0937514	0.0023028	40.71	0.000	0.0892379	0.0982649
Year 2009	0.0783083	0.0024072	32.53	0.000	0.0735901	0.0830264
Year 2010	0.0184861	0.0026878	6.88	0.000	0.0132179	0.0237542
Year 2011	0.0058807	0.0027474	2.14	0.032	0.0004959	0.0112656
Year 2012	-0.0234295	0.0029366	-7.98	0.000	-0.0291851	-0.0176738
Year 2013	-0.1263156	0.0031121	-40.59	0.000	-0.1324152	-0.1202160
Year 2014	-0.1855338	0.0033326	-55.67	0.000	-0.1920656	-0.1790020
Municipality 5	-0.1798455	0.0032100	-56.03	0.000	-0.1861371	-0.1735539
Municipality 7	0.1191200	0.0086300	13.8	0.000	0.1022053	0.1360347
Municipality 9	-0.0704264	0.0126139	-5.58	0.000	-0.0951495	-0.0457033

Municipality 10	0.1176381	0.0089547	13.14	0.000	0.1000870	0.1351892
Municipality 14	-0.0751529	0.0071786	-10.47	0.000	-0.0892228	-0.0610830
Municipality 15	0.3033861	0.0063328	47.91	0.000	0.2909739	0.3157983
Municipality 17	0.1376942	0.0088196	15.61	0.000	0.1204078	0.1549806
Municipality 18	0.3881708	0.0072364	53.64	0.000	0.3739875	0.4023541
Municipality 22	-0.0284743	0.0067312	-4.23	0.000	-0.0416672	-0.0152813
Municipality 24	0.2162091	0.0074378	29.07	0.000	0.2016311	0.2307870
Municipality 25	-0.0068407	0.0097729	-0.7	0.484	-0.0259954	0.0123140
Municipality 37	0.2083423	0.0093926	22.18	0.000	0.1899331	0.2267516
Municipality 39	0.0062385	0.0081393	0.77	0.443	-0.0097143	0.0221914
Municipality 40	-0.0760291	0.0098688	-7.7	0.000	-0.0953719	-0.0566864
Municipality 47	0.0740039	0.0083404	8.87	0.000	0.0576569	0.0903509
Municipality 48	-0.0244139	0.0073122	-3.34	0.001	-0.0387458	-0.0100821
Municipality 51	-0.0037996	0.0086635	-0.44	0.661	-0.0207798	0.0131807
Municipality 52	0.2974457	0.0085752	34.69	0.000	0.2806384	0.3142530
Municipality 53	-0.0353129	0.0078396	-4.5	0.000	-0.0506783	-0.0199474
Municipality 55	0.1050548	0.0074278	14.14	0.000	0.0904965	0.1196132
Municipality 56	0.2444073	0.0091619	26.68	0.000	0.2264501	0.2623645
Municipality 58	0.1010365	0.0074940	13.48	0.000	0.0863484	0.1157245
Municipality 59	0.1305763	0.0104010	12.55	0.000	0.1101905	0.1509622
Municipality 60	0.1897840	0.0084027	22.59	0.000	0.1733148	0.2062533
Municipality 63	0.4649167	0.0169942	27.36	0.000	0.4316083	0.4982251
Municipality 64	0.0275016	0.0161224	1.71	0.088	-0.0040980	0.0591012
Municipality 70	0.2889281	0.0104459	27.66	0.000	0.2684544	0.3094018
Municipality 72	0.0985440	0.0099710	9.88	0.000	0.0790011	0.1180869
Municipality 74	0.2197944	0.0107062	20.53	0.000	0.1988105	0.2407783
Municipality 79	0.2373980	0.0080485	29.5	0.000	0.2216231	0.2531729
Municipality 80	0.0915596	0.0098575	9.29	0.000	0.0722391	0.1108800
Municipality 81	0.1308313	0.0079482	16.46	0.000	0.1152529	0.1464097
Municipality 82	0.1574578	0.0096292	16.35	0.000	0.1385847	0.1763309
Municipality 83	0.3251608	0.0096036	33.86	0.000	0.3063380	0.3439836
Municipality 85	0.1281537	0.0122591	10.45	0.000	0.1041262	0.1521813
Municipality 86	0.1640631	0.0076231	21.52	0.000	0.1491220	0.1790042
Municipality 88	0.2540902	0.0080102	31.72	0.000	0.2383903	0.2697900
Municipality 90	0.6759617	0.0211680	31.93	0.000	0.6344728	0.7174507
Municipality 91	0.1911434	0.0076913	24.85	0.000	0.1760685	0.2062184
Municipality 98	0.2797916	0.0082084	34.09	0.000	0.2637032	0.2958800
Municipality 104	0.2010186	0.0093315	21.54	0.000	0.1827291	0.2193081
Municipality 106	0.2287512	0.0112739	20.29	0.000	0.2066546	0.2508478
Municipality 109	0.1264815	0.0061559	20.55	0.000	0.1144161	0.1385470
Municipality 114	0.0709174	0.0077125	9.2	0.000	0.0558011	0.0860337
Municipality 118	0.0862336	0.0062975	13.69	0.000	0.0738906	0.0985766
Municipality 119	0.1987820	0.0075139	26.46	0.000	0.1840549	0.2135091
Municipality 140	0.3543950	0.0080847	43.84	0.000	0.3385492	0.3702409
Municipality 653	0.1788021	0.0102954	17.37	0.000	0.1586233	0.1989809
Municipality 683	0.3007848	0.0103813	28.97	0.000	0.2804376	0.3211320

Municipality 710	0.3045062	0.0098730	30.84	0.000	0.2851553	0.3238571
Municipality 737	0.2012561	0.0125288	16.06	0.000	0.1766999	0.2258122
Municipality 765	0.1951178	0.0080500	24.24	0.000	0.1793398	0.2108957
Municipality 1651	-0.1006899	0.0093820	-10.73	0.000	-0.1190786	-0.0823013
Municipality 1661	-0.0583963	0.0084499	-6.91	0.000	-0.0749579	-0.0418347
Municipality 1663	-0.2193729	0.0135625	-16.18	0.000	-0.2459552	-0.1927907
Municipality 1680	-0.0063874	0.0101433	-0.63	0.529	-0.0262682	0.0134934
Municipality 1681	0.2098691	0.0078306	26.8	0.000	0.1945214	0.2252169
Municipality 1690	0.0204978	0.0093587	2.19	0.029	0.0021550	0.0388407
Municipality 1699	0.3064132	0.0086328	35.49	0.000	0.2894930	0.3233334
Municipality 1701	0.2544608	0.0072121	35.28	0.000	0.2403252	0.2685964
Municipality 1722	0.2676688	0.0091597	29.22	0.000	0.2497160	0.2856217
Municipality 1730	0.0634667	0.0124401	5.1	0.000	0.0390844	0.0878490
Municipality 1731	0.2924421	0.0074731	39.13	0.000	0.2777949	0.3070893
Municipality 1891	0.1744765	0.0070141	24.88	0.000	0.1607290	0.1882240
Municipality 1987	0.1491611	0.0096818	15.41	0.000	0.1301850	0.1681372
Unemployment	-0.0283226	0.0086690	-3.27	0.001	-0.0453138	-0.0113315
Train	-0.0013715	0.0000267	-51.41	0.000	-0.0014238	-0.0013192
Constant	-0.0479403	0.0016897	-28.37	0.000	-0.0512521	-0.0446284
N	122,908					
R^2	0.7924					
Adjusted R^2	0.7922					
Joint sign. F-test	4936.91***					

PGV 2SLS model (second stage)

	Coefficient	Robust SE	t	Sign.	95% Conf. Interval	
logPrice	-0.2188054	0.1082150	-2.02	0.043	-0.4309029	-0.0067079
logPGV	0.0481722	0.0151929	3.17	0.002	0.0183946	0.0779498
logArea	0.3913940	0.0660089	5.93	0.000	0.2620190	0.5207691
logLot	-0.0588985	0.0167707	-3.51	0.000	-0.0917685	-0.0260285
Rooms	-0.0238753	0.0077556	-3.08	0.002	-0.0390760	-0.0086746
Type corner	0.0331719	0.0103964	3.19	0.001	0.0127954	0.0535484
Type half double	0.1729169	0.0133492	12.95	0.000	0.1467528	0.1990809
Type detached	0.5949517	0.0220066	27.04	0.000	0.5518195	0.6380840
Built 1906-1930	-0.0139751	0.0195792	-0.71	0.475	-0.0523497	0.0243994
Built 1931-1944	0.0224200	0.0208693	1.07	0.283	-0.0184830	0.0633229
Built 1945-1959	-0.0636356	0.0221323	-2.88	0.004	-0.1070142	-0.0202570
Built 1960-1970	0.0323686	0.0203298	1.59	0.111	-0.0074770	0.0722142
Built 1971-1980	0.0325712	0.0189667	1.72	0.086	-0.0046028	0.0697451
Built 1981-1990	0.0445174	0.0196606	2.26	0.024	0.0059833	0.0830516
Built 1991-2001	0.1114089	0.0229408	4.86	0.000	0.0664458	0.1563720
Built >2001	0.1131675	0.0287051	3.94	0.000	0.0569066	0.1694284
Balcony	0.0765461	0.0145566	5.26	0.000	0.0480156	0.1050766
Garage	0.0660844	0.0125531	5.26	0.000	0.0414807	0.0906881

Attic	-0.0672714	0.0074935	-8.98	0.000	-0.0819583	-0.0525845
Loft	0.0312307	0.0105999	2.95	0.003	0.0104553	0.0520060
Barn	-0.0331103	0.0076010	-4.36	0.000	-0.0480079	-0.0182127
Fire	0.0470961	0.0165117	2.85	0.004	0.0147337	0.0794584
Maintenance	0.2801629	0.0202124	13.86	0.000	0.2405473	0.3197785
Density 500-1000	-0.0718774	0.0122434	-5.87	0.000	-0.0958740	-0.0478808
Density 1000-1500	-0.0835152	0.0144291	-5.79	0.000	-0.1117958	-0.0552347
Density 1500-2500	-0.0284188	0.0181405	-1.57	0.117	-0.0639736	0.0071360
Elderly	0.0256212	0.0329308	0.78	0.437	-0.0389219	0.0901643
Year 2004	-0.0385902	0.0076441	-5.05	0.000	-0.0535723	-0.0236081
Year 2005	0.1079790	0.0151171	7.14	0.000	0.0783500	0.1376079
Year 2006	0.1592411	0.0157664	10.10	0.000	0.1283395	0.1901428
Year 2007	0.1266404	0.0164737	7.69	0.000	0.0943525	0.1589283
Year 2008	0.1709351	0.0182020	9.39	0.000	0.1352598	0.2066104
Year 2009	0.2511903	0.0178289	14.09	0.000	0.2162463	0.2861343
Year 2010	0.5034985	0.0179341	28.07	0.000	0.4683483	0.5386488
Year 2011	0.5716572	0.0179057	31.93	0.000	0.5365627	0.6067517
Year 2012	0.4881945	0.0193071	25.29	0.000	0.4503532	0.5260358
Year 2013	0.6281239	0.0234134	26.83	0.000	0.5822346	0.6740133
Year 2014	0.6143451	0.0287822	21.34	0.000	0.5579330	0.6707571
Municipality 5	0.4861402	0.0279271	17.41	0.000	0.4314042	0.5408763
Municipality 7	-0.1819186	0.0592361	-3.07	0.002	-0.2980192	-0.0658180
Municipality 9	0.3152263	0.0672223	4.69	0.000	0.1834730	0.4469795
Municipality 10	-0.0800834	0.0581781	-1.38	0.169	-0.1941104	0.0339435
Municipality 14	0.1291445	0.0456870	2.83	0.005	0.0395996	0.2186893
Municipality 15	-0.3529127	0.0513011	-6.88	0.000	-0.4534609	-0.2523645
Municipality 17	-0.2381672	0.0618541	-3.85	0.000	-0.3593990	-0.1169355
Municipality 18	0.0915195	0.0635087	1.44	0.150	-0.0329553	0.2159942
Municipality 22	0.0298803	0.0434701	0.69	0.492	-0.0553195	0.1150801
Municipality 24	-0.1039312	0.0536932	-1.94	0.053	-0.2091680	0.0013056
Municipality 25	-0.1697557	0.0567080	-2.99	0.003	-0.2809012	-0.0586101
Municipality 37	-0.1320241	0.0701484	-1.88	0.060	-0.2695125	0.0054643
Municipality 39	0.1277764	0.0484558	2.64	0.008	0.0328047	0.2227480
Municipality 40	0.2814226	0.0602251	4.67	0.000	0.1633835	0.3994616
Municipality 47	-0.1951061	0.0512197	-3.81	0.000	-0.2954948	-0.0947174
Municipality 48	0.3054996	0.0452073	6.76	0.000	0.2168950	0.3941043
Municipality 51	-0.0659384	0.0550932	-1.20	0.231	-0.1739190	0.0420423
Municipality 52	0.0008923	0.0619112	0.01	0.989	-0.1204515	0.1222361
Municipality 53	0.3900249	0.0485207	8.04	0.000	0.2949262	0.4851237
Municipality 55	-0.0771660	0.0508696	-1.52	0.129	-0.1768687	0.0225366
Municipality 56	0.2392541	0.0614664	3.89	0.000	0.1187823	0.3597260
Municipality 58	-0.1881389	0.0507188	-3.71	0.000	-0.2875460	-0.0887319
Municipality 59	0.2107056	0.0598495	3.52	0.000	0.0934027	0.3280086
Municipality 60	0.1692589	0.0570830	2.97	0.003	0.0573783	0.2811395
Municipality 63	0.6264067	0.1144740	5.47	0.000	0.4020418	0.8507717
Municipality 64	0.2409784	0.0877072	2.75	0.006	0.0690754	0.4128814

Municipality 70	0.1579343	0.0757541	2.08	0.037	0.0094590	0.3064096
Municipality 72	0.2487118	0.0640779	3.88	0.000	0.1231215	0.3743022
Municipality 74	0.2468859	0.0658997	3.75	0.000	0.1177249	0.3760470
Municipality 79	-0.0159700	0.0569454	-0.28	0.779	-0.1275810	0.0956410
Municipality 80	0.2165826	0.0663988	3.26	0.001	0.0864433	0.3467219
Municipality 81	0.1147558	0.0510490	2.25	0.025	0.0147015	0.2148101
Municipality 82	0.0726853	0.0709400	1.02	0.306	-0.0663544	0.2117251
Municipality 83	0.4325202	0.0677308	6.39	0.000	0.2997702	0.5652701
Municipality 85	0.1328726	0.0682889	1.95	0.052	-0.0009712	0.2667163
Municipality 86	0.0417558	0.0508362	0.82	0.411	-0.0578813	0.1413930
Municipality 88	-0.1127006	0.0558999	-2.02	0.044	-0.2222623	-0.0031388
Municipality 90	0.4031672	0.1706530	2.36	0.018	0.0686935	0.7376409
Municipality 91	-0.0462370	0.0497870	-0.93	0.353	-0.1438176	0.0513436
Municipality 98	0.2258831	0.0596328	3.79	0.000	0.1090049	0.3427612
Municipality 104	0.1377098	0.0623410	2.21	0.027	0.0155237	0.2598959
Municipality 106	0.2642964	0.0715559	3.69	0.000	0.1240493	0.4045434
Municipality 109	-0.0063659	0.0417152	-0.15	0.879	-0.0881262	0.0753944
Municipality 114	0.1519236	0.0469360	3.24	0.001	0.0599306	0.2439165
Municipality 118	0.1537154	0.0403432	3.81	0.000	0.0746442	0.2327866
Municipality 119	0.1007902	0.0525189	1.92	0.055	-0.0021450	0.2037254
Municipality 140	0.2564253	0.0650653	3.94	0.000	0.1288997	0.3839509
Municipality 653	0.1103349	0.0712227	1.55	0.121	-0.0292590	0.2499288
Municipality 683	0.1764479	0.0732167	2.41	0.016	0.0329458	0.3199501
Municipality 710	0.1673923	0.0678964	2.47	0.014	0.0343178	0.3004668
Municipality 737	0.2374197	0.0718021	3.31	0.001	0.0966902	0.3781491
Municipality 765	0.2494479	0.0550864	4.53	0.000	0.1414805	0.3574153
Municipality 1651	0.3282507	0.0583960	5.62	0.000	0.2137965	0.4427048
Municipality 1661	-0.1266134	0.0532466	-2.38	0.017	-0.2309747	-0.0222520
Municipality 1663	0.2937319	0.0786394	3.74	0.000	0.1396015	0.4478622
Municipality 1680	0.0919818	0.0607026	1.52	0.130	-0.0269930	0.2109566
Municipality 1681	0.1871913	0.0507358	3.69	0.000	0.0877509	0.2866317
Municipality 1690	0.1679821	0.0458893	3.66	0.000	0.0780407	0.2579236
Municipality 1699	0.0264390	0.0653472	0.40	0.686	-0.1016391	0.1545170
Municipality 1701	-0.0219346	0.0515689	-0.43	0.671	-0.1230078	0.0791386
Municipality 1722	0.1822280	0.0600246	3.04	0.002	0.0645819	0.2998741
Municipality 1730	0.2895972	0.0770531	3.76	0.000	0.1385759	0.4406184
Municipality 1731	-0.0656819	0.0531569	-1.24	0.217	-0.1698676	0.0385038
Municipality 1891	0.1226803	0.0486311	2.52	0.012	0.0273652	0.2179954
Municipality 1987	0.0215345	0.0661871	0.33	0.745	-0.1081898	0.1512588
Constant	0.0963069	0.0539440	1.79	0.074	-0.0094213	0.2020351
N	122,908					
R^2	0.1156					
Wald Chi^2	16.439.28***					

Appendix IV: Logbook

Total hours: 832

Date	Activity	Outcome	Hours
10-11-2016	Orientation by means of reading course description and last year theses. Construct pre-proposal.	-	2
15-11-2016	Orientation by means of reading course description and last year theses. Construct pre-proposal.	-	3
22-11-2016	Orientation by means of reading course description and last year theses. Construct pre-proposal.	-	2
24-11-2016	Formulating research themes, identifying options for combining thesis with an internship and constructing pre-proposal.	Pre-proposal.	2
25-11-2016	Start meeting Master thesis course, discussed possible topic and time planning. Decided to postpone Master thesis until Semester 2B.	-	2
25-4-2017	Discussing possible Master thesis topics based on student assistentship on mortgage problems and earthquakes in Groningen.	-	1
4-5-2017	Scanning literature for potential topics on earthquakes and the housing market.	-	2
12-5-2017	Searching for literature. Writing research proposal.	-	2
16-5-2017	Telephone meeting with George de Kam on possibilities of using the data from the mortgage problems and earthquakes research. The hours on my student assistantship will also be included in this logbook. Working on research proposal.	-	3
23-5-2017	Skype meeting with George de Kam, decided that I will have a design for my Master thesis and have a clearer idea of what I will do with his dataset the next week.	-	2
26-5-2017	Develop research design.	-	2
28-5-2017	Read literature on earthquakes and the housing market. Develop research design, found out that previous idea was to general, need to find a more specialized topic. Talk about it tomorrow with George de Kam.	-	4
30-5-2017	Skype meeting with George de Kam. Suggested possible structure of Master thesis: first theoretical exploration, second statistical analysis using Qualtrics survey, third interviews.	-	3
2-6-2017	Search literature on housing market of Groningen and effect of earthquakes. Read literature.	-	6
6-6-2017	Search literature on housing market of Groningen and effect of earthquakes. Read literature.	-	6
9-6-2017	Creating literature overview. First mainly focusing on reports written on the housing market in Groningen.	-	6
13-6-2017	Reading reports on housing market in Groningen.	-	7
14-6-2017	Reading reports on housing market in Groningen. Working on literature overview.	-	5
16-6-2017	Reading Harvey and Jowsey (2004), *Urban Land Economics.*	-	4
19-6-2017	Reading Harvey and Jowsey (2004).	-	6
26-6-2017	Finalizing reading Harvey and Jowsey (2004).	-	4
27-6-2017	Meeting with George de Kam, decided to talk about topic thesis in week of 17 July.	-	0.5

18-7-2017	Meeting with the GBB in Loppersum, together with George de Kam. Decided to focus my thesis on an important sub problem following from the broad analysis of the questionnaire.	-	3
4-9-2017	Searching relevant literature in the library.	-	3
7-9-2017	Working on research proposal.	-	3
12-9-2017	Meeting with George de Kam, determined the final topic: relationship between the selling process, earthquakes, mortgage problems and the financial situation. Talking about possible data sources (CBS, Rabobank) and discussing possibilities of doing an internship at the Rabobank.	-	1
14-9-2017	Searching relevant literature on the selling process.	-	8
15-9-2017	Creating structure for thesis.	-	8
18-9-2017	Reading literature on mortgages.	-	4
19-9-2017	Meeting with George de Kam, further brainstormed on the topic, determined the X and Y. Starting to work on realizing internship at Rabobank.	-	4
20-9-2017	Reading literature on mortgages.	-	3
21-9-2017	Working on research proposal, sent it to George de Kam for feedback.	Research proposal (version 1)	8
22-9-2017	Read the feedback on the first version of my research proposal. Used it to improve my research proposal.	-	5
24-9-2017	Finalize research proposal and sent it to George de Kam.	Research proposal (version 2)	2
25-9-2017	Reading relevant literature on the real estate market.	-	5
26-9-2017	Creating relevant themes and structure for theoretical framework. Reading relevant literature on the real estate market.	-	5
27-9-2017	Reading literature on the selling process.	-	3
28-9-2017	Reading literature on the selling process.	-	7
3-10-2017	Meeting with George de Kam, received positive reaction from Rabobank, waiting for their proposal for a meeting.	-	0.5
6-10-2017	Reading literature on the selling process.	-	8
11-10-2017	Meeting with Karl Pladdet and Marieke Siertsema from the Rabobank about possibilities for using their data, George de Kam was also present. Reading literature on the selling process.	-	4
12-10-2017	Constructing email to Rabobank to clarify my research and data that I need. Reading literature on the selling process.	-	8
17-10-2017	Meeting with George de Kam. Decided that I will also make a time planning and sharpen my research proposal, currently I made a research proposal focused on the Rabobank. We also talked about the situation if it will not be possible to use the data from the Rabobank. In that case, we will use CBS microdata which we probably also need when we do have Rabobank data. Therefore, I will start looking who to contact to gain access to the CBS microdata. Finally, we decided that we need an extra supervisor that can support me with the statistical models.	-	0.5
19-10-2017	Rewriting my research proposal.	-	4
20-10-2017	Rewriting research proposal. Contacting people for CBS microdata. Start with writing theoretical framework.	-	7

23-10-2017	Contacting people for CBS microdata. Rewriting research proposal.	Research proposal (version 3)	1
24-10-2017	Contacting people for CBS microdata.	-	0.5
26-10-2017	Writing theoretical framework.	-	5
31-10-2017	Rewriting Research proposal.	Research proposal (version 4).	2
2-11-2017	Meeting with George de Kam. Contacted Christian Lennartz from Rabobank Research on the internship. Contacted Arno van der Vlist about a additional supervisor focusing on the statistical analyses. Rewriting research proposal.	Research proposal (version 5)	4
14-11-2017	Meeting with George de Kam, spoke to Christian Lennartz on the telephone. I still have to write a research proposal in English with a more	-	1
15-11-2017	Working on Research Proposal	-	7
16-11-2017	Working on Research Proposal	-	6
17-11-2017	Working on Research Proposal	-	5
20-11-2017	Working on Research Proposal	-	6
21-11-2017	Meeting with George de Kam, Nicolas Duran and Paul Elhorst. Decided that Elhorst will be my second supervisor.	-	1
22-11-2017	Finalizing research proposal. Creating structure for introduction and theoretical framework.	Research proposal (version 6)	7
23-11-2017	Working on theoretical framework, subchapter on the selling process.	-	7
27-11-2017	Working on theoretical framework, subchapter on the selling process. Read and process feedback on research proposal.	Research proposal (version 7)	6
29-11-2017	Working on theoretical framework, subchapter on the selling process.	-	8
1-12-2017	Working on theoretical framework, subchapter on the selling process.	-	6
4-12-2017	Working on theoretical framework, subchapter on listing price, selling price and TOM.	-	7
5-12-2017	Working on theoretical framework, subchapter on listing price, selling price and TOM.	-	7
6-12-2017	Working on theoretical framework, subchapter on seller motivation and characteristics.	-	6
7-12-2017	Working on theoretical framework, subchapter on seller motivation and characteristics.	-	8
11-12-2017	Working on theoretical framework, subchapter on housing and market characteristics.	-	6
12-12-2017	Working on theoretical framework, subchapter on housing and market characteristics.	-	8
13-12-2017	Working on theoretical framework, subchapter on housing and market characteristics.	-	8
15-12-2017	Working on theoretical framework, subchapter on earthquakes and the housing market.	-	7
18-12-2017	Working on theoretical framework, subchapter on earthquakes and the housing market.	-	8
19-12-2017	Working on theoretical framework, subchapter on earthquakes and the housing market.	-	8
20-12-2017	Working on theoretical framework, subchapter on earthquakes and the housing market.	-	8

21-12-2017	Working on theoretical framework, subchapters on conceptual model and hypotheses.	-	8
22-12-2017	Finalizing preliminary version theoretical framework.	Preliminary theoretical framework (version 1)	6
5-1-2018	Registration Graduate Research Day (GRD).	-	1
8-1-2018	Reading methodologies of recent real estate studies theses. Searching relevant CBS data.	-	8
9-1-2018	Searching relevant CBS data. Reading feedback on preliminary theoretical framework. Writing introduction.	-	7
10-1-2018	Writing introduction.	-	8
11-1-2018	Presentation workshop GRD.	-	2
12-1-2018	Reading feedback on preliminary theoretical framework and preparing points that need to be discussed during the meeting. Meeting met George de Kam. We discussed his feedback on my theoretical framework. General impression was good, it provides a solid foundation for the research. He had several suggestions for adding some articles or discussing others in more detail. Furthermore, I have to evaluate the importance of certain parts and possible shorten them. We also discussed the analysis, we try to do an internship at the Rabobank; however, there might be problems with connecting the NVM data to the Rabobank data. Therefore, the backup-plan is to use NVM data via the University of Groningen together with data from an earlier research of George. Finally, we discussed how to operationalize the effect of earthquakes into variables.	-	1.5
15-1-2018	Preparing presentation for GRD. Emailing Rabobank about status internship. Discovering database George de Kam (called De Kam & Mey data from now on), data preparation.	-	6
16-1-2018	Writing introduction, fully finalize it after empirical analysis is done. Adapting conceptual model. Preparing presentation GRD.	Preliminary introduction (version 1)	7
18-1-2018	Data preparation De Kam & Mey data.	-	3
19-1-2018	Data preparation De Kam & Mey data. Preparing presentation GRD.	-	6
22-1-2018	Preparing presentation for GRD. Writing methodology section. Making map for introduction. Meeting with George de Kam.	-	5
23-1-2018	Data preparation De Kam & Mey data. Preparing presentation GRD.	-	4
24-1-2018	Preparing presentation GRD.	-	2
25-1-2018	Graduate Research Day	Presentation GRD.	1.5
30-1-2018	Meeting with Paul Elhorst, giving me access to NVM data. We discussed my empirical strategy and mainly the problem of the trade-off and simultaneity between TOM and sale price.	-	1
1-2-2018	Exploring the NVM data.	-	4
6-2-2018	Empirical analysis NVM data.	-	6
7-2-2018	Empirical analysis NVM data.	-	6
8-2-2018	Empirical analysis NVM data.	-	6
12-2-2018	Empirical analysis NVM data.	-	8
13-2-2018	Empirical analysis NVM data.	-	8
14-2-2018	Empirical analysis NVM data.	-	8
15-2-2018	Empirical analysis NVM data.	-	7

19-2-2018	Empirical analysis NVM data.	-	8
20-2-2018	Empirical analysis NVM data.	-	6
21-2-2018	Empirical analysis NVM data.	-	8
22-2-2018	Empirical analysis NVM data.	-	7
23-2-2018	Empirical analysis NVM data.	-	5
26-2-2018	Empirical analysis NVM data. Rewriting introduction.	-	8
27-2-2018	Meeting with Paul Elhorst. We discussed my progress with the statistical modeling. The main problems are related to the endogeneity between the earthquake indicators and the fixed effects.	-	2
28-2-2018	Empirical analysis NVM data. Rewriting and finalizing introduction.	Preliminary introduction (version 2).	5
3-3-2018	Rewriting theoretical framework.	-	8
4-3-2018	Rewriting theoretical framework.	-	5
5-3-2018	Rewriting theoretical framework.	-	8
6-3-2018	Empirical analysis	-	7
7-3-2018	Empirical analysis	-	7
8-3-2018	Visited the 'afstudeerkring' from Frans Sijtsma with mainly real estate Master students to gain some feedback. Got the suggestion to work with time series; however, the switch to a new approach and new model would require a large time investment.	-	1
9-3-2018	Empirical analysis	-	8
10-3-2018	Empirical analysis and exploratory analysis.	-	7
11-3-2018	Empirical analysis. Rewriting and finalizing theoretical framework.	Preliminary theoretical framework (version 2)	5
12-3-2018	Writing methodology.	-	7
13-3-2018	Writing methodology.	-	8
14-3-2018	Writing methodology and results.	-	7
15-3-2018	Empirical analysis	-	7
18-3-2018	Empirical analysis and prepare next meeting with supervisors.	-	5
19-3-2028	Prepare next meeting with supervisors.	-	2
20-3-2018	Empirical analysis. Meeting with George de Kam and Paul Elhorst. We discussed the models and the time planning. Three issues to be improved with the models: risk and reference areas, instruments for the 2SLS regression and the cutoff value for the PGV variable. Set a deadline of 6 April for the first concept version.	-	8
21-3-2018	Empirical analysis.	-	7
22-3-2018	Empirical analysis.	-	7
23-3-2018	Writing methodology and results.	-	8
26-3-2018	Empirical analysis.	-	8
27-3-2018	Empirical analysis. Writing methodology. Visited Elhorst for information on the DID model.	-	8
28-3-2018	Empirical analysis. Writing methodology.	-	8
29-3-2018	Writing methodology.	-	8
30-3-2018	Empirical analysis. Writing methodology.	Preliminary methodology.	6
31-3-2018	Writing results.	-	9

2-4-2018	Writing results	-	8
3-4-2018	Writing results	-	8
4-4-2018	Writing results	-	8
5-4-2018	Writing results.	Preliminary results.	7
9-4-2018	Writing conclusion.	-	8
10-4-2018	Writing conclusion. Finalizing concept version thesis.	Preliminary conclusion. Concept version thesis.	8
17-4-2018	Meeting with George de Kam and Paul Elhorst on the concept version of my thesis. We discussed their remarks in the text and the more general improvements that can be made. The latter are related to being more critical on the discussed theory. The results section can be clearer and needs to be reorganized slightly. Furthermore, the societal and academic relevance should be highlighted more clearly. The foundation of the thesis, being the statistical models, was good. We decided that I would need about two weeks to finish the thesis.	-	3
18-4-2018	Rewriting thesis.	-	8
19-4-2018	Rewriting thesis.	-	8
20-4-2018	Rewriting thesis.	-	8
22-4-2018	Rewriting thesis.	-	8
23-4-2018	Rewriting thesis.	-	8
24-4-2018	Rewriting thesis.	-	8
25-4-2018	Rewriting thesis.	-	8
26-4-2018	Rewriting thesis.	-	8
28-4-2018	Rewriting thesis.	-	8
29-4-2018	Rewriting thesis.	-	8
30-4-2018	Rewriting thesis. Meeting with Paul Elhorst on some final issues with the statistical models.	-	8
2-5-2018	Finalizing thesis.	Final version thesis (1).	8
12-5-2018	Finalizing thesis.	Final version thesis (2).	6