CPB Netherlands Bureau for Economic Policy Analysis

Housing Market Effects of a Railroad Tunneling: Evidence from a quasi-experiment

This paper exploits the railroad tunneling in Delft (the Netherlands) as a quasi-experiment to estimate the willingness to pay for the elimination of railroad nuisance. The main objective is to construct a valid counterfactual. To overcome the identification challenges of the standard difference-in-differences approach, we use a two-stage methodology involving the synthetic control method.

Our results indicate large positive effects of the railroad tunneling on property prices. We find that the price elasticity with respect to the distance to the railroad would have been about 5 percentage points lower in case Delft would not have tunneled its railroad.

CPB Discussion Paper

Koen van Ruijven, Joep Tijm

June 2021

Doi: https://doi.org/10.34932/kkq7-br14

Housing Market Effects of a Railroad Tunneling: Evidence from a quasi-experiment

Koen van Ruijven and Joep Tijm[1]

June 2021

Abstract
This paper exploits the railroad tunneling in Delft (the Netherlands) as a quasi-experiment to estimate the willingness to pay for the elimination of railroad nuisance. To overcome the identification challenges of the standard difference-in-differences approach, we use a two-stage methodology involving the synthetic control method. Our results indicate large positive effects of the railroad tunneling on property prices. We find that the price elasticity with respect to the distance to the railroad would have been about 5 percentage points lower in case Delft would not have tunneled its railroad. About half of the effect already capitalized as soon as the tunneling preparations started. Finally, we provide evidence for sorting effects. The railroad tunneling is associated with a significant increase in the socio-economic status of neighborhoods in close proximity to the railroad.

JEL classification: R38, R58

Keywords: Railroad nuisance; Railroad tunneling; Residential property prices; Synthetic control method.

[1] We are grateful to Henri de Groot and Wolter Hassink for numerous insightful discussions. The authors are affiliated with the CPB Netherlands Bureau for Economic Policy Analysis (CPB). Views expressed in this article are our own and do not necessarily reflect those of the CPB. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector. Corresponding author: Koen van Ruijven (k.g.van.ruijven@cpb.nl).

1. Introduction

This paper empirically investigates the impact of the environmental and the social effects of railroads on residential property prices. Ever since the major construction of railroads in the 19th century, scholars from a variety of disciplines have discussed the impact of railroads on several housing market outcomes, such as rents (Ricardo, 1809) and residential segregation (Engels, 1845). The literature predominantly discusses two conflicting effects of railroads. On the one hand, railroads generate accessibility effects, as it enables residents to commute to other places. On the other hand, the use of railroads by trains gives rise to environmental effects in the form of noise and air pollution. Moreover, as it cuts through cities, railroads could lead to social effects such as community severance. In order to assess and quantify the conflicting welfare effects, the capitalization of the effects in residential property prices can be exploited empirically.

Although the positive accessibility effects of railroad stations on residential property prices have been established well in the literature (Debrezion et al., 2011; Donaldson & Hornbeck, 2012; Redding & Turner, 2015), little is known about the environmental and social effects of railroad infrastructure. A number of studies provide evidence for a negative relationship between railroad proximity and residential property values (Al-Mosaind, Dueker, & Strathman, 1993; Debrezion, Pels, & Rietveld, 2010; Nelson, 1992; Poon, 1978). Yet, the relationship found in the literature is generally weak.

The weak results might be explained by a number of econometric identification challenges that could plague the estimation of the environmental and social effects of railroads on residential property prices. First, it is likely that the estimated relationship between the environmental and social effects of railroads, proxied by the distance of a residence to a railway track and residential property prices, is biased due to omitted variables. For instance, neighborhoods close to the railway track may differ from neighborhoods at larger distances in terms of unobserved characteristics, such as safety and historical aspects, that simultaneously affect residential property prices. The second problem is that variation in railway externalities seldomly are exogenous. Specifically, investments to mitigate or even eliminate railway externalities, such as noise barriers and tunnels, may be targeted at urban areas with high or low growth. A third problem is that the accessibility effects of railroads are often hard to separate from the environmental and social effects.

To the best of our knowledge, the only existing empirical study that attempts to address these identification challenges is a study by Diao et al. (2016). These authors have attempted to quantify the environmental and social effects of railroads using a railroad tunneling in Singapore as a quasi-experiment. This study addresses endogeneity issues using a difference-in-differences framework, comparing trends in residential property prices within 400 meters of the railroad, both before and after the tunneling, relative to price trends at larger distances. Diao et al. (2016) find the railroad tunneling in Singapore led to an additional increase in residential properties by 13.7% within 400 meters of the tunneled railroad.

Even with a difference-in-differences strategy, one faces the empirical challenge that the counterfactual of the negative willingness to pay to live near the railroad in absence of the railroad tunneling is unknown. That is, there may be a number of reasons why residential properties at greater distances from the railroad may not necessarily be a valid counterfactual. For example, the negative willingness to pay to live near railroads may have decreased over time because of novel railroad technologies which mitigate the nuisance of railroads. On the other hand, the negative willingness to pay may have increased because of the higher frequency of trains passing the railroads.

The primary contribution of our paper is to provide an alternative empirical strategy that aims to quantify the environmental and social effects of railroads, while addressing these identification challenges. To this end, we employ the railroad tunneling in Delft, a medium sized city in the Netherlands located in between two of the four biggest cities in the Netherlands, Rotterdam and The Hague, and hosting a university. The railroad tunnel was built after reaching a consensus about conflicting interests between the Dutch government and residents in Delft. On the one hand, the Dutch government aimed to expand railroad capacity in Delft since the railroad capacity was insufficient to accommodate the perceived rise in railroad use of over 50 percent (Dutch Railways, 1988). On the other hand, residents in Delft conceived the expected growth in railroad operation to be paired with negative railroad nuisance, which led to "not in my backyard" reactions. In order to fulfil both interests, local authorities decided to construct a 2,300 meter long tunnel, predominantly funded by the Dutch government.

To overcome the empirical challenges, we compare the housing market trends between Delft (the treatment city) and a number of control cities that did not tunnel their railroad. This strategy requires that Delft did not tunnel its railroad for economic concerns, which we will demonstrate both qualitatively and quantitively. In order to develop valid a valid counterfactual city, our identification approach proceeds in two stages. In the first stage, for each city separately, we estimate the willingness to pay to live near the railroad. The output of the first stage is used in our second stage where we use the synthetic control method to construct a counterfactual of the city of Delft. Instead of using an average of control cities, the synthetic control method provides a data-driven algorithm to construct a convex combination of control units based on their similarity to the city of Delft during the pre-tunneling period, both in terms of the pre-treatment trend in the outcome variable (the willingness to pay to live near the railroad), and the covariates relevant to this outcome variable (Abadie et al., 2010; 2015).

Our results indicate large positive effects of the railroad tunneling on property prices. About half of the effect already capitalizes as soon the tunneling preparations started. Our results suggest that the price elasticity would have been about 5 percentage points lower in case Delft would not have tunneled its railroad. To verify the credibility of our results, we show that these results are robust to a set of sensitivity analyses, including the use of cross-validation in the synthetic control methodology, in-time placebo's, and alternate inputs.

As predicted by our conceptual framework, we find that the elimination of environmental effects of railroads also induced social effects in the form of sorting of households. Our results indicate that the railroad tunneling causes a significant increase in the socio-economic status of neighborhoods in close proximity to the railroad. Moreover, the results indicate that the railroad tunneling caused a significant increase in the percentage of residents within the age group of 25 to 44, at the expense of residents of the age group 45 years and older, especially those which have reached the age of retirement (65 years and older). This suggests that residents within the older age groups respond less strong on a change in the environmental effects of railroads.

Our paper contributes to the literature using quasi-experimental research designs to quantify the effects of environmental goods. This literature has used different strategies. For example, Chay and Greenstone (2005) exploit the differential effect of air pollution regulations between different counties in the US. Boes and Nüesch (2011) employ an (unexpected) introduction of German regulation of noise levels in Switzerland (Boes & Nüesch, 2011). Other strategies use actual adjustments in the physical infrastructure, like Ossokina and Verweij (2015) who use the opening of a bypass highway as source of exogenous variation that reduced urban traffic. Our empirical strategy is most closely related to Diao et al. (2016), who use a railroad tunneling in Singapore as a source of exogenous variation in the environmental effects of railroads. We further complement this literature by proposing an alternative empirical strategy that involves a two-stage between-city design. In this way, we are able to account for the possibility that residential properties at larger distance locations from the physical infrastructure are not necessarily a proper counterfactual for those residential properties in close distance locations. Moreover, we show that a railroad tunneling leads to sorting effects which may bias the willingness to pay for environmental effects.

The structure of the paper is as follows. Section 2 provides a discussion of the background and goals of the railroad tunneling in Delft. Section 3 combines the insights of hedonic theory and the Alonso-Muth-Mills framework to provide a conceptual framework of the results of a railroad tunneling. Section 4 describes the data. Section 5 presents the identification strategy. Section 6 discusses the estimation results demonstrating estimates of the willingness to pay to eliminate urban railroad nuisance. Section 7 concludes.

2. The railroad tunneling in Delft:- background

2.1 Railroad tunneling

This paper focuses on the railroad tunneling in Delft, a medium-sized city located in the southwest of the Netherlands. Delft comprises around 100,000 inhabitants (Statistics Netherlands, 2018) and is regarded as one of the centers of technological research in the Europe. It provides home to the Technical University Delft and possesses a considerable amount of historic scenery in its historical city center originating in medieval times (with official city rights being granted in 1246). Since 1965, trains used the elevated railroad at the northern side of the station. The elevated track was built to accommodate travel modes from the eastern and western side of the city. Several passing points were available underneath the elevated track (see left-hand side Figure 1). At the southern side of the station, a small tunnel was built to accommodate within-city travel (Van Duin & Wilms Floet, 2005).

Figure 1: *The quality of the public space* **-** The situation before and after tunneling

Since the 1980s, the passengers' use of railroads in the Netherlands gradually increased, which induced the Dutch Railroad company to explore opportunities to expand the number of railroad tracks in Delft (Nederlandse Spoorwegen, 1988). Historically, the railroad segment north and south of Delft possessed four railroad tracks, while Delft only possessed two tracks. For this reason, Delft was seen as a potential bottleneck for the perceived surge in railroad use during the 2000s, mainly between The Hague and Rotterdam (see Table 1). After the publication of a series of feasibility studies, the main urban planner concluded that "*A four-track railroad tunnel of 2300 meter was the only sustainable solution to accommodate the rise in railroad use* (SOVI, 1993; Witteveen+Bos, 2003)."

Table 1: *The perceived rise in daily railroad use in Delft*

	2001		2015	
	Passenger trains	Cargo trains	Passenger trains	Cargo trains
Day	217	2	358	5
Evening	66	0	103	3
Night	59	1	73	3
Total	342	3	534	11

Source: Akoestisch Spoorboekje (2002)

The decision to construct a four-track railroad tunnel was made in February 2006. The municipality council of Delft demanded funding guarantees by the Dutch government, which it received at the end of 2005 (Ministry of Housing, Spatial Planning and the Environment, 2009). The construction started in 2009 with the first preparations. The tunnel was officially opened in 28 February 2015 and was named after William of Orange (ancestor of the Dutch monarchy and assassinated in Delft in 1584). The remainder of 2015 was used to demolish the remaining railroad tracks left unused at the surface. Figure 2 presents a timeline.

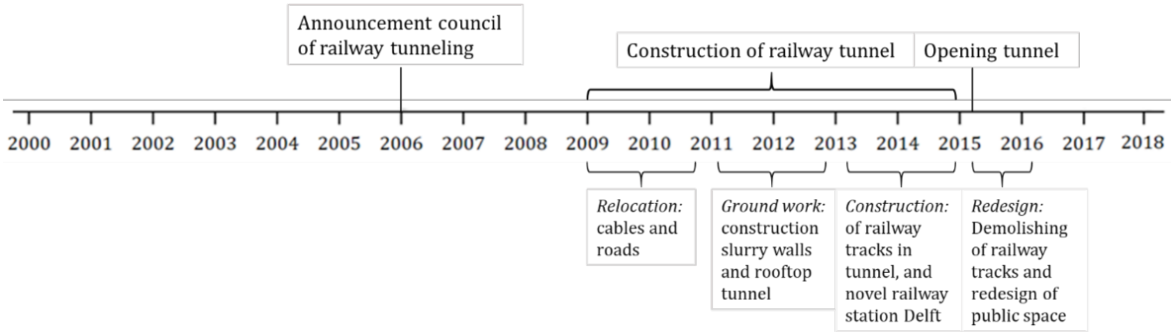

Figure 2*: Timeline:* A chronology of events

2.2 Railroad tunneling goals

According to the environmental impact assessment, the railroad tunneling comprised three main objectives (Witteveen + Bos, 2005). First, the tunnel had to expand Delfts' railroad capacity. To this end, the municipality of Delft received about 500 million euro of the Dutch government. Second, the tunnel had to eliminate the growing nuisance stemming from railroad use. Noise pollution and the associated vibrations were perceived as the most severe railroad externality in Delft. Residential properties in direct proximity to the tunnel were confronted with noise levels as high as 85 decibels. This is shown on the left hand side of Figure 3, which shows the energetic level of railroad noise during 2006 in Delft. The rise in railroad use was expected to increase this level further.[2] The right hand side of Figure 3 presents evidence of the complete elimination of railway noise nuisance along the tunneled railroad in 2016. The effect of the railroad tunneling on air pollution was not considered to be significant. This was because of the transition of diesel fueled towards electricity charged trains during the 2000's. Still, the environmental impact assessment expected pollution to decrease since railroad usage coincided with litter from toilets, and particles that wore of the train wheels. These forms of pollution fell on, and off, the elevated railroad on the cars parked underneath.

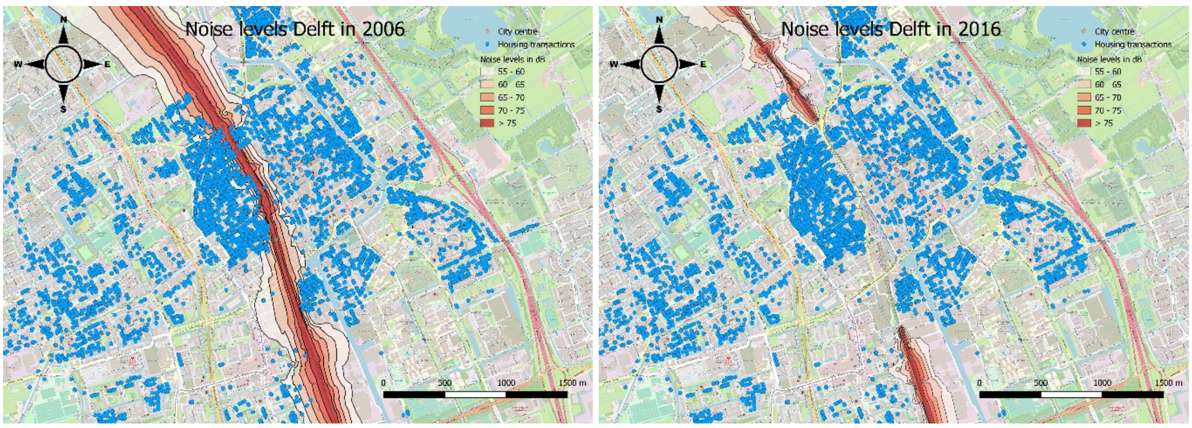

Figure 3: *Railroad noise nuisance:* The situation before and after tunneling
Notes: Both panels in the figure depict the annually averaged energetic level of railroad noise, measured in L_{den} (level day, evening, night). Noise transmitted during the evening and night are 'penalized' with a 5db and 10db increment factor, respectively. The data source is ProRail Netherlands, the depicted background is from OpenStreetMap.

[2] Given a number of railroad innovations, such as rail dampers and the use of trains with lower noise emissions, this expectation turned out to be false. On average, railroad noise decreased by 3 dB during the time period 2006-2016.

Third, the railroad was tunneled to enhance the quality of the public space. The left-hand side of Figure 1 shows the public quality before the tunneling took place. The figure clearly conveys that the railroad could be experienced as a visual and psychological barrier. Residents at either sides of the railroad were unable to see past the railroad and as such, the railroad was the epitome of the urban fabric barrier between the west and east side of Delft. The lack of sight underneath the elevated track also enabled people to meet there, which caused feelings of unsafety for passengers traversing the railroad by night. This changed substantially after the tunnel construction was completed. The freed space was used to expand the canonical canal structure of Delft, complemented with greening facilities. The new situation is exhibited in the right-hand side of Figure 1. In essence, the demolition of the railroad reconnected the urban fabric of Delft.

We note that the railroad was not tunneled in order to improve the travel times *within* Delft. Since the railroad could be traversed easily ex-ante of the tunneling, the travel times remained fairly constant, both for cars, for cyclists, and for public transport (i.e. bus and tram). This is illustrated quantitatively in Appendix Figure A1 and Table A6. The within-city travel time benefits therefore do not constitute part of the valuation for the elimination of railroad nuisance that we estimate in this paper.

3. Conceptual framework

3.1 The theoretical implications of a railroad tunneling in a hedonic price model

This paper examines the effect of a railroad tunneling on the residential property market. The absence of a market for railroad nuisance and the quality of urban public space prohibits a direct assessment of the willingness to pay for a railroad tunneling. To circumvent this problem, economists have traditionally used the hedonic price method in the housing market to infer the value of nonmarket amenities (e.g. Greenstone & Gallagher, 2008; Boes & Nüesch, 2011). Conceptually, the hedonic price method considers the value of a residence to be described by a vector of n characteristics $P_i = P(x_{i1}, x_{i2}, \dots, x_{in})$, which comprises structural characteristics (e.g. floor space), spatial characteristics (e.g. proximity to the central business district), and neighborhood characteristics (e.g. socio-economic status of residents).

The idea that underlies the hedonic price method is that for each of the characteristics there exist a locus $P(\cdot)$ between residence prices and the continuum of the nth characteristic. This locus, the *hedonic price schedule*, is formed between the interactions of consumers and suppliers of residences. Each point along the hedonic price schedule can be interpreted as the point of tangency between the marginal willingness to pay of consumers for the *n-th* characteristic (the bids) and the producers' marginal costs of producing one extra unit of n (the offers).[3] Put another way, one can infer the marginal implicit price for each of the characteristics by taking the partial derivate of P with respect to the nth characteristic, keeping all else constant: $P_{x_n} = \frac{\partial P}{\partial x_n}$.

Now that the building blocks of hedonic pricing are formalized, we can consider the hedonic price schedule in the residential property market for railroad nuisance. Given the

[3] A more detailed description of the maximization problems of consumers and suppliers of housing is listed in Appendix Section A3 (Hedonic price model).

absence of an explicit market for railroad nuisance, economists have traditionally assumed that the distance of residences towards a railroad (d_1) provides a reasonable proxy for the intensity of experienced railroad nuisance (e.g. Nelson, 1992; Diao, 2016). That is, assuming that the intensity of railroad nuisance is denoted by the first characteristic x_1, the marginal price of railroad nuisance implicit to the overall value of the residence is $P_{x_{1(d_1)}}$, *ceteris paribus*. As consumers dislike railroad nuisance, they will only accept more railroad nuisance if and only if they are compensated for the nuisance in terms of lower residential property prices.

Figure 4 conveys this idea in a hedonic price schedule. For the sake of argument, the figure assumes four types of consumers. Each of the four consumers chooses a certain 'intensity' of railroad nuisance to the point at which their indifference curve (or bid function) touches the marginal price of one unit extra nuisance. The variation in preferences and/or incomes translates into the chosen variation in distance towards the railroad. That is, some consumers sort themselves in locations close to the railroad, and in return for the nuisance, get compensated with lower residential costs. Other consumers sort themselves at larger distances because of their preferences for less nuisance. Yet, these consumers pay higher residential prices. The resulting gradient of the hedonic price schedule with respect to the distance could therefore be interpreted as the equilibrium premium that compensates consumers for accepting the nuisance (including the inferior public space).

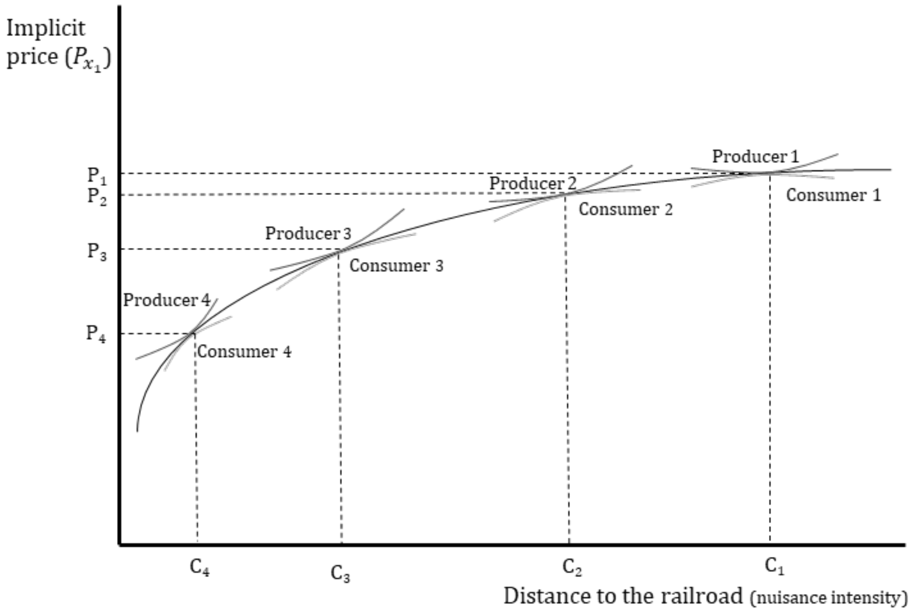

Figure 4: *The Hedonic price schedule in the residential property market for railroad nuisance*

Next, consider a situation that completely eliminates all railroad nuisance: a railroad tunneling. This results in the situation that the nuisance intensity is no longer a function of the distance to the railroad. Assuming that consumer 1 had chosen its nuisance intensity to the point it was almost negligible, the novel, flat gradient shifts upwards just above the point of tangency of consumer and producer 1. The prevailing owners that initially sorted themselves at a point along the hedonic price schedule with nonzero railroad nuisance experience a windfall gain. The price of the residence of consumer 4 for instance, increases from P_4 to P_1. Since the tangency of the hedonic price schedule no longer corresponds to their preferences, some homeowners will choose to migrate towards a residence with an initially chosen optimal level of P and C (i.e., in a location with similar nuisance levels to what they experienced before, keeping utility unchanged). Not all homeowners will migrate, however, due to the presence of moving costs. But overall, the neighborhoods in close proximity to the railroad will constitute a novel residential population with a low(er) tolerance for railroad nuisance.[4]

As theorized in the previous paragraphs, a railroad tunneling has two expected impacts. First, residential property prices will increase near the tunneled railroad. Second, consumers will respond to the railroad tunneling by migrating in, or out, of the affected area near the railroad tunnel (sorting based on preferences).

4. Data and descriptive statistics

The analyses conducted in this paper are based on a dataset that includes information about residential property transactions ranging over the period 1995 to 2018. The set of data are drawn from administrative records of the Dutch Association of Real Estate Brokers and Experts[5] and contains over 200,000 transactions allocated in Delft and in locations that we later use as control cities.[6] Each record comprises information about a dwelling transaction, including the transaction price (in 2015 euros), the exact address, the date of sale, and a large number of characteristics related to the dwelling, such as the dwelling type, the maintenance quality, the number of rooms, the construction year, and living space. We computed the Euclidian distance of each housing transaction to the railroad using geographic information system software (ArcGIS). A similar procedure was used to compute the Euclidean distance of the residence to various amenities (e.g. the CBD) and disamenities (e.g. highway). Table 2 reports the descriptive statistics.

[4] Note that on the longer term, there may be adjustments to the supply of housing which correspondingly will decrease the (positive) effect on residential property prices. In this paper, we allow for supply responses in our empirical strategy.
[5] About 4 out of 5 of all transactions in the Netherlands are conducted by brokers attached to this organization. Despite the incomplete coverage, the dwelling sale price data of this brokers organization turn out to be representative for the remaining 20 percent of the transactions (Statistics Netherlands, 2019).
[6] The data cleaning steps are listed in Appendix Table A1.

Table 2: *Residence-Specific Descriptive Statistics*

	Mean	Standard deviation	Description
Transaction price	224,287	121,576	Transaction price of the residence, deflated in 2018 euros
Structural characteristics			
Floor space (m^2)	108.14	40.22	The number of square meters floor space of the residence
Living space (m^3)	321.78	134.38	The number of cubic meters living space of the residence
Number of rooms	4.09	1.39	The number of rooms in the residence
Number of floors	2.20	0.91	The number of floors in the residence
Residence type			
i) Apartment			Dummy variable that equals one if the residence is an apartment and…
Downstairs	0.05	0.21	located downstairs of a building
Upstairs	0.07	0.25	located upstairs of a building
Porch	0.15	0.35	located in a porch flat
Gallery	0.09	0.29	located in a gallery flat
Other	0.03	0.18	either located in a maisonette, or comprising both the upper and lower floor
ii) House			Dummy variable that equals one if the residence is a house and…
Intermediate	0.37	0.48	located in between other houses
Corner	0.13	0.33	located at a corner
Semi-detached	0.08	0.27	semi- detached from other houses
Detached	0.04	0.20	completely detached from other houses
Dwelling quality			
Maintenance quality inside (1-9)	6.93	1.13	Quality of maintenance inside the dwelling, ranging from bad, bad (1) to excellent (9)
Maintenance quality outside (1-9)	6.97	0.98	Quality of maintenance at the exterior of the dwelling, ranging from bad (1) to excellent (9)
Maintenance of garden (1-5)	3.29	0.72	Quality of maintenance of the garden, ranging from no garden existent (1) to very-well-kept (5)
Insulation quality (0-5)	1.77	1.61	Dwelling has no isolation (0), one-layered isolation, two-layered, three-layered, four-layered, or 5-layered (or full) isolation
Spatial characteristics			
Distance to railroad	772.98	504.82	Logarithmic distance to railroad
Distance to highway	2692.72	1714.53	Logarithmic distance to closest highway
Distance to CBD	1274.47	757.87	Linear distance to CBD
Distance to railroad station	984.53	473.80	Linear distance to railroad station

Notes: The number of observations is 203,845. The non-reported variables included in the $X_{i,t}$ vector of equation (2) include:, whether the dwelling is located next to a park or open water, whether the dwelling has a central heating system, whether it is a listed building, and the building period of the dwelling (in unequally distributed time-periods).

5. Empirical framework

The aim of this study is to develop a valid counterfactual showing the price trajectory near the railroad in case the railroad in Delft would not have been tunneled. In the development of a counterfactual, a conventional approach is to estimate a difference-in-differences strategy. With regards to a railroad tunneling, we would compare trends in residential property prices by affected regions close to a railroad to the trends in residential property prices by non-affected regions located further away, both before and after the tunneling.

There are, however, a couple of reasons why non-affected residential properties located in regions further away may not be a valid counterfactual for affected residential properties close to the railroad. One first reason is that the negative willingness to pay to live near the railroad may have decreased over time because of novel railroad technologies which mitigate the nuisance of railroads. On the other hand, a second reason is that the negative willingness to pay may have increased because of the higher frequency of trains passing the railroads. For both these reasons, non-affected residential properties located at larger distances may not be a valid counterfactual for affected residential properties.

5.1 Two-stage identification approach

We propose an two-stage approach to estimate the causal impact of a railroad tunneling. As a solution, we exploit the notion that cities with otherwise very similar characteristics as Delft may have experienced similar observable (and unobservable) trends that are related to the willingness to pay to live near a railroad. These cities however did not tunnel their railroad, and could therefore be used as potential counterfactual cities. That is, the counterfactual that will illustrate the willingness to pay to live near the railroad in Delft in case the railroad had not been tunneled. We use a synthetic control design *between* cities to construct this counterfactual of the city of Delft. One benign characteristic of the synthetic control method is that it allows for the presence of unobserved confounders to vary over time.[7] As long the synthetic control method is able to construct a counterfactual that matches Delft in observed variables, and assuming that the counterfactual city experiences a similar pre-intervention trend in outcome variables, the synthetic control estimator is unbiased.

Stage 1: Hedonic price method

In stage 1 of the empirical strategy, we model the willingness to pay to live near railroads. We perform this exercise both in Delft $(J = 1)$ and in each of the control cities $(J = 2, \ldots, 28)$, separately.[8] Abstracting from potential sorting effects, we employ the following hedonic price function:

$$\log p_{i,t} = \beta X_{i,t} + \rho S_{i,t} + \gamma_r Z_r + \tau_t T_t + \delta_t \log Distance_i * T_{i,t} + \varepsilon_{i,t} \quad (1)$$

where $X_{i,t}$, $S_{i,t}$, Z_r and T_t again respectively denote a vector of dwelling characteristics, a vector of spatial characteristics, postal code fixed effects and year fixed effects.[9] The treatment variable of interest is the negative of $\log Distance_i * T_{i,t}$. For each city $(J = 1, \ldots, 28)$ it captures the price elasticity with respect to the distance to the railroad. The price elasticity is estimated in each two years using interactions with year dummies $T_{i,t}$. In this way, we are able to identify δ_t over time: the parameter that depicts the willingness to pay for a percentage point change in the distance to the railroad in period t.[10]

The model assumes a log–log relationship between railroad nuisance and the residential property prices. The assumption of a log-log relationship was made both for theoretical and empirical reasons. There are three theoretical reasons. First, noise is commonly expressed in a logarithmic scale (measured as decibels), and therefore the nuisance increasingly becomes smaller as the distance from the railroad increases (see Figure 3). Second, the disamenity of the deficient public space (including the waste) is primarily experienced within a few hundred meters from the railroad. Third, and related to the prior reasons, community severance caused by infrastructure seems to be especially strong within the first few housing blocks, and dissipates

[7] Abadie et al. (2010) demonstrates this result using a motivating model for the synthetic control (pp 494-496).

[8] In section 5.1, we discuss the selection criteria used to exclude cities from the donor pool of control cities.

[9] Section 4 provides a more detailed description of the vector of $X_{i,t}$ and $S_{i,t}$ covariates.

[10] We chose to estimate the price elasticity in two year periods for two reasons. First, it increases the power of the estimation results. Second, it smoothens the pattern in point estimates, which leads to a lower prediction fit (root-mean squared prediction error) of the synthetic control method in stage 2 of our empirical strategy.

as the distance (to the railroad) becomes larger (Anciaes et al., 2016). To this end, it is desirable to use a model where a change in the distance to the railroad has a proportional effect on residential property prices.

Next to the theoretical reasons, we tested for the specification of the functional form using Ramsey reset tests. These tests confirmed that the specification of a log-log relationship suited the data significantly better than a log-linear relationship. Further nonlinearities in the relationship were not detected.[11]

Stage 2: The synthetic control method

In stage 2, we use the model results of stage 1 (the δ_{Jt}'s) as input for the synthetic control method (Abadie et al., 2010; Abadie et al., 2015). Instead of comparing Delft with the average of other control cities, the synthetic control method provides a data driven procedure to construct the missing counterfactual using a convex combination of control units. It allocates nonnegative weights $W = w_j \geq 0 \; \forall \, j \; \in \{2, \dots, J+1\}$ to cities in the donor pool that add up to one ($\sum_{j=2}^{J+1} w_j = 1$). The effect of tunneling φ_{1t} is estimated as follows:

$$\varphi_{1t} = \delta_{1t}^{I} - \sum_{j=2}^{J+1} w_j^* \delta_{jt} \tag{2}$$

In theory, any choice for W could produce a different synthetic control for Delft. However, for the synthetic control estimator to be unbiased, we are required to choose W such that the synthetic control has (a) a similar pre-intervention trend in the outcome variable[12], and (b) the pre-tunneling characteristics Z_{0m} are similar to those of Delft Z_{1m} [13]:

$$(a)\; \hat{\delta}_{1,T^0}. = \sum_{j=2}^{J+1} w_j * \hat{\delta}_{j,T^0} \qquad\qquad (b)\!: Z_{1m} = \sum_{j=2}^{J+1} w_j^* Z_{0m} \tag{3}$$

The synthetic control estimator can be implemented such that condition (3) holds approximately. In practice, the method achieves this by the minimization of the distance between the $1 * k$ vector of pre-tunneling characteristics of Delft Z_{1m} and the $(k * J)$ matrix of the $m = 1, \dots, k$ pre-tunneling characteristics of the synthetic control (X_0). The distance is minimized with respect to W and to the metric V.

$$W^*(V) = arg\,\min(X_1 - X_0 W)' V (X_1 - X_0 W) \tag{4}$$

where W satisfies ($w_j \geq 0 \; \forall \, j \; \in \{2, \dots, J+1\}$, ($\sum_{j=2}^{J+1} w_j = 1$), and V satisfies $v_m \geq 0$, $\sum_1^k v_m = 1$.

[11] Moreover, the assumption of a log-log relationship was confirmed by using a difference-in-differences specification using 200 meter distance bands from the railroad. The analyses compared residential property price trends within these distance bands relative to the residential property trends at a distance larger than 1000 meters from the railroad (tunnel). The point estimates of the distance bands indicated the largest effects were found at close distances from the tunnel (within 200 meters). After 600 meters, the point estimates were still significantly positive, but small.

[12] T^0 list the number of pre-intervention periods.

[13] $m = 1, \dots, k$ refers to the number of pre-intervention characteristics (or predictors).

The V-matrix captures the relative predictive power of each of the pre-tunneling characteristics. It is therefore essential to assign larger v-weights to pre-tunneling characteristics that have a larger predictive power on the outcome. Instead of a subjective choice, we determine the v-weights in a data driven manner such that the mean squared predictor error (MSPE) of the pre-intervention outcomes are minimized. Specifically, following Abadie et al. (2010), we employ a nested iterative optimization procedure that searches among all V-matrices and sets of w-weights in order to select the (convex) combination of control units that has the lowest mean squared prediction error. In this sense, we are able to observe whether the synthetic control estimator adheres to the (parallel trend) assumption (a) in condition (3).

An advantage of the synthetic control method is that the time-varying omitted variables are not restricted to be constant over time. This is a relevant feature since we are unable to effectively control for unobserved spatial trends stage 1 and 2 of our empirical strategy. Abadie et al. (2010) show that as long as the synthetic control satisfies condition (3b), and it matches the pre-intervention periods of the treatment unit during a sufficient matching window (3a), then we may assume the time-varying omitted variables to be similar in Delft and the synthetic control. To this end, we use the following key identifying assumption that underlies the synthetic control estimator: *as long as a convex set of cities is able to approximate Delft during pre-tunneling treatment, then any subsequent discrepancies should reflect the effects of tunneling.*

The standard synthetic control procedure is not well-suited for direct inference. The procedure produces one estimate of the tunneling effect, and as a result, we are not able to infer whether the estimate differs statistically from zero or not. Recently, the literature has developed a number of inferential techniques to circumvent this problem (Abadie et al., 2010; Acemoglu et al., 2016; Saia, 2017) by proposing a number of (falsification) permutation exercises. We use the inferential technique as proposed by Saia (2017), which is based on subsampling methods. To this end, we randomly draw 300 subgroups, each using half of the original donor sample of control cities.[14] Consequently, with these subsamples, we estimate a number of 300 synthetic counterfactual estimates. By averaging the estimates, we are able to infer whether the estimate differs statistically from zero.

The inputs used in the synthetic control estimator

We have made three decisions regarding the inputs used in the synthetic control estimator:

1. We designate the start of the intervention $t = T_0 + 1$ to the first period that the (preparatory) construction work at the tunnel started. We admit, however, that the exact timing of the potential capitalization effect is unsure. On the one hand, it is possible that residents anticipate on the livability effects in the future, and for this reason, the tunneling may capitalize during, or even before the construction work is conducted (i.e. the present discounted values). On the other hand, the tunneling construction may be accompanied by nuisance, which decreases likelihood of a capitalization effect before completion of the

[14] The number of synthetic control subsampling procedures and the sample size of the donor pool were chosen arbitrarily. We however, have performed a number of robustness exercises using different values. The exercises indicate that the variation among results is very small.

tunnel. In section 6.2, we show that our results are robust to changes in the designation of the intervention period.

2. We implement the synthetic matching procedure using a unique predictor dataset of six predictors. As a baseline, we include two characteristics that measure typical disamenities that are associated with living close to a railroad (noise pollution, air pollution). In addition, we include four city-specific characteristics inherent to the neighborhoods close to the railroad in Delft (i.e. the employment density, the amount of cultural heritage, the percentage of 15-24 year olds, and overall expectation about the neighborhood in the future). Table 3 lists the predictors that are used to perform the synthetic matching procedure. The predictors are averaged over the period 2005-2006, a few years before the first tunneling preparations were conducted.[15] The average of the donor pool in Table 3 is based on a sample that contains 27 control cities. The data sources are included in the Appendix. In Section 6.2, we also show that the results are robust to changes in the set of predictors. [16]

Table 3: *Predictor means*

	Delft	Average Donor Pool
Noise level (in decibels)	69.78	66.94
Air quality (particular matter)	27.16	25.89
Density (ln number of jobs)	10.80	10.78
Cultural heritage (in km^2)	1.02	0.44
Percentage residents aged 15-24 years	0.139	0.110
Expectation development neighborhood	1.92	1.82

Notes: The table presents the predictor variable similarly between Delft and average of the donor pool of control cities. The values of the predictors are computed within close distances of the railroad, except for the density and cultural heritage predictor, which are computed within the rectangular shapes as presented in Figure 5. The predictors are averaged over the period 2005-06. Appendix Table A2 provides a detailed description of the predictor variables.

3. We take a number of steps in order to ensure the donor pool constitutes plausible 'comparable' control cities. To start, we discard cities that have conducted activities that may have affected their urban railroad externalities. Since there were no other cities that have tunneled their railroad, this specifically refers to the cities that have conducted transit-oriented developments around transit stations.[17] Subsequently, we restrict the donor pool to cities that have a similar urban structure relative to Delft. The urban structure of Delft is presented in Figure 5. The green line depicts the 2300 meter tunnel. The largest (Euclidean) distance of a dwelling from the railroad at the western side of the city is 1700 meter, while the largest (Euclidean) distance is 1500 meter for the eastern side the city. We use these exact rectangular measurements in our control cities, depicted

[15] In section 6.2, we test for the robustness of our results using different sets of novel predictors in the synthetic control estimator. The inclusion of these predictors does not alter our results substantially.

[16] Each of the used predictors that we use in the synthetic control has been either theoretically or empirically validated in the literature as determinant of the willingness for residential properties.

[17] The following cities were excluded: Amsterdam, Arnhem, Breda, Den Bosch, Den Haag, Maastricht, Rotterdam, Tilburg and Utrecht.

at the right hand side of Figure 5.[18] We thus discard cities that did not experience annual dwelling transactions within these rectangles. Finally, we restrict the donor pool to control cities with (a similar amount of) dwellings at both sides of the railroad.

Using these restrictions helps to ensure cities have similar characteristics in the donor pool. In this way, we are able to avoid interpolation biases (Abadie et al., 2015).

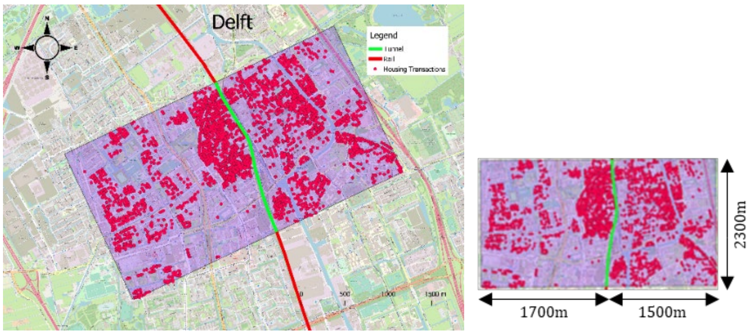

Figure 5: *The rectangular measurements in Delft*
Notes: The depicted background in the figure is from OpenStreetMap.

The exogeneity of the railroad tunneling in Delft

The credibility of our empirical framework hinges on the question whether the railroad tunneling in Delft could be classified as exogenous. Specifically, investments in relocations (tunnels) of transportation infrastructure are usually specifically targeted at urban areas with high (or low) growth. To address this endogeneity concern, we have to ensure economic growth in Delft was not systematically related to the distance to the tunnel.

We build on the conjecture that economic growth in Delft was never considered as an objective (Gemeente Delft, 2005). The principal reason why Delft received funding from the Dutch government was the overall growth in railroad use of the network. The prevailing two track railroad infrastructure in Delft was perceived to be unable to accommodate the rise in railroad traffic. Other cities also experienced similar increases in railroad traffic. Yet these cities already possessed sufficient railroad capacity to accommodate this rise. As a result, Delft was designated as a potential bottleneck that could hinder the growth of the entire network (Nederlandse Spoorwegen, 1988).

Figure 6 shows that the railroad network improved similarly over time in Delft, in all other railroad stations in the Netherlands, and in a selection of (comparable) cities. The figure shows the indexed evolution of generalized travel times over time (the level of service of each railroad station[19]). Even though the index improved slightly more in Delft during the time period 2004-2007, the index improved slightly more in comparable cities during later time periods. Moreover, the network did not experience a sudden spike in its improvement after the tunnel was opened

[18] The (relatively short) Euclidean distances in Delft are driven by the presence of highways at both the eastern and the western side of the railroad. In order to ensure comparability, we therefore control for the closest distance to a highway of a dwelling, both in Delft and in the control cities.

[19] The level of service indicator is a measure of the weighted generalized travel time, comprising a measure of in-vehicle time, the frequency, and a transfer penalty.

(since 2015).[20] This is confirmed in Table 4, which demonstrates whether the index of Delft differs significantly from all other railroad stations in the Netherlands (column 1), and from railroad stations of comparable cities (column 2). None of the indexed level of service differences are significantly different from zero, both before and after the tunneling. This indeed suggests that the investment was not targeted to enhance travel times specifically in Delft.

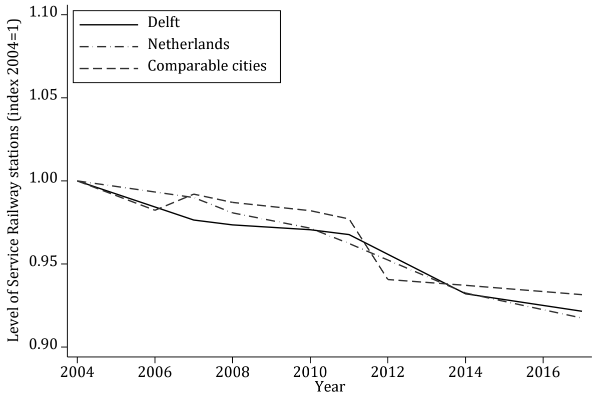

Figure 6: *The Comparable Growth in Level of Service of Railroad Stations*

Table 4: *Does the Level of service of the railroad station of Delft differ significantly?*

	(1) Netherlands	(2) Comparable cities
Delft	-0.012	0.009
	(0.037)	(0.026)
Delft * 2009-10	0.008	-0.007
	(0.062)	(0.051)
Delft * 2011-11	-0.007	-0.004
	(0.065)	(0.052)
Delft * 2013-14	-0.007	-0.017
	(0.064)	(0.051)
Delft * 2015-16	0.016	-0.031
	(0.061)	(0.055)
Delft * 2017-18	0.021	-0.036
	(0.063)	(0.057)

Notes: The dependent variable is level of service of a railroad station. The specification includes year fixed effects. Significance at the 10%, 5% and 1% level respectively denoted by */**/***.

Notes: Figure 6 presents the indexed change in weighted generalized travel times of railroad-stations over the time period 2004-2018. The dash-dotted line includes all transit-stations in the Netherlands, except Delft. The dashed line only includes cities with comparable city-characteristics as Delft, such as the number of inhabitants, the employment density, and other socio-economic characteristics. We describe the computation process of the level-of-service of transit-stations in Appendix table A4.
Data source: Authors' calculations based on data by ProRail Netherlands

Next, to further validate the exogeneity argument, we consider changes in employment levels. In particular, we check whether employment growth was higher in Delft at higher proximity to the tunnel, relative to the employment growth further away. If for example, the growth was larger at lower distance to the tunnel, this would suggest the investment is specifically targeted to promote economic growth. The employment growth comparison is exhibited in Figure 7 over the time period 2000-2018.

A comparison between the solid black line and the dashed black line indicates that the changes in employment levels in Delft were roughly similar across space. However, since 2014 employment levels improved slightly more at larger distances from the railroad. The pattern of a slightly higher employment growth at larger distances from the railroad is discernable in other comparable cities as well. This is shown by the solid and dashed grey lines. Table 5 presents quantitative evidence of the distance-employment growth relationship. The pattern of slightly

[20] In Appendix Table A5, we present the growth in the number of passengers that either board, or leave their train in Delft, and in a number of cities which are comparable in terms of city-characteristics. The table shows that the growth in railroad passengers was higher in Delft, particularly during the time period 2015-17 - a growth that was presumably driven by the increase in the number of students at the university of Delft. This growth in the number of passengers boarding or leaving their trains, however, provides less information that the level of service indicator shown in Table 4. The former namely, does not take into account the frequency and the reliability of the trips by train. Put differently, the table in the appendix provides no insight in the notion that for each railroad station, the travel times to other railroad stations, its frequency and its reliability, is dependent on the entire Dutch train network.

higher employment growth at larger distances is confirmed by the elasticity coefficient of 0.032. The interaction terms for the (post-)tunneling periods with Delft indicate that this pattern does not vary significantly from other cities during the intervention periods.

Overall, we fail to find any relationship between the railroad tunneling of Delft and considerations about economic growth. This suggests that the railroad tunneling can indeed be classified as a rather exogenous event.

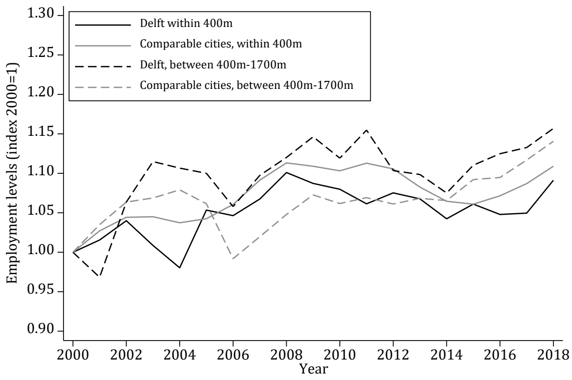

***Figure 7:** The Growth in Employment Levels*

***Table 5:** The relationship between employment growth and distance to the railroad*

Log distance	0.032**
	(0.014)
Log distance* Delft *2009-10	0.011
	(0.006)
Log distance* Delft *2011-12	0.009
	(0.007)
Log distance* Delft *2013-14	0.007
	(0.010)
Log distance* Delft *2015-16 [a]	0.011
	(0.009)
Log distance* Delft *2017-18 [a]	0.014
	(0.009)

Notes: The dependent variable is the log of the employment level. The specification includes year fixed effects and postal code fixed effects. Significance at the 10%, 5% and 1% level respectively denoted by */**/***.

Notes: Figure 7 presents the growth in employment levels, over the time period 2000-2018. The employment level data is based on the total number of working people (i.e. both full and part-time), independent of the working type (i.e. employees for a companies, self-employed, entrepreneurs are all included). Employment growth close to a railroad is measured within a distance of 400 meter of the railroad, while the employment growth for larger distances is measured at distances of 400 meter to 1,700 meter of a railroad. [a] Since 2015, the employment growth in Delft is measured in proximity to the *tunneled* railroad.
Data source: Authors' calculations based on data by LISA (Netherlands).

6. Results

6.1 Effects on residential property prices

Figure 8 plots the trend in willingness to pay to live near the railroad in Delft and its synthetic counterfactual during the period 1995 to 2018. The solid line indicates the price elasticity with respect to the distance to the railroad in Delft of stage 2. The dashed line presents the average price elasticity obtained from 300 synthetic counterfactuals (stage 3). This counterfactual estimate shows how the willingness to pay for railroad externalities in Delft would have developed in case Delft would not have tunneled its railroad. The grey area around synthetic Delft denotes the 95% confidence interval. During the pre-tunneling period, the dashed line closely follows the trajectory of the solid line, which suggests that synthetic Delft may provide a suitable counterfactual.

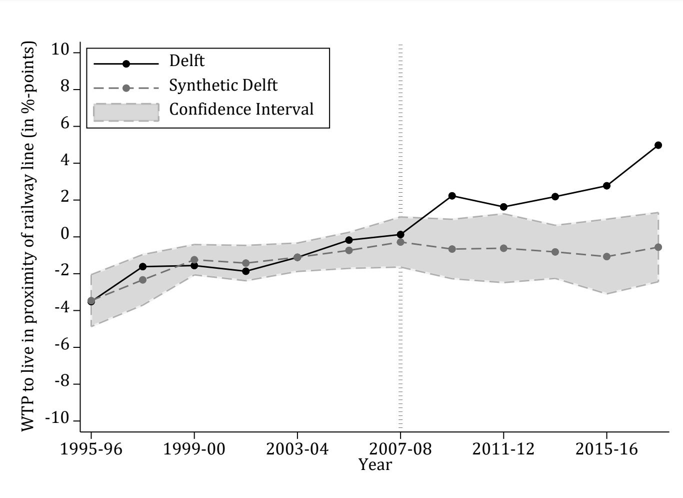

Figure 8: *The impact of a railroad tunneling on residential property prices:* The willingness to pay to live in proximity of the railroad in Delft versus the synthetic counterfactuals

The estimate of the willingness to pay for railroad externalities is exhibited by the difference between the actual price elasticity in Delft and its synthetic counterpart. As soon as Delft imposes the first tunneling preparations, the average trajectory starts to deviate significantly from counterfactual Synthetic Delft. The direct positive effect indicates that about half of the willingness to pay for the railroad tunneling already capitalizes before the urban railroad externalities are eliminated. This resembles what some coin as an *anticipation effect:* residents expect that the forthcoming utility derived from dwellings located close to the railroad will rise as a result of the future railroad tunneling, leading to a surge in the present discounted values of these dwellings.

Over time, the discount factor of the rate of time preference as well as the uncertainty about the future benefits decrease. The result is the diverging pattern observed in Figure 8: as the actual tunneling and the redesign of the public space draw nearer, Delft gradually diverges from the counterfactual estimate. The slight decrease in the WTP effect during the time period 2011-2012 is consistent with the conjecture that the net capitalization effect may decrease in times of the most intensive construction nuisance. Overall, Figure 8 suggests that the price elasticity would have been about 5 percentage points lower in case Delft would not have tunneled its railroad.

Table 6 presents further evidence on whether the average constructed control unit resembles a sensible counterfactual. The table compares the similarity of predictor covariates in Delft relative to its counterfactual. The results in Table 6 show that the pre-tunneling characteristics of synthetic Delft closely match the characteristics of Delft. The V-weights suggest that especially the transmitted noise levels along railroads, and the amount of cultural heritage are important predictors. All in all, synthetic provides a much better comparison unit than the average of the donor pool. We are therefore confident that the identifying assumptions of the synthetic control estimator are satisfied.

Table 6: *Predictor means*

	Delft	Synthetic Delft	Average Donor Pool	V-weights
Noise level (in decibels)	69.78	69.41	66.94	0.235
Air quality (particular matter)	27.16	26.73	25.89	0.167
Density (ln number of jobs)	10.80	10.88	10.78	0.160
Cultural heritage (in km^2)	1.02	0.83	0.44	0.213
Percentage residents aged 15-24 years	13.93	12.31	11.01	0.115
Expectation development neighborhood	1.92	1.89	1.82	0.109

Notes: The table presents the average predictor variable similarly between Delft and the computed synthetic control units. We computed the average upon completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool. Appendix Table A2 provides a description of the predictor variables.

Table 7 displays the w-weights that are assigned the donor pool of control cities. The reported weights indicate that the railroad externalities trajectory in Delft is best approximated by a large set of control cities with moderate weights. Especially Deventer, Dordrecht, Gouda, Haarlem and Nijmegen turn out to be cities in close resemblance to Delft. 15 out of the 27 control cities are designated zero w-weights.

Table 7: *Designated w-weights to donor pool cities*

Alkmaar	0	Dordrecht	0.10	Hilversum	0.07
Almelo	0	Ede	0	Leeuwarden	0.02
Almere	0	Eindhoven	0.02	Leiden	0.08
Alphen ad Rijn	0	Emmen	0	Nijmegen	0.19
Amersfoort	0.04	Enschede	0	Oss	0
Apeldoorn	0	Gouda	0.13	Purmerend	0
Assen	0	Groningen	0.04	Zaandam	0
Bergen op Zoom	0	Haarlem	0.12	Zoetermeer	0
Deventer	0.12	Heerenveen	0.05	Zwolle	0

Notes: The table reports the average w-weights designated to each of the control cities in the donor pool. We computed the average upon completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool.

6.2 Robustness

In this section, we determine the credibility of the results by conducting a number of sensitivity tests.

First, we test whether our results are robust when we adopt cross-validation over different time periods. In other words, we validate whether our predictors possess out-of-sample prediction power. Based on the previous section, this remains unclear, since we optimized the v_m matrix to minimize the pre-intervention outcome differences in-sample. For this reason, we adopt cross-validation and divide the pre-tunneling years into a training period and a validation period (Abadie et al., 2015). The different validation periods are listed in columns (2) and (3) in Table 8. The data of the training period are averaged for the entire period prior to the validation period. In all specifications, the point estimates of the alternative cross-validation windows are quantitatively very similar to the baseline results. This suggests our results are not driven by the prediction procedure and that the predictors are sufficiently able to predict out-of-sample.

Second, we investigate whether our results are dependent on the predictors adopted in the distance minimization algorithm of the synthetic control. We subsequently test the adoption of three types of predictor covariates: (i) socio-economic, (ii) demographic, and (iii) neighborhood characteristics. The idea of this robustness check is to test whether the synthetic controls – and the corresponding estimates – change when they account for additional 'relevant' predictors. The results are reported in columns (4) to (7). It shows that none of the additional predictors leads to significantly different synthetic controls. The estimates in columns are nearly identical to the baseline results.

Third, we conduct a number tests to check whether our results hold when we implement a different adoption date of the intervention. In our baseline results, we set the start of the tunneling preparations as first intervention period. We found that about half of the tunneling effect directly capitalizes during these years. Intuitively, one could argue that part of the effect may already capitalize earlier, for example at the moment the project was announced. To check this notion, we reassign the intervention date 2 and 4 years prior to the announcement. These test are based on the idea that our results may not be credible in case we find significant divergences during an artificially chosen period. Columns (8) and (9) report the results.

The results display a very similar pattern in comparison to the baseline results. This suggests that our results are not driven by an artificially chosen treatment date. The in-time (placebo) coefficients from 2001-2002 to 2007-2008 indicate that the announcement of a tunnel project does not lead to immediate capitalization effects in residence prices. Instead, the first anticipation effects occur when the first preparations of the railroad tunneling are conducted. The timing of the capitalization effects suggests that just the announcement is not sufficient for people to already expect their forthcoming utility to increase from dwellings in close distance to the tunnel. This may be a reflection of the uncertainty; it is unknown when the actual railroad tunneling takes place, and as a result, there are no anticipation effects.

Fourth, we check whether our results hold when we interact a number of robustness checks. In particular, column (10) reports the results when we include the full set of controls in the synthetic control algorithm and adopt the cross-validation exercise of column (3). Column (11) also includes the full set of controls in the synthetic control algorithm, but adopt the in-time placebo procedure of column (9). Again, the results are highly robust.

Table 8: *Robustness checks*

	A. Cross-validation window			*B. Alternate inputs*				*C. In-time placebo's*		*D. Interactions*	
	Baseline	2001-07	2003-07					Treatment-4Y	Treatment–6Y	Col (3)+(7)	Col (7)+(9)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Pre-intervention	0.001	−0.003	−0.004	−0.002	0.001	0.00	0.003	−0.004	−0.004	0.002	−0.006
	(0.006)	(0.012)	(0.012)	(0.006)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008	(0.013)	(0.014)
Placebo estimates											
2001-02									0.001		−0.008
									(0.012)		(0.014)
2003-04								−0.007	−0.006		−0.003
								(0.007)	(0.011)		(0.006)
2005-06								−0.002	−0.003		0.001
								(0.007)	(0.010)		(0.008)
2007-08								0.001	0.000		0.004
								(0.012)	(0.015)		(0.007)
2009-10	0.029	0.025	0.027	0.023	0.022	0.024	0.024	0.029	0.030	0.033	0.028
	(0.008)	(0.010)	(0.011)	(0.008)	(0.009)	(0.008)	(0.008)	(0.010)	(0.011)	(0.009)	(0.011)
2011-12	0.022	0.019	0.023	0.023	0.023	0.023	0.022	0.022	0.023	0.024	0.023
	(0.010)	(0.014)	(0.014)	(0.008)	(0.010)	(0.009)	(0.010)	(0.009)	(0.012)	(0.012)	(0.015)
2013-14	0.030	0.028	0.030	0.033	0.032	0.037	0.035	0.029	0.031	0.037	0.033
	(0.007)	(0.013)	(0.013)	(0.018)	(0.017)	(0.016)	(0.015)	(0.016)	(0.020)	(0.014)	(0.015)
2015-16	0.038	0.038	0.040	0.041	0.037	0.043	0.044	0.038	0.039	0.045	0.044
	(0.010)	(0.014)	(0.014)	(0.015)	(0.015)	(0.014)	(0.013)	(0.017)	(0.017)	(0.016)	(0.013)
2017-18	0.055	0.050	0.051	0.057	0.055	0.055	0.054	0.055	0.056	0.054	0.053
	(0.010)	(0.010)	(0.010)	(0.006)	(0.008)	(0.010)	(0.010)	(0.011)	(0.010)	(0.013)	(0.011)
Baseline inputs	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Socio-economic				Y			Y			Y	Y
Demographic					Y		Y			Y	Y
Neighborhood						Y	Y			Y	Y

Notes: The table reports the average difference between the actual willingness to pay to live near the (tunneled) railway line in Delft versus its synthetic counterpart over the indicated time periods. The average was computed after completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool. The baseline predictors are outlined in Table 6. Columns (4) to (6) separately adds sets of predictors relative to the baseline Column (4) includes a socio-economic status indicator. Column (5) includes the percentage of inhabitants aged between 25-44 and 45-64, respectively. Column (6) includes the satisfaction about the built environment, and available amenities (green and schools). Column (7) includes all of the previously mentioned sets of predictors. For a full description of the alternate input predictors, see Appendix Table A3. Standard errors between parentheses.

Lastly, we conduct a permutation test to check whether the results in Delft are achieved by chance. That is, suppose we randomly reassign the treatment to one of the control cities in the donor pool, what is the chance that these cities achieve a result of a similar or higher magnitude than Delft? The answer to this question is given by a permutation test, akin the placebo test performed by Abadie et al. (2015) and Acemoglu et al. (2016). In our case, we perform a randomization procedure in two domains. First, out of the sample of 27 control cities, we randomly assign one of them as the intervention unit. Second, for the remaining 26 control cities, we randomly draw a subgroup at half of its original size. We perform this procedure 5000 times and thus estimate a number of 5000 synthetic 'placebo' results. The left hand side of Figure 9 displays the histogram of the average post-intervention effects. The light grey indicate the placebo values, the black values reflect the values of Delft.

The distribution of placebo effects is centered around zero, indicating that ''placebos'' do not cause systematic impacts on control units. The 'real' treatment estimates of Delft differ significantly from the placebo estimates at a 95% confidence level. Moreover, the results become even more significant when we divide the treatment estimates by the 'fit' of its pre-treatment counterfactual (RSMPE). This exercise is shown at the right-hand side of figure 9. This suggests that the chance of finding the results of Delft at random are very low.

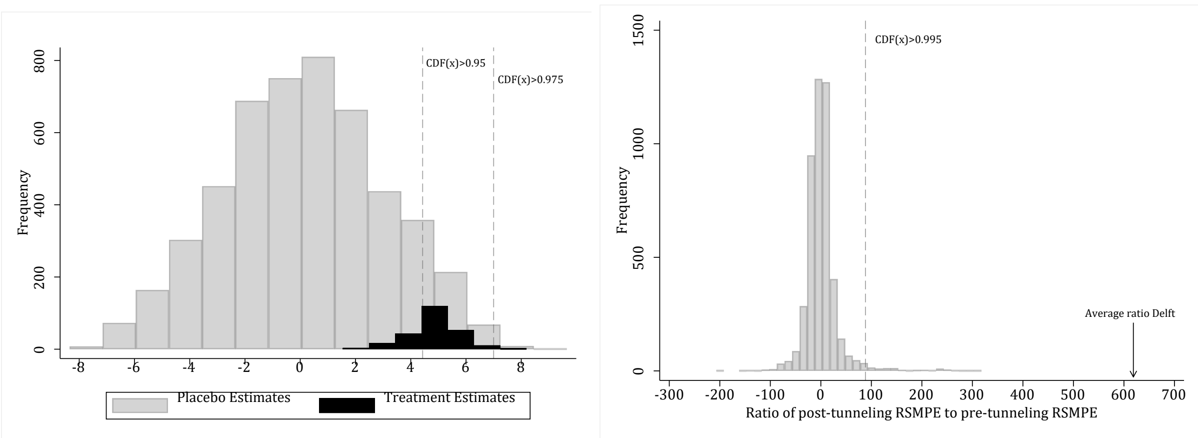

Figure 9: *Histogram of placebo effects*

Notes: The left-hand side of Figure 9 juxtaposes the actual estimates of the city of Delft to the placebo estimates. The values on the x-axis indicate percentage points. The right-hand side of Figure 9 uses the estimated average post-intervention effects and divides them by their pre-intervention fit, or root mean squared prediction error (RMSPE). This exercise is conducted because a large post-intervention gap between a (placebo) unit and its synthetic counterpart may not necessarily provide any information about the impact of the intervention (or placebo) when the pre-intervention counterfactual poorly tracks the unit of interest.

6.3 Heterogeneity in the Treatment Effect

In this subsection, we test whether the willingness to pay effects differs across space. Intuitively, one would expect the impact of the railroad tunneling to be more intense at specific places along the railroad. For instance, it could be that the effects on the willingness to pay are stronger at the area that initially would be labelled as the 'wrong side of the track' – the area at the west of the railroad in Figure 3. Alternatively, it could be that the dominant wind direction may cause residents at the other side (west) of the railroad to experience the negative externalities much more intensely, resulting in stronger effects on residential property prices.

To allow for heterogeneity in treatment effects, we repeat our three-stage empirical strategy, now by dividing the sample of residence units into two districts: one at the western side of the railroad, and one at the eastern side. Figure 10 presents the results. Again, the willingness to pay estimate for railroad externalities is exhibited by the difference between the actual price elasticity at both sides of the railroad of Delft and their synthetic counterpart.

There is an disparate pattern in the trajectories of the effects. While the average trajectory of the east side of Delft directly starts to deviate significantly from its counterfactual as soon as the tunneling preparations start, the average trajectory for west side of Delft does not. The estimate for the latter indicates the railroad tunneling did not have much of an effect during the first 2 years of railroad preparations. From the time-period of the actual tunneling onwards (2011), the negative willingness to pay effect to live in proximity of the railroad gradually disappears. Since its synthetic counterfactual roughly stays at a similar level, the estimate exhibits a gradual diverging pattern as well. This is shown in Table 9, which presents the quantitative analogs of Figure 10.

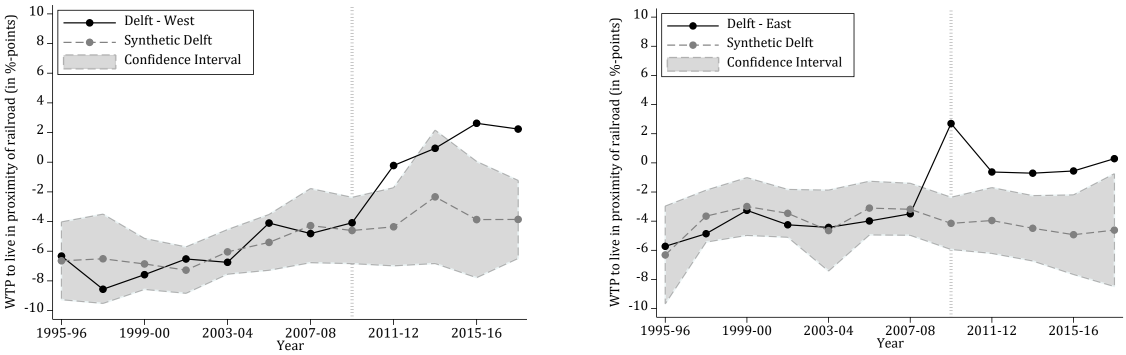

Figure 10: *The impact of a railroad tunneling on residential property prices in different districts:* The willingness to pay to live in proximity of the railroad in Delft versus its synthetic counterfactuals

Interestingly, in contrast to the area west of the railroad, the estimates of the area east of the railroad show a direct capitalization effect of the railroad tunneling as soon as the first railroad tunneling preparations start. However, the trajectory in the effect size (column 3) is peculiar. The positive effect halves as soon as the actual tunneling activities begin (2011), and never fully recovers to the size of the first treatment period (2009-10). This suggests a mismatch in the expectations of residents about the present discounted values of their dwellings. Apparently, residents in and around the city center of Delft were too overly optimistic about the effects of the railroad tunneling. All in all, we find no significant difference in the size of the railroad tunneling effects, although the 'wrong side of the railroad track' has a slightly higher willingness to pay effect during the last intervention period.

To provide further scrutiny into the estimation results, we divide both CBD sides into two areas: south and north. Column (4) and (5) present the estimation results for the area north and the area south at the side west of the CBD. The results indicate that especially the area at the south has experienced a strong effect on the willingness to pay to live near the (tunneled) railroad. This area was traditionally known for its large amounts of inhabitants with a low socio-economic status. The railroad tunneling may have affected this relatively disadvantaged area with an influx of inhabitants with a higher socio-economic status. Lastly, the results in column (6) and (7) do not exhibit strong differences. These locations already possessed larger shares of inhabitants with a relatively high socio-economic status. The next section aims to further elucidate the possible sorting effects of the railroad tunneling.

Table 9: *The heterogenous effects of a railroad tunneling on residential property prices*

	Baseline	West	East	West		East	
				North	South	North	South
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pre-intervention dif.	0.001	−0.003	−0.001	−0.009	0.009	−0.008	−0.012
	(0.006)	(0.012)	(0.013)	(0.017)	(0.017)	(0.015)	(0.018)
2009-10	0.029	0.011	0.052	−0.020	0.028	0.054	0.025
	(0.008)	(0.015)	(0.013)	(0.018)	(0.016)	(0.017)	(0.021)
2011-12	0.022	0.037	0.021	0.034	0.047	0.009	0.017
	(0.010)	(0.014)	(0.013)	(0.015)	(0.021)	(0.019)	(0.019)
2013-14	0.030	0.035	0.032	0.006	0.048	0.036	0.023
	(0.007)	(0.021)	(0.015)	(0.032)	(0.016)	(0.021)	(0.018)
2015-16	0.038	0.062	0.042	0.052	0.076	0.023	0.039
	(0.010)	(0.018)	(0.015)	(0.023)	(0.019)	(0.012)	(0.018)
2017-18	0.055	0.063	0.048	0.048	0.072	0.041	0.047
	(0.010)	(0.012)	(0.021)	(0.015)	(0.014)	(0.015)	(0.023)

Notes: The table reports the average difference between the actual willingness to pay to live near the (tunneled) railroad in Delft versus its synthetic counterpart over the indicated time periods. The average was computed after completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool. For a full description of the input predictors, see Appendix Table A2. Standard errors between parentheses.

6.4 Effects on Sorting

In this section, we test whether the characteristics of neighborhoods in close proximity to the railroad change in Delft after the railroad tunneling. In other words, what types of consumers value the elimination of railroad nuisance highly and sort themselves in close proximity to the tunneled railroad? We test for changes in three sets of sorting characteristics socio-economic, demographic, and immigrant background type.

We adopt a similar identification approach as described in section 5.1. We again adopt a two-stage methodology. In the first stage, for each city separately, we determine the probability of sorting characteristic R being more prevalent at close distance from the railroad *over time*. Specifically, we check whether characteristics of neighborhoods become more or less prevalent within a distance of 800 meters from the railroad relative to the characteristics of neighborhoods at distances of 800 to over 1,500 meters from the railroad.[21] We estimate the following equation function:

[21] There are two reasons for the decision to determine the probability within a distance of 800 meters from the railroad relative to further away. First, as mentioned in footnote x, we conducted a number of tests to determine the functional

$$\log R_{rt} = \gamma_r Z_r + \tau_t T_t + \beta Railroad_r + \vartheta_t (Railroad_r * T_t) + \varepsilon_{i,t} \quad (8)$$

where $\log R_{rt}$ is the log transformed characteristic of neighborhood r in year t ($T = 1995, \ldots 2018$). The Z_r's denote a set of postal code fixed effects, which will capture the time-invariant neighborhood characteristics. The T_t's denote a set of year fixed effects, which will absorb year specific trends. $Railroad_r$ denotes a dummy that equals 1 if a neighborhood is located in close proximity to the railroad (within 800 meters). The treatment variable of interest is the interaction variable $Railroad_r * T_t$, which captures the probability of sorting characteristic R being more prevalent at close distance from the railroad *over time*. Lastly, $\varepsilon_{i,t}$ is the error term.

The regression outputs of the first stage $\vartheta_{t.j}$ are used in the second stage, the synthetic control method. The goal of the second stage is to establish a counterfactual in the sorting characteristics in case Delft would not have tunneled its railroad. We use the similar set of predictors in the synthetic control procedure as used in section 6.1 and again adopt the subsampling method to infer whether the results are statistically different from zero. The results for each of the sorting characteristics are reported in columns 1-11 in Table 10.

There are a few important findings. First, the railroad tunneling causes a significant increase in the socio-economic status of neighborhoods in close proximity to the railroad. During the final evaluation period (2017-2018) the socio-economic status of neighborhoods located within 800 meters from the railroad is 28% higher than its counterfactual.[22] The point estimates in columns 2 and 3 suggest this change is not driven by (average) income levels of the neighborhoods. Rather column 4 indicates at least part of the significant increase in socio-economic status could be attributed to a relative decrease in the percentage of non-active persons (not in the labor market or searching for a job). Another reason for the significant increase could be that average education level of the neighborhood in close proximity to the tunneled railroad has increased, however we do not have data to quantify this hypothesis.

The second main finding is that the railroad tunneling is associated a change in the age demographics of neighborhoods. In particular, column(7) indicates that the railroad tunneling causes a significant increase in the percentage of residents within the age group of 25 to 44. The point estimates exhibited in columns (8) and (9) further reveal this increase comes at the expense of residents of the age group 45 years and older, especially those which have reached the age of retirement (65 years and older). This suggests that residents within the older age groups respond less strong on a change in railroad nuisance. Finally, the railroad tunneling has no significant impact on the percentage of residents with an immigrant background.

form of the willingness to pay relationship of the distance to the railroad and residential property prices. In these tests, we also experienced with 200 meter distance bands from the railroad, using differences-in-difference analyses comparing trends in residential property prices relative to non-affected areas (more than 1km from the railroad). The outcome of these tests suggests the (additional) willingness to pay effects to live near the tunneled railroad largely dissipated after 800 meters. Appendix Figure A2 suggests the distance of within 800 meters holds for the socio-economic status pattern as well.

[22] That is, $(e^{0.247} - 1) * 100 \approx 28\%$.

Table 10: *The impact of a railroad tunneling on sorting*

	Dependent variable										
	Panel A. Socio-economic variables				*Panel B.* Demographic variables					*Panel C.* Immigrant background	
	Socio-economic (1)	Income (working) (2)	Income (average) (3)	Non actives (4)	Aged 0-15 (5)	Aged 15-24 (6)	Aged 25-44 (7)	Aged 45-64 (8)	Aged 65 and older (9)	Western (10)	Non-Western (11)
Pre-intervention	0.001	0.018	−0.005	0.002	0.010	−0.007	0.027	−0.037	−0.002	−0.001	−0.001
	(0.028)	(0.016)	(0.026)	(0.031)	(0.036)	(0.044)	(0.029)	(0.056)	(0.017)	(0.014)	(0.062)
2009-10	0.166	−0.060	−0.051	0.028	0.062	−0.011	0.067	−0.129	−0.006	0.010	−0.007
	(0.145)	(0.029)	(0.044)	(0.050)	(0.057)	(0.073)	(0.059)	(0.081)	(0.038)	(0.030)	(0.096)
2011-12	0.258	−0.029	−0.022	−0.186	0.069	−0.033	0.085	−0.107	−0.020	−0.031	−0.013
	(0.143)	(0.024)	(0.037)	(0.068)	(0.053)	(0.060)	(0.046)	(0.066)	(0.038)	(0.033)	(0.088)
2013-14	0.294	−0.041	−0.001	−0.219	0.075	−0.041	0.090	−0.096	−0.049	−0.019	−0.018
	(0.187)	(0.031)	(0.046)	(0.113)	(0.080)	(0.080)	(0.057)	(0.079)	(0.056)	(0.046)	(0.121)
2015-16	0.277	−0.034	−0.003	−0.253	0.064	0.002	0.091	−0.059	−0.129	0.015	−0.068
	(0.163)	(0.030)	(0.055)	(0.188)	(0.084)	(0.103)	(0.063)	(0.091)	(0.075)	(0.107)	(0.138)
2017-18	0.247	−0.036	−0.024	−0.327	0.052	0.065	0.104	−0.024	−0.204	0.030	−0.119
	(0142)	(0.031)	(0.059)	(0.155)	(0.096)	(0.121)	(0.075)	(0.108)	(0.099)	(0.183)	(0.181)

Notes: The table provides evidence whether characteristics of neighborhoods become more or less prevalent during and after the railroad tunneling within a distance of 800 meters from the railroad relative to the characteristics of neighborhoods at distances of 800 to over 1,500 meters from the railroad of Delft. The table reports the average difference between the relative prevalence for each of the sorting characteristics in Delft versus its synthetic counterpart over the indicated time periods at the left-hand side of the table. The average was computed after completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool. Standard errors between parentheses.

The characteristics on the neighborhood level are defined as follows: (1) the socio-economic status comprising a composite of the education level, the income level and the position on the labor market (non-employed versus employed), (2) the average gross income for residents who are active on the labor market (3) the average gross income for all residents, (4) the percentage of residents that are not active on the labor market, (5-9) the percentage of residents within a specific age group, (10-11) the percentage of residents with a Western/non-Western immigrant background.

The effects of sorting on the estimates on residential property prices

Thus far, we restricted the set of controls to a set of housing characteristics, and fixed effects in stage 1 of our empirical strategy. We therefore did not account for the possibility that sorting of households on itself may affect the willingness to pay to live near the tunneled railroad. For instance, it is likely that households value the presence of other households with a higher socio-economic status positively, as shown by Bayer et al. (2007). The dynamic process in demographic and socio-economic variables may then be themselves potential demand shifters for living close to the tunneled railroad. Another reason could be that part of the willingness to pay effect could be attributed to changes in income levels, instead of differences in preferences. To see whether these demand shifters affect the willingness to pay to live near the railroad, we include three sets of controls on a neighborhood level (postal code) in stage 1 of the empirical strategy.

The estimates are reported in Table 11. Column (1) presents the baseline results of Figure 8. Column (1) only includes housing variables. Columns (2) to (4) include socio-economic variables, demographic variables, and employment variables, respectively.

The inclusion of the socio-economic controls (average income and non-actives) leads to slight reductions in the point estimates. This indeed suggests that part of the willingness to pay effect for the railroad tunneling can be attributed to sorting of households with a relatively higher socioeconomic status., which is consistent with the evidence presented in section 6.3. The inclusion of demographic variables to stage 1 also leads to a reduction in the size of the point estimates – albeit very small.

Column (4) accounts for the possibility that employment levels have increased more favorably at smaller distances from the railroad tunnel, which in turn have positively affected the willingness to pay near the railroad. Earlier in section 5.1, we showed however that the trends in employment growth did not differ significantly across space in Delft. This is confirmed by the point estimates, which are almost identical to those of the baseline. This indicates that employment changes did not impact the willingness to pay results for the railroad tunneling in Delft.

Column (5) includes all potential demand shifters simultaneously. The results are very similar to the baseline estimates, except for the last intervention periods (after the actual railroad tunneling). The point estimates of the time period 2017-18 suggest that about 20% of the total effect on the willingness to pay can be attributed to observable changes on the neighborhood level.

Lastly, we use the post-regularization method in stage 1 of our empirical strategy to determine whether the included controls contribute 'satisfactory' to the fit of the model (Chernozhukov et al., 2015). In other words, we use a LASSO-type of estimator where a penalty term is included for each of the potential control variables. Variables that contribute little to the fit are eventually set equal to zero (for a more detailed description, see Appendix section A3). The results exhibited in column (6) are very similar to those of column (5). This suggests that, at least the majority of, the included controls in column (5) contribute to the prediction accuracy of stage

1 in our empirical strategy. That is, our evidence for the presence of demand shifters – especially due to sorting by socio-economic status – is not driven by the inclusion of multicollinear controls.

Table 11: *The effects of a railroad tunneling on residential property prices* – Demand shifters

Panel A. (baseline scm)	(1)	(2)	(3)	(4)	(5)	(6)
Pre-intervention dif.	0.001	−0.003	−0.004	−0.002	−0.001	−0.004
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)
2009-10	0.029	0.029	0.027	0.020	0.022	0.017
	(0.008)	(0.008)	(0.011)	(0.009)	(0.011)	(0.010)
2011-12	0.022	0.024	0.016	0.023	0.021	0.017
	(0.010)	(0.009)	(0.010)	(0.013)	(0.009)	(0.011)
2013-14	0.030	0.031	0.032	0.028	0.029	0.024
	(0.007)	(0.014)	(0.014)	(0.017)	(0.016)	(0.014)
2015-16	0.038	0.033	0.040	0.039	0.034	0.031
	(0.010)	(0.010)	(0.012)	(0.015)	(0.011)	(0.010)
2017-18	0.055	0.048	0.052	0.054	0.046	0.045
	(0.010)	(0.011)	(0.009)	(0.010)	(0.009)	(0.011)
Housing characteristics	X	X	X	X	X	X
Socio-economic variables		X			X	X
Demographic variables			X		X	X
Employment variables				X	X	X
Lasso						X

Notes: The table reports the average difference between the actual willingness to pay to live near the (tunneled) railroad in Delft versus its synthetic counterpart over the indicated time periods. The average was computed after completion of a loop of 300 synthetic control units for Delft. Each synthetic unit was optimized using a randomly drawn weighted average at half of the donor pool. The table separately adds sets of control variables to stage 1 of our empirical strategy. Column(1) reports the outcomes of baseline set of control variables. Column (2) includes the socio-economic status of the neighborhood (average income by working residents and non-actives). Column (3) includes controls for the percentage of inhabitants aged 15 and 24, 25 and 44, 65 and older and the percentage of residents with a (non)-western immigrant background. Column (4) includes employment levels, ranging over 4 employment types. Column (5) includes each of the previously mentioned sets of control variables. A full description of the controls is provided in Appendix Table A2. Column (6) uses a lasso method (post-regularization) to determine the optimal set of control variables. The lasso method was used for all of the previously mentioned sets of control variables, including the housing characteristics shown in Table 2. For a full description of the Lasso method, see Appendix section A3 (Post-regularization methodology). Standard errors between parentheses.

7. Conclusions

This study has exploited the railroad tunneling in Delft as a quasi-experiment to estimate the willingness to pay for the elimination of railroad externalities. The basis of our analysis is that other cities did not tunnel their railroad. Using this insight, we use a three-stage approach in order to estimate the counterfactual willingness to pay. Our results indicate that the price elasticity with respect to the distance to the railroad in Delft would have been about 5 percentage points lower in case Delft would not have tunneled its railroad. These results are robust to a series of robustness checks, including a number of in-time, and control-to-treatment placebo's.

About half of the effect already capitalizes as soon the tunneling preparations start. We therefore find evidence for a considerable *anticipation effect.* That is, residents already expected that their forthcoming utility derived from dwellings located close to the railroad would rise as a result of the future railroad tunneling, leading to a surge in the present discounted values of these dwellings. The positive housing market effects, however, are not similar across the railroad. There is a significant heterogeneity in the willingness to pay effects, which seem to coincide with the average socioeconomic status of residents that initially inhabited the area before the railroad tunneling. Areas with a high average socio-economic status did not experience large effects on the willingness to pay to live near the railroad, while areas with a low average socio-economic status did experience large positive effects.

Our paper contributes to the literature that use quasi-experiments to quantify the environmental effects of infrastructure. We complement the literature by showing that infrastructure adjustments may also cause sorting effects. Not accounting for these effects may bias the estimated willingness to pay for the environmental effects of infrastructure adjustments.

One caveat of our study is that we did not analyze the effect of the railroad tunneling on the internal structure of the city of Delft. The change in the internal urban structure may be an additional welfare effect of a railroad tunneling – next to its environmental and social effects. For instance, the overall attractiveness of the city may increase due to a railroad tunneling. In the setting of a monocentric city, this is shown by an upward shift in the willingness to pay in the CBD. The presence of this potential welfare effect could be investigated in new research.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495-510.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., & Mitton, T. (2016). The value of connections in turbulent times: Evidence from the United States. *Journal of Financial Economics*, *121*(2), 368-391.

Al-Mosaind, M. A., Dueker, K. J., & Strathman, J. G. (1993). *Light-rail transit stations and property values: a hedonic price approach* (No. 1400).

Alonso, W. (1964). Location and land use. Toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*

Anciaes, P. R., Jones, P., & Mindell, J. S. (2016). Community severance: where is it found and at what cost?. *Transport Reviews*, *36*(3), 293-317.

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1-14.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, *99*(3), 435-448.

Boes, S., & Nüesch, S. (2011). Quasi-experimental evidence on the effect of aircraft noise on apartment rents. *Journal of Urban Economics*, *69*(2), 196-204.

Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Post-selection and post-regularization inference in linear models with many controls and instruments. *American Economic Review*, *105*(5), 486-90.

Debrezion, G., Pels, E., & Rietveld, P. (2011). The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market. *Urban Studies*, *48*(5), 997-1015.

Diao, M., Qin, Y., & Sing, T. F. (2016). Negative externalities of rail noise and housing values: Evidence from the cessation of railway operations in Singapore. *Real Estate Economics*, *44*(4), 878-917.

Duin, L. van, & Floet, W. W. (2005). Spoorzone Delft. *OverHolland*, 91-103.

Engels, F. (1845). Die Lage der arbeitenden Klasse in England.

European Commission (2018). EU Transport in Figures: Statistical Pocketbook.

Gemeente Delft (2005). Bestemmingsplan Spoorzone Delft, Rotterdam, RBOI.

Greenstone, M., & Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the superfund program. *The Quarterly Journal of Economics*, *123*(3), 951-1003.

Muth, R. (1969). Cities and housing: The spatial patterns of urban residential land use. University of Chicago, Chicago, 4, 114-123.

Ministry of Housing, Spatial Planning and the Environment, 2009, MIRT Projectenboek 2009, The Hague: Ministerie van VROM.

Nederlandse Spoorwegen (1988). Rail 21: Sporen naar een nieuwe eeuw, Utrecht.

Nelson, A. C. (1992). Effects of elevated heavy-rail transit stations on house prices with respect to neighborhood income. *Transportation Research Record*, (1359).

Poon, L. C. (1978). Railway externalities and residential property prices. *Land Economics*, 218-227.

Saia, A. (2017). Choosing the open sea: The cost to the UK of staying out of the euro. *Journal of International Economics*, *108*, 82-98.

Statistics Netherlands (2019). Available online at: http://statline.cbs.nl/ [Accessed 9 July 2019].

Stuurgroep Ondergrondse Vervoers-Infrastructuur (SOVI) (1993) Ondergronds Overwegen, Rotterdam: SOVI.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Witteveen+Bos (2003). Milieueffectrapport Spoorzone Delft stedelijke herontwikkeling & aanleg tunnel, Deventer: Witteveen+Bos.

Appendix

This Appendix provides detailed information about the data (section A1), additional figures (section A2) and the used methods (A3).

A1: Data Appendix

Table A1: *Selection process of Residential Property Data*

Selection Criteria	*Number of observations*
1. Initial dataset (1985-2018)	3,721,138
2. Discard observations with no permanent residential function	3,531,802
3. Discard observations with unknown building year	3,425,385
4. Discard observations with missing residential characteristics (see Table 2)	3,397,629
5. Discard cases with less living space (volume) than 12 m2 (20m3) or more than 10000 m2 (10000m3)	3,180,863
6. Discard observations with unreliable characteristics	2,857,165
7. Discard potential outliers (transaction price outside the 0.5 and 99.5 percentile)	2,828,593
8. Discard postal codes with less than 50 transactions (over the period 1985-18)	2,814,342
9. Discard observations sold before 1994 (due to representability issues)	2,499,543
10. Keep treatment city (Delft) and control cities	750,568
11. Discard observations outside rectangular boundaries (see section 5.1 for more details)	203,845

Table A2: *Data Description - Predictors Stage 2*

Variable	Description	Source
Baseline inputs		
Noise level (in decibels)	The average energetic noise level due to railroad use, within 200 meters of the railroad. The energetic level is computed for the entirety of days using a L_{den} measure (level day, evening, night). Noise transmitted during the evening and night are ‘penalized’ with a 5db and 10db increment factor, respectively.	ProRail Netherlands
Air quality (particular matter)	The average air quality level within 200 meters of the railroad, measured in particular matter (PM_{10})	Netherlands National Institute for Public Health and the Environment (RIVM) – Data Nationaal Samenwerkingsprogramma Luchtkwaliteit (NSL)
Density (logarithmic number of jobs)	Natural logarithm of the total number of jobs, measured within the rectangular shape (see Figure 5)	Central Bureau of Statistics (CBS) – LISA employment register
Cultural heritage(in km^2)	The number of km^2 of cultural heritage, which is assigned to groups of real estate to ‘protect’ their status. This may either be due to considerations of the public interest, their beauty, their spatial or structural coherence or their scientific or cultural-historical value. The group consists of at least one monument. The number of km^2 cultural heritage are measured within the rectangular shape (see Figure 5)	Ministry of Education, Culture and Science – Rijksdienst voor het Cultureel Erfgoed
Percentage residents aged 15-24 years	The percentage of inhabitants with an age between 15 and 24 years old.	Statistics Netherlands (CBS) –Woon en buurtkaarten
Expectation development neighborhood (1-3)	Mean indicator of the residents’ opinion on the answer on the question whether the quality neighborhood will decline (1), stay the same (2) or improve (3). The indicator is computed within 400 meter of the railroad.	Ministry of Housing, Spatial Planning and the Environment. Drawn from the three publications: VROM (1998) *WBO1998: release 1.0* VROM (2002) *WBO2002: release 1.0* VROM (2005) *WoON2006: release 1.2*

Table A3: *Data Description – Additional Predictors Stage 2*

Variable	Description	Source
Socio-economic status	An indicator of the average socio-economic status of the neighborhood along the railroad. The socio-economic status is based on the following characteristics: the education level, the income level and the position on the labor market (non-employed versus employed). The indicator is computed within 400 meter of the railroad.	Bureau for Social and Cultural analyses (SCP) – Statusscores
Percentage 25-44 year olds	The percentage of inhabitants with an age between 25 and 44 years old.	Statistics Netherlands (CBS) –Woon en buurtkaarten
Percentage 45-64 year olds	The percentage of inhabitants with an age between 45 and 64 years old.	Statistics Netherlands (CBS) –Woon en buurtkaarten
Satisfaction about quality of schools (1-5)	Mean indicator of the residents' satisfaction on the quality of schooling amenities in the neighborhood, ranging from 1 (not satisfied at all) to 5 (very satisfied). The indicator is computed within 400 meter of the railroad.	Ministry of Housing, Spatial Planning and the Environment. Drawn from the three publications:
Satisfaction about green amenities (1-5)	Mean indicator of the residents' satisfaction on green amenities in the neighborhood, ranging from 1 (not satisfied at all) to 5 (very satisfied). The indicator is computed within 400 meter of the railroad.	VROM (1998) *WBO1998: release 1.0* VROM (2002) *WBO2002: release 1.0*
Satisfaction built environment	Mean indicator of the residents' satisfaction on the built environment of the neighborhood, ranging from 1 (not satisfied at all) to 5 (very satisfied). The indicator is computed within 400 meter of the railroad.	VROM (2005) *WoON2006: release 1.2*

Table A4: *Description computation level of service railroad stations*

For all railroad stations in the Netherlands, we computed a weighted generalized travel time indicator. This indicator captures the 'efficiency' of a particular railroad station to all other individual railroad stations destinations in the Netherlands. The efficiency measure is based on the sum of the in-vehicle time, the frequency, and a transfer penalty (requirement to switch to other trains/modality while travelling from point A to point B). In other words, for each railroad station, we have a measure of how efficient the railroad station is from itself (point A) to a destination (point B). Since there are over 300 railroad stations (combinations) in the Netherlands, we have 300 generalized travel time indicators for each railroad station.

We computed the change in the weighted average of the generalized travel time for each railroad station in two steps:

Step 1	In the first step, and for each railroad station separately, we observe the number of travelers stepping at that railroad station and going to a particular destination. For instance, we observe the number of people travelling from Delft to train towards Utrecht. Therefore, we can compute the percentage of travelers going from Delft towards one of the over 30 potential destinations by train. Using these percentages, we are able to compute a weighted generalized travel time indicator for each railroad station.
Step 2	In the second step, we compute the percentage differences in weighted generalized travel times for certain time periods.

Table A5: *The Growth in Railroad Passengers*

	Delft (1)	Comparable Cities (2)
2007-08	0.94%	2.86%
2009-10	5.15%	5.20%
2011-14	7.42%	4.55%
2015-17	12.09%	5.97%

Notes: The table presents percentage change in the number of travelers that step in, or out, at particular railroad stations for the time periods shown at the left-hand side of the table. Column (2) only includes cities with comparable city-characteristics as Delft, such as the number of inhabitants, the employment density, and other socio-economic characteristics.

Data source: Authors' calculations based on data by Dutch Railroads (NS)

A2: Figures

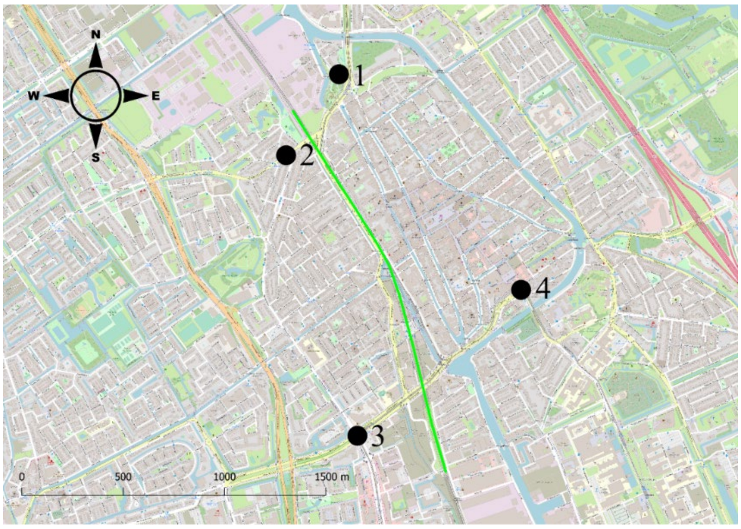

Figure A1: *Map of Delft illustrating the starting and ending points of the simulated travel times in Table A6.* The background is from the OpenStreetMap contributors (available under the Open Database License).

Table A6*: Simulated within-city travel time changes, by car (in minutes)*

From\ To	1 Wateringsevest	2 Ruys de B.str.	3 Papsouwselaan	4 Zuidpoort
1 Wateringsevest	x	0	0	1
2 Ruys de B.str.	0	x	1	1
3 Papsouwselaan	1	0	x	1
4 Zuidpoort	2	1	1	x

Notes: The table presents the difference in within-city travel time by car for each of the starting and ending points, differencing the ex-ante to the ex-post tunneling travel time. The table indicates all of the travel times increased slightly or remained constant. We note however that none of the differences can be deemed significant. All estimates fall within the margin of confidence of the model. This finding also holds for simulations when starting point 1 and 2 were set further from the tunnel, and when they were placed to the south. Hence, the difference in travel time is not dependent whether the starting (and ending point) were placed at close, or at larger distances from the tunnel. All in all, we are confident that the travel times by car did not change significantly over time.

The travel times differences by bicycle, not shown in this table, are quantitatively similar to the ones illustrated above. The only significant change in travel time for bicyclist was achieved for the route from the railroad station towards the technical university of Delft (located east-southwards of point 4). This route however was primarily enhanced for students, and not for residents. Residents do not benefit from this enhanced route.

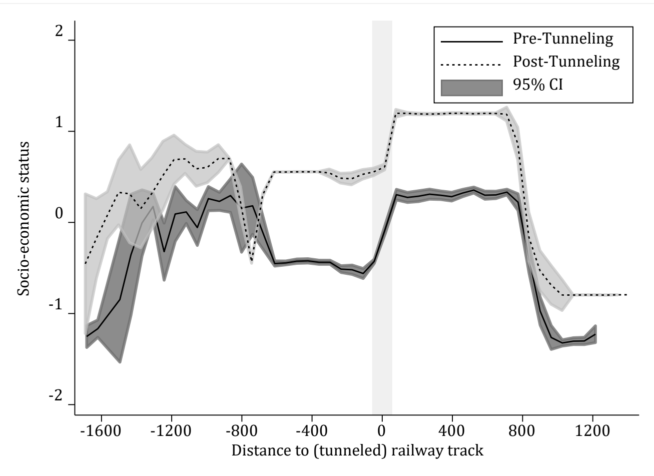

Figure A2: *The socio-economic status gradient.* The relationship between the distance to the railroad and the socio-economic status of neighborhoods

Notes: The figure displays a (Epanechnikov) kernel-weighted local polynomial regression of a socio-economic status indicator of neighborhoods on the Euclidean distance towards the railroad. The grey areas indicate the 95% confidence intervals. The pre-tunneling trajectory is depicted over the time period 2003-2008, the post-tunneling trajectory over the time period 2016-2018.

A3: Methods

This section includes additional information about the theoretical and empirical methods used in this paper. The next paragraph provides a description of the maximization problems of consumers and suppliers of housing within the hedonic pricing model.. Thereafter, the paragraph presents a formal overview of the post-regularization estimation strategy.

Hedonic price model

From a consumers point of view, the marginal willingness to pay for each characteristic $P_i = P(x_{i1}, x_{i2}, \ldots., x_{in})$, is determined by households maximizing their utility function subject to a budget constraint:

$$\max U = u(C, X) \qquad\qquad \text{subject to } C = W - P(X) \qquad (A1)$$

where C is the level of household consumption of a numeraire good, W the income level of the household and P the price of a residence (or housing costs). Putting together the maximization problems yields $\frac{\partial U/\partial x_n}{\partial U/\partial C} = \frac{\partial P}{\partial x_n}$. In other words, along the hedonic price schedule, the marginal rate of substitution between the *n-th* characteristic and consumption must be equal to the marginal cost of one extra unit of *n*.

Suppliers of residential housing maximize their profit subject to a cost function. That is, suppliers choose a point along the hedonic price schedule at which their iso-profit curve touches the marginal rate of substitution for individual consumers of housing. In other words, the suppliers choose a point where the higher cost of producing one extra unit of *n,* and the marginal benefit of the residence price, is equalized.

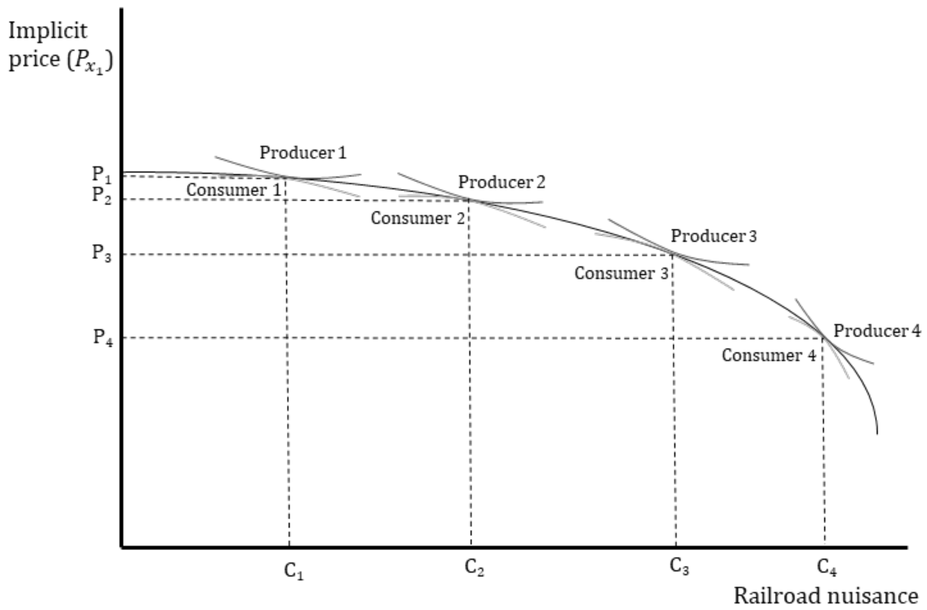

Figure A3: *The Hedonic price schedule in the residential property market for railroad nuisance*

Post-regularization methodology

The post-regularization method builds on the Least Absolute Shrinkage and Selection Operator (Tibshirani, 1996). The intuition behind the method is that a penalty term is included for each of the potential control variables. Variables that contribute little to the fit are eventually set equal to zero. That is, the method is used to achieve a sparse solution.

We use the post regularization methodology (Chernozhukov et al., 2015) to select the set of controls that are relevant for prediction . This selection methodology was used for three vectors: the vector of observable spatial characteristics $S_{i,t}$, the vector of property characteristics $X_{i,t}$, and the vector of year fixed effects interacted with postal code fixed effects. The first vector includes controls such as the distance towards the city center, while the second vector includes (interaction) controls between, for example, the dwelling type (e.g. apartment in a porch flat), the maintenance quality, and the age group. The included controls of the three vectors might provide useful additional information to the model. However, including too many variables might lead to overfitting, and some variables may be multicollinear with the treatment estimator.

The post regularization method proceeds in three steps. In the first step a lasso regression is estimated with $\log p_{i,t}$ as dependent variable and $X_{i,t}$ as an extended set of regressors. Since we are only interested in selecting the set of $X_{i,t}$ variables that are relevant for prediction, we chose not to penalize postal code and year fixed effects, which is shown in the second part of equation (2). The penalty term denoted by λ and the penalty loadings Ψ are used to approximate a sparse solution. (The loadings in Ψ are chosen to normalize the variables in $X_{i,t}$). Equation (2) is minimized in order to generate the residuals of $\log p_{i,t}$.

$$arg\min\left(\frac{1}{N}(\log p_{i,t} - \beta X_{i,t} - \tau_t T_t - \gamma_r Z_{r(i,t)_i})^2 + \frac{\lambda}{N}\|\Psi\beta\|\right) \tag{A2}$$

In the second step, a lasso regression is estimated with $\log Distance_i$ as the dependent variable and again the control variables $X_{i,t}$ as regressors. The method minimizes the underlying function to generate the residuals of $\log Distance_i$.

$$arg\min\left(\frac{1}{N}(\log Distance_i - \beta X_{i,t} - \tau_t T_t - \gamma_r Z_{r(i,t)})^2 + \frac{\lambda}{N}\|\Psi\beta\|\right) \tag{A3}$$

In the final step, the effect of $\log Distance_i$ on $\log p_{i,t}$ is determined by a bivariate regression of the residual log price on the residual distance (the orthogonalized versions).