Thesis

STREEM -VU

(June 2024)

Grounds for Concern:

The Price of Risk in Residential Foundations

Thijmen van Hessen (2715610)

June 29, 2024

Supervisor: prof. dr. Jos van Ommeren

VU VRIJE UNIVERSITEIT AMSTERDAM

Acknowledgements

I would like to express my gratitude to Professor Jos van Ommeren, my thesis supervisor, for his invaluable advice, guidance, and support. Special thanks to Yashvant Premchand, PhD Researcher at the Department of Spatial Economics, for his assistance with questions, data, and general guidance. I am also deeply grateful to my family, friends, and loved ones for their unwavering support throughout this journey. Thank you all for your unique contributions.

ABSTRACT

This thesis investigates the impact of foundation risks on housing prices in the Dutch real estate market. Despite the significant prevalence of foundation-related issues affecting between 750,000 and 1,000,000 homes, the market may not fully account for these risks in property valuations. The study utilizes a comprehensive analysis of approximately 460,000 transactions across ten Dutch municipalities, employing hedonic pricing models and difference-in-differences regressions to assess the influence of foundation risks while controlling for spatial heterogeneity.

Key findings reveal that foundation risks are not consistently reflected in housing prices. Exceptions include properties on peat soils with higher differential settlement risks, showing a minor price decrement, and the interaction between pole rot risk and wooden foundations. However, these effects are modest and highlight the complexity of accurately capturing foundation risks in property values. Additionally, properties on clay soils command a premium, whereas peat soils are associated with lower prices. Substantial subsidence also negatively impacts property values.

The study underscores the need for enhanced market transparency through policy interventions, such as mandatory foundation inspections and standardized assessments, to improve property valuations. By providing empirical evidence on the economic relationship between foundation risks and home values, this research contributes to better understanding and mitigating these risks, ultimately benefiting homeowners, policymakers, and financial institutions.

TABLE OF CONTENTS

1. Introduction

1.1 Context and Research Question

For most, a home is more than just bricks and mortar; it's a cornerstone of stability and comfort. Yet, this foundational security is under threat; as droughts intensify, the environment changes and all over the world urban centers increase not only in size but also in density the structural foundations of many homes are at risk. Foundation damages are caused by external environmental factors such as changes in groundwater levels and subsidence rates, and they represent a critical yet complex challenge in urban development. These risks are influenced by various factors including the foundation type, the weight of the building, and the soil characteristics and effects are unnoticed due to their underground nature. Such complexities make foundation risks a significantly understudied and unacknowledged issue, despite their potential to severely impact property values and urban livability (Costa, Kok & Koff, 2020; Kok & Angelova; 2020). As these risks go unnoticed, foundation damages can lead to a decline in property quality, ultimately resulting in property devaluation and/or increased maintenance costs.

In the Netherlands, climate change complicates matters even further. Increased droughts such as those in 2018 have caused an alarming number of foundation damage notifications even in areas where damages were not expected (KCAF, 2024). Regions such as the Netherlands, characterized by a dense urban development and unique hydrological and geomorphological conditions (Schothorst, 1977), cope with higher foundation risks (Costa, Kok & Koff, 2020; Kok, van der Putten & Kraus, 2021; Leusink, 2018; Willemsen, Kok & Kuik, 2020). A first-of-its kind study by Deltares has highlighted that foundation damages pose a significant risk for the Netherlands, potentially affecting between 750,000 and 1,000,000 (9% to 12%) houses (Kok & Angelova, 2020). Recent studies project that by 2050, costs for residential real estate in the Netherlands due to foundational damages could reach from 20 to 60 billion EUR, emphasizing the urgency of this issue (Hommes et al., 2023; Kok & Angelova, 2020; Kok, van der Putten & Kraus, 2021). The economic ramifications of foundation damages are profound, with repair costs ranging between €50,000 and €100,000 per house, which constitutes 10% to 30% of total house values (Klaassens, 2015). The implications extend beyond direct repair costs, influencing property devaluation and leading to increased insurance premiums.

Economic theory suggests that homes with expected foundation repair costs have a lower value than similar homes where this is not the case. Houses expected to require foundation repairs typically would sell for less than comparable properties without such issues, due to the detriment of anticipated repair expenses. When the disparity in price between a property with foundation risks and a similar one without matches the expense of repairing the foundation, the anticipated costs of damage are entirely priced in. (Hommes et al., 2023) However, where most housing defects can be seen and their impact on house prices measured, foundation problems remain hidden under the ground. Unfortunately, making an accurate assessment of the foundation's quality or forecasting future foundation problems is complicated. This complexity raises the question of whether considerations of potential foundation difficulties are included in the decision-making process for homebuyers.

An additional reason for the current information gap is that a foundation report is not mandatory for selling a home and considering the high cost of such reports, coupled with the potential for a negative impact on sale prices if foundation issues are disclosed (Hommes et al, 2023), many sellers are deterred from investigating the state of their property's foundation. This leads to a lack of transparency and disclosure regarding foundation conditions in most property listings, potentially causing large housing market inefficiencies (RLI, 2024). The complexity of foundation issues, combined with the financial risks they pose to buyers, can result in pronounced market distortions, and have broader economic consequences. Furthermore, the scarcity of foundation quality data complicates the study and analysis of foundation-related problems. This information gap and its potential significant impacts on the housing market have prompted the following research question:

"What is the effect of foundation risks on housing prices?"

Answering this question will provide insights into the relationship between foundation risks and property values. Given the significance of foundation risks, it is crucial for stakeholders, including policymakers, financial institutions, and property owners, to understand and address these challenges proactively. The development and implementation of mitigation and adaptation measures, informed by comprehensive economic risk assessments, are essential for safeguarding the stability of the housing market and ensuring the long-term viability of residential real estate investments and public welfare.

1.2 Contribution and Outline

This study contributes to the literature in several ways. First, it highlights the significant impact of information gaps in the housing market, particularly regarding foundation risks. It demonstrates that these risks are often unknown to buyers and unaccounted for in decision-making, illuminating a critical market distortion and underscoring the need for policy interventions to enhance information dissemination and market efficiency.

Second, the research aims to provide empirical evidence on the economic relationship between foundation risks on home values using the hedonic pricing model. A hedonic pricing method, applied within the Dutch real estate context, is applied to estimate the economic value attached to foundation risks. Hedonic pricing methods work best when the market is fully efficient, and all information is available to everyone (Chau & Chin, 2003). However, it has been established that foundation risks are often unknown. Therefore, contrary to most studies that seek to confirm significant effects, this research aims to demonstrate that although foundation risks are present, they do not impact prices due to a lack of foundation knowledge of both buyers and sellers. The hypothesis of the research is therefore that prices are not reflecting foundation risks. An important note is that proving the absence of an effect is impossible in statistics. Instead, it is only possible to determine if there is sufficient evidence to reject the null hypothesis.

In addition to using foundation risk scores, this research incorporates the use of additional environmental risk drivers that contribute to these risks to account for the complex and diverse process of foundation deterioration. A cross-sectional dataset, which includes sale prices and indicators of foundation risks, will be used to prove the hypothesis. This evidence is valuable for policymakers and researchers who aim to work on this common yet underacknowledged problem.

Third, this research extends the application of hedonic pricing theory to include foundation risks, subsidence and soil types, an area previously underexplored in the literature. By applying this method to assess the impact of foundation risks and its drivers on prices, the study not only showcases the versatility and utility of hedonic pricing models in capturing the economic value of non-market attributes but also enhances the methodological toolkit available to researchers and policymakers for evaluating economic impacts of environmental characteristics.

After this introductory section, the paper proceeds to outline the theoretical groundwork in section 2, incorporating discussions and existing literature on the origins and characteristics of foundation risks and frameworks for evaluating the economic impacts of housing characteristics. In section 3, the data will be described, along with the introduction of a comprehensive empirical framework and strategy specific to foundation risks. Section 4 consists of the results and extra analysis while section 5 presents a discussion of the results, highlighting further research directions and policy implications concerning the economic evaluation of foundation risks.

2. Literature Review

2.1 Foundations and the Dutch Housing Market

The Netherlands' unique landscape poses significant challenges to building foundations due to the intensive artificial drainage of peat and clay soils, historically leading to land subsidence from peat oxidation and the consolidation of clay particles (Schothorst, 1977). Contrary to popular belief, the country's current low-lying state is largely a result of human activities such as land drainage. As a result, continued subsidence is expected and common as clay and peat soils settle following drainage and drought periods. At the same time, projected climate change is likely to increase both the severity and frequency of these conditions (RLI, 2024; Kok, van der Putten & Kraus, 2021).

In response to these conditions, a variety of foundation types have been developed, tailored to different soil types, building loads, and regulatory requirements. Historically, up until the 1970s, shallow and wooden pile foundations were commonly employed. Shallow foundations are placed directly on the load-bearing soil layer, sometimes enhanced to increase its bearing capacity. This method is cost-effective and suitable for lighter structures where the load-supporting soil layer is near the surface. Wooden and concrete pile foundations, where piles are driven or screwed into the ground to reach a stable soil layer, are used for heavier structures or where the weak soil layers are thicker (Kok & Angelova, 2020). Since 1970, concrete pile foundations have become the standard due to their durability and safety. Currently, researchers estimate that approximately 70% (5 million) of all properties, including non-housing, have shallow foundations, while 5-6% (400,000) utilize wooden foundations, and the remainder are supported by concrete piles (RLI, 2024).

Wooden pile foundations, particularly those on clay or peat substrates, are vulnerable to both differential settlement, and pole rot (bacterial degradation and fungal decay) if the timber dries out. Such damage often only becomes evident after decades, as the degradation processes are slow (Kok & Angelova, 2020). Pole rot[1] can lead to a loss of load-bearing capacity after an average cumulative dry period of 10-20 years, causing buildings to sink. Damage to buildings with shallow foundations can especially occur if the building does not settle uniformly but tilts.

[1] Pole and piles can be used interchangeably, from this point onward the term pole rot will be used for clarity.

This can be caused by variations in subsidence rates, soil composition or groundwater levels, and irregular property weight dispersion.

The unique geological conditions and evolving climate of the Netherlands create a dynamic and often challenging environment for maintaining building foundations, which directly influences real estate practices and property values. Given the country's centuries-long experience with water and soil-related risks, Dutch regulations require homeowners to disclose any significant structural information about their houses. This legal mandate ensures transparency in the housing market, affecting property prices by disclosing defects and promoting repairs.

The study by Hommes et al. (2023) delves into the prevalence of disclosures and repairs related to foundations in property listings and examines the implications of these practices on market perceptions and real estate values. The authors examined the prevalence of property advertisements disclosing the state of foundations and the impact of such disclosures. They discovered that only 2.2% of properties constructed before 1975 include information about their foundations, and even then, the condition of the foundations remains ambiguous in half of these cases. Properties identified as having 'bad' foundations exhibited a 12% decrease in value (€47,000) compared to expected prices without foundation damage. Conversely, disclosures of repaired foundations led to an average price increase of only 2% (€13,500). Thus, the average price disparity amounts to over €60,000, likely reflecting the costs of foundation renovations, which are estimated to range from €50,000 to €100,000. The authors analyzed how foundation issues affect home prices, finding that the impact is smallest for apartments (~10% depreciation) and slightly greater for other home types. Due to higher valuations, the same percentage decrease translates into a larger absolute loss in value for detached homes. Conversely, the price increase from repaired foundations is higher for detached homes (~6%) compared to other house types (~2%). This difference is likely because repair costs for detached homes are covered by fewer owners, whereas for apartments and other connected homes, costs are shared among multiple owners, reducing the per-homeowner expense. Overall, Hommes et al. (2023) confirm that foundation issues are factored into prices when their condition is disclosed.

Adding to these findings, Clayton (1997) highlight the cognitive biases that further complicate market dynamics, especially in assessing risks like those associated with foundations. Clayton

(1997) points out that rational price expectations often diverge from actual outcomes, reflecting the imperfect nature of the housing market. This imperfection is partly due to human behavior as Salzman and Zwinkels describe (2017), where over-optimism and over-confidence skew buyers' perceptions of risk. This cognitive bias suggests that even when foundation issues are known and disclosed, the actual impact on property values might be underappreciated by the market. Sellers might choose to underreport or omit foundation issues due to fear of significant devaluation, while buyers are overconfident and look away from potential problems. The lack of transparency and communication exacerbates the problem of foundation risk underpricing.

2.2 Hedonic Pricing

The research on housing markets and measuring housing preferences is extensive. Kok and Costa (2021) performed a literature review and highlight significant gaps in the research on the economic cost assessment of subsidence, noting the absence of a standardized framework which hampers effective decision-making and knowledge sharing across studies. Despite the complexities and potential for market distortions, hedonic pricing stands out for its ability to derive value estimates directly from market behaviors and real transactions.

The price of a house is often determined by a combination of factors and the marginal willingness to pay for those attributes. The quantitative effects of these characteristics are often measured through hedonic pricing models. The method is based on the fact that the price of a house often consists of a combination of structural (e.g. size, quality and construction year), locational (e.g. nearby parking facilities, greenspace, distance to highway) and neighborhood (e.g. average neighborhood income, age, and education level) attributes. The hedonic price function can be described as the equilibrium pricing of different varieties of a heterogeneous good, influenced by the dynamics of supply and demand. The relative price of a house results from the combined effect of its different attributes and on the property market demand and supply determine the characteristics' marginal contributions to the total value of the bundle of characteristics. The values of these attributes, estimated through regression analysis, together determine the house's position in the housing market.

The basic form of a hedonic pricing model, as conceptualized by Rosen (1974), posits that the price of a house can be expressed as a function of its property attributes:

$$P = P(X)$$

Here, X represents a vector of housing characteristics, which includes structural variables such as size and age, as well as locational variables like proximity to the city center and highways, all of which collectively describe the quality of the housing. The model can further be extended to encompass neighborhood characteristics such as average income, crime levels, and school quality. In this context, household utility U would then be a function of both the consumption of housing X and a composite good C:

$$U = u(X, C)$$

Each consumer selects values of X and C to maximize their utility, subject to a budget constraint M, where M equals the total expenditure on housing and composite goods, and assuming the unit price of the composite good is 1:

$$M = P(X) + C$$

To solve this a Lagrangian function is applied, which then gives the first order condition:

$$\frac{\frac{\partial U}{\partial X}}{\frac{\partial U}{\partial C}} = \frac{\partial P(X)}{\partial X}$$

This equation indicates that the change in utility from a marginal increase in X is equivalent to the change in property price for the same increase in X, assuming all other factors remain constant. In essence, this condition shows that the marginal willingness to pay (MWTP) for a specific housing quality X is equal to the first derivative of the price with respect to that quality. It implies that the rate at which consumers are willing to trade housing quality for a composite good, holding utility constant, is reflected in the slope of the hedonic price function.

The hedonic method was pioneered by Court (1939) to model the prices of cars as a function of their characteristics. The method was then popularized by Griliches (1961) and Rosen (1974) formalized the method by linking the method with microeconomic supply and demand theories. The method has since then been adapted in many research areas and has been used extensively for scientific studies on multidimensional commodities such as housing. Ridker and Henning (1967) were the first researchers who used the method for residential properties by measuring the relationship between air quality and home values. Nevertheless, Freeman (1979) was the first who also supplied the theoretical justification for hedonic approach, using the hedonic price equations to calculate marginal implicit prices and the marginal willingness to pay for environmental benefits. Researchers have since then used the method to infer the effects of property attributes such as age (Goodman & Thibodeau, 1997; Rubin, 1993), location (Heyman, Law & Berghauser, 2018; Kiel & Zabel, 2008) size (Kagie & Wezel, 2007), open space (Brander & Koetse, 2011; Sander & Polansky, 2009), and green space (Czembrowski & Kronenberg, 2016; Morancho, 2003; Panduro & Veie, 2013). According to Sirmans, Macpherson and Zietz (2005) the most studied characteristics in hedonic pricing are age, size, garage space, fireplaces and air conditioning. Malpezi (2003) denotes that the most common variables in hedonic price analyses are; number of rooms, floor area, house type category, availability of heating and cooling, age, structural features, and structural materials used and their respective quality. Although there has been a lot of discussions on the implicit value of structural characteristics of housing (16 studies in 2010) most papers focus on neighborhood characteristics (178 papers in 2010) (Herath & Maier, 2010). This paper will aim to use both neighborhood risk levels and individual property risk as estimates for house prices.

There are also papers using the hedonic estimation to investigate the effects of structural attributes by looking at depreciation and maintenance (Armengot, Williams & Padial, 2021; Billings, 2015; Francke & van de Minnde, 2017; Harding, Rosenthal & Sirmans, 2007; Wilhelmsson, 2008). A common proxy measure for depreciation in hedonic pricing models is property age. Through an extensive literature review Malpezzi et al. (1987) showed that on average home prices decrease with age at a declining rate although they vary widely between studies based on method, region and time periods. Billings (2015) contributes to the discussion by noting that structural renovations in housing can introduce a bias in hedonic models designed to assess attribute impacts on housing values. Including detailed renovation data from Charlotte, North Carolina, revealed this bias, which tends to skew hedonic indices positively. However, the author considers this renovation bias to be minor, impacting research outcomes only slightly and categorizing it as a secondary concern in research design. Wilhelmsson (2008) studied the effects of maintenance on housing values in Stockholm and found a price difference of 13% between well maintained houses and those that were not. Using a hedonic analysis, Francke & van de Minne (2017) observed that after half a century of neglecting residential housing maintenance in ‘s Hertogenbosch, a depreciation of 48% was noted, indicating an annual decrease of around 1%. At the same time, Harding, Rosenthal and Sirmans (2007) used a repeated sales model to find that without maintenance houses depreciate on average 2.5% annually while maintained houses depreciate only 2% per year. Armengot, Williams and Padial (2021) show that the relationship between a building's depreciation and its maintenance requirements is not straightforward. Structures from various time periods and construction eras have distinct maintenance needs. This variation aligns with the different foundation types used in various eras, which likely results in qualitative differences. These findings corroborate that age is an important factor to consider when analyzing structural characteristics such as foundations.

Hedonic pricing methods have also been used to assess the impact of environmental processes such as floods (Bui, Wen & Sharp, 2022), earthquakes (Koster & van Ommeren, 2015), and wildfires (Donovan, Champ & Butry, 2007) on property values. Donovan, Champ and Butry (2007) stated that when publishing wildfire risk ratings, the increased awareness had negative impacts on prices. Koster and van Ommeren (2015) found that even with full compensation of damages house prices were still lower, which the authors interpret as being an indicator for discomfort in the future.

While the hedonic pricing literature is extensive, empirical research on the effects of foundation damages and/or risks on house prices is scarce. The most important scientific papers on using hedonic pricing to discover the impact of foundation risks and subsidence are by Willemsen, Kok & Kuik (2020), and Yoo and Perring (2017). Yoo and Perring (2017) find that subsidence due to water depletion leads to a lowering of property values between 5% to 11% for homes that are located close to land subsidence features in Arizona (USA). Willemsen, Kok & Kuik (2020) were the first to have used hedonic pricing analysis for subsidence risks in the Netherlands and found a negative and significant effect (-6% and -7%) of (uniform) subsidence on real estate values. However, critique can be issued over their data generalizability, dealing with omitted variables, and reliance on historical data without addressing future risks comprehensively.

While the economic effects of subsidence have been acknowledged, the relationship of different soil types with house prices is unknown as no previous high-quality studies were found in the literature. Westerveld and van den Hurk (1973) do recognize that soil types played a large role in urban planning and development in the Netherlands. They also denote that the rapid increase of cities in the 20th century required many previously unsuitable lands to be improved through adding sand layers on top of soil and peat soils. Unfortunately, the exact relationships between soil types, city development and house prices remain an understudied phenomenon.

3. Methodology

3.1 Study Area

This research will focus on ten municipalities in the Netherlands as shown in Figure 1. This sample was selected due to data accessibility and in agreement with the thesis supervisors. While a lot of regions in the Netherlands deal with foundation damages and risks these municipalities were selected based on known foundation issues and a preference for a diverse group of municipalities. All are in regions with dense urban areas while some also include rural area (e.g. Alphen aan den Rijn).

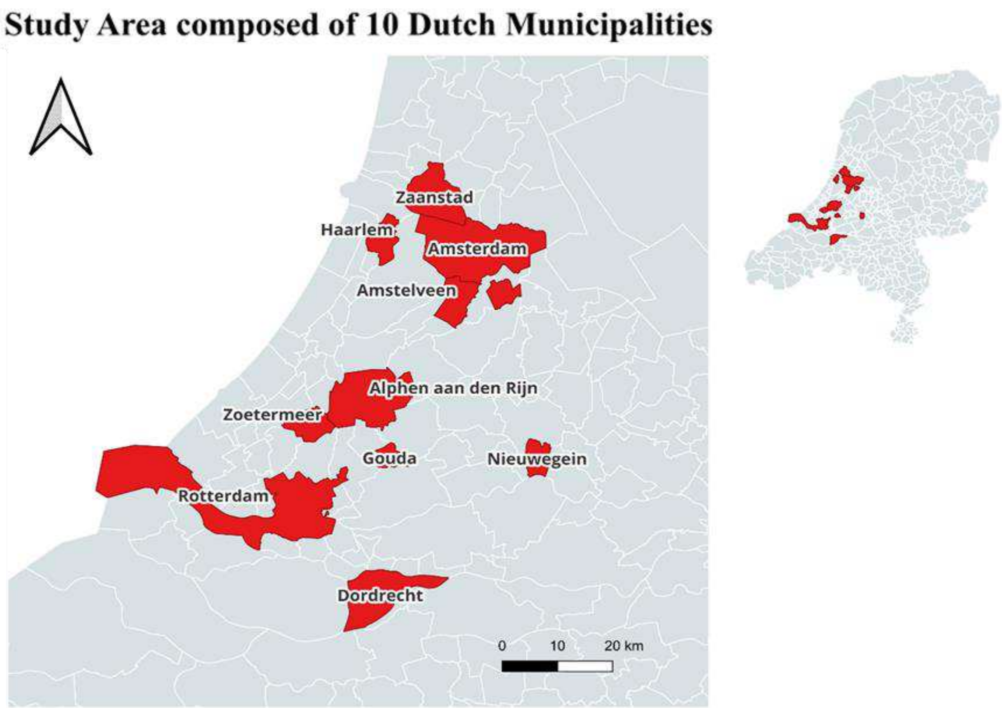
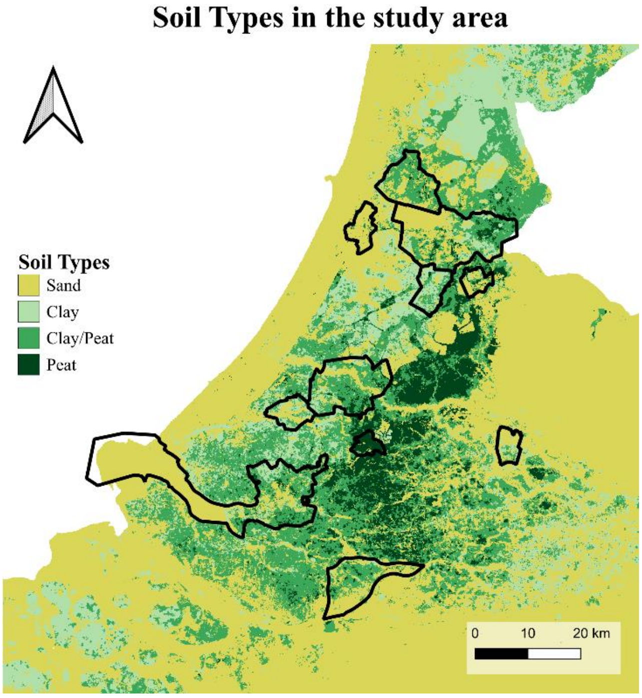
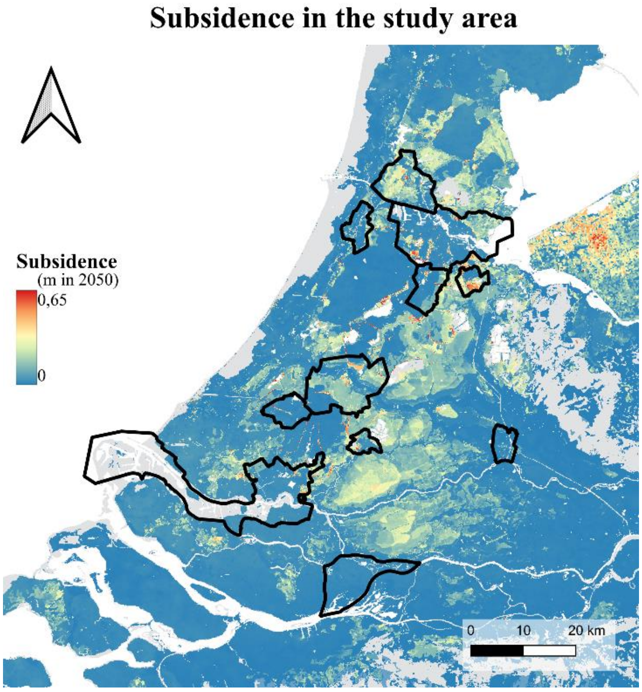


Figure 1. Study area.

Figures 2 and 3 show that soil types and subsidence rates vary widely between and within the municipalities. The western parts of the Netherlands typically have a wide variety of soil types and subsidence rates due to their history of Holocene marine and fluvial sediments. They often suffer from widespread ground settlements which is primarily driven by the presence of highly compressible fine-grained 'soft soils' such as clay and peat (Den Haan & Kruse, 2007). These types of soils are extra vulnerable to subsidence and droughts due to their high permeability and the settling and oxidations of particles (van den Born et al., 2016).





Figures 2 and 3. Soil types and subsidence.

3.2 Data

The data used in this paper comes from a combination of sources[2] and was edited and pre-processed through STATA 17 (StataCorp, 2024) and QGIS (QGIS.org, 204). The main variables that are used in the regression are the neighborhood risks. The risks for Dutch neighborhoods concerning foundations were calculated by Deltares (Kok & Anglova, 2020) and published on the Climate Impact Atlas website (Kok, 2021). The method used by Kok & Anglova (2020) and based on the paper by Costa, Kok and Koff (2020). Their method involved calculating the minimum and maximum damages per property based on five classified damage categories and multiplying these with the probabilities of a foundation type. At the basis of their risk assessment lies the conceptual framework of the UNISDR (2016) based on hazard, exposure, and vulnerability. The advantage of their method is that it is suited for large-scale applications due to its ability to establish correlations between varying hazard levels, building characteristics, and expected damages, allowing for spatial extension and flexibility with incremental data improvements.

Their methodology for assessing pole rot risks consisted of three steps. Initially, for each property, the likelihood of it being built on wooden foundations was calculated. Then,

[2] The additional research portfolio contains an expansive summary of the data used in this study.

properties are assigned to a sensitivity class (low, medium, high), which was dependent on groundwater levels, the depth of the uppermost part of the foundation pole, and soil type at the pole head level. Based on the sensitivity level, property age, and soil type the damage levels per property were calculated.

The differential settlement risk analysis used a similar method, also using the probabilities a house has shallow foundations. To assess sensitivity, subsidence data and factors influencing differential settlement were analyzed. These factors include the thickness of the fill layer, soil heterogeneity within neighborhoods, and the increased susceptibility of properties on clay to differential settlement. Additionally, structural characteristics such as the presence of a cellar were considered. Based on the settlement speed per year (corrected for property and soil sensitivity) each property was assigned to a damage class. These damage classes per property (both for pole rot and differential settlement) were multiplied with the foundation type probabilities to calculate expected damages. The researchers did not know current damage levels or past repairs, so results are the cumulative (nominal) damages from construction year till 2050 (assuming repairs are done before 2050). For both risks, damages were calculated with mild and strong climate change under the WH-climate scenario (Klein Tank et al., 2014).

While the risk data was calculated per property the available online dataset is aggregated on neighborhood level. The neighborhood risk scores were calculated by:

$$Risk_{neighborhood} = \frac{\Sigma(\#properties_{damageclass} * weight_{damageclass})}{\#\, properties}$$

This calculation implicates that a high-risk score can signal that a neighborhood has many properties on a sensitive foundation with a low average damage class, or a smaller number of properties on a sensitive foundation but with high expected damage. To account for this, the method will apply property specific controls and risk driver interaction terms.

The subsidence map was obtained from Y.R. Premchand (VU affiliated PhD student) and was developed by van der Meulen et al (2007) and Fokker et al (2019). The map (Figure 3) shows the expected subsidence in 2050 in meters on a 100m x 100m grid. Parts of the map that had missing values (denoted with -1) were excluded. To account for subsidence occurring not just at the location of the point layer of the house, and subsidence being an increased problem if there are varying subsidence rates occurring under the house, a buffer was created for each property. This buffer was calculated using the size variable and dividing it by 0.9 to achieve average ground space (Regiocontainer, 2019). Using zonal statistics on the subsidence layer, mean and median subsidence was determined per property polygon. An additional transformation was done dividing the meters expected subsidence in 2050 by 30 (years) and multiplying these with 1000 to find the expected millimeter subsidence per year.

To investigate the relationship between soil types and prices, a 2022 soil quality map was obtained from TNO (Appendix F). The soil types are shown in Figure 2 and are composed of safe (or sandy) soils, clay, clay/peat and peat soils. The soil type layer is structured as a 100 x 100m grid and covers the entirety of the Netherlands. For this map and the other previous datasets, the data was prepared to only encompass the ten municipalities in the study area using administrative boundaries shapefiles from the non-departmental public body Kadaster (PDOK, 2024). All the previously mentioned spatial datafiles were adjusted and reprojected to the same spatial index in QGIS (*RD New*).

In hedonic models, one common and straightforward method to account for spatial factors is to include the distance from each property to the central business district (CBD) as an explanatory variable (Herath & Maier, 2010). However, since Dutch cities typically lack distinct CBDs but feature historic centers that attract residential preference, the distance to these centers were used as a control. Consequently, a point layer was created[3], marking a central point in each municipality to represent these centers. Using this spatial framework, the distance to the nearest center for each property was calculated employing the Euclidean distance formula[4] based on their x and y coordinates.

The property dataset of transactions and related property variables was obtained from the Dutch Association of Real Estate Agents (NVM) through the VU (Appendix F). The dataset contains information and variables on 468,361 house transactions between 2001 and 2022 in the study area.

[3] The creation of center points was done based on the QGIS Open Street Map, using the centroids feature.
[4] The Euclidian distance is the straight-line segment distance between two points (Danielsson, 1980).

3.3 Pre-processing and Descriptives

As the data was gathered from several sources multiple steps were taken to clean the data and prepare it for analysis. As most of the transaction data cleaning had been done by the VU staff before sharing it, not a lot of steps were necessary. First, all observations that had missing values for any of the risk indicators, missing neighborhood ID's or negative rooms were removed (this led to 10,497 transactions being deleted). Secondly, residuals from a basic regression model were examined. Standardized residuals greater than 3 or less than -3 were flagged as outliers and dropped as these extreme values can potentially skew the analysis (which led to 1,728 dropped transactions). Due to house prices being highly skewed these were log-transformed to further ease the computation and interpretability. Other steps that were done included calculating the subsidence values to mm per year, normalizing the foundation probabilities, and adding a pre-1970 dummy. In Table 1 the descriptive statistics of the final sample are described, Appendix F contains the scripts for the data cleaning and further information on risk variables per municipality can be found in Appendix B.

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Price (in euro)	304048	256657	25000	8500000
Price (in euro per m2)	3055	1725	307	24285
Log Price	12.42	.59	10.13	15.95
Pole Rot Risk (in 2050)	1.8922	3.7514	0	37.43
Differential Settlement Risk (in 2050)	1.521	5.0742	0	90
Probability of Shallow Foundations	.0611	.1748	0	1
Probability of Wooden Foundations	.1694	.3331	0	1
Probability of Safe Foundations	.7688	.3887	0	1
Size (in m2)	101.26	46.76	25	858
Number of Rooms	3.8	1.5	1	24
Distance to Center (in km)	3.0671	2.2187	.0288	25.211
Subsidence (in m in 2050)	.0697	.1154	0	.6511
Subsidence (in mm/year)	2.324	3.8463	0	21.70
Subsidence Variety	1.4526	.7126	1	4
Sandy Soil	.4445	.4969	0	1
Clay Soil	.0989	.2985	0	1
Clay/Peat Soil	.3234	.4678	0	1

Variable	Mean	Std. Dev.	Min	Max
Peat Soil	.1332	.3398	0	1
Apartment	.6058	.4887	0	1
Terraced House	.2542	.4354	0	1
Semi-Detached House	.1177	.3222	0	1
Detached House	.0223	.1477	0	1
Garden	.7099	.4538	0	1
Maintenance Outside	.7611	.1093	0	1
Maintenance Inside	.7514	.1416	0	1
Good Maintenance	.8513	.3557	0	1
Listed Building	.0202	.1406	0	1
Constructed before 1906	.1066	.3085	0	1
Constructed between 1906-1930	.1853	.3885	0	1
Constructed between 1931-1944	.1073	.3095	0	1
Constructed between 1945-1959	.0652	.2468	0	1
Constructed between 1960-1970	.1218	.3271	0	1
Constructed between 1971-1980	.1121	.3155	0	1
Constructed between 1981-1990	.1156	.3197	0	1
Constructed between 1991-2000	.0987	.2983	0	1
Constructed between 2001-2010	.0686	.2527	0	1
Constructed between 2011-2020	.0184	.1345	0	1
Constructed in 2021	.0004	.0206	0	1
Constructed before 1970	.5922	.4914	0	1
Property Age (inferred)	63.15	60.11	0	1016

Note: The number of observations is 456,136.

Table 1 summarizes the descriptive statistics, showing that the average house price is €304,048 and the average price per square meter is €3,055. While the average property price in 2021 was €387,000 (CBS, 2022) the reason the average sample price is lower because the transactions span from 2001 to 2021. For the analysis, risk variables were added to the NVM dataset. First, the predicted (yearly) median subsidence values to each house, giving an expected ~0.07-meter subsidence for houses in 2050 and ~2.3 millimeter annually. This number is much lower than the average subsidence rates in the western part of the country (8 millimeters) (Stouthamer & van Asselen, 2015). The possible reason for this dispersion is probably that in rural areas the subsidence rates are much higher than for urban regions (RLI, 2020).

Because the neighborhood risk is the most important variable in this study, this will be discussed in more detail. The neighborhood risk variables were obtained from the Climate Impact Atlas website (Kok, 2021) and are based on the study by Kok & Anglova from Deltares (2020). These data provide neighborhood risk levels for both pole rot and differential settlement across various time periods and climate scenarios. For both risk indicators, the average risk in 2050 was calculated using the mild and strong climate scenario scores per property, yielding an average risk of 1.89 for pole rot (with a standard deviation of ~3.8) and 1.5 (with a standard deviation of ~5.1) for differential settlement. Although the mean values for this variable are close to 0, the maximum values are quite high. The distribution of the risk variables in Figure 4 show they are highly skewed.

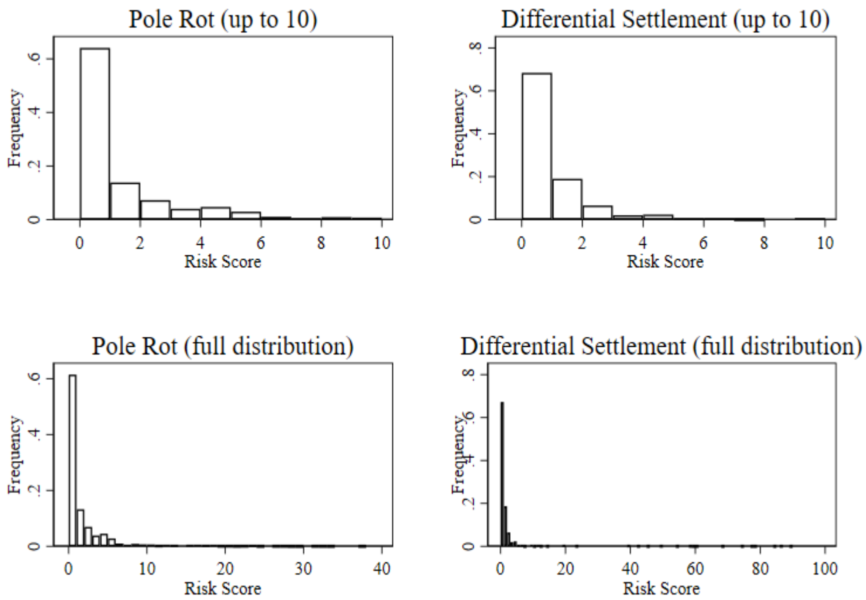


Figure 4. Distribution of neighborhood risk scores.

The average risk scores per construction year were plotted in Figure 5 and show pole rot risks are much higher for properties build before 1950. This pattern of older properties having higher pole rot risk is expected seeing that property age was used as an indicator of pole rot risk. Differential settlement risks show higher and more irregular risk patterns between ~1960 and ~2010 compared to pole rot risk.

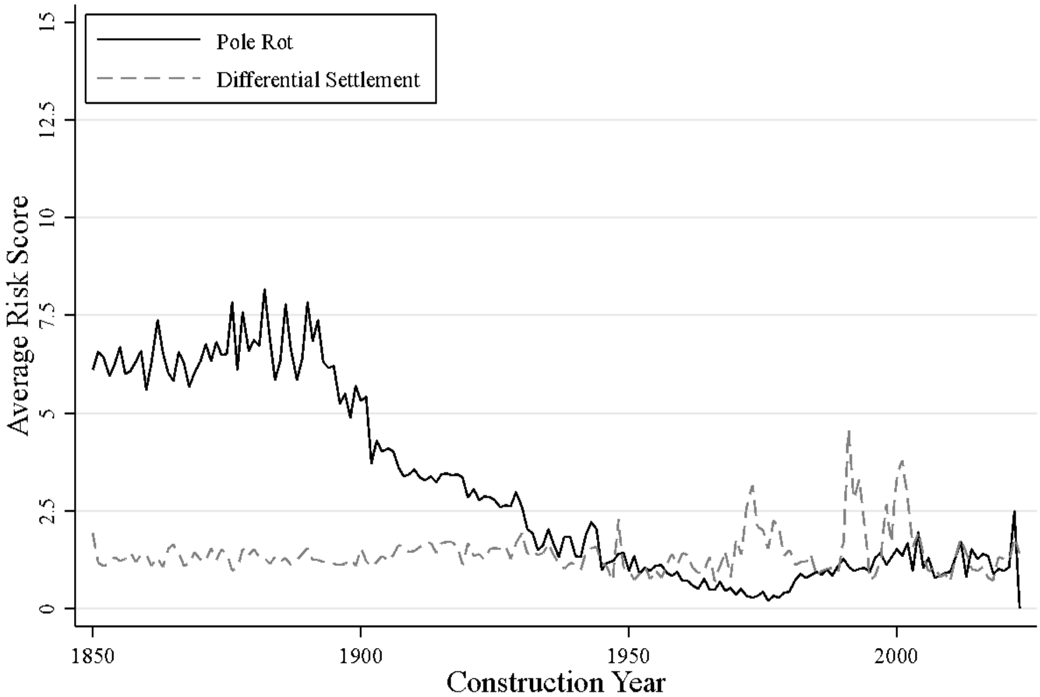


Figure 5. Average risk scores by construction year[5].

The pole rot risk variables seem to be partly correlated with construction age, which is also what the correlation matrices in Appendix E and literature seem to suggest (Klaassen, 2008). The distribution of the risk variables across municipalities (Table B1) also shows pole rot risks are most common in the older cities of Amsterdam, Zaanstad and Haarlem. Differential settlement risks seem to be most prominent in Zaanstad, Amsterdam and Zoetermeer.

Additionally, this research utilized tables from Deltares that provided probabilities for specific foundation types based on age category, soil type, and region (Kok & Anglova, 2020). These tables were manually transformed into a foundation probabilities dataset (Appendix F) and merged with the transaction data to assign foundation probabilities to each house. The mean probabilities for houses were approximately 0.06 for shallow foundations (6%), 0.17 for wooden foundations (17%), and 0.77 for concrete/safe foundations (77%). The soil type

[5] The inferred construction years from 1850 onward were used due to not every construction year being available and visibility reasons. The oldest property in the dataset stems from 1005.

descriptives show that the most common soil occurring under houses is sandy/safe soil (44%) with clay, peat/clay and peat soil types composing of 9%, 32% and 13% of the sample.

A notably high prevalence of apartments at ~60% makes sense considering dense cities such as Amsterdam and Rotterdam are part of the sample. Conversely, the relatively low prevalence of detached houses at only ~2% is also a result of the highly urbanized municipalities that were chosen. A significant majority of properties exhibit good maintenance with ~85% reported as in good condition and ~76% showing good exterior maintenance. These high percentages suggest that the housing stock is generally well-maintained and a majority of houses do not show foundation damages like wall cracks. Moreover, the data demonstrates that a substantial portion of properties, approximately 59%, was constructed before 1970, meaning they have a higher chance of foundation risks and damages. The reason the construction year was not described in this table is because ~70 thousand observations do not have a construction year listed and this research used construction age categories.

The correlations between the variables are presented in Appendix E. The first matrix shows the correlations between the property specific characteristics and log price. Most correlations were expected such as between *size* and *rooms*, *apartment*, and *terraced/size/rooms*, *listed* and *distance to the center*, *maintenance inside* and *outside*. Unexpected correlations include a negative relationship between gardens and price, possibly due to gardens occurring in more suburban cheaper areas. The second correlation matrix shows the risk variables together with the age of a building when sold and a pre-1970 dummy. Most strong correlations are expected because pole rot and differential settlement risk were (partly) calculated using property age, soil types, subsidence, and the foundation probabilities. Foundation probabilities were also calculated using property age and soil types, so these relationships were also expected. Not surprisingly, subsidence is correlated quite strongly with soil types as subsidence occurs more peat soils and less on sandy soils. However, the negative correlation between clay and subsidence is surprising as theory would suggest subsidence to be occurring in clay soils. This relationship could partly be explained by the fact that subsidence of clay layers is dependent on the type of clay, type of built-up area and thickness of the layer (Koster, Stafleu & Stouthamer, 2018; RLI, 2024). The spatial patterns of the clay diversity and its relationship with subsidence are not studied in this paper but could be in further research. Prices are mainly correlated with age of properties and pole rot risks. This relationship probably exists because older houses are often located in city centers which have higher prices and pole rot risks are partly dependent on building age. Additionally, older homes are sold for a premium above a certain age and due to vintage effects (Koster, van Ommeren and Rietveld, 2016; Rolheiser, van Dijk & van de Minne, 2020).

3.3 Empirical Framework

This paper aims to estimate the causal effect of foundation risks on house prices. Theoretically, homebuyers can know beforehand if a property has suffered foundation damages or is at risk of them occurring in the future. Following hedonic pricing theory, it can be assumed that riskier houses will have a discounted value compared to those that have safe/high quality foundations. However, due to a lack of knowledge, the relationship between risks and prices is likely nonexistent or very small. To investigate this assumption, the risk indicators described in the previous section will be used as determinants of property prices to see if they affect prices at all.

Although using hedonic pricing for real estate analysis has been criticized due to nonlinearity, multicollinearity and heteroskedasticity problems it is still the most common method due to its flexibility, easiness, and few restrictions (Kuminoff, Parmeter & Pope, 2010; Owusu-Ansah, 2011; Palmquist, 2005). The most prevalent statistical model used in hedonic pricing is the ordinary least squares (OLS) regression and is therefore also used in this study (Mayer et al., 2019: Yoo & Perring, 2016).

For hedonic pricing methods, estimation issues could arise because houses have an almost limitless number of independent variables and if they are not all included in the model, the coefficients can become biased (Kuminoff, Parmeter & Pope, 2010; Mayer et al., 2019). It is therefore of great importance to choose the correct functional form and check for multicollinearity issues and omitted variable bias (OVB). There is no consensus on the appropriate functional form as economic theory generally doesn't provide guidance on the correct specification (Cassel & Mendelsohn, 1985; Palmquist, 2005), although models that included all variables and simpler functional forms such as linear, log-linear, log-log and linear Box-Cox functions are recommended (Cropper et al., 1988; Kuminoff, Parmeter & Pope, 2010).

The log-linear form of the OLS model was chosen to allow for easier interpretation of marginal effects and it is performing well to uncover marginal implicit prices even if mis specified

(Cropper et al., 1988; Mayer et al., 2019; Yoo & Perring, 2016). Additionally, log-linear models are particularly well-suited for analyzing housing price data, which is often heavily skewed by outliers. These models not only provide coefficients that are intuitive and interpretable as (semi-)elasticities but also help minimize the problem of heteroscedasticity, leading to residuals that are closer to normal distribution. Because of these advantages it is also the most common form of specification in real estate analysis (Herath & Maier, 2010). Given the logarithmic nature of the dependent risk variable, the following equation is used to convert the coefficients into percentages:

$$(exp(\beta) - 1) \times 100 \tag{1}$$

3.4 Empirical Strategy

The model used in this research is a hedonic regression of the log house price on explanatory variables. The hedonic price equation to be estimated can be expressed in the most general form as:

$$ln(P_{it}) = a + \beta X_i + \theta_t + \varepsilon_{it} \tag{2}$$

where a, β, and θ are the parameters to be estimated. The dependent variable $ln(P_{it})$ represents the log house price for property *i* in year *t*, a represents the intercept, X_i includes a vector of attributes for property *i*, θ_t captures the year (21) fixed effects to control for annual price trends, and ε_{it} is the identically and independently error term. This setup provides a baseline understanding of how different property attributes influence the price without any foundation risks. The property controls in X_i are composed of *size*, *rooms*, *property types*, *garden*, *maintenance inside* and *outside* and the neighborhood variable *distance to the center*. Obviously, for categorical dummy variables one category is omitted to avoid the dummy trap.

An important consideration in the hedonic pricing method in combination with house qualities and maintenance is the impact of property age. While the dataset includes information on both interior and exterior maintenance of a house, the quality of the foundations is related to age and unrelated to maintenance. As a property ages, the risks associated with its foundation also increases, especially for pole rot risks (section 2.1). However, due to age being correlated to many property and neighborhood variables and foundation (quality) data not being available, age cannot be used as a direct measure for foundation risk impacts. Still, it is important to include property age as a factor in the hedonic equation. Goodman and Thibodeau (1997) showed that housing deprecation is non-linear and “dwelling age-induced heteroskedasticity is prevalent in hedonic house price equations”. They explain that as the age of a property increases, the price can vary more due to varying rates of deterioration of older houses and the longer time available for modifications such as renovations and expansions.

At the same time, historical vintage effects can positively increase the prices of older houses and surrounding buildings (Koster & van Ommeren, 2016; Wilhelmsson, 2008). Thus, to control for the vintage/historical age effects the variable *listed* is added to vector X_i. An additional control for property age are construction year decade dummies that are represented by vector D_i representing 11 construction year dummies for property i.

One of the main issues for hedonic methods is the issue of OVB, especially because of spatial and temporal correlation in the error term. The residual correlation is often caused by misspecification of spatially delineated variables, systematic mismeasurement of regressors or spatial covariates not being included (Von Graevenitz & Panduro, 2015). It is crucial to account for other spatial factors that might correlate with both foundation risks and house prices to avoid biased estimates. For instance, certain risk values might be more prevalent in areas with superior amenities or higher environmental quality. Failing to control for these factors could obscure the true effect of risk on house prices, resulting in biased coefficients and standard errors (Anselin, 2010). To mitigate these issues, researchers often employ spatial weight matrices, spatial fixed effects, or quasi-experimental methods such as difference-in-differences techniques. Despite the potential benefits of spatial models like spatial error, spatial lag, or spatial Durbin models, this research avoids these due to the unknown functional forms and spatial weights, leading to identification problems (Gibbons and Overman, 2012). A literature review by Kuminoff, Parmeter, and Pope (2010) revealed that approximately 60% of studies using hedonic pricing methods apply spatial fixed effects to address OVB (and 40% use temporal fixed effects). Implementing spatial fixed effects significantly reduces bias from spatially omitted variables by absorbing the price effects of spatially clustered unobserved factors (Kuminoff, Parmeter, & Pope, 2010; Zabel, 1999). These spatial fixed effects account for unobserved, time-invariant spatial attributes that influence house prices, thus enhancing the robustness of the analysis. In the study area there are 255 postal code areas, which are larger than the neighborhood-level aggregations of the risk variable, allowing for a more comprehensive control of spatial heterogeneity in the analysis[6]. The model will therefore include spatial fixed effects μ at the postal code area k.

$$ln(P_{itk}) = a + \beta X_i + \lambda D_i + \theta_t + \mu_k + \varepsilon_{itk} \tag{3}$$

The first and main specification used to assess foundation risks relies on the neighborhood risk variables provided by Deltares. Using a neighborhood risk index as a dependent variable is justified for several reasons. First, the main reason for using this index is that it is the only national scale quantitative measure of foundation risks. Additionally, most properties do not have foundation quality stated in the advertisements and homeowners rarely provide a foundation quality report (Hommes et al., 2023). The assumption here is that due to a lack of information, the only way people know the foundation risk is either through online sources, own knowledge or through damages that have already occurred and are visible. Because the *maintenance outside* variable is used as a control, the risk coefficients will just capture the effects of expected risk and/or the expected unseen foundation damages. This risk variable provides an objective and data-driven approach to evaluating the influence of foundation risks, thus enhancing the robustness and reliability of the analysis. For these reasons, $Risk$ per property (i) is added as a variable[7] (based on in which neighborhood the house is) to assess the impact of foundation risks on house prices. This leads to the base model:

$$ln(P_{itk}) = a + \beta X_i + \lambda D_i + \theta_t + \mu_k + \gamma Risk_i + \varepsilon_{itk} \tag{4}$$

Where γ represents the coefficient for a one standard deviation change in neighborhood risk. In addition to the neighborhood risk level other foundation risk indicators are subsequently included as interactions. Although two types of risk are investigated (pole rot and differential settlement) both models will use the same interactions with the property specific variables; foundation probabilities, subsidence rates and soil types. The model with added interactions will thus look like:

$$ln(P_{itk}) = a + \beta X_i + \lambda D_i + \theta_t + \mu_k + \gamma Risk_i + \phi Z_i + \delta(Risk_i \times Z_i) + \varepsilon_{itk} \tag{5}$$

In this equation the added term Z_i represents the property risk variable, ϕ the coefficient to be estimated for the property risk variable and δ the parameter to be estimated for the interaction effect. The other coefficients and terms remain the same as the previous equations.

[6] Sensitivity checks using different spatial fixed effects will be used to investigate the relationships further.
[7] For easier interpretability of the coefficients, the risk values were standardized.

Due to risk values being assigned based on neighborhood aggregated risk, the property-specific interactions are necessary to capture more granular effects and provide a more detailed understanding of how foundation risks specifically impact individual property prices. Unfortunately, these property risk variables can be somewhat correlated with the neighborhood risk variables because they form part of the original neighborhood risk calculation (section 3.2). However, multicollinearity in the regressors does not bias the coefficient estimates, although it can decrease the efficiency by increasing the standard errors (Dormann et al., 2013). Further specifications (in sections 4.2 and 4.3) will aim to estimate the separate effects of these risk indicators to provide a more comprehensive analysis.

Including these interactions is logical because it acknowledges that while neighborhood risk gives a general overview, the specific characteristics of each property can lead to more detailed variations in how risk impacts house prices. For instance, two houses in the same neighborhood might be located on different soils and have different subsidence rates, which can alter their vulnerability to risks like pole rot or differential settlement. By incorporating property-specific risk variables and their interactions with neighborhood risk, the model can differentiate between these scenarios, offering a more precise estimation of the risk impacts. Additionally, this approach helps to mitigate potential aggregation bias that could occur if only neighborhood-level risks were considered. Aggregation bias can obscure the true relationship between risk and property prices by averaging out individual differences. By using property-specific variables, the model ensures that these differences are accounted for, leading to more accurate and reliable results. Furthermore, foundation risks are a result of a combination of many different drivers that all vary in impact and patterns. To accustom for these complexities, it is reasonable to include a multitude of risk indicators.

The foundation variables range from 0 to 1 and give a probabilistic indication whether a house has wooden, shallow, or concrete foundations based on its age, region, and soil type. While these foundation probabilities are not certain, it is the best indicator available. Houses with wooden foundations have a much higher pole rot risk while houses on shallow foundations are at risk for differential settlement. If the neighborhood risk increases with one unit and a house is highly likely (probability = 1) on wooden/shallow foundations it can be expected that this will negatively impact the price compared to houses in neighborhoods with lower risks.

Although one might argue predicted subsidence cannot be related to house prices, Yoo and Perrings (2016) found expected subsidence has a similar effect on price as historical subsidence. The downside of this is that when using predicted subsidence as a regressor it will likely also capture past subsidence and damages to houses. The subsidence variable used in the model will be the expected median[8] subsidence per property in millimeter per year. While regressing with the continuous subsidence variable, a dummy will also be created to inspect the effect of being above a certain subsidence damage class. The threshold value of 3 millimeter/year was chosen based on the guidelines compiled by the Organisation for Independent Foundation Research that state subsidence under 3 millimeter/year is negligible and subsidence above 4 millimeter/year causes large damages (KCAF, 2012). Important to note here is that subsidence can cause more than just foundation damages, as it also can influence surrounding structural damages (e.g. sidewalks), although these effects are not as important for prices (Willemsen, Kok & Kuik, 2020).

3.5 Assumptions and Limitations

To address any issues of correlated errors within neighborhoods, the standard errors are clustered at the neighborhood level. This is done to account for the fact that observations within the same neighborhood are not independent and have the same neighborhood risk levels, leading to underestimation of standard errors. Additionally, for models where neighborhood risk is not used, the standard errors are clustered robustly to account for heteroscedasticity complications. This will ensure the differences in the variance of the errors will be taken into account (Hayes & Cai, 2007). Normality issues will be invalid due to the large sample size and the central limit theorem (Rosenblatt, 1995). Multicollinearity among control variables is less concerning when the primary focus of the analysis is on the risk indicators themselves, not on the precise coefficients of the control variables. Therefore, multicollinearity does not reduce or impact the predictive power of the whole model (Xiao & Xiao, 2017). A further exploration into the variables using regression combinations (Appendix F) showed there were no critical multicollinearity problems (VIF > 5) (Hair, Ringle & Sarstedt, 2011).

While there is potential for reverse causality in the model, several counterarguments address these concerns. More affluent neighborhoods might attract more public or private investment

[8] The median subsidence was used to account for the fact that this is a better representative of expected subsidence than the mean value which could be distorted by a small part of the house being on a very low or high subsidence grid cell.

in foundation assessments and damage mitigation, potentially reducing foundation risks. This is addressed by including spatial fixed effects, which capture unobserved, time-invariant spatial attributes, thus preventing bias in the estimated effects. House prices do not influence soils or subsidence rates because these are considered natural phenomena (Yoo & Perrings, 2017). Similarly, house prices do not affect the probability of having a certain foundation type, as these probabilities are determined by construction years, regions, and soil types. Therefore, the way neighborhood risk was calculated is based entirely on natural phenomena and historical construction practices, which are exogenous to current house prices. Even if higher house prices could lead to more investment in maintenance and renovations, thereby improving foundation quality and reducing risks, these improvements are not reflected in the risk index used in this model. Therefore, all risk variables are exogenous to house prices, ensuring the validity of the risk assessment in the regression analysis.

Another possible issue for the model is that the risk variable does not consider renovations and mitigation measures. If more affluent households and neighborhoods gather more information and invest more in mitigating foundation risks, these areas will have higher prices and lower risks compared to less affluent areas where such investments are not made. However, since these risks are constructed exogenously, this could potentially lead to an overestimation of the true foundation risk in affluent areas and an underestimation in less affluent areas, potentially biasing the predicted relationship between house prices and foundation risks. Although this might seem like a problem, foundation risks have only begun to get attention since 2018 (RLI, 2024), indicating that most homeowners have not inspected or repaired their foundations. Furthermore, renovations typically occur only after foundations have begun causing damage, and many risks have yet to result in damage due to the long-term nature of pole rot and differential settlement (RLI, 2024; Kok & Angelova, 2021).

3.6 Difference-in-differences

As further analysis and to address endogeneity issues, a temporal difference-in-differences (DiD) strategy has been developed. Before the sixties, a lack of regulations and advanced techniques meant lot of houses were being built on shallow and wooden foundations, especially on risky soils like clay and peat (section 2.1). These properties are now at higher risk than similar houses built on the same soil types after 1970, when foundation regulations and techniques improved, and all houses were built on concrete piles. Consequently, houses on risky soils can be considered the treatment group, and houses on safe soils can be the control group. The analysis thus compares price differentials between soil types from before 1970 and after 1970. The underlying assumption of this regression is that while various housing improvements occurred around this time (RLI, 2024), the only qualitative housing differences between houses on different soil types can be attributed to changes in foundation types.

First, $UnsafeSoil$ is defined as a dummy that equals 1 for houses built on risky soils (clay clay/peat and peat) and 0 otherwise (sandy soil). *Post1970* is a dummy variable indicating with a 1 if a house was built after 1970 and 0 if otherwise.

$$ln(P_{itk}) = a + \beta X_i + \theta_t + \mu_k + \eta UnsafeSoil_i + \kappa Post1970_i + \chi(UnsafeSoil_i \times Post1970_i) + \varepsilon_{itk} \quad (6)$$

The coefficients to be estimated are similar to that of the previous equation (5) but now include treatment, and a time period dummy and excludes construction year dummies due to multicollinearity issues. The coefficients η, κ, and χ estimate the impacts of the treatment and time periods. Specifically, η represents the average difference in prices between houses on risky and safe soil, κ reflects the effect of houses built after 1970 on price irrespective of soil type. The interaction term χ measures the additional impact on house prices for houses built on risky soils and after 1970. Thus, χ directly measures how much the change in foundation practices affected the price of properties on unsafe soils compared to safe soils post-1970, providing a clear indication of whether these changes effectively mitigated risks associated with unsafe soils or not.

An important requirement for the DiD strategy to work is the parallel trends assumption (Ryan et al., 2019). It is possible to see the trends for houses on different soils in Figure X. To test for the parallel trends assumption, the method developed by Riveros-Gavilanes (2023) was adapted. As the authors note, it is necessary to include a reference period (*T – 1*) before the time period dummy. The time period dummies are a function of time that are used to identify pre and post periods relative to the reference decade, which is a necessary condition for this specification to work (Riveros-Gavilanes, 2023). Additionally, the direct effects of unsafe soils and time periods need to be omitted for the model to work. Thus, *Pre1960* is a dummy variable indicating if a house was built before 1960 and 0 otherwise. Houses built between 1960 and 1970 serve as the reference group for the testing of the parallel trends. The equation therefore becomes:

$$ln(P_{itk}) = \beta X_i + \theta_t + \mu_k + \alpha(UnsafeSoil_i \times Pre1960_i) + \tau(UnsafeSoil_i \times Post1970_i) + \varepsilon_{itk}$$

where the coefficients β, θ, and μ are related to the property (i) attributes, time (t) and postal code (k) fixed effects. These covariates need to be included because they can influence the potential trends (Roth et al., 2023) The coefficient α is the coefficients of interest, which tests for the slopes between houses on unsafe and safe soils in the pre-intervention period relative to the reference decade (1960-1970). The parameter τ estimates the generic average treatment effect on the treated (unsafe soils) after the intervention where the $Post1970_i$ dummy identifies the post-treatment periods. The equation aims to test whether there are differential slopes in the treatment and control groups in the pre-1960 period. The null hypothesis (H_0: $\alpha = 0$) represents the absence of different slopes between the groups in the pre-1960 period and signals parallel trends. The alternative (H_A: $\alpha \neq 0$) implies the existence of differential pre-trends before the treatment (better foundation techniques) came into effect. This would indicate that houses constructed on unsafe soils have different price trends over time before 1960 compared to houses on sandy soils.

4. Results

4.1 Risk Interactions

The results in Table 2 provide a detailed analysis of how two types of foundation risks and their interactions with property-specific characteristics influences log house prices.

Table 2: Risks and Interactions on Log Price

	Pole Rot					Differential Settlement				
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Risk (1SD) × Wooden Foundations	-0.007*					0.000				
	(0.0041)					(0.0062)				
Risk (1 SD) × Shallow Foundations		-0.004					0.015			
		(0.0077)					(0.0149)			
Risk (1 SD) × Subsidence (in mm/year)			0.000					-0.000		
			(0.0005)					(0.0003)		
Risk (1 SD) if Subsidence (> 3 mm/year)				-0.002					-0.003	
				(0.0059)					(0.0039)	
Risk (1 SD) × Clay Soil					-0.002					-0.007
					(0.0088)					(0.0077)
Risk (1 SD) × Clay/Peat Soil					-0.002					-0.001
					(0.0041)					(0.0020)
Risk (1 SD)× Peat Soil					0.000					-0.008**
					(0.0039)					(0.0036)
Wooden Foundations	0.001					-0.004				
	(0.0086)					(0.0076)				
Shallow Foundations		-0.008					-0.011			
		(0.0101)					(0.0115)			
Subsidence (in mm/year)			-0.002***					-0.002***		
			(0.0006)					(0.0006)		
Subsidence (> 3 mm/year)				-0.015**					-0.013**	
				(0.0059)					(0.0059)	
Clay Soil					0.018***					0.017***
					(0.0072)					(0.0063)
Clay/Peat Soil					0.001					0.001
					(0.0044)					(0.0044)
Peat Soil					-0.009					-0.006
					(0.0086)					(0.0079)
Risk (1 SD)	0.005	0.002	0.001	0.002	0.003	-0.005	-0.005	-0.000	-0.002	-0.003
	(0.0044)	(0.0041)	(0.0043)	(0.0042)	(0.0046)	(0.0041)	(0.0038)	(0.0038)	(0.0056)	(0.0040)
Constant	11.559***	11.561***	11.562***	11.562***	11.562***	11.567***	11.566***	11.569***	11.567***	11.567***
	(0.0461)	(0.0458)	(0.0456)	(0.0457)	(0.0454)	(0.0444)	(0.0442)	(0.0445)	(0.0443)	(0.0445)
Year FE (21)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE (987)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Property Controls (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136	456,136	456,136	456,136	456,136	456,136
R-squared	0.885	0.885	0.885	0.885	0.885	0.885	0.885	0.885	0.885	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: We cluster standard errors at the neighborhood level.

The main hypothesis that risks do not significantly influence house prices because buyers generally lack detailed knowledge about foundation conditions is largely supported by the findings. The variables of interest are the coefficients of the interaction terms between neighborhood risk and the property specific variables (foundation probabilities, subsidence, and soil types). The specifications for the regression are all the same and are based on equation (5). The high R-squared of the models is mainly caused by the inclusion of the controls, construction year dummies and transaction year and postal code fixed effects as shown by Table 5 and 6 in the robustness analysis and Table C1 in Appendix C. As for the constant, the intercept α represents the log house price when all regressors are set to zero. The constant is relatively stable across all specifications in all further regressions.

In general, most risk interacted coefficients in Table 2 show insignificant effects, suggesting the hypothesis that risks are not well known, and buyers do not take them into account when acquiring a house, is confirmed. The interaction estimates are not suitable for drawing inferential conclusions, as they do not achieve statistical significance. However, there are two exceptions, the first being the coefficient for the interaction term between pole rot risk and wooden foundations (-0.007) in column (1), which is significant at the 10% level. This coefficient indicates that for houses suspected to be on wooden foundations (probability of 100%), the price is estimated to decrease by approximately 0.7% for each standard deviation increase in pole rot risk. This decrease is over and above the general effect of pole rot risk in the neighborhood captured by γ and the effect of having wooden foundations captured by ϕ. The interaction term specifically captures the additional negative impact on house prices when both conditions (higher pole rot risk and wooden foundations) are present. This was somewhat expected from the literature (section 2.1 and 3.2), as houses on wooden foundations are the only type of foundations at risk for pole rot. Despite the significant result, it is complicated to actually infer if this effect is correct. As foundation probabilities are dependent on soil types, construction years and municipality, it is likely that this variable is a distorted proxy and does not show the actual impact wooden foundations have. A further explanation and analysis of the foundation probabilities will be done in section 4.3.

The other significant interaction coefficient (-0.008) in column (10) indicates that for houses on peat soils, each additional standard deviation increase in differential settlement risk is associated with an approximately 0.8% decrease in house prices (significant at the 5% level). This is additional to the general effect of neighborhood differential settlement risk captured by

γ and the effect of being on peat soil compared to sandy soils captured by ϕ. This finding contrasts with the literature, which suggests that clay soils are most vulnerable to differential settlement due to their shrink-swell behavior. It is interesting that while being on peat soil doesn't seem to directly affect prices (no statistical significance for this coefficient) in combination with differential settlement it does appear to have an effect.

The direct effect of subsidence in millimeter/year, captured by ϕ in the model, is consistently negative (-0.002) and highly statistically significant, indicating that a one mm per year increase in subsidence is associated with ~0.2% lower house prices. For properties experiencing subsidence greater than 3 millimeter/year, the coefficient (ϕ) is also statistically significant and negative, with the model suggesting that there is a small (approximately -1.5% and -1.3%) but economically significant impact of severe subsidence on house prices. Table C2 in Appendix C shows that when interacting different subsidence categories and risk only the houses with subsidence above 4 millimeter/year have statistically significant coefficients that implicate these houses have around ~1.6% lower prices compared to the reference category (houses with subsidence between 0 and 2). This means the threshold coefficient ϕ is probably capturing the effects of subsidence that occur above 4 millimeter/year[9].

While the interaction terms for clay/peat and peat soils (δ) do not show significant impacts, the direct effects of soil types (ϕ) reveal that houses on clay soils seem to have a positive and significant higher price. The coefficient for clay soil is ~0.017 at the 1% significance level, indicating an economically relevant relationship between clay soil and property prices. This could be partially explained by the fact that clay soils are most common in highly attractive and expensive cities like Haarlem, Amsterdam, and Rotterdam (Table B3). Although postal code fixed effects are used to control for spatial patterns, the high concentration of clay soil in these desirable urban areas might still contribute to the observed positive price effect, reflecting the premium associated with properties in these locations.

[9] The statistically significant interaction coefficients of the medium and large subsidence with pole rot risk categories in Table C2 should be interpreted with caution due to the small number of observations within these categories and their concentration in three municipalities (Table B1 in Appendix B).

The direct effects of risk scores (γ) on house prices are insignificant for all modelled interactions and show positive associations with pole rot risk and prices while differential settlement seems to indicate a negative effect on price. The lack of statistical significance suggests that foundation risks at the neighborhood level do not noticeably impact house prices, aligning with the initial hypothesis.

At the same time, the regression results in Table C7 (Appendix C) indicate that the coefficients for pole rot risk vary, showing both positive and negative influences on house prices. The additional regression of risk categories on prices was conducted to explore potential nonlinear relationships. Interestingly, houses with pole rot risk scores between 2 and 3 are priced significantly (at the 5% level) higher (~2.4%) than those in the reference category with scores of 0 to 2, while other risk categories show insignificant and fluctuating results. These findings highlight the challenges in inferring the true effects of risk on prices when using categorical dummies. Older houses, often located in city centers like Amsterdam, Haarlem, and Zaanstad, feature historical amenities and vintage elements. These houses typically have higher property values but are more susceptible to pole rot (section 3.2). The model does attempt to control for these patterns by using distance to the center control and using spatial fixed effects at the postal code level. However, because the model uses dummy variables to indicate which category of risk scores a house is in, the price effect being measured in this categorical regression is likely due to the spatial pattern of risk instead of an actual effect.

In contrast, differential settlement risk shows a consistent negative (but mostly insignificant) relationship with house prices. Only houses with risk levels between 3 and 5 exhibit statistically significant lower prices, approximately -3.5%, compared to those with minimal risk. possibly due to observations with higher differential settlement risk being concentrated in outer parts of Amsterdam and Rotterdam (Figure A2) while being most prominent in Zoetermeer and Zaanstad (Table B1).

4.2 Soils and Subsidence

To investigate the results of the main regression output even further, the following sections discuss the direct effects of property specific risk variables and a robustness check for the controls and fixed effects. The analysis of the direct impact of different risk driving variables adds context and nuance to the main regression while the addition of controls and varying use of fixed effects and thresholds, examines how the risk indicators change under different specifications.

Soil types and subsidence rates were assigned based on a detailed grid layer, with each cell measuring 100 x 100 meters, suggesting that properties within proximity share similar soil characteristics and subsidence rates. This spatial resolution introduces a granularity that significantly influences the effectiveness of fixed effect models at different geographic scales. In this context, neighborhood fixed effects, which control for unobserved heterogeneities within small areas (median size of neighborhoods[10] is 0.33 km^2) might not be capturing the effects of soils, subsidence, and foundation probabilities. Properties within the same neighborhood are likely to exhibit minimal variation in soil types due to the precise method of soil type assignment. Consequently, this could diminish the observed impact of foundation types on property prices, as the model adjusts for these highly localized factors. Essentially, the neighborhood fixed effects could be too restrictive, leaving little room for foundation types to exhibit a substantial influence on property prices within these small areas. The diversity in soil types and subsidence rates across a postal code likely introduces sufficient variability, which the fixed effects model can utilize to predict the impact more robustly on property prices.

Therefore, the following tables present the direct effects of the risk drivers with postal code fixed effects while Table C3, Table C4 and Table C5 show the same regressions but with neighborhood fixed effects. The neighborhood fixed effect specification for subsidence and soil types demonstrates the significance of the coefficients remains the same and subsidence coefficients remain very similar. The only change is that the effect of clay and peat soils is a little lower (0.013 and -0.004). The other difference is that the neighborhood fixed effects model has a slightly higher R-squared, indicating higher predictive power.

[10] Table B4 in Appendix B summarizes the area statistics for neighborhoods.

Table 3: Subsidence and Soil Types on Log Price (Postal code FE)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Subsidence (mm/year)	-0.001***					
	(0.0002)					
Subsidence (> 3 mm/year)		-0.015***				
		(0.0058)				
Sandy Soil			-0.003***			
			(0.0009)			
Clay Soil				0.013***		
				(0.0013)		
Clay/Peat Soil					-0.000	
					(0.0009)	
Peat Soil						-0.004***
						(0.0014)
Constant	10.915***	10.915***	10.918***	10.915***	10.915***	10.915***
	(0.0491)	(0.0490)	(0.0492)	(0.0493)	(0.0491)	(0.0490)
Year FE (21)	Yes	Yes	Yes	Yes	Yes	Yes
Postal code FE (255)	Yes	Yes	Yes	Yes	Yes	Yes
Property Controls (11)	Yes	Yes	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136	456,136
R-squared	0.894	0.894	0.894	0.894	0.894	0.894

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The coefficients for subsidence in millimeter/year and subsidence above the 3 millimeter/year threshold are similar to the ones found in the main regression. Both are negative and highly significant, indicating that each additional millimeter of subsidence per year significantly lowers house prices by ~0.1%, and subsidence above 3 millimeter/year is associated with ~1.5% lower prices. This suggests that buyers do slightly recognize the impact of subsidence on property values.

In the main regression, sandy soils serve as the reference category for soils. Adding the coefficient of sandy soils (-0.003) to the clay soil coefficient of 0.013 results in a similar positive effect for clay soils found in Table 2 (-0.0018), providing further evidence that clay soils are positively related to house prices. The effects of clay/peat soils are around 0 and insignificant in both the main regression and this specification. For peat soils, the combined effect of the sandy soil and peat soil coefficients is similar to the main regression's estimate of -0.009 and -0.006 (Table 2). When directly analyzing the influence of peat soil on prices the coefficient is statistically significant while in Table 2 they are not (likely due to larger standard errors). At the same time, the coefficient of peat interacted with differential settlement risk is positive and statistically significant. These results point to a negative relationship between peat soils and prices, especially in neighborhoods with higher differential subsidence risks (and likely more subsidence). The discrepancy between the coefficient's significance highlights the impact of model complexity on the precision and interpretability of the estimated coefficients. The overall significance of the direct soil coefficients (excluding clay/peat soils) does support the argument that soil types are related to house prices and needs further research attention. Important to not here is that subsidence is negatively correlated with clay soils and positively with peat soils (Appendix E), possibly indicating subsidence is affecting house prices through soil types.

4.3 Foundations

The assignment of foundation probabilities is determined by several factors: municipality, construction year, and soil type. The effects of municipality variations are controlled through fixed effects, while differences due to construction years are partly adjusted via construction year dummies. Table C4 shows the same model but without construction year dummies and stronger coefficients, indicating age is indeed correlated with foundation probabilities. This control strategy ostensibly isolates soil type as the main source of variation in foundation probabilities within fixed regions. This correlation is however not perfect since the soil type map consisted of four soil types while the foundation probability tables (Appendix F) only identified three soil types. Nevertheless, this relationship highlights that the results in Table 4 may not fully capture the actual effect of foundation types on property prices but rather a proxy effect for soil types.

Table 4: Foundation Probabilities on Log Price (Postal code FE)

VARIABLES	(1)	(2)	(3)
Wooden Foundations	-0.004***		
	(0.0015)		
Shallow Foundations		-0.008***	
		(0.0024)	
Concrete Foundations			0.005***
			(0.0013)
Constant	10.915***	10.915***	10.914***
	(0.0491)	(0.0491)	(0.0491)
Year FE (21)	Yes	Yes	Yes
Neighborhood FE (987)	No	No	No
Postal code FE (255)	Yes	Yes	Yes
Property Controls (11)	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes
Observations	456,136	456,136	456,136
R-squared	0.894	0.894	0.894

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Robust standard errors are used.

Table 4 shows that the foundation probabilities have significant effects in line with the literature. The effect of wooden and shallow foundations on prices seems to be negative and statistically significant while the effect of concrete foundations is positive. At the same time, when using neighborhood fixed effects, only the shallow foundations effect remains significant and around the same strength (-0.007), indicating this is a robust relationship over various spatial scales. The results of the main regression (Table 2) show a significant relationship between pole rot risk and prices when interacted with wooden foundations and no strong or significant relationship for wooden foundations directly. The coefficient for wooden foundations, when interacting with differential settlement is similar to the one in the Table 4, albeit insignificant. This all suggest that only for houses on wooden foundations a higher pole rot risk score influences prices. However, the effect of foundation probabilities (also when interacted) is likely just a distorted proxy for soil types, thus the coefficients for foundation probabilities do not prove any real effect of foundation quality on price.

4.4 Control variables and fixed effects

The Tables 5 and 6 present the addition of controls combined with neighborhood risk levels to show how the predictive power of the model increases.

Table 5: Pole Rot Risk on Log Price (with added controls)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Pole Rot Risk (1 SD)	0.096***	0.083***	0.048***	0.048***	0.002
	(0.0125)	(0.0096)	(0.0088)	(0.0086)	(0.0041)
Constant	12.425***	11.024***	11.991***	11.162***	11.561***
	(0.0163)	(0.0457)	(0.0691)	(0.0482)	(0.0458)
Year FE (21)	No	No	No	Yes	Yes
Postal code FE (255)	No	No	No	No	Yes
Property Controls (11)	No	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	No	No	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136
R-squared	0.026	0.445	0.494	0.710	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: We cluster standard errors at the neighborhood level.

Table 5 presents a comprehensive view of the impact of various controls and fixed effects on the relationship between neighborhood foundation risks and property values. In the first column (1), the coefficient for pole rot is positive and highly significant, indicating that higher pole rot risks are initially associated with increased house prices. This finding is counterintuitive, as one would expect higher foundation risks to negatively impact property values. However, the R-squared value in this model is relatively low (0.026), suggesting that risk alone only explains a small fraction of the variance in house prices and the coefficient is probably capturing omitted variables. A limitation of this model is that the adjusted R-squared could not be used due to the use of clustered standard errors at the neighborhood level, therefore the model does not adjust for complexity and the number of predictors.

When property controls are added in column (2), the coefficient for pole rot risk decreases slightly but remains significant. The R-squared jumps to 0.445, indicating that property characteristics like *size*, *rooms*, and *maintenance* significantly explain the variance in house prices. In column (3), including construction year dummies further reduces the risk coefficient by almost half while maintaining its significance, but the R-squared increases only slightly to 0.494 This provides evidence that the age effect of houses is significantly correlated with pole rot risk, capturing a substantial part of the risk's impact on prices. The small increase in R-squared indicates that while construction year is important, much of the price variation is already explained by the other property controls, and the age effect is not adding a substantial new dimension of explanatory power.

Adding year fixed effects in column (4) significantly increases the R-squared to 0.710, indicating that temporal factors (between 2000 and 2021), such as market trends and economic conditions, play a critical role in explaining house prices. In column (5), incorporating spatial fixed effects results in the pole rot risk coefficient dropping to zero and becoming insignificant, with the R-squared reaching 0.885. This suggests that spatial fixed effects capture most of the previously observed relationship between risk and house prices, indicating that the risk coefficient might have been proxying for other spatially correlated attributes like neighborhood desirability and local amenities, rather than having a direct effect on foundation risks.

Table 6: Differential Settlement Risk on Log Price (with added controls)

VARIABLES	(1)	(2)	(3)	(4)	(5)
Differential Settlement Risk(1 SD)	0.007	0.006	0.002	0.003	-0.005
	(0.0092)	(0.0064)	(0.0049)	(0.0045)	(0.0037)
Constant	12.425***	11.062***	12.024***	11.194***	11.566***
	(0.0170)	(0.0470)	(0.0668)	(0.0486)	(0.0442)
Year FE (21)	No	No	No	Yes	Yes
Postal code FE (255)	No	No	No	No	Yes
Property Controls (11)	No	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	No	No	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136
R-squared	0.000	0.428	0.489	0.705	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: We cluster standard errors at the neighborhood level.

Table 6 illustrates the effects of various controls and fixed effects on the relationship between differential settlement risk and property values. The R-squared initially shows that differential settlement risk alone explains none of the variance in house prices, with this trend continuing similar to the previous pole rot table. Regarding the coefficients of differential settlement risk, they start at 0.007 in column (1) and decrease to 0.006 upon adding property controls in column (2), further drop to 0.002 with construction year dummies in column (3), and slightly rise to 0.003 with year fixed effects in column (4). With the inclusion of postal code fixed effects in column (5), the coefficient turns negative (-0.005). This trend suggests that differential settlement risk might negatively impact house prices when thoroughly controlling for other (spatial) factors, though all coefficients are statistically insignificant, indicating caution in interpreting the effects.

A comparison between this outcome and the coefficients in Table 5 shows that differential settlement risk does not seem to be correlated with variables and spatiotemporal patterns that influence house prices. The pole rot risk exhibited a more substantial and consistent association with house prices until spatial fixed effects were included, which absorbed the effect. This indicates that neither risk variables have a significant direct impact on house prices when accounting for controls and spatiotemporal patterns.

Additional specifications of the last column (5) but with varying spatial fixed effects (neighborhood and municipality) in Table C6 show that the effects of risks are insignificant except for when using municipality fixed effects (pole rot becomes 0.023 at the 1% level and differential settlement 0.011 at the 10% level). However, municipality fixed effects do not account for the finer spatial heterogeneity within municipalities, which can lead to spatial correlation in the residuals. This suggests that while municipality-level fixed effects capture broader regional trends, they probably overlook more localized influences on house prices that are correlated with risk as the postal code fixed effects columns show.

4.5 Difference-in-differences

Table 7 presents the outcome of the Difference-in-Differences (DiD) regression (equation 6). The Table D1, Figure D1 and Figure D2 in Appendix D present evidence the parallel trend assumption holds, and the null hypothesis of no differential trends is not rejected. Also, this regression shows that pre-1960 the houses on unsafe soils had lower prices than those on sandy soils while post-1970 the houses on unsafe soils were priced higher, indicating some relationship between soils, foundation changes around that time and prices.

Table 7: Unsafe Soil and Post-1970 dummies on Log Price

VARIABLES	(1) Neighborhood FE	(2) Postal code FE
Unsafe Soil	0.001	-0.000
	(0.0012)	(0.0059)
Post-1970	0.083***	0.079***
	(0.0014)	(0.0074)
Unsafe Soil × Post-1970	0.006***	0.008
	(0.0015)	(0.0072)
Constant	10.668***	11.370***
	(0.0084)	(0.0221)
Year FE (21)	Yes	Yes
Neighborhood FE (987)	Yes	No
Postal code FE (255)	No	Yes
Property Controls (11)	Yes	Yes
Construction Year Dummies (11)	No	No
Observations	456,136	456,136
R-squared	0.891	0.880

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first coefficient (η) represents the price differential between houses built on clay or peat soils compared to sandy soils for properties independent of construction year in both specifications. This finding suggests that there is no significant difference between sandy soils and clay, clay/peat and peat soils (combined). However, seeing that soil types are related to prices differently (i.e. clay positively and peat negatively) it could be these effects cancel each other out. The coefficient for houses constructed after 1970 (κ) indicates newer houses are bought at higher prices (about 8% on average) with statistical significance at the 1% level. The coefficient of interest (χ) is however that of the interaction, showing a small effect on price (~0.8%) suggesting that houses on unsafe soils are relatively more valued post-1970 above and beyond the general time trend and any initial differences between the soil types. The coefficient suggests that regulatory and technical foundation improvements have mitigated some of the risks associated with foundations and home buyers have recognized this.

Despite this significant effect, as explained in the previous section, soil types were assigned based on a 100m x 100m grid cell, indicating spatial variation in soil is only present between houses if the fixed effect region encompasses at least one or more grid cells. Therefore, the difference-in-difference regression was also done using postal code fixed effects for robustness. Even though the R-squared for the neighborhood fixed effects specification (1) is higher it is important to be careful with inference due to the forementioned limitations of using neighborhood fixed effects.

The postal code fixed effects column (2) demonstrates that while the interaction coefficient direction and strength is similar, it is not statistically significant. This indicates that when controlling for more different soil types per fixed effects region, the effect of foundation regulations and techniques introduced around 1970 do not seem to have had an impact on property prices. Due to this model using soil types as a proxy for foundation treatment and subsidence being correlated with soils it could also be possible that, seeing the significant impacts of soil types on prices, the results of this regression are a result of these soil and subsidence effects rather than the changes in foundation types. The varying results and model construction does not provide a definitive answer on the question whether foundation risks play a role in home buyer decisions. Although the results suggest a negative impact on prices, the limitations[11]of the model and the insignificance of the coefficients in the second specification make it difficult to assert a definitive relationship between improved foundations and property values.

[11] Limitations of the model are among others; use of aggregated soil types for unsafe soil, the shift in regulation and foundation techniques not occurring specifically around 1970 but as a more gradual process between 1950 and 1970 (staggered DiD should be used in future research) and the simplified parallel trends controls.

5. Discussion

5.1 Key Findings and Academic Context

In contextualizing the results of this study within the broader academic literature, it is important to note the parallels and divergences with prior research. The study by Hommes et al. (2023) represents one of the few existing investigations into the relationship between foundation risks and property prices in the Netherlands, identifying a significant price decrement of 12% for houses with compromised foundations. This effect was observed in a limited dataset where foundation quality was explicitly listed. In contrast, the more comprehensive analysis of approximately 460,000 transactions in ten Dutch municipalities in this study has indicated that foundation risks are not consistently reflected in market prices. This analysis includes both regression models with interactions among neighborhood risk levels and foundation risk drivers, as well as difference-in-differences regressions. This discrepancy indicates the variability in how foundation risks are capitalized into property values, suggesting that the broader market may not always recognize or react to these risks, particularly in the majority of settings where specific foundation details are not transparently disclosed.

Among the findings, two exceptions were noted regarding the negative price effects on houses with specific foundation risks, with caution needed in interpreting these results due to potential distortions in the variables used. First, houses on peat soils in neighborhoods with higher differential settlement risks displayed a statistically significant negative relationship of around 0.8% in property values, translating to an average depreciation of €2432 on properties valued at €304,048 in the sample. Considering the substantial average repair costs ranging from €50,000 to €100,000, these findings suggest a mispricing of foundation risks in the market, providing evidence that these risks are not fully accounted for in property valuations.

Second, the interaction between pole rot risk and wooden foundations showed a 0.7% decrease per standard deviation increase in pole rot risk. However, this latter finding should be approached with caution. Despite the significance of the results, the actual impact of wooden foundations on prices is complex to ascertain. The probabilities of foundation types being influenced by soil types, construction years, and municipalities make foundation probabilities a likely distorted proxy that does not accurately reflect the true impact on prices. This complexity underscores the importance of using nuanced approaches to understand how specific foundation risks interact with other environmental and structural factors in influencing property values.

The results of the direct regressions of risk drivers showed that properties on clay soils command a premium of around 1.7% relative to those on sandy soils while peat soils are also associated with lower prices (~0.7%) unrelated to neighborhood risks. Additionally, the difference-in-difference regressions highlight price variations between houses on sandy soils and those on other soil types. The outcome of the difference-in-difference regression necessitates further exploration to check if the significant effects are caused by an actual change in foundations or the impact of different soil types and spatial patterns on prices. Summarizing, these results do not only enhance the understanding of how soil factors can drive economic outcomes in housing but also underscore the importance of incorporating such variables in risk assessments and mitigation strategies. The consistent effects of different soil types on property prices signal a critical area for further exploration, particularly given the sparse research landscape surrounding the interplay between soil conditions, subsidence, foundation stability, and housing market dynamics.

The direct risk regressions further indicate subsidence is associated with negative prices, and that houses expected to experience subsidence rates above 4 millimeter/year are sold at prices approximately 1.5% lower than those with less subsidence. These findings suggest that buyers recognize the risks associated with subsidence, or that damages have already occurred, thus depreciating home values. These effects are smaller than those found in previous studies by Willemsen, Kok, and Kuik (2020) and Yoo and Perring (2017), who reported that uniform subsidence impacts prices by -7% and -9.9%, respectively. However, Willemsen, Kok, and Kuik also observed that differential subsidence has a milder effect of around -2%, similar to the findings of this study. Therefore, the consistent negative impact of (future) subsidence aligns with expectations set by previous research.

5.2 Limitations and Further Research

This research provides important insights, showing that risks in foundations are not accounted for in property prices. However, it is crucial to acknowledge certain limitations and offer guidance to researchers, policymakers, and market participants.

Firstly, the primary objective of the research was to demonstrate that risk does not influence housing prices. However, proving the absence of an effect poses inherent challenges that may not be as convincing as demonstrating a clear effect. Issues such as errors in model specification, data limitations, and underlying assumptions are more likely to impact the outcome in a way that obscures effects, rather than revealing them. This underscores the fundamental challenge in providing robust evidence for the absence of an effect, as opposed to demonstrating its presence.

Secondly, this study relies on large-scale aggregated data, which incorporates various uncertainties and assumptions. On the individual level, buyers are likely to inform themselves extensively about potential foundation damages and their quality. This suggests that individual buyers may have a more precise understanding of risk than what is reflected by the aggregated risk scores used in this study, highlighting a gap between macro-level data and micro-level realities. Future research and initiatives should focus on gathering detailed data on foundation conditions, particularly in high-risk areas, to close the gap between broad, aggregated data and individual buyer knowledge.

Thirdly, the tight housing market conditions from 2014 to 2021 in the Netherlands likely pressured buyers into making quick decisions, often without the opportunity to thoroughly investigate foundation issues. Additionally, Hommes et al. (2023) note that in such a market, buyers frequently lack the ability to negotiate reductions in price that reflect the anticipated costs and efforts of necessary repairs. Indicating that risks might have affected prices were it not that market conditions override these concerns, compelling buyers to prioritize purchase opportunities over potential risks. Therefore, it is crucial to incorporate how sustained periods of tight housing market conditions alter the traditional dynamics of risk assessment and price negotiation in future hedonic pricing models. Practically, policymakers and real estate professionals could benefit from developing tools and strategies to better inform buyers about potential foundation risks, especially under the tight Dutch housing market conditions expected in the coming decade. This could include enhanced disclosure requirements to increase transparency or the introduction of standardized assessments for foundation risks, which could help balance the negotiation power between buyers and sellers, ensuring that property prices more accurately reflect underlying risk factors.

Additionally, the correlation between subsidence and soil types warrants caution, as soil types may be reflecting the effects of subsidence. Part of the reason why subsidence has such a consistent effect on prices is likely because expected subsidence is correlated with historic subsidence, leading to lower prices in areas already being affected by subsidence. The correlations between foundation probabilities, soil types, and subsidence indicate that directly measuring the effects of future risks and risk drivers on house prices is challenging. These risks are likely correlated with past damages resulting from subsidence, which include not only foundation problems but other related issues as well.

Finally, another limitation of this study is the minimal to negligible risk levels across many neighborhoods and regions within the study area, which influenced the locations with concentrations of risks driving the results. For instance, the prevalence of pole rot in city centers and older houses may create an apparent increase in price associated with risk, complicating accurate inference. This issue is further exemplified by attempts to measure nonlinear effects using categorical distinctions, which inadvertently led to comparisons between groups situated in distinct regions and, consequently, different housing markets. To overcome this problem future studies should look at more regional effects, hopefully providing robust and useful results for local governments and organizations.

6. Conclusion

This study embarked on a detailed exploration of the relationship between foundation risks and housing prices, focusing on the Dutch real estate market, which has a significant prevalence of foundation-related issues. Between 750,000 and 1,000,000 homes in the Netherlands are either currently affected or at risk of experiencing foundation damage, highlighting a significant issue that could profoundly impact property values. The primary goal of this thesis was to determine whether and to what extent foundation risks are incorporated into housing prices, with the overarching expectation that such risks do not significantly influence house prices. This hypothesis stems from an assumption that the market's understanding and incorporation of foundation risks into property valuations are limited, likely due to buyers' inadequate knowledge about the specific foundation conditions of properties and the depreciation risks if foundation problems are discovered. The complexities associated with detecting foundation risks such as pole rot and differential settlement, which often develop gradually and remain undetectable without thorough inspections, further exacerbate this lack of awareness.

The results of the study underscore but do not prove the hypothesis to be true. The coefficients of neighborhood risk variables demonstrate that foundation risks are not related to adjustments in property prices. This finding was observed using a hedonic pricing model and difference-in-differences specification to estimate the impact of foundation risks. Interestingly, this study uncovered small but significant relationships between soil types and house prices, necessitating further research. Notably, properties situated on peat soils were associated with lower prices, whereas those on clay soils commanded higher prices. Additionally, substantial levels of subsidence were found to negatively impact prices, a result that aligns with existing literature.

The discrepancy in market behavior underscores the need for comprehensive policy interventions to enhance market transparency and improve the accuracy of property valuations. This can be addressed by improving disclosure requirements, mandating foundation inspections prior to sales, and introducing standardized foundation assessments, possibly supported by governmental or industry-led schemes. Such measures would rectify market inefficiencies by providing all parties with critical information, enabling more informed decision-making and ensuring fairer transaction conditions. These changes would not only increase buyer knowledge but also lead to a clearer understanding of property values, benefiting stakeholders across the real estate market.

In conclusion, this study not only highlights the significant impact of information gaps in the housing market regarding foundation risks, but also illuminates a critical market distortion and underscores the need for policy interventions to enhance information dissemination and market efficiency. By providing empirical evidence on the economic relationship between foundation risks and home values, and demonstrating the versatility and utility of hedonic pricing models, this research contributes significantly to the field, enhancing the methodological toolkit available for evaluating economic impacts of environmental characteristics and improving the long-term viability of residential real estate investments and public welfare.

References

Anselin, L. (2010). Thirty years of spatial econometrics. Papers in regional science, 89(1), 3-26.

Armengot, J., Williams, B., & Padial, J. F. (2019). Spatial and temporal impacts on building depreciation. Journal of Property Research, 36(2), 206-225.

Billings, S. B. (2015). Hedonic amenity valuation and housing renovations. Real Estate Economics, 43(3), 652-682.

Brander, L. M., & Koetse, M. J. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. Journal of environmental management, 92(10), 2763-2773.

Bui, N., Wen, L., & Sharp, B. (2022). House Prices and Flood Risk Exposure: An Integration of Hedonic Property Model and Spatial Econometric Analysis. The Journal of Real Estate Finance and Economics, 1-32.

Cassel, E., & Mendelsohn, R. (1985). The choice of functional forms for hedonic price equations: comment. Journal of Urban Economics, 18(2), 135-142.

Clayton, J. (1997). Are housing price cycles driven by irrational expectations?. The Journal of Real Estate Finance and Economics, 14, 341-363.

Costa, A. L., Kok, S., & Korff, M. (2020). Systematic assessment of damage to buildings due to groundwater lowering-induced subsidence: Methodology for large scale application in the Netherlands. Proceedings of the International Association of Hydrological Sciences, 382, 577-582.

Court, A. (1939). “Hedonic Price Indexes with Automotive Examples”. In: The Dynamics of Automobile Demand. New York: General Motors, pp. 98–119 (cit. on p. 2)

Cropper, M. L., Deck, L. B., & McConnell, K. E. (1988). On the choice of funtional form for hedonic price functions. The review of economics and statistics, 668-675.

Czembrowski, P., & Kronenberg, J. (2016). Hedonic pricing and different urban green space types and sizes: Insights into the discussion on valuing ecosystem services. Landscape and Urban Planning, 146, 11-19.

Danielsson, P. E. (1980). Euclidean distance mapping. Computer Graphics and image processing, 14(3), 227-248.

Den Haan, E. J., & Kruse, G. A. M. (2007). Characterisation and engineering properties of Dutch peats. Characterisation and engineering properties of natural soils, 3, 2101-2133.

Donovan, G. H., Champ, P. A., & Butry, D. T. (2007). Wildfire risk and housing prices: a case study from Colorado Springs. Land Economics, 83(2), 217-233.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46.

Fokker, P. A., J. L. Gunnink, K. Koster, and G. de Lange (2019). "Disentangling and parameterizing shallow sources of subsidence: application to a reclaimed coastal area, Flevoland, the Netherlands".

Journal of Geophysical Research: Earth Surface 124.5, pp. 1099–1117.

Francke, M. K., & Van de Minne, A. (2017). The hierarchical repeat sales model for granular commercial real estate and residential price indices. The journal of real estate finance and economics, 55, 511-532.

Freeman, A. M. (1979). Hedonic prices, property values and measuring environmental benefits: A survey of the issues, Scandinavian Journal of Economics, vol. 81, pp. 154-171

Gibbons, S., & Overman, H. G. (2012). Mostly pointless spatial econometrics?. Journal of regional Science, 52(2), 172-191.

Goodman, A. C., & Thibodeau, T. G. (1997). Dwelling-age-related heteroskedasticity in hedonic house price equations: An extension. Journal of Housing Research, 299-317.

Griliches, Z. (1961). Hedonic price indexes for automobiles: an econometric analysis of quality change. The price statistics of the federal government. NBER General Series, 73(3), 173-196.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. Journal of Marketing theory and Practice, 19(2), 139-152.

Harding, J. P., Rosenthal, S. S., & Sirmans, C. F. (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. Journal of urban Economics, 61(2), 193-217.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. Behavior research methods, 39, 709-722.

Herath, S., & Maier, G. (2010). The hedonic price method in real estate and housing market research: a review of the literature.

Heyman, A. V., Law, S., & Berghauser Pont, M. (2018). How is location measured in housing valuation? A systematic review of accessibility specifications in hedonic price models. Urban Science, 3(1), 3.

Kagie, M., & Wezel, M. V. (2007). Hedonic price models and indices based on boosting applied to the Dutch housing market. Intelligent Systems in Accounting, Finance & Management: International Journal, 15(3-4), 85-106.

KCAF (2024, January 26). Stand van het Land; Funderingsproblematiek in Nederland. Retreived from: https://www.kcaf.nl/wp-content/uploads/2024/02/2024-0003-001-Standvanhetland_1.2.pdf

KCAF. (2012). F3O richtlijn: onderzoek en beoordeling van funderingen op staal (ondiepe funderingen). Retrieved from: https://www.kcaf.nl/site/images/uploads_pdf/2012.01.17_richtlijn_onderzoek_funderingen_op_staal.pdf

Kiel, K. A., & Zabel, J. E. (2008). Location, location, location: The 3L Approach to house price determination. Journal of Housing Economics, 17(2), 175-190.

Klein Tank, A., Beersma, J., Bessembinder, J., Van den Hurk, B., & Lenderink, G. (2014). KNMI 14: Klimaatscenario's voor Nederland. KNMI publicatie.

Kok, S. & Angelova, L. (2020). Impact droogte op funderingen. Deltares. Retrieved from: https://www.verzekeraars.nl/media/7875/20200930-rapport-impact-droogte-op-funderingen.pdf

Kok, S. (2021). New maps of foundation risk at neighborhood level in the Climate Impact Atlas. Deltares. Retrieved from: https://www.deltares.nl/en/news/new-maps-of-foundation-risk-at-neighbourhood-level-in-the-climate-impact-atlas

Kok, S., Van der Putten, S., & Kraus, J. (2021). Naar een kennisagenda funderingsproblematiek. Duiding van de omvang en aard van de problematiek en kennisvragen daarbij.Retrieved from: https://publications.deltares.nl/11206466_002.pdf

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

Koster, H. R., van Ommeren, J. N., & Rietveld, P. (2016). Historic amenities, income and sorting of households. Journal of Economic Geography, 16(1), 203-236.

Koster, K., Stafleu, J., & Stouthamer, E. (2018). Differential subsidence in the urbanised coastal-deltaic plain of the Netherlands. Netherlands Journal of Geosciences, 97(4), 215-227.

Kuminoff, N. V., Parmeter, C. F., & Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities?. Journal of environmental economics and management, 60(3), 145-160.

Malpezzi, S. (2003). Hedonic pricing models: a selective and applied review. Housing economics and public policy, 1, 67-89.

Mayer, M., Bourassa, S. C., Hoesli, M., & Scognamiglio, D. (2019). Estimation and updating methods for hedonic valuation. Journal of European Real Estate Research, 12(1), 134-150.

Morancho, A. B. (2003). A hedonic valuation of urban green areas. Landscape and urban planning, 66(1), 35-41.

Owusu-Ansah, A. (2011). A review of hedonic pricing models in housing research. Journal of International Real Estate and Construction Studies, 1(1), 19.

Palmquist, R. B. (2005). Property value models. Handbook of environmental economics, 2, 763-819.

Panduro, T. E., & Veie, K. L. (2013). Classification and valuation of urban green spaces—A hedonic house price valuation. Landscape and Urban planning, 120, 119-128.

PDOK. (2024). Dataset: Bestuurlijke grenzen. Kadaster. Retrieved from: https://www.pdok.nl/atom-downloadservices/-/article/bestuurlijke-grenzen#af08a124b73f59267bc0b9e9f36eb5eb

QGIS.org. (2024). QGIS Geographic Information System. QGIS Association. http://www.qgis.org

Regiocontainer. (2019). Onderzoeksanalyse: wat egebeurt er met de woon- en perceeloppervlakte van woningen?. Retrieved from: https://www.regiocontainer.nl/woon-en-perceeloppervlakte/

Ridker, R. G., & Henning, J. A. (1967). The determinants of residential property values with special reference to air pollution. The review of Economics and Statistics, 246-257.

Riveros-Gavilanes, J. M. (2023). A simple test of parallel pre-trends for Differences-in-Differences.

Rosenblatt, M. (1971). Curve estimates. The Annals of Mathematical Statistics, 42(6), 1815-1842.

Roth, J., & Sant'Anna, P. H. (2023). When is parallel trends sensitive to functional form?. Econometrica, 91(2), 737-747.

Ryan, A. M., Kontopantelis, E., Linden, A., & Burgess Jr, J. F. (2019). Now trending: Coping with non-parallel trends in difference-in-differences analysis. Statistical methods in medical research, 28(12), 3697-3711.

RLI. (2020). Stop bodemdaling in veenweidegebieden. Raad voor de Leefomgeving en Infrastructuur. Retrieved from: https://www.rli.nl/sites/default/files/advies_stop_bodemdaling_in_veenweidegebieden_-_def.pdf

RLI. (2024, February). Goed gefundeerd: advies om te komen tot een nationale aanpak van funderingsproblematiek. Raad voor de Leefomgeving en Infrastructuur. Retrieved from:

https://www.rli.nl/sites/default/files/Rli%20advies%20Goed%20gefundeerd_Advies%20om%20te%20komen%20tot%20een%20nationale%20aanpak%20van%20funderingsproblematiek_2.pdf

Rolheiser, L., van Dijk, D., & van de Minne, A. (2020). Housing vintage and price dynamics. Regional Science and Urban Economics, 84, 103569.

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

Rubin, G. M. (1993). Is housing age a commodity? Hedonic price estimates of unit age. Journal of Housing Research, 165-184.

Salzman, D., & Zwinkels, R. C. (2017). Behavioral real estate. Journal of real estate literature, 25(1), 77-106.

Sander, H. A., & Polasky, S. (2009). The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. Land Use Policy, 26(3), 837-845.

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

StataCorp. (2023). Stata 18 Base Reference Manual. College Station, TX: Stata Press.

Stouthamer, E., & Van Asselen, S. (2015). Potential of Holocene deltaic sequences for subsidence due to peat compaction. Proceedings of the International Association of Hydrological Sciences, 372(372), 173-178.

van den Born, G. J., Kragt, F., Henkens, D., Rijken, B., Van Bemmel, B., Van der Sluis, S., ... & ten Brinke, W. B. M. (2016). Dalende bodems, stijgende kosten: mogelijke maatregelen tegen veenbodemdaling in het landelijk en stedelijk gebied: beleidsstudie. PBL. Retrieved from: https://research.wur.nl/en/publications/dalende-bodems-stijgende-kosten-mogelijke-maatregelen-tegen-veenb

Van der Meulen, M. J., A. J. van der Spek, G. de Lange, S. H. Gruijters, S. F. van Gessel, B. - L. Nguyen, D. Maljers, J. Schokker, J. P. Mulder, and R. A. van der Krogt (2007). “Regional

Sediment Deficits in the Dutch Lowlands: Implications for Long-Term Land-Use Options (8 pp)". Journal of Soils and Sediments 7, pp. 9–16.

UNISDR. (2016).Report of the Open-Ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction. Retrieved from: https://www.preventionweb.net/files/50683_oiewgreportenglish.pdf

Von Graevenitz, K., & Panduro, T. E. (2015). An alternative to the standard spatial econometric approaches in hedonic house price models. Land Economics, 91(2), 386-409.

Westerveld, G. J. W., & Van Den Hurk, J. A. (1973). Application of soil and interpretive maps to non-agricultural land use in the Netherlands. Geoderma, 10(1-2), 47-65.

Wilhelmsson, M. (2008). House price depreciation rates and level of maintenance. Journal of Housing Economics, 17(1), 88-101.

Xiao, Y., & Xiao, Y. (2017). Hedonic housing price theory review. Urban morphology and housing market, 11-40.

Zabel, J. E. (1999). Controlling for quality in house price indices. The Journal of Real Estate Finance and Economics, 19, 223-241.

Appendices:

Appendix A: Maps

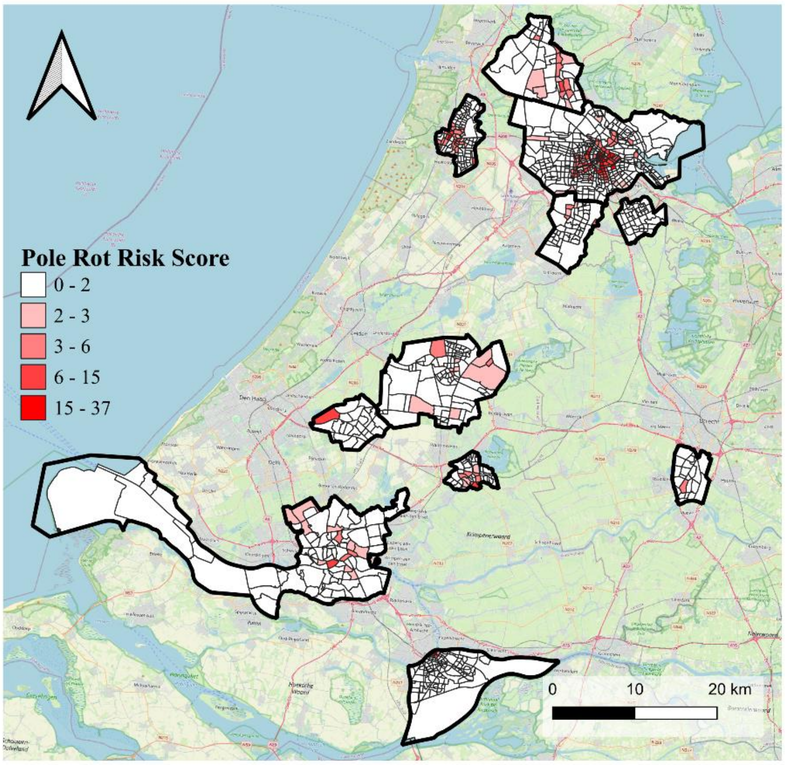


Figure A1. Pole Rot Risk map of the study area.

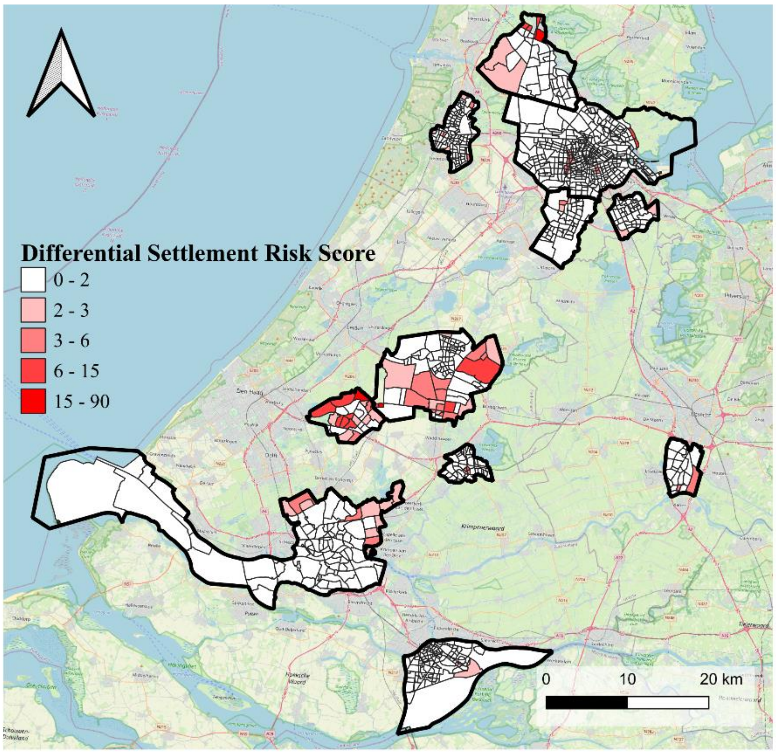


Figure A2. Differential Settlement Risk map of the study area.

Appendix B: Descriptives

Table B1: Risk Distribution over Municipalities

	Pole Rot				Differential Settlement				
Municipality	**0-1**	**1-3**	**3-8**	**>8**	**0-1**	**1-3**	**3-8**	**>8**	**Total**
Nieuwegein	13,449 (89.4%)	1,592 (10.6%)	3 (0.02%)	0 (0.0%)	12,748 (84.8%)	1,844 (12.3%)	452 (3.0%)	0 (0.0%)	15,044
Amstelveen	14,505 (73.3%)	4,684 (23.7%)	598 (3.0%)	0 (0.0%)	18,001 (91.0%)	1,465 (7.4%)	321 (1.6%)	0 (0.0%)	19,787
Amsterdam	100,456 (58.3%)	17,272 (10.0%)	34,505 (20.0%)	20,252 (11.7%)	115,395 (66.9%)	41,278 (23.9%)	14,559 (8.4%)	1,253 (0.7%)	172,485
Haarlem	19,455 (42.9%)	14,039 (31.0%)	8,553 (18.9%)	3,309 (7.3%)	27,842 (61.4%)	16,615 (36.6%)	899 (2.0%)	0 (0.0%)	45,356
Zaanstad	15,123 (48.5%)	6,934 (22.2%)	7,862 (25.2%)	1,263 (4.1%)	18,150 (58.2%)	8,485 (27.2%)	24 (0.1%)	4,523 (14.5%)	31,182
Alphen aan den Rijn	14,342 (67.1%)	6,642 (31.1%)	400 (1.9%)	0 (0.0%)	18,599 (87.0%)	1,044 (4.9%)	1,268 (5.9%)	473 (2.2%)	21,384
Dordrecht	17,091 (65.9%)	7,328 (28.2%)	458 (1.8%)	1,055 (4.1%)	21,398 (82.5%)	4,534 (17.5%)	0 (0.0%)	0 (0.0%)	25,932
Gouda	9,519 (51.0%)	6,548 (35.1%)	2,585 (13.8%)	21 (0.1%)	16,596 (88.9%)	674 (3.6%)	1,403 (7.5%)	0 (0.0%)	18,673
Rotterdam	50,732 (62.8%)	27,052 (33.5%)	2,923 (3.6%)	0 (0.0%)	56,370 (69.9%)	22,123 (27.4%)	2,214 (2.7%)	0 (0.0%)	80,707
Zoetermeer	25,556 (99.9%)	9 (0.04%)	21 (0.1%)	0 (0.0%)	1,660 (6.5%)	16,564 (64.7%)	2,671 (10.4%)	4,691 (18.3%)	25,586
Total	280,228 (61.4%)	92,100 (20.2%)	57,908 (12.7%)	25,900 (5.7%)	306,759 (67.3%)	114,626 (25.1%)	23,811 (5.2%)	10,940 (2.4%)	456,136

Table B2: Subsidence Distribution over Municipalities

Municipality	Small (0-2 mm/year)	Medium (2-3 mm/year)	Large (3-4 mm/year)	Very Large (>4 mm/year)	**Total**
Nieuwegein	13,73	375	394	545	15,044
Amstelveen	17,585	187	272	1,743	19,787
Amsterdam	136,698	2,04	1,629	32,118	172,485
Haarlem	37,616	1,408	1,476	4,856	45,356
Zaanstad	15,569	1,915	1,559	12,139	31,182
Alphen aan den Rijn	10,303	332	484	10,265	21,384
Dordrecht	25,877	2	0	53	25,932
Gouda	12,08	869	913	4,811	18,673
Rotterdam	71,112	895	747	7,953	80,707
Zoetermeer	24,936	14	20	616	25,586
Total	365,506	8,037	7,494	75,099	456,136

Table B3: Soil Type Distribution over Municipalities

Municipality	Sandy Soil	Clay Soil	Clay/Peat Soil	Peat Soil	Total
Nieuwegein	9,658	2,612	2,473	301	15,044
Amstelveen	4,449	10,599	4,435	304	19,787
Amsterdam	83,751	14,916	47,937	25,881	172,485
Haarlem	42,647	1	1,7	1,008	45,356
Zaanstad	7,665	272	19,22	4,025	31,182
Alphen aan den Rijn	7,973	2,675	9,511	1,225	21,384
Dordrecht	6,018	1,359	17,385	1,17	25,932
Gouda	921	0	1,058	16,694	18,673
Rotterdam	26,228	4,293	40,064	10,122	80,707
Zoetermeer	13,464	8,363	3,742	17	25,586
Total	202,774	45,09	147,525	60,747	456,136

Table B4: Neighborhood Area Statistics

Statistic	Value	Unit
Count	987	entries
Unique Values	345	different areas
Missing Values	0	entries
Minimum Value	0	km^2
Maximum Value	56.67	km^2
Range	56.67	km^2
Sum	1,865.87	km^2
Mean Value	1.29	km^2
Median Value	0.33	km^2
Coefficient of Variation	0.28	-

Appendix C: Regressions

Table C1: Controls on Log Price

VARIABLES	(1)	(2)	(3)	(4)
Size (in m2)	0.006***	0.007***	0.006***	0.006***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Number of Rooms	0.073***	0.025***	0.025***	0.025***
	(0.0008)	(0.0008)	(0.0006)	(0.0006)
Distance to Center (in km)	-0.006***	-0.010***	-0.025***	-0.007***
	(0.0003)	(0.0002)	(0.0009)	(0.0014)
Terraced House	-0.141***	-0.108***	0.100***	0.099***
	(0.0018)	(0.0014)	(0.0011)	(0.0012)
Semi-Detached House	-0.086***	-0.062***	0.173***	0.168***
	(0.0022)	(0.0018)	(0.0015)	(0.0015)
Detached House	0.089***	0.060***	0.344***	0.333***
	(0.0053)	(0.0049)	(0.0040)	(0.0039)
Garden	-0.066***	-0.047***	0.000	0.004***
	(0.0016)	(0.0013)	(0.0008)	(0.0008)
Maintenance Outside	0.346***	0.462***	0.172***	0.155***
	(0.0088)	(0.0071)	(0.0046)	(0.0044)
Maintenance Inside	0.291***	0.353***	0.278***	0.277***
	(0.0082)	(0.0065)	(0.0040)	(0.0038)
Good Maintenance	0.040***	0.031***	0.050***	0.051***
	(0.0026)	(0.0020)	(0.0012)	(0.0012)
Listed Building	0.257***	0.162***	0.082***	0.069***
	(0.0053)	(0.0043)	(0.0031)	(0.0030)
Constructed before 1906	-0.609***	-0.273***	-0.138***	-0.127***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1906-1930	-0.747***	-0.369***	-0.172***	-0.150***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1931-1944	-0.890***	-0.517***	-0.146***	-0.137***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1945-1959	-0.977***	-0.606***	-0.209***	-0.198***
	(0.0642)	(0.0475)	(0.0479)	(0.0484)
Constructed between 1960-1970	-1.051***	-0.669***	-0.265***	-0.240***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1971-1980	-1.070***	-0.701***	-0.205***	-0.189***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1981-1990	-0.948***	-0.580***	-0.173***	-0.155***
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 1991-2000	-0.803***	-0.429***	-0.050	-0.029
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 2001-2010	-0.686***	-0.466***	-0.033	-0.005
	(0.0641)	(0.0474)	(0.0479)	(0.0484)
Constructed between 2011-2020	-0.487***	-0.485***	-0.045	-0.026
	(0.0643)	(0.0475)	(0.0480)	(0.0485)
Constructed in 2021	-0.218***	-0.524***	-0.066	-0.076
	(0.0681)	(0.0526)	(0.0504)	(0.0509)
Constant	12.023***	11.194***	11.567***	10.915***
	(0.0643)	(0.0477)	(0.0483)	(0.0491)
Year FE (21)	No	Yes	Yes	Yes
Neighborhood FE (987)	No	No	No	Yes
Postal code FE (255)	No	No	Yes	No
Observations	456,136	456,136	456,136	456,136
R-squared	0.489	0.705	0.885	0.894

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C2: Neighborhood risks and subsidence categories (interacted)

VARIABLES	(1) Pole Rot	(2) Differential Settlement
Risk (1 SD) × Medium (2-3 mm/year)	-0.016**	-0.001
	(0.0066)	(0.0046)
Risk (1 SD) × Large (3-4 mm/year)	-0.025***	-0.002
	(0.0068)	(0.0050)
Risk (1 SD) × Very Large (>4 mm/year)	-0.001	-0.003
	(0.0062)	(0.0042)
Risk (1 SD)	0.003	-0.002
	(0.0043)	(0.0056)
Medium (2-3 mm/year)	-0.002	-0.006
	(0.0076)	(0.0079)
Large (3-4 mm/year)	-0.009	-0.013
	(0.0084)	(0.0089)
Very Large (>4 mm/year)	-0.016**	-0.014**
	(0.0063)	(0.0063)
Constant	11.563***	11.568***
	(0.0453)	(0.0443)
Year FE (21)	Yes	Yes
Postal code FE (255)	Yes	Yes
Property Controls (11)	Yes	Yes
Construction Year Dummies (11)	Yes	Yes
Observations	456,136	456,136
R-squared	0.885	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: We cluster standard errors at the neighborhood level.

Table C3: Subsidence and Soil Types on Log Price (Neighborhood FE)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Subsidence (mm/year)	-0.002***					
	(0.0001)					
Subsidence (> 3 mm/year)		-0.015***				
		(0.0011)				
Sandy Soil			-0.003***			
			(0.0008)			
Clay Soil				0.018***		
				(0.0012)		
Clay/Peat Soil					-0.001	
					(0.0008)	
Peat Soil						-0.008***
						(0.0013)
Constant	11.570***	11.568***	11.568***	11.566***	11.567***	11.570***
	(0.0483)	(0.0482)	(0.0484)	(0.0484)	(0.0483)	(0.0482)
Year FE (21)	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE (987)	Yes	Yes	Yes	Yes	Yes	Yes
Property Controls (11)	Yes	Yes	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136	456,136
R-squared	0.885	0.885	0.885	0.885	0.885	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C4: Foundation Probabilities on Log Price (excluding construction year dummies)

VARIABLES	(1)	(2)	(3)
Wooden Foundations	-0.010***		
	(0.0015)		
Shallow Foundations		-0.026***	
		(0.0025)	
Concrete Foundations			0.011***
			(0.0012)
Constant	10.656***	10.656***	10.645***
	(0.0087)	(0.0087)	(0.0087)
Year FE (21)	Yes	Yes	Yes
Neighborhood FE (987)	No	No	No
Postal code FE (255)	Yes	Yes	Yes
Property Controls (11)	Yes	Yes	Yes
Construction Year Dummies (11)	No	No	No
Observations	456,136	456,136	456,136
R-squared	0.889	0.889	0.889

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C5: Foundation Probabilities on Log Price (Neighborhood FE)

VARIABLES	(1)	(2)	(3)
Wooden Foundations	-0.001		
	(0.0017)		
Shallow Foundations		-0.007***	
		(0.0024)	
Concrete Foundations			0.002
			(0.0013)
Constant	10.676***	10.676***	10.674***
	(0.0081)	(0.0081)	(0.0083)
Year FE (21)	Yes	Yes	Yes
Neighborhood FE (987)	Yes	Yes	Yes
Postal code FE (255)	No	No	No
Property Controls (11)	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes
Observations	456,136	456,136	456,136
R-squared	0.894	0.894	0.894

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table C6: Neighborhood Risks on Log Price with Varying Spatial Effects

	Neighborhood FE		Postal code FE		Municipality FE	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES						
Pole Rot Risk (1 SD)	0.008		0.002		0.023***	
	(0.0055)		(0.0041)		(0.0052)	
Differential Settlement Risk (1 SD)		-0.000		-0.005		0.011*
		(0.0049)		(0.0037)		(0.0064)
Constant	10.912***	10.915***	11.560***	11.567***	10.891***	10.905***
	(0.0414)	(0.0413)	(0.0464)	(0.0443)	(0.0396)	(0.0426)
Year FE (21)	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood FE (987)	Yes	Yes	No	No	No	No
Postal code FE (255)	No	No	Yes	Yes	No	No
Municipality FE (10)	No	No	No	No	Yes	Yes
Property Controls (11)	Yes	Yes	Yes	Yes	Yes	Yes
Construction Year Dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456,136	456,136	456,136	456,136	456,136	456,136
R-squared	0.894	0.894	0.885	0.885	0.828	0.827

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: We cluster standard errors at the neighborhood level.

Table C7: Neighborhood risks categories

	(1)	(2)
VARIABLES	Pole Rot	Differential Settlement
Low (2-3)	0.024**	-0.001
	(0.0097)	(0.0086)
Medium (3-6)	0.007	-0.035*
	(0.0127)	(0.0192)
High (6-15)	-0.005	-0.000
	(0.0197)	(0.0253)
Very high (>15)	0.006	-0.004
	(0.0213)	(0.0473)
Constant	11.562***	11.569***
	(0.0464)	(0.0446)
Year FE (21)	Yes	Yes
Postal code FE (255)	Yes	Yes
Property Controls (11)	Yes	Yes
Construction Year Dummies (11)	Yes	Yes
Observations	456,136	456,136
R-squared	0.885	0.885

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix D: Parallel trends

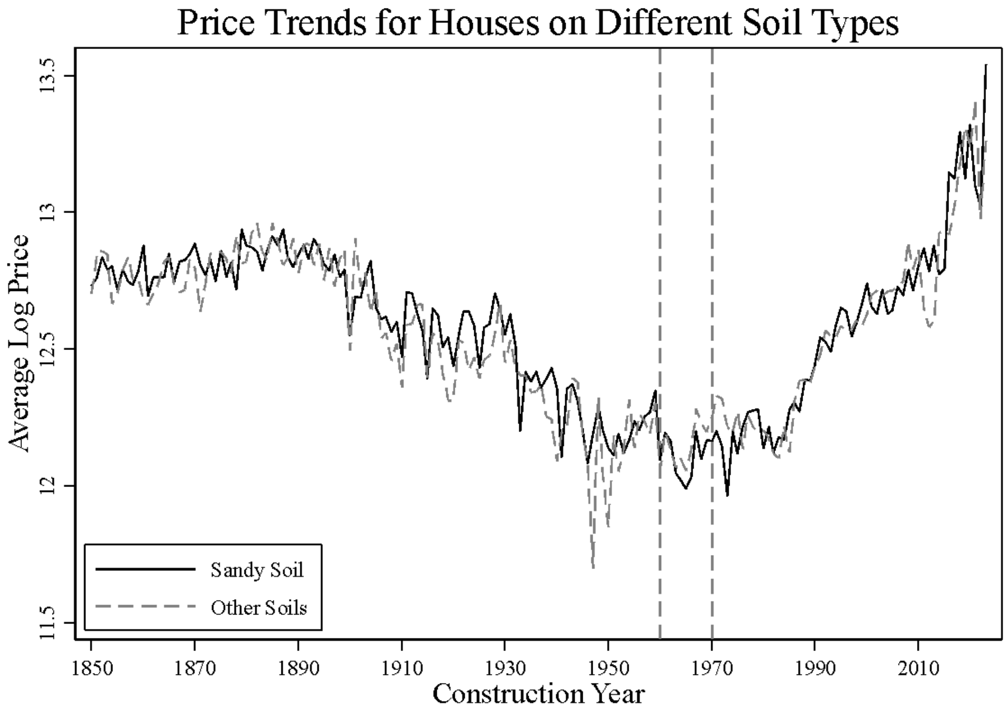


Figure D1. Price trends for houses on sandy soils and other soils.

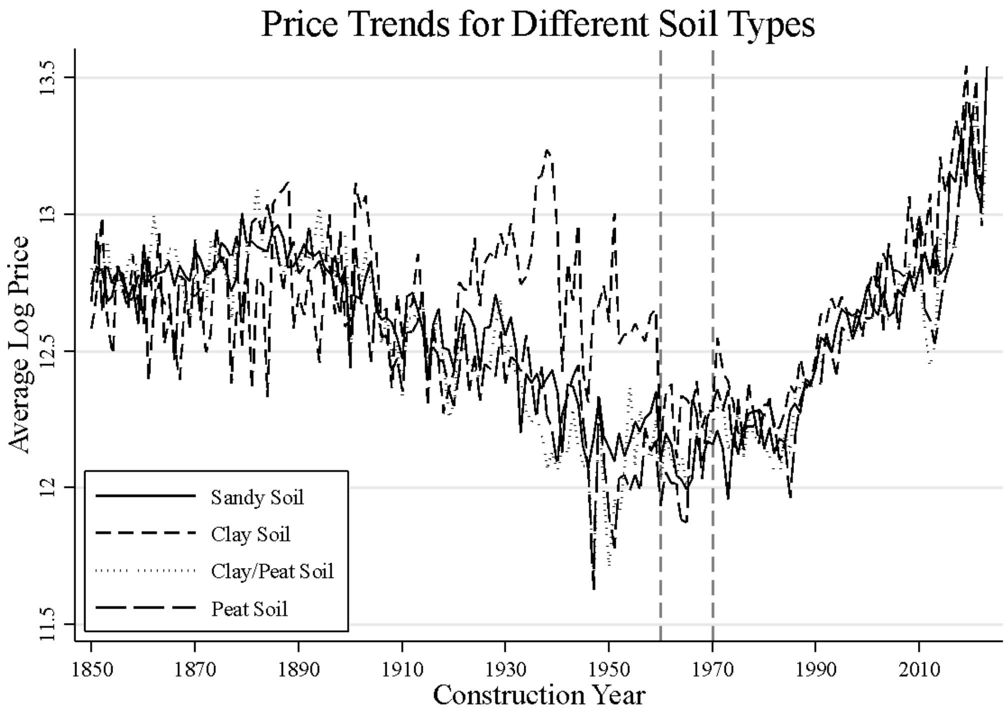


Figure D2. Price trends for houses on sandy, clay, clay/peat and peat soils.

Table D1: Soil and time trends for the PTA

VARIABLES	(1)
Unsafe Soil × Pre-1960	-0.003**
	(0.0012)
Unsafe Soil × Post-1970	0.052***
	(0.0011)
Constant	10.659***
	(0.0086)
Postal code FE (255)	Yes
Property Controls (11)	Yes
Year FE (21)	Yes
Observations	456,136
R-squared	0.889

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

interaction_pre = 0

F(1,455114) = 6.63
Prob > F = 0.0100

Appendix E: Correlation matrices

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Log Price	1.0000													
(2) Subsidence Median	-0.0038	1.0000												
(3) Subsidence (in mm/year)	-0.0038	1.0000	1.0000											
(4) Subsidence Variety	0.0832	-0.0023	-0.0023	1.0000										
(5) Pole Rot Risk (in 2050)	0.1627	0.0560	0.0560	-0.0061	1.0000									
(6) Differential Settlement Risk (in 2050)	0.0116	0.1447	0.1447	0.0039	-0.0048	1.0000								
(7) Probability of Shallow Foundations	-0.0479	-0.0666	-0.0666	0.0143	-0.0033	0.0227	1.0000							
(8) Probability of Wooden Foundations	0.0845	0.0947	0.0947	0.0001	0.3741	0.0338	0.0832	1.0000						
(9) Probability of Concrete Foundations	-0.0500	-0.0512	-0.0512	-0.0065	-0.3185	-0.0390	-0.5211	-0.8938	1.0000					
(10) Sandy Soil	0.0411	-0.2009	-0.2009	-0.0177	-0.0976	-0.0203	-0.3126	-0.4549	0.5322	1.0000				
(11) Clay Soil	0.0503	-0.1120	-0.1120	0.0033	-0.0836	-0.0429	0.2400	-0.0397	-0.0739	-0.2963	1.0000			
(12) Clay/Peat Soil	-0.0680	0.0815	0.0815	0.0200	0.0033	0.0261	0.1011	0.2641	-0.2737	-0.6185	-0.2290	1.0000		
(13) Peat Soil	-0.0108	0.2799	0.2799	-0.0046	0.2117	0.0314	0.1072	0.3366	-0.3366	-0.3507	-0.1298	-0.2710	1.0000	
(14) Age of Property	0.1891	-0.0817	-0.0817	-0.0019	0.3468	-0.0143	0.0066	0.4129	-0.3557	0.1178	-0.1518	-0.0863	0.0800	1.0000

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Log Price	1.0000											
(2) Size (in m2)	0.5901	1.0000										
(3) Number of Rooms	0.4648	0.7602	1.0000									
(4) Distance to Centre (in km)	-0.0641	0.0309	0.0537	1.0000								
(5) Apartment	-0.1883	-0.4857	-0.5542	-0.1127	1.0000							
(6) Terraced	0.0719	0.2858	0.3858	0.0479	-0.7238	1.0000						
(7) Semi-Detached	0.1060	0.2361	0.2499	0.0733	-0.4527	-0.2132	1.0000					
(8) Garden	-0.1583	-0.0246	0.0269	0.1258	-0.1376	0.1374	0.0306	1.0000				
(9) Maintenance Outside	0.1944	0.0461	-0.0381	-0.0118	0.0935	-0.0702	-0.0382	-0.0123	1.0000			
(10) Maintenance Inside	0.1994	0.0485	-0.0439	-0.0191	0.0725	-0.0575	-0.0279	-0.0210	0.7112	1.0000		
(11) Good Maintenance	0.1311	0.0246	-0.0412	0.0039	0.0446	-0.0337	-0.0148	0.0191	0.4995	0.7046	1.0000	
(12) Listed Building	0.1455	0.0692	0.0203	-0.1003	0.0300	-0.0171	-0.0236	-0.0507	0.0174	0.0125	0.0031	1.0000

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Construction Year	1.0000															
(2) Constructed pre-1906	-0.5966	1.0000														
(3) Constructed 1906	-0.3723	-0.1250	1.0000													
(4) Constructed 1931	-0.1827	-0.1004	-0.1477	1.0000												
(5) Constructed 1945	-0.0241	-0.0799	-0.1175	-0.0944	1.0000											
(6) Constructed 1960	0.0537	-0.1109	-0.1630	-0.1309	-0.1042	1.0000										
(7) Constructed 1971	0.1543	-0.1102	-0.1620	-0.1301	-0.1035	-0.1436	1.0000									
(8) Constructed 1981	0.2406	-0.1117	-0.1643	-0.1320	-0.1050	-0.1457	-0.1447	1.0000								
(9) Constructed 1991	0.3064	-0.1036	-0.1523	-0.1223	-0.0973	-0.1350	-0.1342	-0.1361	1.0000							
(10) Constructed 2001	0.3175	-0.0848	-0.1247	-0.1001	-0.0797	-0.1105	-0.1098	-0.1114	-0.1033	1.0000						
(11) Constructed 2011	0.1889	-0.0425	-0.0625	-0.0502	-0.0399	-0.0554	-0.0551	-0.0558	-0.0518	-0.0424	1.0000					
(12) Constructed 2021	0.0301	-0.0060	-0.0088	-0.0071	-0.0057	-0.0078	-0.0078	-0.0079	-0.0073	-0.0060	-0.0030	1.0000				
(13) Sandy/Safe Soil	-0.1016	0.0655	0.0462	0.0604	0.0233	-0.0682	-0.0027	-0.0299	-0.0840	0.0058	-0.0026	-0.0023	1.0000			
(14) Clay Soil	0.1373	-0.0720	-0.1018	-0.0781	-0.0191	0.0663	0.0460	0.0444	0.0592	0.0467	0.0116	0.0112	-0.2995	1.0000		
(15) Clay/Peat Soil	0.0683	-0.0476	-0.0464	-0.0322	0.0034	0.0771	-0.0084	-0.0055	0.0513	0.0017	0.0081	-0.0024	-0.6148	-0.2459	1.0000	
(16) Peat Soil	-0.0727	0.0368	0.0910	0.0281	-0.0216	-0.0689	-0.0265	0.0109	-0.0028	-0.0539	-0.0181	-0.0035	-0.3336	-0.1334	-0.2738	1.0000

Appendix F: Research portfolio

Link to the research portfolio.