Gabriel M. Ahlfeldt[♣], Volker Nitsch[⊗], Nicolai Wendland[♦]

Online Appendix to Ease vs. noise: Long-run changes in the value of transport (dis)amenities

Version: June, 2019

1 Introduction

This appendix complements the main paper by providing additional information and complementary results not reported in the main paper for brevity. We begin with a short review of the related capitalization literature in Section 2. In Section 3, we provide additional detail regarding the data used. Section 4 adds to the historical difference-in-differences analysis, providing additional detail on the construction and distribution of weights, robustness checks, and complementary analyses. Sections 5 provides a complementary analysis of the historical noise capitalization effect using a spatial differences approach. Section 6 complements the contemporary spatial differences analyses. Section 7 describes in detail how we compute the income elasticities and the station accessibility measures discussed in section 5.1 in the main paper. Section 8 explains how we estimate the extra costs for constructing an underground line instead of an elevated line, followed by an analysis of the aggregate effect of the reduction in noise emission on land values in Section 9. In Section 10, we examine the long-run change in land prices in Berlin. Finally, Section 11 provides additional background material on our calculations of property taxation.

2 Review of related capitalization research

A vast literature has inferred the value of non-marketed goods such as clean air (Chay and Greenstone, 2005), health risk (Currie et al., 2015; Davis, 2004), proximity to hazardous waste sites (Greenstone and Gallagher, 2008), crime risk (Linden and Rockoff, 2008), public school quality

[♣] London School of Economics and Political Sciences (LSE) & Centre for Economic Policy Research (CEPR). g.ahlfeldt@lse.ac.uk, www.ahlfeldt.com

[⊗] Darmstadt University of Technology. nitsch@vwl.tu-darmstadt.de

[♦] Touro College Berlin. wendland@urbancontext.org

(Cellini et al., 2010; Gibbons et al., 2013), high-speed broadband (Ahlfeldt, Koutroumpis, et al., 2016) or building externalities related to design and maintenance (Ahlfeldt and Holman, 2017; Rossi-Hansberg et al., 2010) from spatial variation in property prices. This approach is derived from the spatial equilibrium assumption in bid-rent theory, one of the workhorse tools in urban economics (Alonso, 1964; Mills, 1967; Muth, 1969). Essentially, it is argued that the value of (urban) land must offset all utility and productivity enhancing or depreciating factors, including noise and accessibility, if households are mobile and markets are competitive. The revealed preference approach is a popular tool in social cost-benefit analyses, which are, in many settings, the preferred method to evaluate welfare effects of public policies (Osborne and Turner, 2010).

Reviewing the literature, a number of studies have analyzed the property price effects of transportation infrastructure (e.g. Bajic, 1983; Baum-Snow and Kahn, 2000; Bowes and Ihlanfeldt, 2001; Damm et al., 1980; Dewees, 1976; McDonald and Osuji, 1995; Voith, 1993). Recent applications focus, in particular, on the property price effects of transport innovations, e.g. improvements of a road or rail network, to achieve better identification (Ahlfeldt, Moeller, et al., 2015; Billings, 2011; Gibbons and Machin, 2005; Hurst and West, 2014; McMillen and McDonald, 2004; Xu et al., 2015). The literature is surveyed in, among others, Mohammad et al. (2013), Bartholomew and Ewing (2011), Debrezion, Pels, and Rietveld (2007), Gibbons and Machin (2008), and Wrigley and Wyatt (2001). Overall, the findings suggest that transport infrastructures (and railways in particular) are typically associated with an increase in local property values. Quantitatively, the results in the literature are more heterogeneous, but based on the more robust evidence (exploiting variation over time) it seems fair to conclude that a one-kilometer reduction in station distance tends to increase house prices by about 2-7%. Cross-sectional hedonic estimates tend to be larger.

On transport-related disamenity effects, there is cross-sectional evidence that aircraft noise depreciates property prices (see J. P. Nelson, 2004 for a meta-analysis). Recent studies have also made use of quasi-experimental methods to identify aircraft noise effects (Ahlfeldt and Maennig, 2015; Boes and Nüesch, 2011; J. P. Nelson, 2004; Pope, 2008). The consensus in this literature is that a one-decibel increase in aircraft noise depreciates house prices by 0.5-0.6%. This is somewhat less than the mean of 0.92 % (median 0.74 %) across 24 earlier cross-sectional studies reviewed by J. P. Nelson (2008). As for road noise, Graevenitz (2018) reports that a one-decibel increase in noise above 55 db leads to a reduction in house prices in the range of 0.1 to 1.4 %. These results are similar to what Day et al. (2007) find. J. P. Nelson (2008) concludes that across 25 reviewed studies, the mean estimate for the effect of a one-decibel increase in house prices was -0.57 %. The evidence on other noise sources and, in particular, rail noise (A. C. Nelson, 1992) is somewhat less complete and robust (Navrud, 2002). Still, there is some evidence suggesting that railway lines may have negative property price effects at a highly localized level, possibly due to noise (e.g. Al-Mosaind et al., 1993; Debrezion et al., 2010; A. C. Nelson, 1992). Other dimensions of environmental quality, e.g., clean air or water, are typically associated with positive capitalization effects (Harrison and Rubinfeld, 1978; Leggett and Bockstael, 2000; J. P. Nelson, 1978), as are unspoilt natural spaces (Gibbons, 2015; Tyrväinen and Miettinen, 2000).

3 Background and data

3.1 Real GDP growth

In modern industrial economies, steady economic growth subject to some cyclicality has become the norm. As a result, an average consumer today can spend a budget that is more than seven times as large as that of their ancestors a century ago. This rise in income has important implications for consumer demand. With an income elasticity of demand below unity, the US consumer expenditure share on the necessities food and clothing has declined from 56.6% in 1900 to 17.3% in 2000 (U S Department of Labor 2006). At the same time, the historical increase in real income has more than proportionately freed up budget for the consumption of non-necessities. For some goods, including a clean, quiet or safe environment, quick access to jobs, or consumption amenities such as retail and entertainment, consumers pay indirectly via the cost of housing. It is, thus, no surprise that the consumer expenditure share on housing has increased by about 50% (from 23.3% to 32.8%) over the 20th century (U S Department of Labor 2006).

These changes are in line with a steady increase in real GDP per capita in the United States, Western Europe and the world as a whole. We compute the rate at which real GDP grew using the 2013 version of the Maddison Project data set (Bolt and van Zanden, 2014).[1] The data set represents a unique collection of real GDP per capita indices by country and world regions, brought together by a group of scholars who continue Angus Maddison´s work on measuring economic performance for different regions and time periods.

Because the data set is an unbalanced panel, it is empirically convenient to estimate the average annual growth rate by regressing the natural logarithm of real GDP per capita against a yearly trend variable. In Table A1, we show the results of such regressions for different countries and world regions. In column (2), we conduct a panel analysis to estimate the average annual growth rate

---

[1] To access the data set, visit http://www.ggdc.net/maddison/maddison-project/home.htm.

across about 170 countries and world regions. In each case, we include all available years since 1900. For the world as an aggregate unit of observation, we find an average annual growth rate of about 2%. The average annual growth rate across all available countries is only marginally smaller. This is about the rate at which the US, Western Europe, and Germany grew. Other world regions such as Latin America, Africa and Asia had slightly lower growth rates of about 1%-1.5% per year.

Tab A1. Real GDP per capita growth since 1900

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ln real GDP per capita (index)							
Year	0.019***	0.018***	0.020***	0.021***	-0.021***	0.015***	0.009***	0.016**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Country effects	-	Yes	-	-	-	-	-	Yes
Unit	World	Coun-tries	US	Western Europe	Ger-many	Latin America	Africa	Western Asia, Eastern Asia
N	63	11,856	111	111	111	63	62	129
r2	.973	.898	.969	.954	.907	.944	.819	.856

Notes: The data set is an unbalanced panel of country year observations covering the years from 1900 to 2010 from the Maddison Project. “World”, “Western Europe”, “Latin America”, “Africa”, “Eastern Asia” and “Western Asia” are aggregated series provided in the data set. Standard errors robust or clustered on countries where fixed effects included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2 Rail noise diffusion

As discussed in the main paper, we use a highly disaggregated map, containing 2007 estimates of the continuous sound level by the source of noise at a 10×10 meter grid from the Berlin Senate Department for Urban Development and the Environment (2013). The noise measure reflects the weighted average noise exposure over one year and all times of a day (L_{den}) at a reception point of four meters above the ground. Following the rules defined by the EU Environmental Noise Directive, the micro geographic noise map is the result of a simulation using a 3D model that is fit to actual noise measurements. The model incorporates features of the track design (e.g. speed, squeaking noises in curves, the presence of lubrication facilities) and the terrain geography (e.g. elevation of the track, built-up structure, bridges) that affect noise dissemination. We note that the data are provided for rail corridors extending 300 meters in either direction from an elevated rail line. Outside these corridors, data are missing as noise levels are deemed generally too low to be relevant. To avoid missing values, we expand the coverage by gradually deflating noise levels outside the corridors using a regression-based extrapolation approach. With this approach, we estimate the noise decay in track distance within the noise corridors and, using the estimated rate of decay, predict the noise levels outside the corridors. Because the noise levels at the margin of the noise corridors are generally low, this manipulation hardly affects the data as we measure noise in terms of decibel exceeding 40 decibels.

In Table A2, we analyze the spatial pattern of rail noise dissemination. The results in the first column reveal that a 0/1 dummy indexing parcels that immediately face the elevated rail line (those with an unobstructed view) already explains more than half of the spatial variation in rail noise (in excess of 40 decibels). In line with intuition, rail noise is highly localized within an area close to the viaduct. In the second column, we replace the view dummy with two sets of distance dummies. The first set consists of dummy variables that index mutually exclusive buffer areas drawn around the elevated rail line. We define the size of these buffers progressively, i.e. we increase the size as we move further away from the line (where there is less variation in noise). The second set consists of a similarly defined set of indicator variables indexing distance from station rings.

Relative to the residual category (800-1000 meters, where excess noise is essentially zero), rail noise levels increase by up to 22.7 db within the first 25 m buffer. Noise levels then decline steeply in distance from the track so that beyond 200 m, noise levels are economically marginal and beyond 400 meters statistically indistinguishable from the residual category. Conditional on the orthogonal diffusion from the track, there is also some variation along the track. Noise levels are significantly lower very close to stations, in line with the low speeds with which trains enter and exit stations. Adding the view dummy to the model in column (3) reveals some variation within the first distance-from-track categories, but has otherwise little impact. In columns (4) and (5), we distinguish between straight and curved line segments. Within the former, there is no conditional front-row effect, which is the expected result given that a 25-meter buffer along a straight section normally covers just about exactly those parcels (see Figure 2 in the main paper). In contrast, along the curved sections where the building structure is less regular, there is a sizable front-row effect conditional on distance, revealing that buildings represent significant obstacles to noise diffusion and protect areas in the background.

Overall, our analysis confirms that the empirically calibrated 3D noise model employed by the Senate Department for Urban Development and the Environment (2013) produces significant and plausible spatial variation in noise.

Tab A2. Noise diffusion along Line A

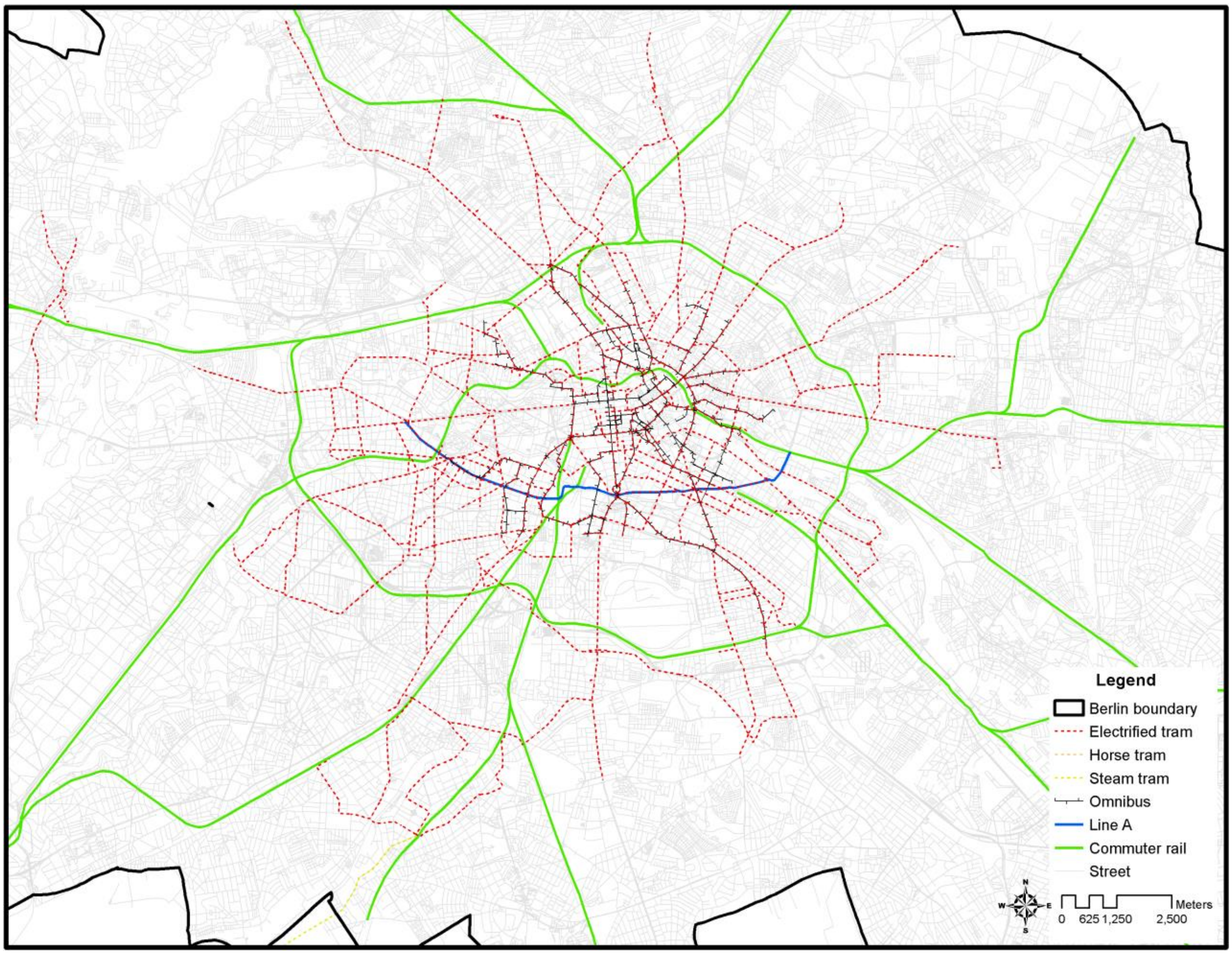
	(1)	(2)	(3)	(4)	(5)
	Noise (decibels) exceeding 40 decibels				
View (dummy)	19.664***		4.308***	0.243	5.415***
	(0.311)		(0.693)	(1.338)	(0.742)
0 m < Track distance <= 25 m		22.792***	18.763***	24.510***	17.091***
		(0.489)	(0.884)	(1.417)	(1.074)
25 m < Track distance <= 50 m		20.901***	17.941***	21.539***	17.291***
		(0.465)	(0.723)	(1.151)	(0.891)
50 m < Track distance <= 100 m		12.525***	11.778***	11.269***	12.102***
		(0.458)	(0.483)	(0.851)	(0.579)
100 m < Track distance <= 200 m		4.063***	4.096***	3.110***	4.819***
		(0.254)	(0.254)	(0.429)	(0.321)
200 m < Track distance <= 400 m		0.833***	0.834***	1.266***	0.655***
		(0.095)	(0.095)	(0.197)	(0.096)
400 m < Track distance <= 800 m		0.007	0.004	-0.000	0.004
		(0.014)	(0.014)	(0.000)	(0.018)
0 m < Station distance <= 25 m		-5.007***	-5.287***	-7.953***	-2.756
		(1.900)	(1.893)	(1.175)	(4.828)
25 m < Station distance <= 50 m		-3.293***	-3.427***	-5.578***	-3.557**
		(1.271)	(1.194)	(1.410)	(1.576)
50 m < Station distance <= 100 m		-1.344**	-1.456**	-5.030***	-0.771
		(0.639)	(0.627)	(1.275)	(0.694)
100 m < Station distance <= 200 m		-0.055	-0.175	0.122	-0.588
		(0.331)	(0.325)	(0.627)	(0.389)
200 m < Station distance <= 400 m		0.237*	0.233*	-0.264	0.374***
		(0.127)	(0.127)	(0.299)	(0.120)
400 m < Station distance <= 800 m		-0.010	-0.007	0.000	-0.011
		(0.013)	(0.013)	(0.000)	(0.016)
Constant	1.439***	-0.002	-0.001	0.000	-0.001
	(0.058)	(0.003)	(0.003)	(0.000)	(0.005)
Sample	All	All	All	Straight	Curved
N	5,456	5,456	5,456	1,651	3,805
r2	.554	.786	.793	.837	.783

Notes: Straight and curved distinguish between parcels along straight or curved line segments. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

3.3 Transport networks

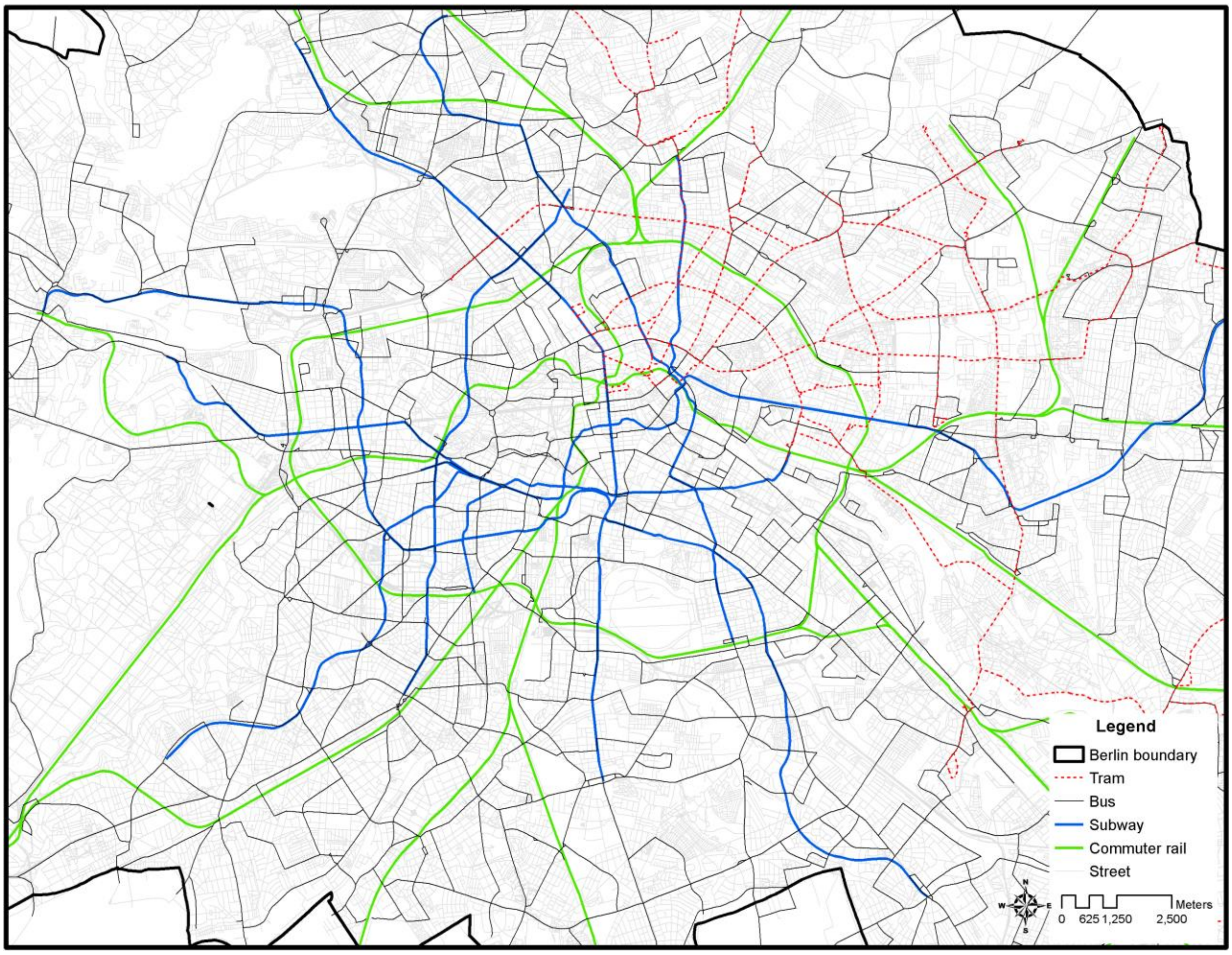
In the figures below, we illustrate the historical and contemporary transport geography of Berlin. The networks and modes illustrated are those which underlie the construction of the transport accessibility measures discussed in section 7.2 of this appendix. The figures show how the commuter rail network, despite significant technological upgrades (e.g. electrification from 1924 onwards) has remained roughly constant in terms of its coverage. In contrast, the subway network has since the opening of Line A developed into one of the densest networks in Europe. In line with the general settlement pattern, there was a dense network of complementary transport modes such as various tram systems and omnibuses within the central city around 1900, but the coverage was less complete in the suburbs. In contrast, the contemporary bus and tram (almost exclusively in the area of former East Berlin) networks cover a much broader area, reflecting the typical 20th century process of urban decentralization.

Fig A1. 1902 Transport geography



Notes: Own data collection. Own illustration based on Senatsverwaltung für Stadtentwicklung Berlin (2006).

Fig A2. 2006 Transport geography



Notes: Own illustration. Data from Ahlfeldt, Redding, et al. (2015) and Senatsverwaltung für Stadtentwicklung Berlin (2006).

4 Historical difference-in-differences models

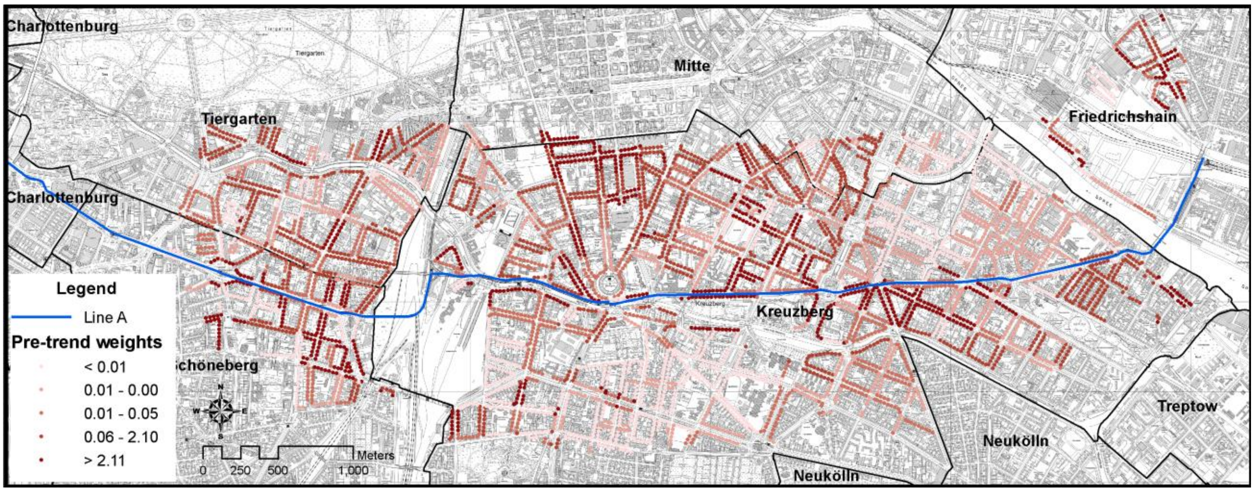
4.1 Weighted-parallel-trends difference-in-differences

It is well known that causal inference using difference-in-differences models relies on the untestable assumption of parallel counterfactual trends. The idea of the weighted estimator discussed in Section 3.2 of the main paper is to reweight observations in a way that one or multiple treatment measures become orthogonal to observable trends in an outcome over the pre-treatment period. The implicit assumption underlying the estimator is that if the weighting removes non-parallel trends successfully during the pre-treatment period (which can be tested), it will likely mitigate a potential non-parallel trends problem during the post-treatment periods (which cannot be tested). For a more formal introduction and evaluation of the estimator in the context of a Monte Carlo study, we refer to a companion paper (Ahlfeldt, 2018). For better accessibility, there is some overlap between the material presented in this appendix and in the companion paper.

4.2 Distribution of DD weights

The algorithm described in Section 3.2 of the main paper finds a vector of parcel weights, which ensures that the partial correlations between our two treatment measures, noise and station distance, with the 1881 to 1890 property price trend are minimized. The resulting weights are plotted in Figure A3. Overall, parcels with relatively high weights are distributed relatively evenly across the study area. The most notable finding are areas with relatively low parcel weights in the central southern section and the north-eastern section of the study area.

Fig A3. Spatial distribution of pre-trend weights



Notes: Classes defined based on quintiles. Own illustration using the Urban Environmental Information System of the Berlin Senate Department (Senatsverwaltung für Stadtentwicklung Berlin, 2006).

Table A3 compares descriptive statistics of the weighted sample to the unweighted parcel population. The distributions are fairly similar. In line with Figure A3, the mean parcel in the weighted sample is somewhat closer to the CBD (Stadtmitte, in the north) and the sub-centre (Kurfürstendamm in the west). But, overall, the weights inspection suggests that the results in the weighted DD will not be driven by a small number of non-representative parcels, so the estimates are hopefully not too far from average effects. Most likely, the DD will have greater internal validity than the historical spatial differences estimate, which is identified from a small number of parcels around the tunnel entrance.

Tab A3. Descriptive statistics in weighted vs. non-weighted sample

	Non-weighted			Weighted		
	Mean	Median	S.D.	Mean	Median	S.D.
Ln land price 1881	4.213	4.094	0.605	4.388	4.094	0.615
Ln land price 1914	5.854	5.768	0.521	6.058	5.991	0.591
Station distance (km)	0.502	0.491	0.237	0.467	0.486	0.226
Noise (10 db)	0.229	0.010	0.553	0.321	0.013	0.665
Distance from CBD	2.018	2.061	0.797	1.764	1.733	1.033
Distance from sub-centre	4.212	4.258	1.725	3.999	3.703	1.712
Distance from Line A track	0.543	0.517	0.265	0.559	0.503	0.310

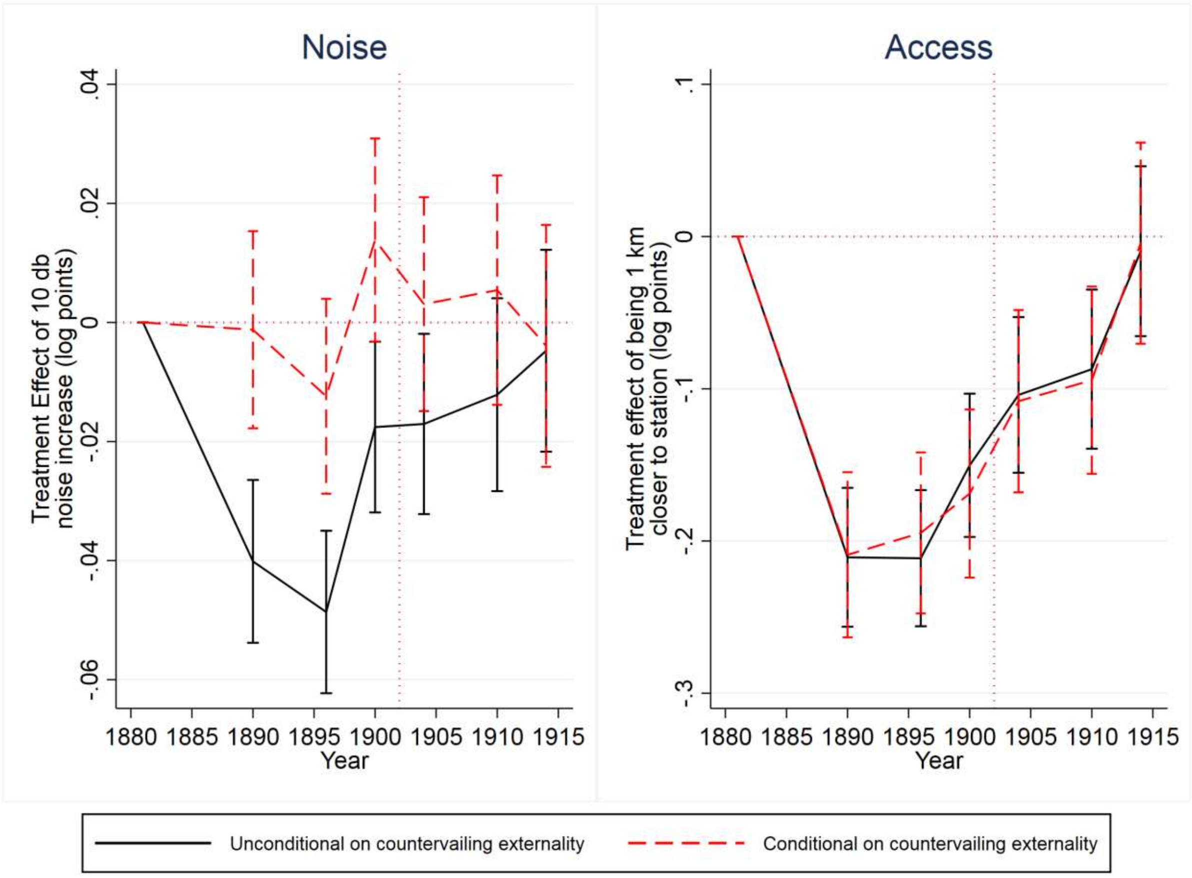
Notes: Source: Ahlfeldt (2018). Weights are constructed using the algorithm described in section 3.2 in the main paper and a Gaussian transformation of the mean 1881 to 1890 land price growth, the distance from the CBD and the distance from the most important sub-centre.

4.3 Time-varying OLS estimates

In section 3.2 of the main paper, we focus on our preferred weighted-parallel-trend (WPT) models. For comparison, we present the OLS-equivalent to Figure 1 below. The OLS results turn out to be somewhat difficult to interpret. According to our estimates, parcels located closer to to-be-opened stations experienced significantly lower land price growth, which points to a violation of the common trend assumption. As shown, the trend is flat from 1890 to 1896 and positive afterwards. To infer the effect of the rail line, a judgement has to be made on a baseline period that provides a counterfactual trend. Because the relative trends are flat, it may be tempting to choose the 1890 to 1896 trend as a baseline, implying a price effect of a one-kilometer change in station distance of about 0.2 log points over the subsequent 20 years. However, given that the concession for the line was granted in 1895, it is possible that the change in trend between 1881-1890 and 1890-1896 is attributable to the rail line, in which case the rail effect would be considerably larger. Another, not particularly conclusive feature of the estimated OLS station effects is the insensitivity of the point estimates to controlling for rail noise effects.

The estimated OLS rail noise effects are even less conclusive. Not controlling for station distance effects, parcels which later become exposed to rail noise experience a relative decline in prices up until 1896, when, shortly after the concession was granted, the trend reverses. Controlling for station distance effects, the land price trends do not seem to depend on the degree to which parcels become exposed to rail noise. This pattern is not in line with rail noise being a disamenity. If anything, the unconditional OLS estimates suggest that rail noise is an amenity.

Fig A4. Difference-in-differences: Time-varying treatment effects (OLS models)



Note: Time-varying treatment effects (α_z^S and α_z^N) based on baseline DD equation (1) and treatment function (2)in the main paper. Access parameters (effects of distance from station) multiplied by -1 so that positive shifts indicate positive economic effects. Vertical error bars indicate the 95% confidence interval based on standard errors that are clustered on parcels. Solid vertical lines denote the year of opening of the metro line (1902).

4.4 Alternative covariates and objective functions

In the models reported in section 3.2 in the main paper, the DD weights are constructed as a mix of parcels that are normal with respect to distance from the CBD, distance from the sub-centre, and land price growth over the 1881 to 1890 period. Ideally, weighted DD results will be replicable using different sets of uncorrelated weights as this suggests that identification is not driven by a limited number of units receiving high weights. Therefore, we have generated two alternative set of weights, which we use in Table A4 throughout models (3) to (6) (columns (1) and (2) replicate the baseline model for comparison). We stress that the weights in (5) and (6), which use distance from the Line A rail track instead of the 1881-1890 land price growth as a covariate, are virtually uncorrelated with the baseline weights used in Table 1 in the main paper (correlation coefficient: 0.076). Given this, it is reassuring that the estimates remain within the same ballpark.

Tab A4. Weighted DD: Varying predictors

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln land price (1881-1914)					
Distance x (km) x (t > 1900)	-0.174***	-0.191***	-0.183***	-0.214***	-0.256***	-0.315***
	(0.030)	(0.039)	(0.031)	(0.040)	(0.044)	(0.061)
Noise (10 db) x (t > 1900)	-0.034***	-0.046***	-0.039***	-0.051***	-0.018*	-0.037***
	(0.008)	(0.011)	(0.008)	(0.011)	(0.010)	(0.014)
Parcel effect	Yes	Yes	Yes	Yes	Yes	Yes
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
Anticipation effects	-	Yes	-	Yes	-	Yes
Predictors	Land price growth, distance from CBD, distance from sub-centre	Land price growth, distance from CBD, distance from sub-centre	Land price growth, distance from station, rail noise	Land price growth, distance from station, rail noise	Distance from rail track, distance from CBD, distance from sub-centre	Distance from rail track, distance from CBD, distance from sub-centre
N	37,933	37,933	37,898	37,898	38,192	38,192
r2	.931	.931	.929	.93	.915	.916

Notes: Source: Ahlfeldt (2018). Unit of observation is parcel-year (balanced panel). Weighted DD models use weights constructed to minimise the conditional correlations between noise and the 1881–1890 land price trend as well as access (distance from station) and the 1881–1890 land price trend. Weights are constructed using the algorithm described in section 2.4.1 and a Gaussian transformation of the listed covariates. Land price growth is the deviation from the mean 1881 to 1890 land price growth. Announcement effects are distance and noise variables interacted with 1896 and 1900 effects. Balanced panel of repeated parcel observations for 1881, 1890, 1896, 1900, 1904, 1910 and 1914. Standard errors in parentheses clustered in parcels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We similarly evaluate the sensitivity of the weighted DD estimates to using alternative objective functions in the weight-generating algorithm. As described in the main paper, we search over a parameter space defined by $q_1 = 0, 0.01, 0.02, \dots, 1$, $q_2 = 0, 0.01, 0.02, \dots, 1$, $q_3 = 0, 0.01, 0.02, \dots, 1$ to identiy the parameter vector $Q(q_1, \dots, q_m)$ in equation (4) in the main paper. To this end, we run r regressions of the form $\Delta\ln(P_{i,1890}) = c_r^0 + c_r^S \tilde{S}_i + c_r^N \widetilde{N}_i + \varepsilon_{ri}$, where $\Delta\ln(P_{i,1890})$ is the change in log land price from 1881 to 1890 and tilde denotes normalization by standard deviation. In each regression, observations are weighted by W_i, which depends on the vector $Q(q_1, \dots, q_m)$. In the baseline approach, we select the combination of parameters that minimizes the additive objective function $\sum_{V=(S,N)} \left(\widehat{c_r^V}\right)^2$. As alternatives, we consider a function $\max\left(|\widehat{c_S^1}|, |\widehat{c_S^2}|\right)$, to which we refer as min-max objective function, and a multiplicative function $\prod_{=(S,N)} \left(\widehat{c_m^q}\right)^2$.

In Table A5, we evaluate how the weighted DD estimates change as we alter the objective function in the algorithm. Evidently, the results are not particularly sensitive to the choice of the selection criterion.

Tab A5. Weighted DD: Varying objective functions

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln land price (1881-1914)					
Distance x (km) x (t > 1900)	-0.174***	-0.191***	-0.182***	-0.211***	-0.175***	-0.194***
	(0.030)	(0.039)	(0.031)	(0.040)	(0.030)	(0.039)
Noise (10 db) x (t > 1900)	-0.034***	-0.046***	-0.038***	-0.050***	-0.034***	-0.047***
	(0.008)	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)
Parcel effect	Yes	Yes	Yes	Yes	Yes	Yes
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
Anticipation effects	-	Yes	-	Yes	-	Yes
Objective function	Additive	Additive	Multipli-cative	Multipli-cative	Min-max	Min-max
N	37933	37933	38052	38052	37933	37933
r2	.931	.931	.93	.93	.93	.93

Notes: Source: Ahlfeldt (2018). Unit of observation is parcel-year (balanced panel). Weighted models use weights constructed to minimise the conditional correlations between noise and the 1881–1890 land price trend as well as access (distance from station) and the 1881–1890 land price trend. Weights are constructed using a Gaussian transformation of the 1881 to 1890 land price growth, the distance from the CBD and the distance from the most important sub-centre. Announcement effects are distance and noise variables interacted with 1896 and 1900 effects. Balanced panel of repeated parcel observations for 1881, 1890, 1896, 1900, 1904, 1910 and 1914. Standard errors in parentheses clustered in parcels. * p < 0.10, ** p < 0.05, *** p < 0.01.

We note that we have selected the covariates and the objective function used in our baseline approach following an inspection of how the weights address the non-parallel-trends problem during the pre-treatment period. In Table A6, we provide two tests of the conditional correlations between the treatment variables and pre-treatment outcome trends. Models (1–6) regress the change in ln land price over the 1881–1890 period (the period targeted by the algorithm) against both treatment variables. Models (7–12) replicate the exercise using the change in ln land price over the 1890–1900 period as a dependent variable. This (non-targeted) pre-treatment period has not been used in the computation of the weights, so it can be used in an overidentification test.

Models (1) and (7) present OLS estimation results. There is a significant correlation between station distance and land price growth over the targeted period. Compared to prices right next to a to-be-constructed station, prices at a 1km distance grow at a 0.221 log points higher rate (24%). There is also a significant correlation during the non-targeted period, however, with the opposite sign, suggesting the presence of unobserved effects that interact non-linearly with time. Conditional on the station-distance effect, the noise effect is insignificant. However, station distance and noise are correlated, which explains why the unconditional correlation between noise and the change in price is significant (to keep the presentation compact, we do not report the results of formal tests). The main takeaway from these results is that the parallel-trends assumption is violated during the pre-treatment period, thus, it seems likely that it does not hold during the post-treatment period.

The remaining models use weights to address this problem, which are constructed using different algorithms, objective functions and covariates. All approaches succeed in achieving their formal objective of reducing the correlation among treatments and trends during the targeted period (models 2–6). In several instances, the effects of both treatment variables are close to and not statistically distinguishable from zero. The models using the Gaussian transformation of land price growth as a covariate perform best in terms of the overidentification tests reported throughout models (8–12). Apparently, the treatment-trend correlation is low among parcels that experienced "normal" growth over the targeted period.

Tab A6. Marginal treatment effects on pre-outcome trends (placebos)

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln land price 1890 – ln land price 1881 (targeted period)					
Distance (km)	0.221***	-0.007	-0.024***	-0.006	-0.022**	-0.009
	(0.028)	(0.010)	(0.009)	(0.076)	(0.009)	(0.010)
Noise (db)	0.008	-0.004	0.001	-0.036**	0.000	-0.004
	(0.009)	(0.004)	(0.003)	(0.015)	(0.003)	(0.003)
r2	.0146	.0005	.0051	.0071	.0031	.0004
	(7)	(8)	(9)	(10)	(11)	(12)
	Ln land price 1900 – ln land price 1890 (not targeted period)					
Distance (km)	-0.052***	-0.038	-0.054	-0.172***	-0.051	-0.040
	(0.015)	(0.033)	(0.033)	(0.058)	(0.033)	(0.033)
Noise (db)	0.007	-0.011	-0.014	-0.012	-0.014	-0.011
	(0.006)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
r2	.0045	.0011	.0023	.0120	.0021	.0013
Objective	-	Additive	Additive	Additive	Multi.	Min-max
Covariates	-	Land price growth, distance from CBD, distance from subcentre	Land price growth, distance from station, rail noise	Distance from rail track, distance from CBD, distance from subcentre	Land price growth, distance from CBD, distance from subcentre	Land price growth, distance from CBD, distance from subcentre
N	5,456	5,456	5,456	5,456	5,456	5,456

Notes: Ahlfeldt (2018). Unit of observation is parcel. Columns (1) and (7) show results of separate OLS regressions of land price growth over the first (1) and second (2) period in the data against the treatment measures. The subsequent columns show results of weighted regressions, where the weights are recovered using objective functions, and a Gaussian transformation of the covariates indicated in the bottom of the table. Robust standard errors in parentheses. Additive minimises/multi./min-max minimises the sum/product/the largest of squared standardised coefficients on distance and noise. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

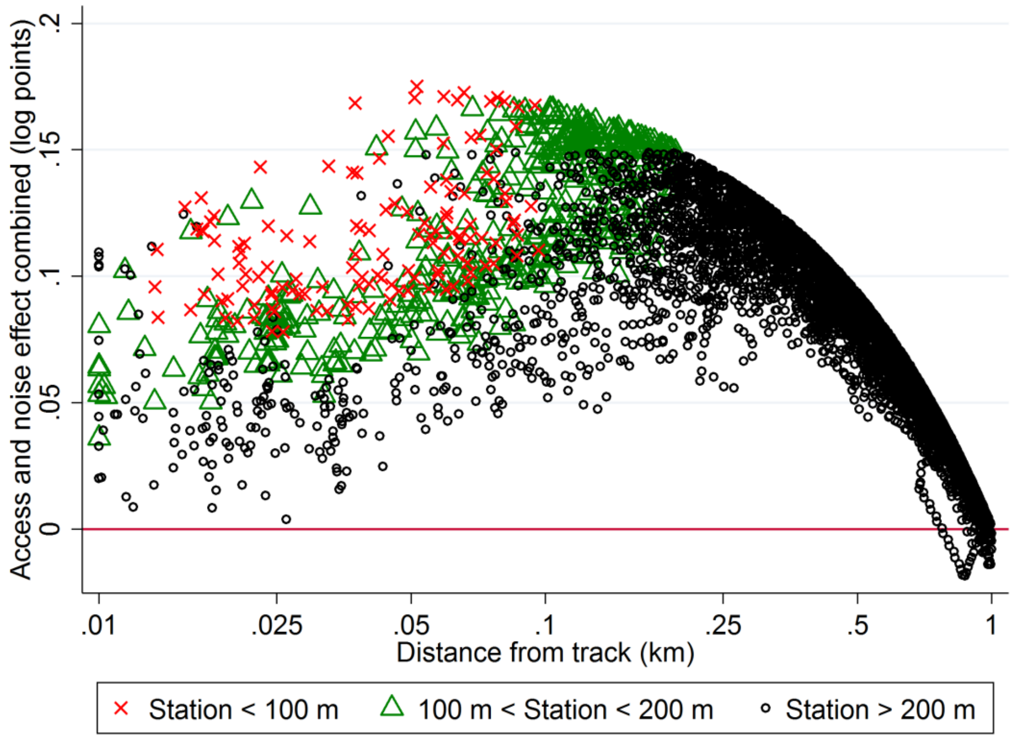
The weights used in models (2) and (8) are the most promising in terms of addressing non-parallel trends in the data, as they minimise the treatment variables' effects on outcome trends over the targeted and the non-targeted period. This is why we use these weights in the main paper.

4.5 Countervailing externalities

In the figure below, we explore the countervailing nature of rail externalities. To illustrate the net benefit from locating close to the elevated rail line, we plot the predicted joint station access and rail noise effect ($\hat{\alpha}^S S_i + \hat{\alpha}^N N_i$) from model (6) in Table 1 in the main paper against the straight-line distance from the elevated track. The figure illustrates that, for the clear majority of parcels, being located closer to the elevated line is associated with net benefits relative to locations at the outer margin of our study area. Beyond 100 m, the rail effect tends to be positive as reflected by the expected negative relationship between rail effect and track distance. At shorter distances, the net proximity effect tends to be negative, reflecting an increasing noise disamenity. This inverse U-shaped relationship is the expected pattern for a densely-developed area where noise tends to be highly localized. Some further interesting features of the countervailing nature of rail externalities are evident from the figure. As long as a location is sufficiently close to a station, the net effect of the line is positive, suggesting that the benefits from access to the line are relatively large. Land prices of parcels within 100 m of a station increase by at least 5% relative to those located at the margin of our study area. For parcels within a 100-200 m distance to a station, the effect is about half the size. Among the parcels further away from the nearest station, there are at least a handful for which the negative rail noise effect exceeds the positive station access effect.

As a plausibility check, we illustrate this negative net effect with a numerical example. The largest distance between two stations along the elevated line is about 1 km, implying that a parcel can be located at most 500 m from a station while still being located directly at the track. At 500 m, the benefit from rail access compared to the outer margin of the study area amounts to some $(0.5 \times 1.84 =)0.092$ log points. At this location, a parcel will be exposed to a very high noise level. Multiplying the 99th percentile in the distribution of rail noise (exceeding 50 db) of 26.1 db by the per-decibel noise effect of (-0.036/10) yields an effect of -0.93 log points, which indeed more than compensates for the accessibility effect.

Fig A5. Net benefit of proximity to elevated rail line



Notes: Figure illustrates the joint effects of station distance and rail noise predicted by model (6) in Table (1), formally: $\hat{\alpha}^s S_i + \hat{\alpha}^N N_i$. All effects are expressed relative to the outer margin of our study area. Therefore, we do a normalization by the mean across the predicted effects within the outmost 50 meters. Station indicates distance from the nearest station.

4.6 Time-varying implicit prices and treatment trends

In table A7, we provide a number of robustness checks on our preferred empirical model, reported in column (6) of Table 1. We begin by estimating an extended version of specification (1), allowing for time-varying implicit prices for various characteristics throughout columns (1-5). The interaction between time-invariant covariates and year effects are demanding controls, creating concerns of over-controlling. Some changes in implicit prices, e.g., distance from CBD or distance from the Kurfürstendamm, could be caused by the elevated line, implying a potential bad control problem (Angrist and Pischke, 2009). Yet, the station distance effect remains significant throughout all models, although it is reduced considerably. The noise effect becomes insignificant once we allow for time-varying effects for distance from rivers, lakes, or canals. Since the elevated track was partially built along a canal, however, it is difficult to separately identify the time-invariant effect of time-varying noise and the time-varying effect of time-invariant distance from rivers, lakes, or canals.

Tab A7. Weighted DD estimates: Robustness I

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln land price (1881-1914)					
Distance (km) × after	-0.130***	-0.094**	-0.129***	-0.114***	-0.073**	-0.097**
$(S_i \times (t > 1902)_t)$	(0.040)	(0.039)	(0.033)	(0.033)	(0.032)	(0.040)
Noise (10 db) × after	-0.036***	-0.030***	0.010	-0.007	-0.004	-0.014
$(N_i \times (t > 1902)_t)$	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.011)
Distance × (year – 1902)						-0.000
						(0.001)
Distance × (year – 1902) x $(t > 1902)$						-0.010***
						(0.003)
Noise × (year – 1902)						-0.000
						(0.000)
Noise × (year – 1902) × $(t > 1902)$						-0.002**
						(0.001)
Parcel effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Anticipation effects	Yes	Yes	Yes	Yes	Yes	Yes
Distance from CBD effects	Yes	Yes	Yes	Yes	Yes	-
Distance from Kudamm effects	-	Yes	Yes	Yes	Yes	-
Distance from water body effects	-	-	Yes	Yes	Yes	-
Distance from main street effects	-	-	-	Yes	Yes	-
Tram density effects	-	-	-	-	Yes	-
N	37,933	37,933	37,933	37,933	37,933	37,933
r2	0.934	0.936	0.942	0.944	0.944	0.931

Notes: Weighted DD models use weights constructed to minimize the conditional correlations between rail noise and the 1881-1890 land price trend as well as station distance and the 1881-1890 land price trend. After is a dummy variable indicating years after the line opening (1902). Announcement effects are distance and noise variables interacted with 1896 and 1900 effects. All other effects are time-invariant covariates interacted with year effects. Distance from CBD is defined as distance from the underground station "Stadtmitte" (downtown). Distance from Kudamm (slang for Kurfürstendamm) is defined as distance from Breitscheidplatz. Tram density is defined as kernel smoothed density of tram tracks within 2 km (bandwidth according to Silverman (1986)). Data is a balanced panel of repeated parcel observations for 1881, 1890, 1896, 1900, 1904, 1910, 1914. Standard errors in parentheses are clustered on parcels. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In column (6), we add interaction terms between our treatment measures and time trends (year – 1902) and the same interacted with an after-period dummy $(t > 1902)$. With this specification, we test for an effect of the treatments on levels and trends in land prices. The near to zero and insignificant pre-trend effects [Distance × (year – 1902) and Noise × (year – 1902)], once again, confirm that the weights achieve their purpose of eliminating the conditional correlations between pre-intervention price trends on the one hand and rail noise and station access on the other. The estimated station distance effect on land price levels $(S_i \times (t > 1902)_t)$ about halves in magnitude compared to the benchmark specification (column 6 of Table 1), but remains significant. The post-intervention trend in the distance treatment effect [Distance x (year – 1902) x after], however, reveals that ten years after the opening of the line the treatment effect has increased to some $-0.097 - 10 \times 0.01 = -0.197$ log points, which is remarkably close to the baseline effect reported in column (6) of Table 1. The post-intervention noise level [Noise × (year – 1902)] and trend [Noise

× (year – 1902) × ($t > 1902$)] effects are both negative as expected, though not individually significant. The cumulated effect of -0.037 after ten years, however, is not only close to the baseline estimate, but also statistically significant at the 1% level.[2]

4.7 View effects and semi-parametric station distance effects

In table A8, we further investigate the spatial pattern of the effect of the opening of Line A on nearby land prices. For comparison, column (1) replicates the baseline model from Table 1, column (6) in the main paper. In column (2), we replace the noise variable with a dummy indexing parcels with an unobstructed view on the elevated line. This dummy variable should also capture disamenity effects from rail vibrations as these tend to be highly localized. There is a negative effect associated with a direct view, however, at about -4.5%, the effect is significantly smaller than the noise effect implied by the baseline model for parcels exposed to very high noise levels (-9.3%, see discussion in section 4.4 in this appendix). The station distance effect is also substantially reduced, possibly because of the confounding effects of unobserved rail disamenities. Compared to the noise measure, the view dummy appears to be a less efficient disamenity measure. In column (3), we estimate the view effect conditional on the noise effect. The noise effect remains close to the baseline model, but the view effect is close to and statistically indistinguishable from zero. Because noise is highly localized, our noise and view measures are highly correlated, raising concerns about the separability of the effects in a multivariate analysis. To address this concern, we replicate the baseline model (including the noise measure, excluding the view measure) restricting the sample to parcels that do not offer a direct view on the elevated line because the view is obstructed by other buildings in column (4). In this model, we identify the noise effect excluding the parcels exposed to the highest noise levels. Yet, the noise effect remains close to the baseline model. Together, the evidence suggests that the disamenity effect of the rail line is primarily driven by noise and not by an unpleasant view.

---

[2] The standard error is computed as follows: $\exp\left(var\left(\widehat{\alpha^N}\right) + 10^2 \times \text{var}(\widehat{\alpha^{NT}}) + 2 \times (10) \times \text{cov}(\widehat{\alpha_A^N}, \widehat{\alpha_A^{NT}})\right) - 1$, where $\widehat{\alpha^N}$ is the estimated noise treatment level effect (as defined in equation (3) and $\widehat{\alpha^{NT}}$ is estimated trend effect [Noise × (year – 1902) × ($t > 1902$)].

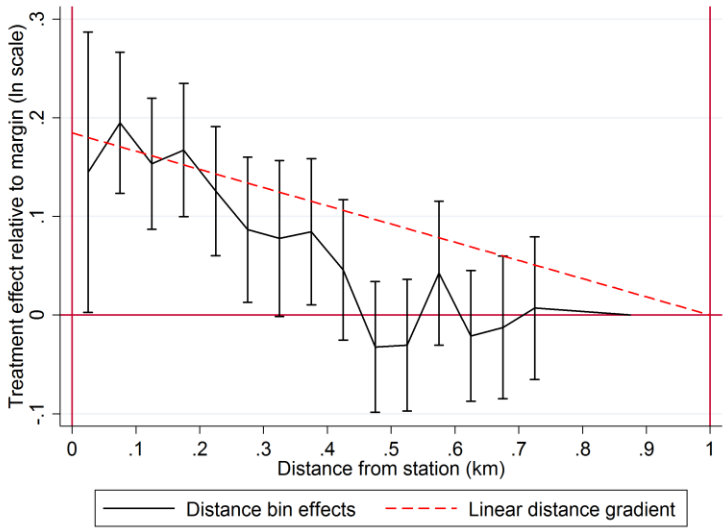
Tab A8. Difference-in-differences estimates: Robustness II

	(1) Ln land price	(2) Ln land price	(3) Ln land price	(4) Ln land price	(5) Ln land price	(6) Ln land price
Dist (km) x Post	-0.184***	-0.138***	-0.203***	-0.194***		
	(0.040)	(0.034)	(0.040)	(0.042)		
Noise (10 db) x Post	-0.036***		-0.039***	-0.040***	-0.060***	-0.064***
	(0.010)		(0.012)	(0.014)	(0.009)	(0.011)
View (0,1) x Post		-0.044***	0.001			0.009
		(0.016)	(0.021)			(0.021)
0 m < Station distance <= 50					0.145**	0.147**
					(0.072)	(0.073)
50 m < Station distance <= 100					0.195***	0.199***
					(0.036)	(0.037)
100 m < Station distance <= 150					0.153***	0.156***
					(0.034)	(0.034)
150 m < Station distance <= 200					0.167***	0.171***
					(0.034)	(0.035)
200 m < Station distance <= 250					0.125***	0.130***
					(0.033)	(0.034)
250 m < Station distance <= 300					0.087**	0.090**
					(0.038)	(0.038)
300 m < Station distance <= 350					0.078*	0.081**
					(0.040)	(0.041)
350 m < Station distance <= 400					0.085**	0.092**
					(0.038)	(0.038)
400 m < Station distance <= 450					0.046	0.065*
					(0.036)	(0.036)
450 m < Station distance <= 500					-0.032	-0.018
					(0.034)	(0.036)
500 m < Station distance <= 550					-0.031	-0.041
					(0.034)	(0.040)
550 m < Station distance <= 600					0.042	0.018
					(0.037)	(0.039)
600 m < Station distance <= 650					-0.021	-0.036
					(0.034)	(0.036)
650 m < Station distance <= 700					-0.012	-0.033
					(0.037)	(0.038)
700 m < Station distance <= 750					0.007	0.005
					(0.037)	(0.038)
Parcel effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Anticipation effects	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	Excluding direct view	All	All
N	37,933	37,933	37,933	37,933	37,933	37,933
r2	.93	.93	.93	.93	.932	.931

Notes: Model (1) is the baseline model. Weighted DD models use weights constructed to minimize the conditional correlations between the treatment variables and the 1881-1890 land price trend. Weights are constructed specifically for each combination of treatment variables (distance, noise, view). Announcement effects are distance and noise variables interacted with 1896 and 1900 effects. Balanced panel of repeated parcel observations for 1881, 1890, 1896, 1900, 1904, 1910, 1914. Standard errors clustered on parcels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In column (5) we address the question of whether the accessibility effect is sufficiently localized to justify a restriction to 1 km distance buffer. Therefore, we replace the linear distance measure with a set of distance bin dummies defined for mutually exclusive 50-meter rings up to 750 meters. The remaining distances are the residual category. We find that the station effect decays quickly, flattening out already at about 400-500m. Compared to the baseline category, the land prices within the innermost rings increased by about 0.19 log points, which is very close to the effect implied by the linear distance gradient estimate for a 1 km change in station distance. A graphical comparison is provided in the figure below. The results support the baseline model in that they suggest that there are unlikely station effects beyond one kilometer and that the one-kilometer distance effect is in line with a less parametric specification. In column (6), we add the view dummy to the model from column (5). Once more, we do not find evidence for a view effect.

Fig A6. Historical PTW-DD models: Distance from station gradient vs. distance bin effects



Notes: Figure compares the linear distance effect from the baseline model (Table A8, column 1) to the distance bin effects estimated in Table A8, column 5. Distance bins are dummy variables indicating mutually exclusive 50-meter rings defined for 0-50, ..., 700-750 meters. The residual category is 750-1000 meters. Error bars indicate the 95% confidence interval.

4.8 Stability of the hedonic price function

The interpretation of our difference-in-difference parameters as hedonic implicit prices hinges on the assumption that the hedonic function remained approximately constant over the study period (Kuminoff and Pope, 2014). In table A9, we provide a series of cross-sectional estimates of a simple hedonic model in which the land price is expressed as a function of some of the arguably most conventional location attributes in the hedonic literature. We find that the marginal effect of distance from the CBD remained approximately constant over the period from 1896 to 1910. The marginal effect of distance from the nearest park remained approximately constant from 1890 to 1910. In contrast, there is more variation in the effect of distance from rivers and canals, reflecting an increasing discount on the price of land close to waterways. However, it is likely that the variation in the water proximity effect is driven by an actual increase in proximity cost rather than a change in the hedonic implicit price of a time-invariant location factor. During our historical period, Berlin experienced sizable economic growth and a doubling of its population. Economic growth was fueled by rapidly increasing domestic cargo shipping, facilitated by significant investments into the regional waterway infrastructure. Between 1880 and 1914, several new canals (Oder-Spree-Kanal, Teltowkanal, Neuköllner Schifffahrtskanal, Hohenzollernkanals) and harbors (Urbanhafen, Südhafen Spandau, Tegeler Hafen, Osthafen, Hafen Britz, Tempelhofer Hafen, Steglitzer Hafen, Hafen Lichterfelde, Nordhafen Spandau, Westhafen) were constructed and a sizable fraction of the Spree river (Unterspree) was channeled. Moreover, in 1900, a large power plant (Heizkraftwerk Charlottenburg) opened at the Spree River shore close to our study area which was supplied with coal via the river (Natschka, 1971). Naturally, the growing traffic generated noise and pollution, rationalizing a land price discount close to waterways at a constant implicit price for amenities. Thus, overall, we view the evidence provided in the below table as supportive of a stable hedonic function around the years when Line A opened (1902).

Tab A9. Hedonic estimates by year

	(1) Ln land price	(2) Ln land price	(3) Ln land price	(4) Ln land price	(5) Ln land price	(6) Ln land price	(7) Ln land price
Distance from the CBD (km)	-0.381***	-0.353***	-0.319***	-0.298***	-0.276***	-0.272***	-0.234***
	(0.008)	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
Distance from parks (km)	-0.092***	-0.146***	-0.135***	-0.142***	-0.160***	-0.163***	-0.190***
	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
Distance from rivers and canals (km)	-0.043**	0.057***	0.138***	0.205***	0.247***	0.288***	0.314***
	(0.021)	(0.015)	(0.015)	(0.014)	(0.015)	(0.017)	(0.017)
Year	1881	1890	1896	1900	1904	1910	1914
N	5456	5456	5456	5456	5456	5456	5456
r2	.373	.662	.61	.574	.559	.505	.483

Notes: Standard errors in parentheses. Robust standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.9 Varying levels of spatial clustering

It is conventional to address serial autocorrelation by clustering standard errors (Bertrand, Duflo, Mullainathan, 2004). Following the convention, we cluster standard errors at the level of parcels, the unit at which we repeatedly observe our outcome of interest, throughout our empirical analyses. Here, we evaluate the effects of accounting for a spatial structure in the error term by clustering at higher spatial levels. To this end, we generate grids based on geographic coordinates of varying grid size. Table A10 presents the results of our baseline model when clustering standard errors at the level of those grid cells. We find that the estimated noise and distance effects remain significant when clustering up to the level of 200x200 meter grid cells. That said, we note that we already control for unobserved spatial heterogeneity at the finest possible level by means of parcel fixed effects. We have complete coverage of parcels within our study area, so we do not expect a spatial clustering problem in the sampling. And we have a parcel-specific assignment to treatment. Therefore, following Abadie, Athey, Imbens, and Wooldridge (2017), we keep the parcel-clustered model as our baseline.

Tab A10. Difference-in-differences estimates: Varying levels of spatial clustering

	(1) Ln land price	(2) Ln land price	(3) Ln land price	(4) Ln land price	(5) Ln land price	(6) Ln land price
Dist (km) x Post	-0.171***	-0.171***	-0.171***	-0.171**	-0.171**	-0.171*
	(0.031)	(0.050)	(0.065)	(0.078)	(0.085)	(0.093)
Noise (10 db) x Post	-0.028***	-0.028***	-0.028**	-0.028*	-0.028*	-0.028
	(0.008)	(0.010)	(0.014)	(0.016)	(0.017)	(0.018)
Parcel effect	Yes	Yes	Yes	Yes	Yes	Yes
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
Clustering grid (in m)	25 x 25	50 x 50	100 x 100	150 x 150	200 x 200	250 x 250
N	37933	37933	37933	37933	37933	37933
r2	.93	.93	.93	.93	.93	.93

Notes: Pre-trend weighted (PTW) models use weights constructed to minimize the conditional correlations between noise and the 1881-1890 land price trend as well as access (distance from station) and the 1881-1890 land price trend. Balanced panel of repeated parcel observations for 1881, 1890, 1896, 1900, 1904, 1910, 1914. Standard errors clustered on spatial grid cells as indicated in the table. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

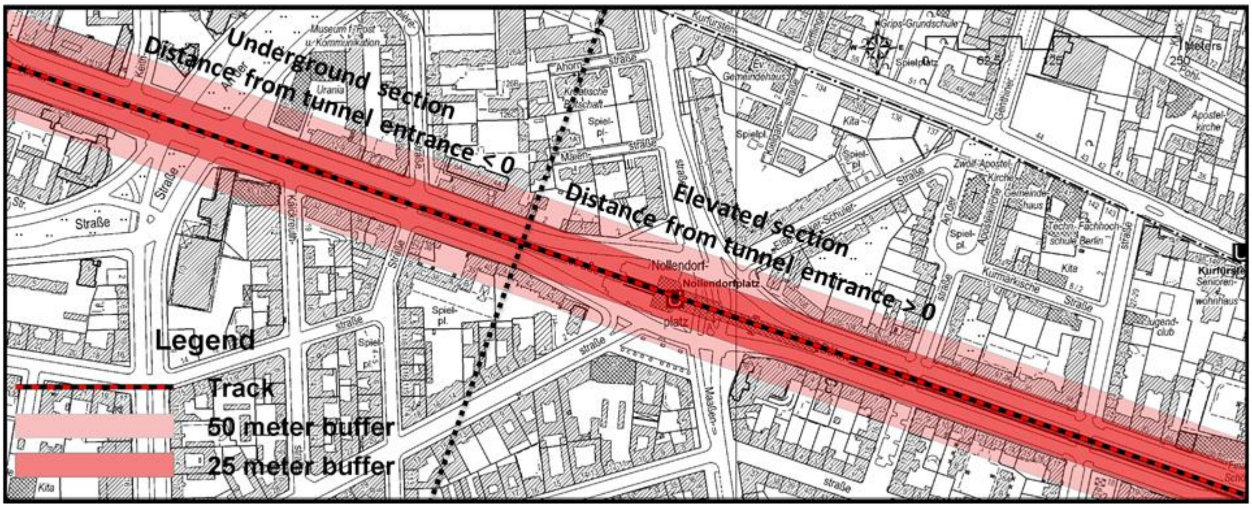
5 Historical spatial differences models

5.1 Empirical strategy

The specific character of Line A, in combination with the spatially highly disaggregated data available to us, enables us to identify the effect of the noise disamenity using a relatively sharp change in the spatial distribution of rail noise at the tunnel entrance where the line switches from being elevated to running underground and vice versa. Our SD approach to exploiting this feature is inspired by the regression discontinuity designs, in particular the fuzzy version (Hahn et al. 2001).

We note that the agreement to construct the line as an underground line within the boundaries of the city of Charlottenburg, whose authorities opposed the erection of an elevated line, was reached not earlier than three years before the inauguration. Therefore, for the change in noise at the tunnel entrance, anticipatory effects are unlikely. The idea of our SD approach is to wash out any effect of accessibility and other location characteristics that can be assumed to be similar within a very small area, thereby generating a precise estimate of the pure rail noise effect. Most notably, our land price data allows us to identify the effect using very small spatial windows from the rail track and the tunnel entrance. The figure below illustrates the micro geography around the tunnel entrance, which is right at the intersection of the two dotted lines. Evidently, a 50-meter buffer drawn around the track comfortably covers the boulevard under which the line is routed as well as the front rows of buildings framing the boulevard.

Fig A7. Micro geography at tunnel entrance



Notes: Dotted line is the orthogonal intersecting with the track at the tunnel entrance. Own illustration using the Urban Environmental Information System of the Berlin Senate Department (Senatsverwaltung für Stadtentwicklung Berlin 2006).

Our baseline SD specification takes the following form:

$$\Delta\ln(P_{ic}) = \alpha^N \Delta N_i + \rho K_c + X_i b + \varepsilon_i,$$

where $\Delta \ln P$ is the change in ln land price from 1900 to 1904, ΔN_i is a measure of change in rail noise (equal to N_i as the rail noise level in the initial period is zero), and K_c is a dummy variable indexing parcels within a spatial window from the track and the orthogonal that intersects with the track at the tunnel entrance (the black dotted line in the above figure). In our baseline specification, we set the window to 50 meters from the track and 500 meters from the orthogonal. The noise effect is then identified conditional on all unobserved effects on levels and trends that are common to this corridor. Notably, the corridor excludes the boundary between Berlin and Charlottenburg to the west of the tunnel entrance, so administrative boundary effects do not interfere with the within-corridor identification of noise effects. In the spirit of the regression discontinuity literature, we define a running variable D_i, which is the distance from the orthogonal, taking negative values within the underground section (to the left of the dashed orthogonal in the above figure) and positive values within the elevated section (to the right of the dashed orthogonal).

While we observe large variation in noise levels over a short distance around the tunnel entrance, the variation is not discrete in space since noise dissipates gradually in space. The positive noise values along a fraction of the underground segment of Line A correspond to non-compliers in a fuzzy discontinuity design. We use the interaction term $(D > 0)_i \times K_c$ as an instrumental variable for N_i, where $(D > 0)_i$ is a dummy variable that takes the value of one if the condition is true. The model is then estimated using 2SLS. To further strengthen the identification, we add a vector of control variables X_i, which captures trend heterogeneity with respect to observable characteristics. The coefficient of interest is α^N and provides a causal estimate of the extent to which the exposure of noise emitted by an elevated rail depreciates land prices under the identifying assumption that the conditional counterfactual trends are homogenous within the corridor (indexed by K_c).

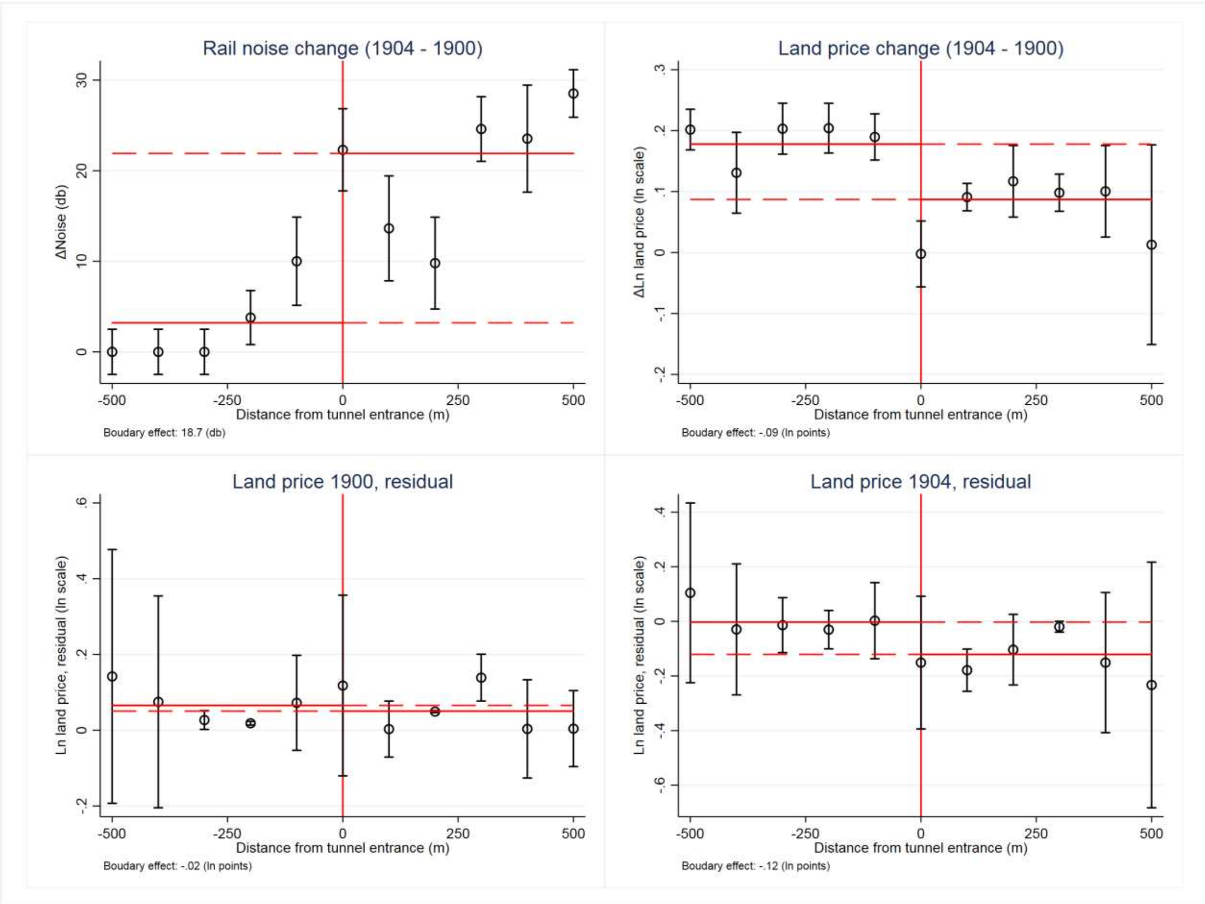
5.2 Baseline results

The tunnel entrance between the stations Nollendorfplatz and Wittenbergplatz, where Line A turns from an elevated line into an underground line, provides a source of sharp variation in rail disamenities. In the figure below, we illustrate the distributions of rail noise emitted by Line A around the tunnel entrance, as well as the distributions of land prices in levels and changes (1900-1904, the line opened in 1902). We restrict the sample to plots within close proximity to the track (50 meters), because this is where the noise disamenity of an elevated line is concentrated in this densely developed urban setting. We group parcels into 100-m-bins for which we then illustrate the mean value of an outcome as circles. The error bars allow for a quick evaluation of whether or not a within-bin mean is statistically different (at the 90% level) from the mean across all observations on the other side of the tunnel entrance.

Considering a rail corridor covering 500 meters in either direction of the tunnel entrance, the noise level (in excess of 40 decibels) along the elevated sections exceeds the noise level along the underground section by about 18 decibels on average (upper-left panel). Average noise levels are relatively low at about 100-200 meters from the tunnel entrance within the elevated section because parcels are somewhat further away from the track at the square *Nollendorfplatz*. There are some noise spillover effects onto the underground section of the line within the first 200 meters of the tunnel entrance, which is intuitive given that the rail line vanishes underneath a boulevard and there are no structures that would impede diffusion along the track. The average land price growth along the elevated section is 0.09 log points lower, implying a 5% noise effect for a 10-decibel increase that appears quite stable and significant (upper right panel). The bottom panels show that, controlling for other factors, a significant difference in land price levels exists after the opening of

Line A, but not before, which serves as a useful placebo test. A positive outlier in 1904 land prices at 300 meters (bottom-right panel) is also present in 1900 (bottom left panel) and, therefore, disappears in the time differenced SD model (upper right panel). The models in changes (upper right) and levels (bottom right) produce boundary effects that are similar (-0.09 vs. 0.12), suggesting that the noise estimates discussed here are comparable to the contemporary SD estimates in section 4 of the main paper.

Fig A8. Historical spatial differences in noise and land prices



Notes: Each circle illustrates the mean value of a dependent variable within a grid cell. One dimension of the grid cells are 100-m bins defined based on the distance from the orthogonal line intersecting with the track at the tunnel entrance (the dotted line in Figure A7). The other dimension is a 50-m-distance buffer around the track. Negative distances from the tunnel refer to the underground section. Solid horizontal lines indicate the means (weighted by the number of observations) within the underground (negative distance) and elevated (positive distance) segments. Error bars are the 90% confidence intervals based on robust standard errors from separate parcel-level regressions (within the buffer). For each outcome, we run one regression of the outcome against dummies indicating positive distance (≥ 0) bins, and another regression of the outcome against dummies indicating negative distance (<0) bins. For each bin, the error bar represents a test if the mean within the bin is different from the spatial counterfactual (the dashed line). The boundary effect corresponds to the difference between the two horizontal lines. Rail noise change from 1900 to 1904 is approximated by rail noise in 2007 (in excess of 40 db) since there was no rail noise in the study area prior to Line A (this assumes that noise levels did not change over time, see Section 2.3 for a discussion). Residual land prices (in the bottom panels) are from regressions of ln land prices against locational characteristics (distance from the CBD, Kurfürstendamm, the nearest major road, the nearest river, canal or lake, 1900 tram density, 1900 to 1904 change in tram density, dummies for residential land use and commercial land use) and lagged ln land prices (1890 and 1896).

In the table below, we report parametric estimates of the noise effect. For comparison, we begin with a parsimonious specification where we compare 1900-1904 land price growth rates across all parcels within the underground section and the elevated section of the line, i.e. there is no restriction to a specific source of variation in noise changes. As shown in column (1), there is a significantly negative noise effect of just about one fifth of the boundary effect displayed in the figure above. Once we implement the restriction of the identification to the difference in noise within the Line A corridor, however, the effects are well within the same range (columns 2 and 3). The model controlling for noise spillover effects on the underground section (column 3) yields a noise effect on land prices (-4.1% for 10-decibel increase) that is very close to the weighted DD estimate from Table 1, column (6) in the main paper. This finding is particularly reassuring because this model is closest to the weighted DD specification as it controls for unobserved heterogeneity in levels, but not in trends. In the next columns (4-5), the models become even more demanding by including kitchen sink controls that capture heterogeneity in land price trends with respect to observables. For instance, in the final column, we add lagged land prices (1890 and 1896) to control for the effect of unobserved characteristics on land prices in levels and trends. These extensions moderately increase the magnitude of the estimated noise effect.

Tab A11. Noise effects: Historical boundary discontinuity models

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln land price 1904 - ln land price 1900					
Noise (10 decibel)	-0.012***	-0.052***	-0.041***	-0.062***	-0.052***	-0.049***
	(0.003)	(0.014)	(0.012)	(0.015)	(0.013)	(0.013)
Noise spillover effect	-	-	Yes	-	Yes	Yes
Corridor effect	-	Yes	Yes	Yes	Yes	Yes
Controls	-	-	-	Yes	Yes	Yes
Lagged ln land prices	-	-	-	-	-	Yes
Noise IV	-	Yes	Yes	Yes	Yes	Yes
N	7,869	7,869	7,869	7,869	7,869	7,869
r2	.0019	-	-	-	-	-

Notes: Corridor effect is a dummy variable taking the value of one for parcels within a tunnel distance of 500 m (either side of the entrance) and track distance of <= 50 m, and zero otherwise. Noise instrument is a dummy variable taking a value of one for parcels along the elevated section of the corridor and zero otherwise. Noise spillover effect is a dummy variable taking the value of one for parcels within the corridor and within the first 250 m from the entrance along the underground section. Controls include distance from the CBD, distance from Kurfürstendamm (sub-centre), distance from canal, river or lake, distance from main street, distance from 1904 station, 1900 tram density, and change in tram density from 1900 to 1904. IV models estimated using 2SLS. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To substantiate these findings, we have conducted several robustness tests. First, we replicate the analysis of the simplest and least demanding SD specification for two periods before (1890-1896 and 1896-1890) and two periods after (1904-1910 and 1910-1914) the actual opening period of the line (1900-1904). In all four "placebo" models the point estimates of the noise effect are close to and statistically indistinguishable from zero (precisely estimated zeros), further indicating that our SD estimates reported in Table 2 are not driven by unobserved trends. Next, we estimate a reduced-form version of the SD model (using the instrument as explanatory variable), weight observations by their distance from the tunnel entrance, and add controls for spatial trends in the distance from the tunnel entrance. In another sensitivity analysis, we experiment with various combinations of track distances and tunnel entrance distances that define the rail corridor as well as different polynomial orders of distance trend controls. The results, presented and discussed in more detail in the next sub-sections, support our baseline findings.

5.3 Placebo treatment periods

In the table below, we replicate the SD model from Table A11, column (2) using land price growth during periods before and after the intervention as dependent variables. We find economically marginal and statistically insignificant effects for all periods, suggesting that the noise disamenity effect around the tunnel entrance capitalized into land prices within a relatively short period of time. Also, the absence of similar effects in the other periods makes it unlikely that the noise effects found in section 3 in the main paper are driven by unobserved trends that are correlated with, but unrelated to, the noise disamenity. In this context, we note that we use model (2) from Table A11 as the baseline model because it is the least demanding specification, presumably generating small standard errors. This imposes a harder hurdle for a falsification test, making the statistical insignificance of the estimates reported in the table below more meaningful.

Tab A12. Discontinuity in differences estimates: Placebo periods

	(1)	(2)	(3)	(4)
	Log land price 1896 - log land price 1890	Log land price 1900 - log land price 1896	Log land price 1910 - log land price 1904	Log land price 1914 - log land price 1910
Noise (10 db)	-0.006	-0.007	-0.003	-0.001
	(0.004)	(0.008)	(0.016)	(0.020)
Corridor effect	Yes	Yes	Yes	Yes
Noise IV	Yes	Yes	Yes	Yes
N	11,353	11,353	11,353	11,353
r2	-	-	-	-

Notes: 2SLS estimates. Corridor effect is a dummy variable taking the value of one for parcels within a tunnel distance of 500 meters (either side of the entrance) and track distance of <= 50 meters, and zero otherwise. Noise instrument is a dummy variable taking a value of one for parcels along the elevated section of the corridor and zero otherwise. Noise instrument is a dummy variable taking a value of one for parcels along the elevated section of the corridor and zero otherwise. Robust standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

5.4 Reduced-form analysis and local identification

In the table below, we alter the SD baseline specification in that we report reduced-form estimates using the dummy variable indexing the elevated segment of the rail corridor (the instrument K_c in the baseline model) as the explanatory variable. We apply this model to explain the spatial variation in noise as well as land price growth in columns (1) and (2). In line with Figure A8, we find noise levels are, on average, 17.5 decibels higher within the elevated segment of the rail corridor while land prices are 0.09 log points lower. Because column (1) reports the first stage of the model reported in Table A11, column (2), the noise effect implied by columns (1) and (2) in Table A13 (-0.09/1.75=-0.051) is mechanically the same as the result in Table A11, column (2).

In the next columns, we estimate the change in ln land price as one moves from the underground to the elevated section of the rail corridor, restricting the identification to observations that are closer to the tunnel entrance using a similar approach as in e.g. Ahlfeldt, Maennig, et al. (2016) and Ahlfeldt and Holman (2017). In columns (3-4), we assign weights TW to observations that decline in distance from the tunnel entrance TD as determined by a Gaussian kernel function:

$$TW_i = \frac{1}{\lambda_T \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{TD_i}{\lambda_T}\right)^2\right),$$

where λ_T is a bandwidth that determines the degree of smoothing. We set the optimal bandwidth of $\lambda_T = 133$ meters in column (3) per the Silverman (1986) rule.[3] In column (4), we use half the optimal bandwidth, which improves the local fit at the expense of a greater variance. In columns (5-6), we employ an alternative approach to estimating the discontinuity at the boundary using the following model:

$$\ln(\Delta P_{ic}) = \beta(D_i > 0)_i + \textstyle\sum_o \vartheta^o D_i^o + \sum_o \gamma^o((D > 0)_i \times D_i)^o + \varepsilon_i,$$

where ΔP is the change in ln land price from 1900 to 1904 and D_i is the distance from the tunnel entrance (the orthogonal) as in section 3.3 in the main paper (with negative values within the underground segment). $(D_i > 0)_i$ is a dummy variable that is one if the condition is true (within the elevated segment) and zero otherwise. This specification allows for separate distance trends on either side of the tunnel entrance of polynomial order *O* and provides an estimate of the change in land prices right at the tunnel entrance. We use a linear trend specification in column (5) and a quadratic trend specification in column (6). The estimated boundary effects in land prices across

[3] Formally, the bandwidth is chosen as $\lambda = 1.06 \times \sigma N^{-1/5}$.

columns (3-6) are consistently close to the baseline model in column (2). If anything, further narrowing the identification to variation close to the tunnel entrance marginally increases the boundary effect.

We note that the models in Table A13 differ from those in Table A11 in that we restrict the sample to the rail corridor rather than controlling for the rail corridor and using the full sample. In Table A11, we opt for the latter option because the additional observations help with the identification of the effects of the various control variables that we add in Table A11, columns (5-6). Here, we opt for the former option without any cost to keep the models simple and transparent. A similar control for distance trends within the rail corridor would otherwise require a full set of interactions between the rail corridor dummy and all distance variables (and their interactions with the elevated segment dummy).

Tab A13. SD in differences estimates: Reduced form estimates

	(1) Noise (10 decibels)	(2) Ln land price 1904 - ln land price 1900	(3) Ln land price 1904 - ln land price 1900	(4) Ln land price 1904 - ln land price 1900	(5) Ln land price 1904 - ln land price 1900	(6) Ln land price 1904 - ln land price 1900
Elevated track (Distance from tunnel > 0)	1.746***	-0.090***	-0.108***	-0.093**	-0.093*	-0.104*
	(0.192)	(0.024)	(0.033)	(0.043)	(0.048)	(0.054)
Track buffer (m)	500	500	500	500	500	500
Tunnel buffer (m)	50	50	50	50	50	50
Distance weights	-	-	Yes	Yes	-	-
Bandwidth	-	-	Optimal	1/2 x opt.	-	-
Linear distance trends	-	-	-	-	Yes	-
Quadratic distance trends	-	-	-	-	-	Yes
N	84	84	84	84	84	84
r2	.508	.157	.213	.167	.174	.26

Notes: Track buffer defines the sample of parcels included in terms of distance from the track. Tunnel buffer defines the sample of parcels included in terms of distance to the orthogonal intersecting with the track at the tunnel entrance (the vertical line in Figure 5 in the main paper). Distance weights decline in distance from the tunnel entrance and are constructed using a Gaussian kernel. Optimal bandwidth (133 meters) set per the Silverman (1986) rule. Linear trends are distance from the orthogonal and distance from the orthogonal interacted with being on the elevated section of the track. Quadratic trends are the same and the same variables squared. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.5 Sensitivity to corridor definition

Throughout the results reported above and in the main paper we focus on a rail corridor that covers 50 meters from the track and 500 meters from the tunnel entrance. The chosen threshold distances are preferred because they contain a reasonable number of observations (87) while, at the same time, ensuring that the included parcels are within the narrow (potential) noise impact area and sufficiently close to each other so that other locational factors can be reasonably assumed to be similar. In the table below, we present the results of a sensitivity analysis in which we experiment with different values for the two distance thresholds. We consider all combinations of 25 / 50 / 100 meters from the track, 250 / 500 / 1000 meters from the tunnel entrance excluding distance trends as well as controlling for linear and quadratic trends on each side of the tunnel entrance.

The pattern of results is generally comprehensive and reassuring. Excluding distance trends, we consistently find results within a relatively close range of our benchmark estimates. Including distance trends, the results become more volatile. With linear trends, we tend to find relatively larger effects when using shorter distance bands, and insignificant effects with the largest distance from track band. This is in line with the linear functional form being too restrictive to account for the trends around the tunnel entrance if the sample becomes too wide. With quadratic trends, we find the opposite pattern. This is in line with quadratic trends being a too flexible functional form if limited observations (shorter distance from tunnel entrance) are available. Because we allow the slopes of the trend to vary across both sides of the tunnel entrance, it is not surprising that higher order polynomials lead to somewhat instable results, either overestimating or underestimating the true discontinuity. Reassuringly, despite the increase in volatility of the estimate, the mean across the estimates conditional on linear as well as quadratic trends is very close to our benchmark results.

Tab A14. Discontinuity estimates: Sensitivity analysis

Distance from track	Distance from tunnel entrance	No trends Coefficient	S.E.	Linear trends Coefficient	S.E.	Quadratic trends Coefficient	S.E.
25	250	-0.18***	0.04	-0.16***	0.06	0.03	0.04
50	250	-0.11***	0.03	-0.07	0.05	-0.03	0.05
100	250	-0.10***	0.03	-0.06	0.06	0.00	0.06
25	500	-0.07**	0.03	-0.19***	0.04	-0.08	0.06
50	500	-0.09***	0.02	-0.09*	0.05	-0.10*	0.05
100	500	-0.04	0.02	-0.15***	0.05	-0.07	0.07
25	1000	-0.12***	0.03	-0.04	0.04	-0.22***	0.05
50	1000	-0.13***	0.02	-0.03	0.03	-0.16***	0.04
100	1000	-0.07***	0.02	0.01	0.03	-0.18***	0.05
Mean		-0.10		-0.09		-0.09	

Notes: Table summarizes results of variants of the column (2, no trends), (5, linear trends), (6, quadratic trends) in Table A11 (the baseline models are within the dotted lines). * / ** / *** denotes significance at the 10% / 5% / 1% level.

6 Contemporary spatial differences models

6.1 Housing capital

The theoretical framework outlined in Section 2.6 in the main paper implies that building capital is a linear transformation of housing value per land unit $K/L = \delta\psi H/L$. It follows, that the latter should be positively correlated with observable features of capital. Moreover, such features should be negatively correlated with station distance and rail noise since these are disamenties. In Tables A15 and A16, we put these predictions to an empirical test.

In Table A15 we stick to the natural log of the ratio of the transaction price over the parcel area as the dependent variable. Given the Cobb-Douglas production function, this variable in log terms is proportionate to the building capital per land unit. We regress this dependent variable against various observable features of building capital, controlling for space-time fixed effects and focusing on distinct parts of the study area. We find that the price per land unit is positively correlated with housing space. Conditional on housing space, the price per land unit is positively correlated with the quality of the housing stock, which we measure as two indicator variables encoded for buildings in good and poor condition. These variables are encoded by members of the committee of valuation experts who maintain the official transaction records and conduct onsite examinations where indicated. The price per land unit is also positively correlated with features of the building such as an elevator, a basement, or an underground car park. In line with intuition, building capital depreciates as a building ages, albeit at a relatively low rate of about 0.2% per year. This is in line with an old fabric in Berlin (median construction year in the sample is 1935) that is being maintained through regular investments into building capital.

Tab A15. Capital density vs. housing features

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln (transaction price / parcel area)					
Transaction year - construction year	-0.002***	-0.003***	-0.001***	-0.002***	-0.004***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln (floor space / parcel area)	0.571***	0.582***	0.814***	0.553***	0.561***	0.896***
	(0.002)	(0.004)	(0.014)	(0.007)	(0.006)	(0.031)
Building is in good condition (dummy)	0.414***	0.360***	0.566***	0.417***	0.269***	0.432***
	(0.005)	(0.007)	(0.024)	(0.013)	(0.008)	(0.036)
Building is in poor condition (dummy)	-0.480***	-0.324***	-0.220***	-0.336***	-0.357***	-0.227***
	(0.006)	(0.008)	(0.011)	(0.014)	(0.012)	(0.022)
Building has an elevator (dummy)	0.212***	0.017	0.094***	0.025	-0.073	0.103***
	(0.011)	(0.016)	(0.022)	(0.027)	(0.049)	(0.035)
Building has a basement (dummy)	0.191***	0.113***	0.003	0.154***	0.097***	-0.051
	(0.006)	(0.007)	(0.019)	(0.014)	(0.009)	(0.049)
Building has an underground car park (dummy)	0.262***	0.198***	0.214	0.235*	0.108*	0.550**
	(0.057)	(0.063)	(0.132)	(0.140)	(0.064)	(0.234)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Station x year effects	-	Yes	Yes	Yes	Yes	Yes
Sample	All	Berlin	Distance from CBD < 5 km	5km < distance from CBD < 10 km	Distance from CBD > 10 km	1 km elevated Line A buffer
N	71,231	70,584	14,462	20,539	35,321	3,228
r2	0.648	0.768	0.658	0.747	0.694	0.680

Notes: Unit of analysis is property transaction. Line A buffer is a dummy variable indexing properties within the one-kilometer (elevated) Line A buffer used in the historical DD analysis. Standard errors robust in (1) and clustered on station year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table A16, we use various variables that capture observable features of building capital in a specification that is otherwise identical to the baseline model in column (6) in Table 3 in the main paper. We find that that the density of housing space decreases in distance from the nearest station and in rail noise. The marginal effects are roughly within the range of the estimated effects on prices per land unit. This is in line with the theoretical framework in Section 2.6 of the main paper, which predicts $dln\left(\frac{K}{L}\right) = dln\left(\psi\frac{H}{L}\right)$.

Other features of building capital follow similar trends. The propensity of a building being in good condition decreases in station distance, while the propensity of a building being in poor condition increases in rail noise. Buildings further away from stations are less likely to have an elevator or a basement while buildings in areas with higher noise levels are less likely to have underground car parking.

Tab A16. Contemporary analysis: Other outcomes

	(1) Ln (floor space / parcel area)	(2) Building is in good condition (dummy)	(3) Building is in poor condition (dummy)	(4) Building has an elevator (dummy)	(5) Building has a basement (dummy)	(6) Building has an underground parking (dummy)
Distance (km)	-0.218***	-0.020*	0.011	-0.021***	-0.004	-0.000
	(0.024)	(0.012)	(0.011)	(0.006)	(0.010)	(0.001)
Rail noise (10 db)	-0.096**	-0.005	0.071**	-0.016	0.035	-0.006*
	(0.043)	(0.027)	(0.036)	(0.012)	(0.023)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	-	-	-	-	-	-
Station x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor x running variable	Yes	Yes	Yes	Yes	Yes	Yes
Noise instrument	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Station distance < 1 km	Station distance < 1 km	Station distance < 1 km	Station distance < 1 km	Station distance < 1 km	Station distance < 1 km
N	46,089	46,143	46,143	46,143	46,143	46,143
r2	.815	.414	.336	.403	.72	.255

Notes: Unit of analysis is property transaction. Controls include structure age, dummies for location within a block (corner lot, street front, backyard, etc.), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (excluding rail noise). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Running variable is distance from the tunnel entrance, taking negative values within the underground section (as in Figure 7). Noise instrument is a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise in models (4-6). Standard errors in parentheses are robust in (1) and (4), clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Rail effects within one kilometer of elevated Line A segment

The spatial scope of the contemporary analysis is not consistent with the historical analysis as we focus on a particular – newly constructed – rail line segment in the former, but cover the entire metro rail network in the latter period. To allow for a better comparability with the historical estimates, we interact the contemporary rail noise and station distance measures with a dummy variable denoting locations within a one-kilometer buffer from the elevated section of Line A. Summing over the baseline noise and distance effects and the respective interaction effects then gives the marginal effects within the buffer area.

In the table below, we replicate all models from Table 3 in the main paper in the same order, adding the interactions with the historical study area buffer. Once we control for station × year effects, none of the interaction effects is significant, i.e. contemporary rail amenity and disamenity effects within the area covered in the historical analysis are not significantly different from the rest of the city area. The interaction effects are also economically small. Our preferred noise (-0.125 vs. -0.122) and station distance (-0.177 vs. 0.144) estimates within the historical study area are marginally larger than the city-wide effects reported in Table 3 in the main paper.

Tab A17. Contemporary analysis: Line A buffer interactions

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln property transaction price / lot size					
Distance (km)	-0.125***	-0.125***	-0.141***	-0.146***	-0.140***	-0.150***
	(0.003)	(0.007)	(0.021)	(0.006)	(0.009)	(0.022)
Rail noise (10 db)	0.010	-0.025	-0.034**	-0.206***	-0.136***	-0.118**
	(0.011)	(0.015)	(0.015)	(0.043)	(0.051)	(0.051)
Noise x Line A buffer	-0.005***	0.000	0.001	-0.026**	-0.013	-0.007
	(0.001)	(0.001)	(0.001)	(0.013)	(0.016)	(0.015)
Distance x Line A buffer	-0.366***	-0.113	-0.035	-0.320***	-0.096	-0.016
	(0.057)	(0.072)	(0.074)	(0.057)	(0.072)	(0.074)
Dist. effect within buffer	-.49	-.239	-.177	-.466	-.237	-.166
Noise effect within buffer	.005	-.024	-.033	-.232	-.149	-.125
Line A buffer	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	-	-	Yes	-	-
Station x year effects	-	Yes	Yes	-	Yes	Yes
Corridor x year effects	-	-	-	Yes	Yes	Yes
Noise instrument	-	-	-	Yes	Yes	Yes
Sample	All	All	Station distance < 1 km	All	All	Station distance < 1 km
N	71,313	71,313	46,143	71,313	71,313	46,143
r2	.268	.584	.608	.272	.586	.61

Notes: Unit of analysis is property transaction. Line A buffer is a dummy variable indexing properties within the one-kilometer (elevated) Line A buffer used in the historical DD analysis. Controls include structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (excluding rail noise). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meters in both directions from a tunnel entrance. Running variable is distance from the tunnel entrance, taking negative values within the underground section (as in Figure 6). Instruments for noise and noise x Line A buffer are a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise in models and the same interacted Line A buffer. Standard errors in parentheses are robust in (1) and (4), clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Distance effects by rail system

In the historical analysis, we focus on the analysis of the opening of the first subway in Berlin, Line A, which represented a sizable transport innovation. The empirical design used in the historical analysis implies that we hold the effects of existing commuter rail network constant, i.e. we estimate a pure subway accessibility effect. As discussed in section 2 in the main paper, the commuter rail network, which was largely developed before the inauguration of Line A, used an older technology (steam trains). Because both networks, today, are comparable in terms of technology (all electrified metro rail), speed and frequency (at least in the central sections), we treat subway and commuter rail stations as perfect substitutes in our baseline analysis.

To distinguish the subway effect from the commuter rail accessibility effect in the contemporary analysis, we allow for an interaction effect between station distance and a dummy variable denoting whether a station belongs to the commuter rail network exclusively, i.e. does not offer access to subway services, in the table below. The non-interacted baseline distance term then reveals the distance effect for stations that belong to the subway network. As with Table A17, we replicate all models from Table 3 in the main paper adding the interaction.

In all models using the full sample of observations, there is an increase in magnitude of the baseline station distance effect (capturing subway effects) and a positive S-Bahn interaction effect, suggesting that a commuter rail station adds less value than a station that (also) offers access to the subway network. A pure subway station still offers sizable positive accessibility effects. In all models including station catchment area × year effects, the sum of the base line (Distance) and the interaction distance (Distance × S-Bahn) points to an effect of a station distance reduction by one kilometer of about 10%. Once we restrict station catchment areas to not exceed one kilometer, the differential distance effect of commuter rail stations is substantially reduced. Likely, the interaction effect is driven by station catchment areas that are, on average, larger for commuter rail stations because these are more frequently located in peripheral parts of the city (see Figure 3 in the main paper). Yet, even in our preferred model for the interpretation of the distance effect (column 3), the magnitude of the subway station distance effect, at -0.198, is larger than in the respective model of Table 3 in the main paper (-0.144). The baseline station distance effect reported in Table A18 (subway stations, including stations that also are served by commuter rail) makes for an interesting comparison to the historical analysis because Line A also included stations that offered access to commuter rail services (e.g. Warschauer Brücke). While these results suggest that our baseline station distance effect may be a lower bound estimate, they do not necessarily violate our assumption of subway and commuter rail stations being perfect substitutes because the sample of subway stations includes stations that offer access to both subway and commuter rail services and these stations are presumably particularly valuable.

Tab A18. Contemporary analysis: Distance effects by system

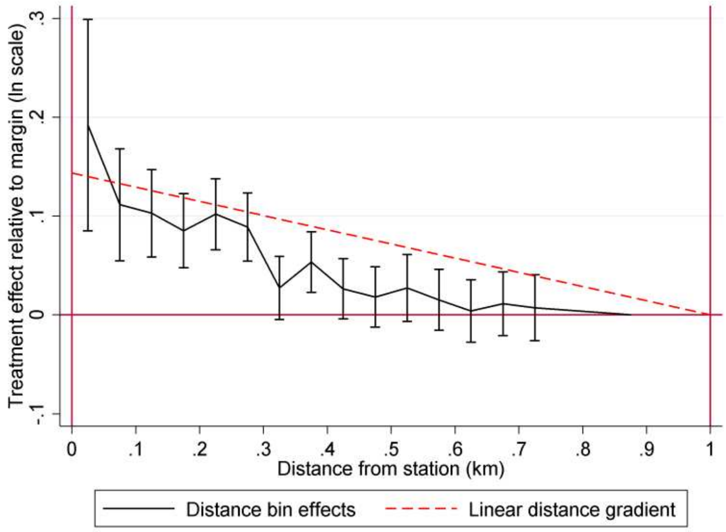
	(1)	(2)	(3)	(4)	(5)	(6)
	Ln property transaction price / lot size					
Distance (km)	-0.291***	-0.238***	-0.198***	-0.294***	-0.242***	-0.221***
	(0.006)	(0.016)	(0.029)	(0.006)	(0.016)	(0.032)
Rail noise (10 db)	-0.051***	-0.035**	-0.037**	-0.167***	-0.140***	-0.121**
	(0.011)	(0.015)	(0.015)	(0.032)	(0.049)	(0.049)
Distance x S-Bahn	0.231***	0.141***	0.097**	0.235***	0.146***	0.124***
	(0.006)	(0.017)	(0.040)	(0.006)	(0.017)	(0.043)
S-Bahn	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	-	-	Yes	-	-
Station x year effects	-	Yes	Yes	-	Yes	Yes
Corridor x year effects				Yes	Yes	Yes
Noise instrument				Yes	Yes	Yes
Sample	All	All	Station distance < 1 km	All	All	Station distance < 1 km
N	71,313	71,313	46,143	71,313	71,313	46,143
r2	.296	.586	.608	.299	.588	.61

Notes: Unit of analysis is property transaction. S-Bahn is a dummy variable indexing properties whose nearest station offers access to commuter rail (S-Bahn) exclusively. Controls include structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (excluding rail noise). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Running variable is distance from the tunnel entrance, taking negative values within the underground section (as in Figure 7). Noise instrument is a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise in models (4-6). Standard errors in parentheses are robust in (1) and (4), clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.4 Semi-parametric station distance effects

To investigate the station distance effect in a more flexible manner, we replace the linear distance variable in Table 3, column (3) in the main paper with a set of dummy variables denoting mutually exclusive 50-meter distance rings up to 750 meters. The remaining distances are the residual category. The results are illustrated in Figure A9 using a similar format as in Figure A6 (which presents a similar analysis based on the historical weighted DD models). The station effect decays quickly, flattening out already at about 400-500 meters. The results support the baseline model in that they suggest that there are unlikely station effects beyond one kilometer and that the one-kilometer distance effect is in line with a less parametric specification. Compared to the baseline category, however, the land prices within the innermost rings increased by about 0.19 log points, which is somewhat more than implied by the linear distance gradient estimate for a one-kilometer change in station distance. This effect is close to the maximum relative capitalization effect near stations found in the historical analysis (see Figure A6).

**Fig A9. Contemporary hedonics models:
Distance from station gradient vs. distance bin effects**



Notes: Figure compares the linear distance effect from the baseline model (Table 3, column 3) to distance bin effects. Distance bin effects are estimated using a model in which we replace the linear distance variable by a set of dummy variables indexing mutually exclusive distance rings defined for 0-50m, …, 700-750m. The residual category is 750-1000m. Error bars indicate the 95% confidence interval.

6.5 Reduced-form analysis: Varying corridor width

Compared to the historical analysis, we have increased the width of the rail corridor segments from 50 to 100 meters because the density of transactions in the contemporary period is smaller than the density of parcels in the historical analysis. In the table below, we evaluate the sensitivity of the results to a restriction to narrower rail corridors using a reduced-form version of the baseline empirical specification (we use the instrument as the explanatory variable). In columns (1) and (2), we estimate the conditional difference in noise and property prices per land unit between the underground and elevated segments of the rail corridors. Since the model in column (1) is the first stage of Table 3, column (6) model, the implied noise effect by columns (1) and (2) in the table below of $-0.177/1.445 = -0.122$ is mechanically the same as in the 2SLS baseline model. In columns (3-6) we reduce the width of the buffer to 75 (3-4) and 50 (5-6) meters. The implied noise effects remain within the same range, although the point estimates are somewhat smaller. The standard errors increase, resulting in insignificant price effects. These results substantiate the impression that the contemporary transactions data requires a slightly more generous definition of the rail corridor than the historical parcel data.

Tab A19. Reduced-form analysis with varying corridor width

	(1) Rail noise (10 decibels)	(2) Ln property transaction price / lot size	(3) Rail noise (10 decibels)	(4) Ln property transaction price / lot size	(5) Rail noise (10 decibels)	(6) Ln property transaction price / lot size
Elevated corridor segment (dummy)	1.821***	-0.177**	1.960***	-0.119	2.088***	-0.154
	(0.097)	(0.071)	(0.103)	(0.092)	(0.113)	(0.107)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Station x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor width	100 m	100 m	75 m	75 m	50 m	50 m
Sample			Station distance < 1 km			
N	46143	46143	46143	46143	46143	46143
r2	.664	.61	.664	.61	.663	.61

Notes: Unit of analysis is property transaction. Controls include station distance, structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (excluding rail noise). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Running variable is distance from the tunnel entrance, taking negative values within the underground section (as in Figure 7). Noise instrument is a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise in models (4-6). Standard errors in parentheses are robust in (1) and (4), clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.6 Local identification

In columns (1) and (2) of table A20, we further narrow the identification of the noise effect to properties closer to the tunnel entrances. We weight observations by their distance from the respective tunnel entrance using a similar approach as in section 5.4 in this appendix. To ensure that we use all observations for the identification of the effects of the various control variables we use a two-step estimation procedure. We first create adjusted property prices as the residuals plus the block fixed effect component from regressions of the natural log of the transaction price per land unit against a host of hedonic controls, year effects, and block fixed effects. Next, we run a weighted regression using the adjusted property prices as dependent variable, keeping observations within the rail corridors exclusively.

In an alternative approach, we add distance from the tunnel entrance trends (taking negative values within the underground section) interacted with a dummy indicating all rail corridors (columns 3 and 4). These models estimate a discontinuity in property prices conditional on a continuous spatial trend. In a further alteration, we allow for separate trends on both sides of the tunnel entrances by also interacting the trends with a dummy variable denoting the elevated parts of the rail corridors (column 5 and 6). These models estimate the change in property prices right at the tunnel entrance.

These different approaches to further restricting the identification to properties close to the tunnel entrances result in significantly larger property price effects, supporting the presence of a price discontinuity. The results suggest that our baseline model produces a rather conservative contemporary noise effect (see for comparison the 0.28 log points difference in the right panel of Figure 4 in the main paper).

Tab A20. Contemporary analysis: Reduced-form analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Adjusted ln property price / lot size	Adjusted ln property price / lot size	Ln property transaction price / lot size	Ln property transaction price / lot size	Ln property transaction price / lot size	Ln property transaction price / lot size
Elevated corridor segment (dummy)	-0.631*** (0.221)	-0.612** (0.245)	-0.333** (0.138)	-0.537*** (0.171)	-0.334** (0.138)	-0.333** (0.138)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Station x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor x year effects	Yes	Yes	Yes	Yes	Yes	Yes
Corridor x trends	-	-	Linear	Quadratic	Linear continuous	Quadratic continuous
Distance weights	Optimal bandwidth 182 m	1/2 optimal bandwidth 91 m	-	-	-	-
Sample			Station distance < 1 km			
N	463	463	46143	46143	46143	46143
r2	.851	.882	.61	.61	.61	.61

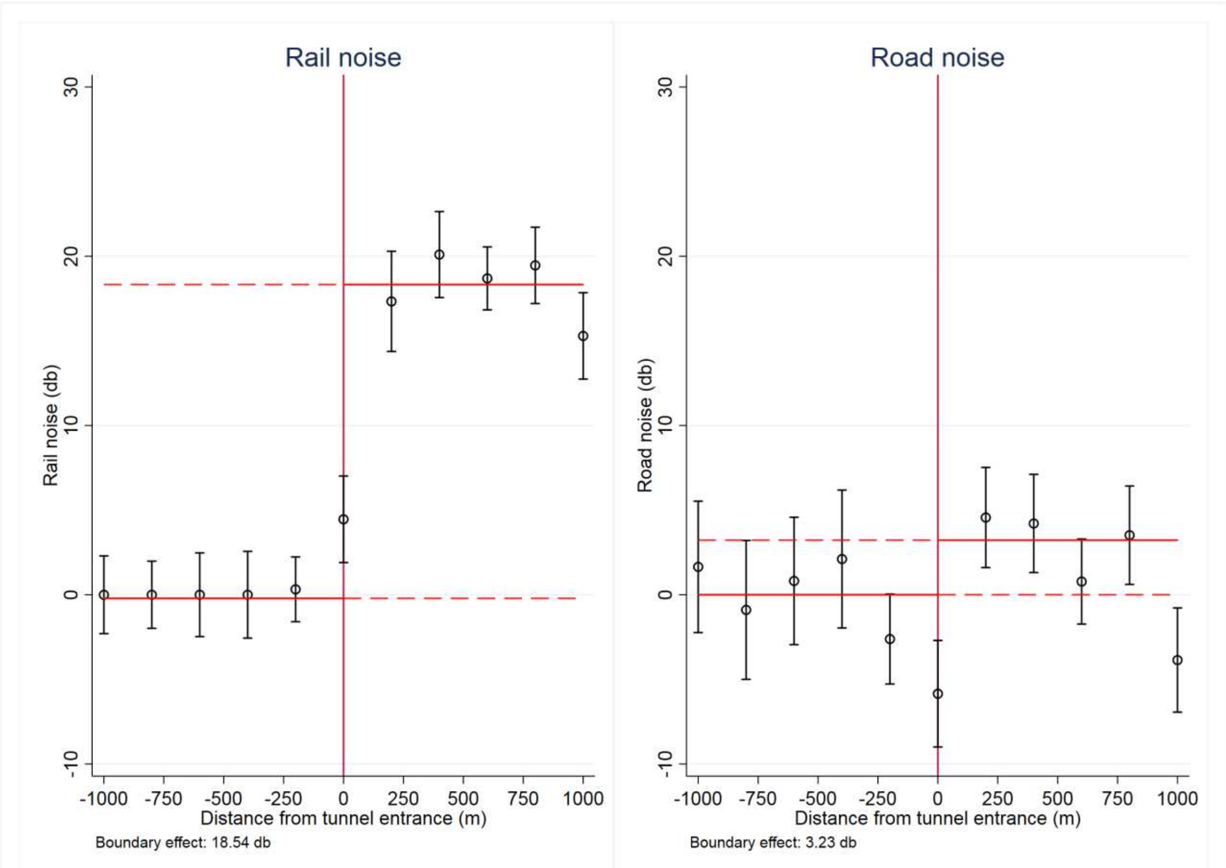
Notes: Unit of analysis is property transaction. Adjusted property prices are the residuals plus the block fixed effect component from regressions of the natural log of the transaction price normalized by lot size against a host of hedonic controls, year effects, and block fixed effects. Controls include station distance, structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (excluding rail noise). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Trends are based on the running variable, which is the distance from the tunnel entrance, taking negative values within the underground section (as in Figure 7). Common trends polynomial trends in the running variable of given order. Separate trends are the same, adding an interaction between trends and a dummy denoting the elevated section. Standard errors in parentheses are clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.7 Variation of road noise within rail corridors

One natural concern with the spatial difference design we employ is that factors other than rail noise may change at the tunnel entrance. A natural candidate is road noise. While we control for road noise when estimating the effect of rail noise on contemporary property prices, it is still interesting to evaluate how road noise changes as we cross the spatial boundary. If there was a large difference in road noise between the two sides of the tunnel entrance, this might indicate that other factors such as pollution, congestion, etc. that are difficult to control for, could also differ. Figure

A10 compares the change in rail noise at the tunnel entrance to the change in road noise. While road noise, on average, is higher along the elevated section (positive distance) than the underground section (negative distance), the difference is just a fraction (about one sixth) of the respective difference in rail noise. In table A21, we estimate the boundary effect conditional on observables, linear corridor-specific distance trends, and corridor-specific time effects. The boundary effect in rail noise remains within close range of the unconditional effect in Figure 4 and is highly statistically significant. In contrast, the boundary effect in road noise drops by about two thirds and is not statistically significant. These results substantiate our interpretation that our spatial difference strategy reveals a capitalization effect of rail noise that is not confounded by road noise effects.

Fig A10. Contemporary spatial differences in rail and road noise



Notes. Each circle illustrates the mean value of a dependent variable within a grid cell. One dimension of the grid cells are 200-m bins defined based on the distance from the tunnel entrance. The other dimension is a 100-m-distance buffer around the track. Negative distances from the tunnel refer to the underground section. Solid horizontal lines indicate the means (weighted by the number of observations) within the underground (negative distance) and elevated (positive distance) segments. Error bars are the 90% confidence intervals based on robust standard errors from separate parcel-level regressions (within the buffer). For each outcome, we run one regression of the outcome against dummies indicating positive distance (≥ 0) bins, and another regression of the outcome against dummies indicating negative distance (<0) bins. For each bin, the error bar represents a test if the mean within the bin is different from the spatial counterfactual (the dashed line). The boundary effect corresponds to the difference between the two horizontal lines.

Tab A21. Contemporary analysis: Conditional boundary effect in rail noise and road noise

	(1) Rail noise (1 db)		(2) Road noise (1 db)	
Elevated segment corridor (dummy)	18.332***	(0.631)	-1.163	(0.855)
Controls	Yes		Yes	
Year effects	Yes		Yes	
Corridor x year effects	Yes		Yes	
Corridor x running variable	Yes		Yes	
N	71,313		71,313	
r2	.241		.301	

Notes: Unit of analysis is property transaction. Controls include structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street, street noise (in model 1), and rail noise (in model 2). Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Running variable is distance from the tunnel entrance, taking negative values within the underground section (as in Figure 7). Noise instrument is a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise in models (4-6). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Comparison of historical to contemporary estimates

7.1 Income elasticities

In section 5.1 of the main paper, we define the willingness to pay (WTP) for a noise reduction of a representative individual at period t as the product of the percentage noise capitalisation effect in house price terms $(1-\delta_t)\alpha_t^N$ ($1-\delta_t$ is the land share as defined in section 2.6 of the main paper), the average income I_t, and the expenditure share on housing η_t:

$$WTP_t^N = -(1-\delta_t)\alpha_t^N \times I_t \times \eta_t$$

Taking log-differences and rearranging the equation we obtain the longitudinal income elasticity of the marginal cost of noise:

$$\frac{\Delta \ln WTP^N}{\Delta \ln I} = 1 + \frac{\Delta \ln(-\alpha^N)}{\Delta \ln I} + \frac{\Delta \ln(1-\delta)}{\Delta \ln I} + \frac{\Delta \ln \eta}{\Delta \ln I}$$

We use our baseline estimates from sections 3.2 (Table 1, column 3) and 4.2 (Table 3, column 6) in the main paper transformed into percentage terms to compute $\ln(-\alpha_{2000}^N) - \ln(-\alpha_{1900}^N)$. For the change in real income $\ln I_{2000} - \ln I_{1900}$ we use the German index of real GDP per capita from the Maddison Project (Bolt and van Zanden, 2014). As discussed in section 3.1 of this appendix, real GDP per capita in Germany since 1900 grew at rates of about 2% per year, in line with the general trend in the world. This corresponds to an aggregated increase by about 650%.

To account for changes in the land share in the value of housing, we make use of estimates reported by Knoll et al. (2017). Accordingly, the share of land at the total value of housing in Germany increased from 0.18 to 0.32 over the period from 1900 to 2000, which, in levels, is roughly in line with recent contemporary estimates of 0.25 for Berlin (Ahlfeldt, Redding, et al., 2015). For the expenditure share on housing we consider contemporary data from the Federal Statistical Office of Germany (2013) and historical data from Hoffmann (1965 [2006]). To obtain a consistently defined category in both periods, we define housing expenditures as the sum of expenditures on rent, utilities, and furniture. This expenditure share increased from 0.21 to 0.30 over the period from 1900 to 2000. This increase by about 50% is in line with the average increase across 14 countries over the same period reported in the working paper version of Knoll et al. (2014) and the increase in the respective U.S. share from 0.23 to 0.33 over the same period (U.S. Department of Labor, 2006). The longitudinal income elasticity of the marginal cost of noise is:

$$\frac{\Delta \ln(WTP^N)}{\Delta \ln I} = 1 + \frac{\ln 0.122 - \ln 0.036}{\ln 634 - \ln 100} + \frac{\ln 0.32 - \ln 0.18}{\ln 634 - \ln 100} + \frac{\ln 0.30 - \ln 0.21}{\ln 634 - \ln 100} = 2.2,$$

where the first term captures the effect of the change in land price capitalization effects, the second term captures the effect of the change in land share, and the last term captures the effect of the change in housing expenditure share. If we were to assume constant share parameters ($\Delta \ln(1-\delta) = \Delta \ln \eta = 0$), the pure effect originating from the change in land price capitalization would imply an income elasticity of 1.61, still greater than unity.

In the same way, we can compute the willingness to pay for a reduction in station distance and the respective long-run income elasticity using our baseline estimates of the station distance effect from Table 1, column (6) and Table 3, column (3):

$$\frac{\Delta \ln(WTP^S)}{\Delta \ln I} = 1 + \frac{\ln 0.144 - 0.184}{\ln 634 - \ln 100} + \frac{\ln 0.32 - \ln 0.18}{\ln 634 - \ln 100} + \frac{\ln 0.30 - \ln 0.21}{\ln 634 - \ln 100} = 1.4$$

We note that the income elasticity increases to 1.5 if we assume, as suggested by several robustness checks in section 6, a contemporary land price capitalization effect that equates to the historical land price capitalization effect.

In the per-capita accounting above, we have implicitly assumed that the value of non-marketed goods is the same to all members of a household. It is theoretically possible that the willingness to accept higher rents is skewed towards the valuation by selected household members, e.g. adults or wage earners. In an extreme scenario, a household's willingness to pay will be driven by the head of household alone. This scenario potentially leads to a different income elasticity because households have become smaller over time, so income has increased less in head-of-household terms than in per-capita terms. The average head-of-household willingness to pay for an amenity A is defined as:

$$HWTP_t^{A,} = (1 - \delta_t)\alpha_t^A \times \eta_t \times I_t \times n_t,$$

where n_t is the average number of persons per household and $I_t \times n_t$ is the real average head-of-household income. Since the log-change in head-of-household income is $\Delta \ln I + \Delta \ln n$, the income elasticity is defined as:

$$\frac{\Delta \ln HWTP^A}{\Delta \ln I + \Delta \ln n} = 1 + \frac{\Delta \ln(\alpha^A) + \Delta \ln(1 - \delta) + \Delta \ln \eta}{\Delta \ln I + \Delta \ln n}$$

Household size in Germany decreased from 2.0 in 1900 (Hoffmann, (1965 [2006])) to 1.5 in 2000 (Federal Statistical Office of Germany, 2017). In per head-of-household terms the longitudinal income elasticity of the marginal cost of noise increases to 3.1. The long-run income elasticity of the value of access increases to 1.7.

7.2 Interaction of rail and road noise

One of the limitations of our data is that we do not observe road noise – the predominant noise type in cities – during the historical period. Theoretically, it is possible that road noise and rail noise are mutually reinforcing, or the marginal disutility of rail noise might be lower if another noise source is present. Given that road noise was likely lower a century ago due to the absence of mass-produced cars, this can have important implications for the long run comparison of rail noise capitalization effects we conduct. If the noise sources were mutually reinforcing in their disutility effects, the increase in the capitalization effect of rail noise over time could be rationalized with the presence of higher levels of baseline road noise in the contemporary period. However, in the table below, we find evidence for the opposite. Higher levels of road noise are associated with lower rail noise capitalization effects. One way to interpret the interaction effect quantitatively is that if we reduce the level of road noise by 10 db, the marginal effect rail noise doubles. The implication is that in the absence of the presumably increased road noise, contemporary rail noise capitalization effects would be higher.

Tab A22. Contemporary analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln property transaction price / lot size					
Distance (km)	-0.050***	-0.096***	-0.130***	-0.052***	-0.097***	-0.141***
	(0.003)	(0.007)	(0.021)	(0.003)	(0.007)	(0.020)
Rail noise (10 decibel)	-0.019*	-0.029*	-0.041***	-0.167***	-0.185***	-0.160***
	(0.011)	(0.015)	(0.015)	(0.033)	(0.049)	(0.048)
Road noise (10 decibel)	-0.029***	0.005	0.002	-0.027***	0.004	0.000
	(0.003)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)
Rail noise x road noise	0.057***	0.018	0.019	0.121***	0.169***	0.159***
	(0.012)	(0.013)	(0.014)	(0.033)	(0.043)	(0.042)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	-	Yes	-	-
Station x year effects	-	Yes	Yes	-	Yes	Yes
Corridor x year effects				Yes	Yes	Yes
Rail noise instruments				Yes	Yes	Yes
Sample	All	All	Station distance < 1 km	All	All	Station distance < 1 km
N	71,313	71,313	46,143	71,313	70,665	45,364
r2	.377	.586	.609	.378	.106	.0538

Notes: Unit of analysis is property transaction. Controls include structure age, dummies for location within a block (corner lot, street front, backyard, etc.), dummies for building condition (poor, good), distance from nearest lake, river or canal, distance from nearest park or forest, distance from nearest landmark building, distance from nearest playground, distance from nearest main street. Station effects identify groups of properties which have the same nearest rail station. Corridor effects identify groups of properties within 100-meter buffers along a rail line, spreading 1000 meter in both directions from a tunnel entrance. Road noise rescaled to have a zero mean before generating the rail noise x road noise interaction. Instruments for rail noise and rail noise interacted with road noise are a dummy variable taking the value of one with the elevated segment of any rail corridor and zero otherwise and the same interacted with road noise. Standard errors in parentheses are robust in (1) and (4), clustered station x year effects in all other models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3 Network accessibility

In section 5.1 of the main paper, we provide a discussion of the effective accessibility that the stations analysed in the historical and contemporary periods offer. This is an important consideration because the station distance effects will not be comparable if the considered stations differ substantially in the connectivity offered. For this purpose, we compare the effective accessibility at a station location in the actual historical and contemporary scenario to the counterfactual scenarios that we establish in our empirical models, i.e. the absence of the considered stations.

To assess the loss of effective accessibility in either counterfactual, we compute a measure of accessibility for each station s, in every period t, and scenario z (actual vs. counterfactual). Following Ahlfeldt, Redding, et al. (2015), we aggregate the population (POP) at all potential destinations j that can be accessed from a station weighted by the bilateral transport cost c_{sjz} to create a measure of effective accessibility:

$$A_{stz} = \sum_j POP_{jt} e^{-\tau c_{sjtz}},$$

where c_{sjtz} is the mean of the travel times by automobile c_{sjtz}^{CAR} and public transport c_{sjtz}^{PUB}, weighted by the bilateral mode share of the car χ_{sjtz}:

$$c_{sjtz} = \chi_{sjtz} c_{sjtz}^{CAR} + (1 - \chi_{sjtz}) c_{sjtz}^{PUB}$$

For the historical period, we set $\chi_{sjtz} = 0$ because the automobile was virtually non-existent in 1900. For the contemporary period, we model the car share as a logit function of the relative travel time advantage of the automobile $\Delta c_{sjtz} = c_{sjtz}^{PUB} - c_{sjtz}^{CAR}$.

$$\ln\left(\frac{\chi_{sjtz}}{1 - \chi_{sjtz}}\right) = \zeta_1 + \zeta_2 \Delta c_{sjtz} <=> \chi_{sjtz} = \frac{\exp(\zeta_1 + \zeta_2 \Delta c_{sjtz})}{1 + \exp(\zeta_1 + \zeta_2 \Delta c_{sjtz})}$$

In computing c_{sjtz} we consider station locations as origins (s) and the geographic centroids of 93 historical city districts (*Ortsteile*) as potential destinations (j). These city districts are the smallest geographic unit for which 1900 population data are available. We compute the least-cost connections in terms of travel time between all origins and destinations taking into account the entire transport geography and all available public transit modes in GIS. As discussed in section 2 of the main paper, these modes include walking, buses, trams, subway (*U-Bahn*) and commuter rail (*S-Bahn*) in the contemporary period, as well as walking, horse-powered buses, horse-powered trams (one line), steam-powered trams (one line), electrified trams (the great majority of tram lines), and commuter rail (powered by steam) in the historical period. For the contemporary period, we also compute travel times by car. In each case, travel times are computed as the sum over the products of network segment lengths and mode-specific speed parameters along the fastest given route.

All contemporary speed parameters as well as the model parameters τ, ζ_1 and ζ_2 are borrowed from Ahlfeldt, Redding, et al. (2015). All historical transport cost parameters are average velocities derived from the study of historical timetables. The effective accessibility premium a station s offers in a period t is then simply defined as:

$$A_{st} = A_{s,t,z=actual} - A_{s,t,z=counterfactual},$$

where in the counterfactual we exclude the respective line segments (Line A in the historical period, the entire rail network in the contemporary period) when computing c_{sjtz}.

In the table below, we summarize the distribution of station accessibility premiums by period. We also provide an accessibility measure in which we aggregate the shares of the population of the 93 *Ortsteile* at the total population in a procedure that is otherwise identical to the one laid out above.

We find that the 11 elevated Line A stations in the historical period (first panel), in their effective accessibility effect, resembled the 275 subway and commuter rail stations of the contemporary network (second panel). While the distributions of absolute premiums in terms of accessible population are very similar, the relative premiums in terms of accessible population shares are somewhat larger in the historical period. This is intuitive given that the population of Berlin in 1900 was smaller than today (approx. 2m vs. 3.5m). If we focus on the segment of the contemporary rail network that belonged to the elevated Line A, the accessibility premia are somewhat larger, reflecting the central position of this segment within the contemporary network. This is in line with the somewhat larger than average point estimate of the station distance effect in this area reported in section 6.1 of this appendix.

Tab A23. Station accessibility premium

Variable	Obs.	Mean	Std. Dev.	Min	Max
Historical: Elevated Line A stations					
Accessibility premium: Population	11	100,134	78,217	17,146	234,579
Accessibility premium: Share of population	11	0.0389	0.0304	0.007	0.0911
Contemporary: All stations					
Accessibility premium: Population	275	100,564	47,551	13,139	234,374
Accessibility premium: Share of population	275	0.0302	0.0143	0.004	0.0703
Contemporary: Elevated Line A stations					
Accessibility premium: Population	10	174,694	43,043	104,116	234,374
Accessibility premium: Share of population	10	0.0524	0.0129	0.0312	0.0703

Notes: Accessibility premium is the accessibility index in the actual scenario minus the accessibility index in the counterfactual scenario. Accessibility index is either the transport cost weighted sum of population or of population shares across potential destinations. Actual scenario includes the entire network. Counterfactual scenario excludes Line A in the historical period and the whole metro rail network in the contemporary period.

7.4 Sorting

Sorting is a well-known phenomenon within cities. Different types of household live spatially segregated because they demand different types of locational amenities. For our long-run comparison it is critical to understand how the incomes of the marginal renter driving our capitalization results compare to the average renter.[4] The change in income of the average renter will be a reasonable approximation for our purposes if the marginal renter is representative for their cohort in both periods (historical and contemporary) or if they rank similarly within the distribution of incomes in both periods. Otherwise, the change in real income of the average renter will underestimate or overestimate the change in real income of the marginal renter. It is generally difficult to observe the

[4] Land prices are determined by the willingness-to-pay of renters and home buyers. For simplicity, and because the Berlin housing market has been dominated by renter occupiers, we refer to renters here.

income of the marginal renter. Historical data of this kind are virtually impossible to collect. To understand how the relative incomes of the relevant marginal renters have changed over time, we rely on indirect evidence.

It is likely that the renters who drive rents and ultimately land prices close to metro rail stations are frequent users of the system. To understand how the incomes of metro rail users compare to those of car users and public transit users more generally, we analyze the 2008 edition of the German micro commuting survey.[5] From this representative survey, we use information on slightly more than 8,000 trips within Berlin in 2008 for which we know the modes used, the distance travelled, the household income as well as the origin and destination city district (*Bezirk*) of the trip. In Figure A11, we summarize the distribution of income by mode. In keeping with intuition, the main takeaway is that those with higher incomes tend to travel by car more often.

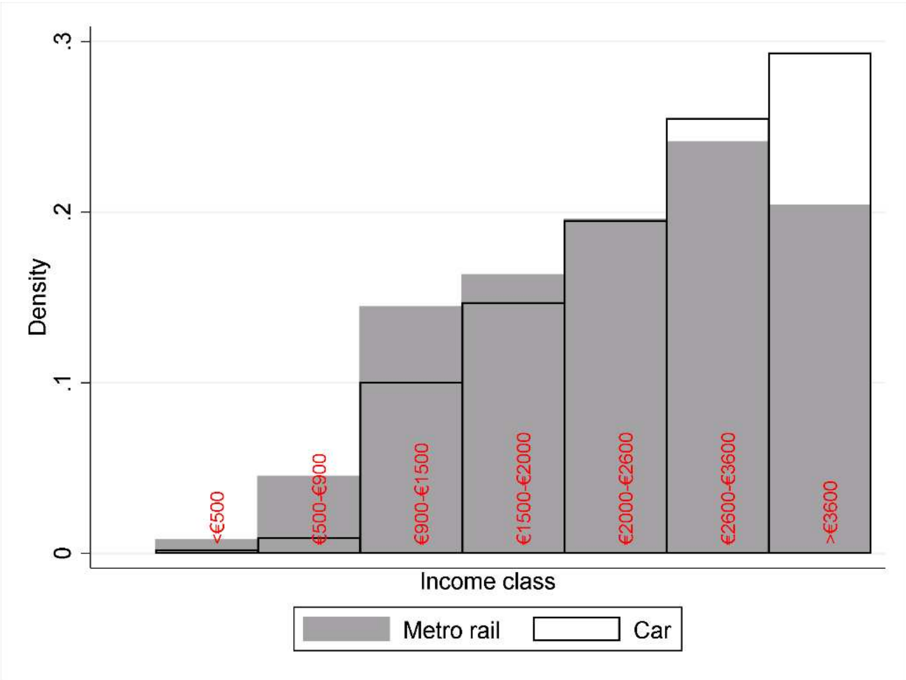
As with most surveys, income in this data set is given by category. For an econometric analysis, we construct a continuous income variable that, for each of the seven income categories reported in Figure A11, takes the value that corresponds to the mean over the category bounds. For the highest category, we assign the mean over the minimum value and twice the minimum value because no upper limit is given. The results of a Logit regression analysis reported in Table A24 confirms that users of the urban rail system, on average, belong to lower income groups. Controlling for trip length and origin and destination effects, the probability of using metro rail for a trip decreases by 0.276% for a one-percent increase in net-household income (column 2).

The results in Table A24 also confirm the strong intuition that the negative correlation between income and the use of public transport is driven by the availability of an attractive, but somewhat pricy alternative, the automobile. Before World War II, cars played a subordinate role as a means of transportation in Berlin. In relative terms, metro rail, therefore, was more attractive to higher-income groups as it was by far the fastest available mode of urban transportation (Leyden, 1933). To tailor to the needs of wealthier income groups, the historical trains operated on Line A featured special coaches that offered higher comfort at higher rates (Schmiedeke, 1997). More generally, some trains operated on several lines of the emerging metro rail network were casually referred to as banker trains ("*Bankierzüge*") due to their popularity among wealthy commuters (Reinhardt,

5 See for details, http://daten.clearingstelle-verkehr.de/224/1/Staedtepegel_SrV2008.pdf.

2015). Overall, it seems fair to conclude that metro rail users during the historical period were, on average and in relative terms, likely richer than metro rail users today.

Fig A11. Distribution by of trips by income category and transport mode



Notes: Income class refers to the net monthly household income. Metro rail includes trips where part of the journey is taken by U-Bahn (subway) or S-Bahn (suburban railway). Car includes trip where part of the journey is taken by car. Raw data are micro survey data from Ahrens et al. (2009).

Tab A24. Mode choice analysis

	(1) Metro rail for part of the trip (0,1)	(2) Metro rail for part of the trip (0,1)	(3) Car for part of the trip (0,1)	(4) Car for part of the trip (0,1)	(5) Other modes (no car and no metro rail) (0,1)	(6) Other modes (no car and no metro rail) (0,1)
Net income (€/month)	-0.093*** (0.017)	-0.129*** (0.019)	0.168*** (0.015)	0.170*** (0.016)	-0.080*** (0.015)	-0.071*** (0.017)
Distance travelled (km)		0.118*** (0.004)		0.029*** (0.003)		-0.178*** (0.006)
Mode elasticity	-.199	-.276	.31	.312	-.151	-.133
Origin effects	-	Yes	-	Yes	-	Yes
Destination effects	-	Yes	-	Yes	-	Yes
N	8,043	8,043	8,043	8,043	8,043	8,043

Notes: Unit of analysis is individual response in survey. Data from a 2008 representative travel survey Ahrens et al. (2009). Results from Logit estimations. Mode elasticity is the elasticity of the probability of selecting a model (over all alternatives) with respect to income, computed at the means of the distributions. Origin and destination effects are at the Bezirke level (12 city districts). Other modes include walking, cycling, bus, tram, and other shared transport. Robust standard errors. * p < 0.10, ** p < 0.05, *** p < 0.01

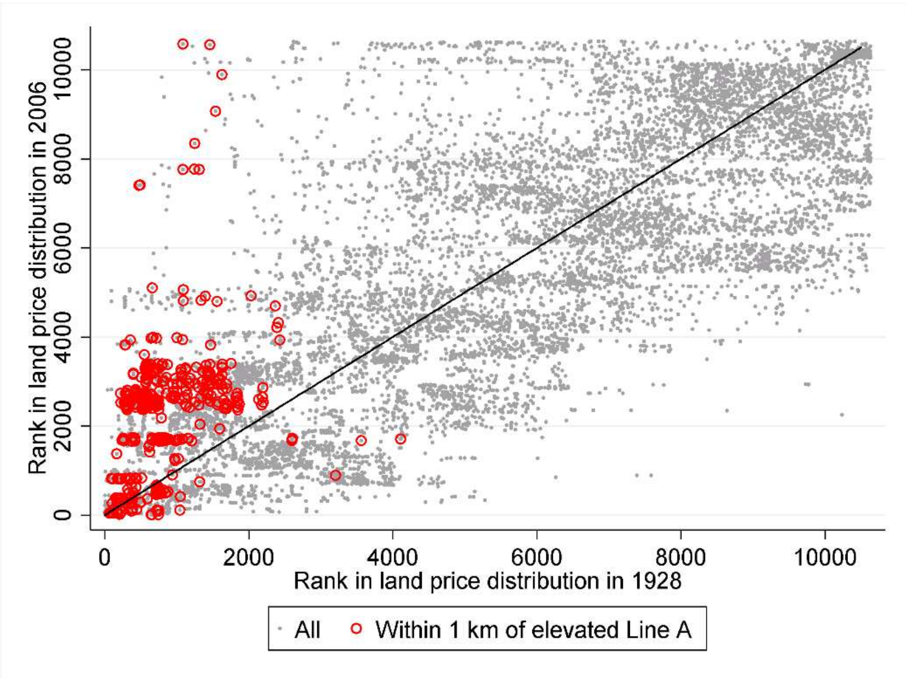
Income patterns within metro areas can vary substantially in space and time, following major trends such as suburbanization, white flight and gentrification. It is therefore possible that the residents living near metro stations today, in relative terms, are more or less wealthy than one hundred years ago, irrespectively of the mode of transportation they choose. Unfortunately, spatially disaggregated income data is not available for the historical period. Therefore, we cannot directly assess if the incomes of residents living near the areas considered in our capitalization studies increased above or below the average rate. As an imperfect approximation, we consider the change in land prices over time. Given that the built-up structure remained similar within the central parts of the city (where Line A is routed through), we would expect relative trends in land prices to be correlated with relative trends in incomes. To gain insights into differential trends, we compare land prices within the 1-km buffer around the elevated part of Line A relative to the rest of the city during the historical as well as the contemporary period. While in our contemporary capitalization study we also include other parts of the network, we have shown in Section 6.1 that the capitalization effects within a 1-km buffer around the elevated part of Line A are roughly representative for the capitalization effects along all elevated parts of the subway network.

The historical land price data from the Müller maps we use in our capitalization studies are available for the central parts of the city only. The first summary of land prices for approximately the entire area within today's city boundaries is available for 1928 (Kalweit, 1929). Using the land price data set compiled by Ahlfeldt, Redding, et al. (2015), we focus on a comparison of 1928 to 2006. In Figure A12, we compare the rank a block occupies in the distribution of land prices in 1928 to its rank in 2006 distribution, where rank one refers to the block with the highest land price. Both rank measures are positively correlated, revealing some degree of persistency in the internal structure of the city. Most of the blocks within the Line A buffer, however, have a high rank (low number, high land price) in 1928, but a low rank (high number, low land price) in 2006. In relative terms, these blocks are perceived as being less attractive during the contemporary period.

In Table A25, we subject the descriptive evidence to some simple econometric tests. We begin by regressing the long-difference in log land prices against a dummy variable that indicates the Line A buffer (column 1). To rule out that changes in land prices are driven by changes in economic density instead of locational attractiveness we control for long-differences in log population, log employment, and log floor area ratio (the ratio of total floor space over land lot size) in column (2). In column (3), we control, in addition, for a range of lagged variables in levels to control for correlated long-run trends. In columns (4-6), we estimate similar models using long-differences in the rank measure introduced in Figure A12 as a dependent variable. The estimates confirm the descriptive evidence from Figure A12. Our preferred estimate from column (2) suggests that in relative terms, land prices in the buffer area decreased by more than 60% (=exp(-0.953)-1).

One interpretation is that this area close to Line A came out as a loser from the long-run cycle of sub-urbanization and gentrification that has been typical for many cities during the 20th century (McMillen, 1996). An alternative explanation is that the area has not yet recovered from the potentially detrimental effects of being close to the former Berlin Wall during the division period. In any case, it seems likely that such a remarkable decrease in the relative price of land is associated with a decrease in the relative income of the local population.

Fig A12. Ranks in the distributions of land prices in the historical vs. the contemporary period



Note: Unit of analysis is housing blocks. Data from Ahlfeldt, Redding, et al. (2015). Rank one corresponds to the highest land price within a period. Sample restricted to a balanced panel.

Tab A25. Long-run change in relative land price close to elevated Line A

	(1) Ln 2006 land price - ln 1928 land price	(2) Ln 2006 land price - ln 1928 land price	(3) Ln 2006 land price - ln 1928 land price	(4) Rank 2006 - Rank 1928	(5) Rank 2006 - Rank 1928	(6) Rank 2006 - Rank 1928
Within 1 km of elevated Line A (0,1)	-1.262***	-0.953***	-0.290***	1463.820***	401.594***	192.957***
	(0.031)	(0.033)	(0.032)	(72.271)	(68.061)	(67.693)
Difference controls	-	Yes	Yes	-	Yes	Yes
Level controls	-	-	Yes	-	Yes	Yes
N	10641	10641	10641	10641	10641	10641

Notes: Unit of analysis is housing blocks. Data from Ahlfeldt, Redding, et al. (2015). Differenced controls are change in ln floor area ratio (FAR) from 1928 to 2006, change in population from 1936 to 2006, and change in employment from 1936 to 2006. Level controls are ln land price in 1928, ln FAR in 1928, ln population in 1936, ln employment in 1936, distance from the CBD, distance from the nearest lake, river, or canal, and distance from the nearest park. Robust standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Briefly summarized, the indirect evidence presented in this section suggests that the marginal renter driving our estimated capitalization effects were, in relative terms (within their cohorts), richer during the historical period than during the contemporary period. The change in real income of the average renter, thus, overestimates the change in real income of the marginal renter, suggesting that the income elasticities we infer from the capitalization studies are lower-bound estimates.

8 Extra cost for underground sections

For the back-of-the-envelope calculations reported in Section 4.5 of the main paper, we require an estimate of the extra cost associated with an underground line (as opposed to an elevated line). To obtain such an estimate, we make use of data compiled by Bousset (1935), who reports per kilometer construction costs for 31 segments of the Berlin underground network opened until 1930. In the table below, we present results of regressions of the natural log of per-kilometer construction costs against a dummy indicating underground sections. In column (1), we control for the opening year using a linear trend. In column (2), we replace this trend by five-year period effects. In column (3), we additionally control for the track width. The results are reasonably consistent across specifications. According to our preferred estimate in column (3), an underground section in the early 20th century in Berlin was about three times as expensive as an elevated section. A collateral finding from column (1) is that per-kilometer metro rail construction costs in Berlin increased by about 4% per year from 1900 to 1930.

Tab A26. Underground extra costs

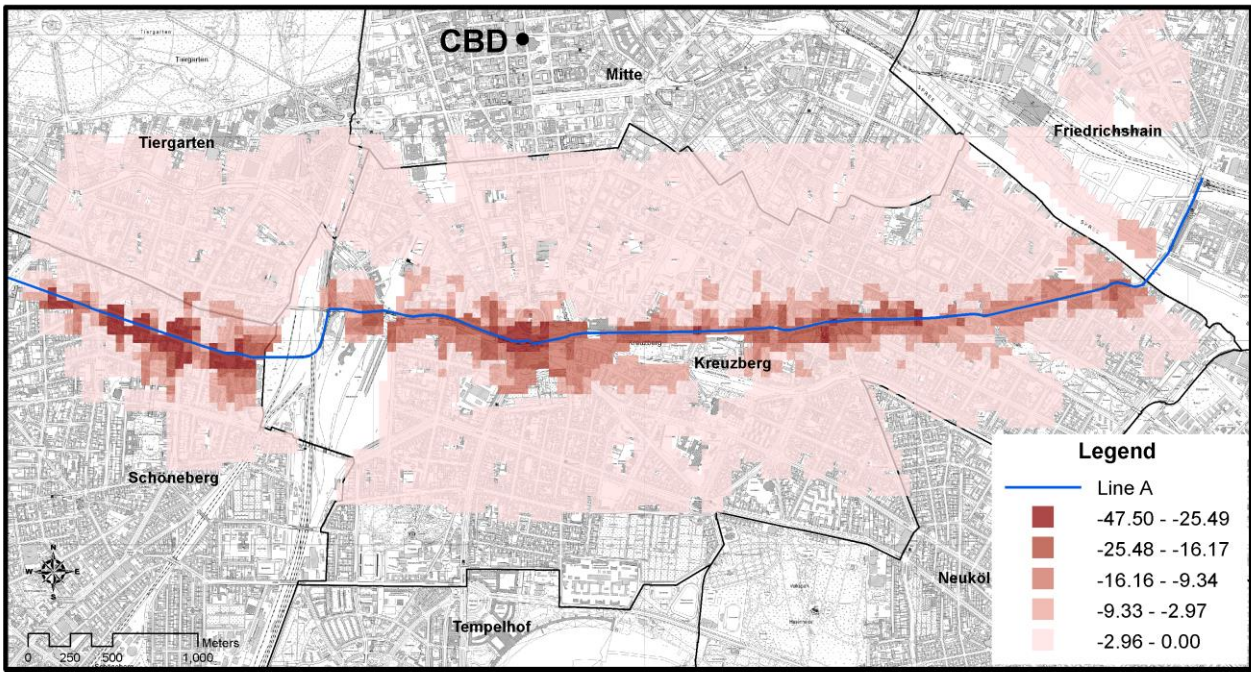
	(1)		(2)		(3)	
	Ln cost per km (million RM)					
Underground section (dummy)	0.985***	(0.264)	1.190***	(0.184)	1.149**	(0.333)
Opening year	0.039***	(0.009)				
Broad gouge (dummy)					0.064	(0.335)
Percent extra underground cost	168		229		216	
Period effect (five years)	-		Yes		Yes	
N	31		31		31	
r2	0.598		0.664		0.664	

Note: Standard errors (in parenthesis) are robust in (1) and clustered on year bins in (2) and (3). * p < 0.1, ** p < 0.05, *** p < 0.01

9 Aggregate land price effects

As described in the main paper, we aggregate parcels (the unit of observation in our regressions) to 50×50-meter grid cells before computing the aggregate effect of rail noise on land prices. The grid size is chosen to ensure that we cover all developed areas and allow for sufficient spatial detail to account for the localized nature of noise emissions. Below, we illustrate the resulting noise effects by grid cells. Figure A13 shows how we only cover parts of the city that were developed in 1900. Figure A13 is also reflective of the typical features of noise emission. Noise is contained to relatively narrow corridors in densely developed areas, but spreads further along open spaces.

Fig A13. Estimated noise effects and land prices



Notes: Plots are aggregated to 50x50 m grid cells. Noise estimate $\widehat{\alpha^N}$ from Table 1 column (1) in the main paper. The noise effect per grid cell g is $P_{g,1900}\left(1-\exp\left(\widehat{\alpha^N} \times N_{g1904}\right)\right)$, where P_g and N_g indicates the average land price and rail noise within a grid cell. The background map shows the situation in 2006, which corresponds to the situation in 1900 in most, but not all areas. Own illustration using the Urban Environmental Information System of the Berlin Senate Department (Senatsverwaltung für Stadtentwicklung Berlin, 2006).

10 Land price appreciation vs. interest rates

In this section, we compare long-run land price growth rates and central bank interest rates to support the back-of-the-envelope calculations presented in Section 5.2 of the main paper. To our knowledge, no price index tracking real estate prices over the 19th and 20th century exists for Berlin. Therefore, we combine our data with a data set on block-level land values in Berlin compiled by Ahlfeldt, Redding, et al. (2015). Consistently using the one-kilometer buffer around Line A as a study area, we regress the log of nominal land prices against block fixed effects and a year trend to obtain an estimate of the average yearly price appreciation during several historical periods. We note that in the results reported in Table A27 we exclude the 1914-1928 period because of the hyperinflation in the aftermath of WWI which complicates the comparison of nominal prices. For the later currency reforms (reichsmark to Deutsche Mark, 1948 and Deutsche Mark to euro, 1998) we apply the official conversion factors (10:1 and 1.95583:1).

As evident from Table A27, growth rates in nominal land prices fluctuate around 5%, with peaks during the major economic boom periods such as the "*Gründerzeit*" (2) and the post-WWII (pre-unification) period (4 and 5). These rates are in line with Knoll et al. (2017) who report a long-run average growth rate of 4% for Germany (4.3% and 3.7% for the post-WWII and the pre-WWII periods).

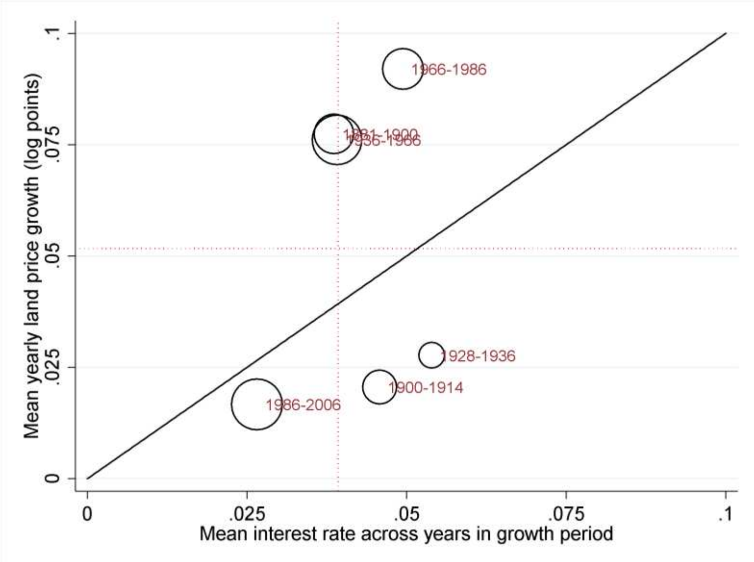
In Figure A14, we compare the estimated land price growth rates to interest rates (central bank discount and base rates). We find that the weighted (by year) average growth rates and interest rates over the period from 1881 to 2006 are roughly within the same range (about 4-5%). The correlation between the two variables is positive, with half of the observations being located above the 45-degree line and the other half below. It seems fair to conclude that over the course of about 130 years, nominal land prices in Berlin appreciated roughly at a rate that reflects the opportunity cost of capital.

Tab A27. Land price appreciation

	(1) Ln land price	(2) Ln land price	(3) Ln land price	(4) Ln land price	(5) Ln land price	(6) Ln land price
Year	0.077***	0.021***	0.028***	0.076***	0.092***	0.017***
	(0.001)	(0.001)	(0.006)	(0.002)	(0.002)	(0.003)
Fixed effect	Blocks	Blocks	Blocks	Blocks	Blocks	Blocks
Period	1881-1900 (4 years)	1900-1914 (4 years)	1928-1936 (2 years)	1936-1966 (2 years)	1966-1986 (2 years)	1986-2006 (2 years)
Area	1 km from Line A	1 km from Line A	1 km from Line A	1 km from Line A	1 km from Line A	1 km from Line A
N	1,348	1,348	614	584	562	576
r2	0.924	0.951	0.845	0.949	0.970	0.757

Notes: Sample are is a 1 km buffer drawn around Line A. Standard errors (in parenthesis) are clustered on fixed effects level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Fig A14. Land price appreciation vs. central bank interest rate



Notes: Mean yearly land price growth are the estimated year effects in Table A14. Interest rate is the central bank discount rate from 1881 to 1998, the base rate as per Discount Rate Transition Act from 1999 to 2002, and the base rate as per civil code thereafter as published by Rahlf (2015) and the Deutsche Bundesbank (interest statistics). Dotted lines are weighted (by year) averages. The black solid line is the 45-degree line.

11 Property taxation

In the back-of-the envelope calculations reported in Table 4 in the main paper, we consider fiscal revenues from property transaction taxes. Here, we provide a comparison in terms of revenues from property taxes, which are internationally more popular..

11.1 Real property tax rates in Germany

In Germany, the property tax is determined as the product of the tax base (the assessed value of the property, the so called *Einheitswert*), a tax rate (*Grundsteuermesszahl*) and a tax factor (*Hebesatz*). The tax rate depends on the property type (e.g. single family houses) while the tax factor varies across federal states. One specific feature of the German property tax system is that the *Einheitswert* is based on an assessment that took place as early as in 1961 (in the states belonging to the former German Democratic Republic, the *Einheitswert* refers to 1935). The *Einheitswert*, thus, substantially underestimates the current market value of a property. The legal tax rate (*Grundsteuermesszahl*), therefore, does not directly correspond to a real property tax rate.

To approximate the real property tax rate in Table A28, we first compute the ratio of the *Einheitswert* over the market value as the inverse of a factor that captures the price inflation over fifty years since 1961. We get to this factor using the weighted (by year) average of the yearly land price growth rates from 1966 to 1986 and 1986 to 2006 reported in the Table A27 (columns 5 and 6). This appreciation rate implies that the *Einheitswert* after 50 years, on average, corresponds to 10.55% of the market value. For the tax rate, we consider values of 0.27%, which applies to single-family houses, and a rate of 0.35% which applies to larger structures. For the tax factor, we consider values of 333% (Hesse, the lowest in Germany), 410% (the German average) and 810% (Berlin, the highest in Germany), reported by the Federal Statistical office (Statistisches Bundesamt Fachserie 14 Reihe 10.1 – 2010).

Under the assumptions made, it is then straightforward to approximate a real property tax rate for the different scenarios by multiplying the ratio of the *Einheitswert* over market value by the tax rate and the tax factor. The typical tax rate in central Berlin is 0.35% (non-single-family houses) and the tax factor is 810%, thus the real property tax is 0.3%. In other parts of Germany, the real property tax rate is likely to be lower because the tax factors are much lower. Moreover, property price appreciation was, on average, higher at 7%, implying a ratio of *Einheitswert* over market value of just about 5% (Bundesministerium der Finanzen, 2011).

The real property tax rate that we estimate for Berlin is low by international standards. According to a Property Tax Comparison Study by the Minnesota Center for Fiscal Excellence (2014), the average property tax in US urban areas was 1.5%. Across urban areas, tax rates vary from 0.61% (Columnia, SC) to 4.1% (Bridgeport, CT).

Tab A28. Real property tax in Germany and Berlin

	(1)	(2)	(3)	(4)	(5)	(6)
Long-run yearly price inflation				4.6%		
Ratio "Einheitswert" / market value				10.55%		
Tax rate (Grundsteuermesszahl)	0.27%	0.35%	0.27%	0.35%	0.27%	0.35%
Tax factor (Hebesatz)	333%	333%	410%	410%	810%	810%
Real property tax	0.09%	0.12%	0.12%	0.15%	0.23%	0.30%

Notes: The yearly price inflation is from an auxiliary regression of the natural log of 1966-2006 (Berlin) land price on block fixed effects and a year trend. The ratio of the “Einheitswert” over the market value is the inverse of a factor that captures the price inflation over fifty years since 1961 (the year of the “Einheitswert” assessment). A tax rate (Grundsteuermesszahl) of 0.27% applies to single-family houses whereas a rate of 0.35% applies larger structures. The “Hebesatz” values are from the Federal Statistical Office (Statistisches Bundesamt Fachserie 14 Reihe 10.1 – 2010) and refer to Hesse (333%, the lowest in Germany), the German average (410%) and Berlin (810%, the highest in Germany). The real property tax is obtained by multiplying the ratio of "Einheitswert" / market value (a measure of the undervaluation of the tax base) by the Grundsteuermesszahl (the tax rate) and the Hebesatz (the tax factor).

11.2 Property tax vs. property transaction tax revenues

In the context of the back-of-the-envelope calculations reported in Table 4 in the main paper, it is noteworthy that, in reality, a public investment will not refinance via the property tax in Berlin because, as described above, the tax base is fixed to 1961 (or 1935) assessed values (the *Einheitswert*). However, as summarized below in Table A29, revenues from property taxes (*Grundsteuer*) and property transaction taxes (*Grunderwerbssteuer*) tend to be within the same range in Berlin. Unlike property taxes, property transaction taxes are based on actual transaction prices and are responsive to increases in real estate prices. The fiscal returns listed in Table 4 can, thus, be thought of being incurred via property taxes instead of property transaction taxes in Berlin, leaving all interpretations and conclusions unaffected. For comparison, we replicate Table 4 for a hypothetical and internationally more conventional scenario – in which the cost of an underground line are recovered in terms of property taxes. Given the results from A28, it is no surprise that the results are similar.

Tab A29. Property transaction tax revenues vs. property tax revenues in Berlin

	2012	2013	2014	2015	Mean
Property transaction tax (million)	578.0	735.4	796.0	960.0	767.3
Property tax (million)	756.7	763.7	776.9	780.8	769.5

Notes: Data are from the State Statistical Office Berlin (available from the website of Berlin Senate Department (https://www.berlin.de/sen/finanzen/steuern/steuereinnahmen/)

Tab A30. The fiscal case for an underground line

	(1)	(2)	(3)	(4)	(5)	(6)
Noise preferences		Historic			Contemporary	
Noise effect on land price (per decibel)	0.41%	0.41%	0.41%	3.32%	3.32%	3.32%
Property tax rate	0.25%	0.75%	1.5%	0.25%	0.75%	1.5%
Estimated total cost (million 1900 RM)				15.94		
Estimated underground extra cost (million 1900 RM)				34.36		
Aggregated noise effect on land value (million 1900 RM)	18.6	18.6	18.6	151	151	151
Yearly tax revenue (million 1900 RM)	0.05	0.14	0.28	0.38	1.13	2.26
Years to recover underground extra costs	738	246	123	91	30	15

Notes: Contemporary land price effect adjusted for changes in land share and housing expenditure share (land price capitalization effect inflated by the ratio of contemporary over historical shares). Cost estimates based on Bousset (1935). Estimated total cost result from multiplying the reported 1902 per km costs of over elevated sections by 8 km (the length of the elevated sections of the Line A). The estimated underground extra cost result multiplying the total cost by the percentage extra costs for underground segments obtained from an auxiliary regression reported in Section 5 of the appendix. Years to recover extra costs are calculated under the assumption that land values grow at a rate similar to cost of capital (see appendix 9 for a justification).

Literature

Abadie, Alberto, Athey, Susan, Imbens, Guido W., Wooldridge, Jeffrey: When Should You Adjust Standard Errors for Clustering? NBER Working Paper No. 24003.

Ahlfeldt, Gabriel M. (2018). Weights to Address Non-parallel Trends in Panel Difference-in-differences Models. *CESifo Economic Studies, 64*(2), 216-240.

Ahlfeldt, Gabriel M., & Holman, Nancy. (2018). Distinctively Different: A New Approach to Valuing Architectural Amenities. *The Economic Journal128,* 1-33.

Ahlfeldt, Gabriel M., Koutroumpis, Pantelis, & Valletti, Tommaso. (2016). Speed 2.0: Evaluating Access to Universal Digital Highways. *Journal of the European Economic Association, 15(3).* 586–625.

Ahlfeldt, Gabriel M., & Maennig, Wolfgang. (2015). Homevoters vs. leasevoters: A spatial analysis of airport effects. *Journal of Urban Economics, 87*, 85-99.

Ahlfeldt, Gabriel M., Maennig, Wolfgang, & Richter, Felix J. (2016). Urban renewal after the Berlin Wall: a place-based policy evaluation. *Journal of Economic Geography, 17*(1), 129-156.

Ahlfeldt, Gabriel M., Moeller, Kristoffer, & Wendland, Nicolai. (2015). Chicken or egg? The PVAR econometrics of transportation. *Journal of Economic Geography, 15*(6), 1169-1193.

Ahlfeldt, Gabriel M., Redding, Stephan J., Sturm, Daniel M., & Wolf, Nikolaus. (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica, 83*(6), 2127-2189.

Ahrens, G.-A., Liesske, F., Wittwer, R., & Hubrich, S. (2009). *Endbericht zur Verkehrserhebung Mobilität in Städten - SrV 2008 und Auswertungen zum SrV-Städtepegel*. Dresden: Technische Universität Dresden.

Al-Mosaind, Musaad A., Dueker, Kenneth J., & Strathman, James G. (1993). Light-Rail Transit Stations and Property Values: A Hedonic Price Approach. *Transportation Research Record, 1400*, 90-94.

Alonso, William. (1964). *Location and land use*. Cambridge, MA: Harvard.

Angrist, Joshua D., & Pischke, Jörn-Steffen. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.

Bajic, V. (1983). The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto. *Urban Studies, 20*(2), 147-158.

Bartholomew, Keith, & Ewing, Reid. (2011). Hedonic Price Effects of Pedestrian- and Transit-Oriented Development. *Journal of Planning Literature, 26*(1), 18-34.

Baum-Snow, Nathaniel, & Kahn, Matthew E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics, 77*(2), 241-263.

Billings, Stephen B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics, 41*(6), 525-536.

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

Bolt, Jutta, & van Zanden, Jan Luiten. (2014). The Maddison Project: collaborative research on historical national accounts. *The Economic History Review, 67*(3), 627-651.

Bousset, E. H. J. (1935). *Die Berliner U-Bahn*. Berlin: Wilhem Ernst & Sohn.

Bowes, David R., & Ihlanfeldt, Keith R. (2001). Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics, 50*(1), 1-25.

Bundesministerium der Finanzen. (2011). *Reform der Grundsteuer*. Berlin: Bundesministerium der Finanzen,.

Cellini, Stephanie Riegg, Ferreira, Fernando, & Rothstein, Jesse. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *The Quarterly Journal of Economics, 125*(1), 215-261.

Chay, Kenneth Y., & Greenstone, Michael. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy, 113*(2), 376-424.

Currie, Janet, Davis, Lucas, Greenstone, Michael, & Walker, Reed. (2015). Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings. *American Economic Review, 105*(2), 678-709.

Damm, David, Lerner-Lam, E., & Young, J. (1980). Response of Urban Real Estate Values in Anticipation of the Washington Metro. *Journal of Transport Economics and Policy, 14*(3), 315-336.

Davis, Lucas W. (2004). The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review, 94*(5), 1693-1704.

Day, Brett, Bateman, Ian, Lake, Iain, May (2007). Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model. *Environmental and Resource Economics*. 37 (1), 211–232.

Debrezion, Ghebreegziabiher, Pels, Eric, & Rietveld, Piet. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *Journal of Real Estate Finance & Economics, 35*(2), 161-180.

Debrezion, Ghebreegziabiher, Pels, Eric, & Rietveld, Piet. (2010). The Impact of Rail Transport on Real Estate Prices: An Empirical Analysis of the Dutch Housing Market. *Urban Studies*.

Dewees, D. N. (1976). The Effect of a Subway on Residential Property Values in Toronto. *Journal of Urban Economics, 3*(4), 357.

Federal Statistical Office of Germany. (2013). *Volkswirtschaftliche Gesamtrechnungen: Private Konsumausgaben und Verfügbares Einkommen*. Wiesbaden, Germany: Federal Statistical Office of Germany.

Federal Statistical Office of Germany. (2017). Privathaushalte, Haushaltsmitglieder: Deutschland, Jahre; Mikozensus. *Genesis Online*.

Gibbons, Stephen. (2015). Gone with the wind: Valuing the visual impacts of wind turbines through house prices. *Journal of Environmental Economics and Management, 72*, 177-196.

Gibbons, Stephen, & Machin, Stephen. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*(1), 148-169.

Gibbons, Stephen, & Machin, Stephen. (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economics, 24*(1), 99-119.

Gibbons, Stephen, Machin, Stephen, & Silva, Olmo. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics, 75*(0), 15-28.

Graevenitz, Katherine (2018). The amenity cost of road noise. *Journal of Environmental Economics and Management, 90. 1-22*.

Greenstone, Michael, & Gallagher, Justin. (2008). Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *The Quarterly Journal of Economics, 123*(3), 951-1003.

Harrison, David Jr, & Rubinfeld, Daniel L. (1978). Hedonic housing prices and the demand for clean air. *Journal of Environmental Economics and Management, 5*(1), 81-102.

Hoffmann, Walther G. ((1965 [2006])). *Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts: Der Verbrauch*. Cologne, Germany: GESIS.

Hurst, Needham B., & West, Sarah E. (2014). Public transit and urban redevelopment: The effect of light rail transit on land use in Minneapolis, Minnesota. *Regional Science and Urban Economics, 46*, 57-72.

Kalweit, Ferdinand. (1929). *Die Baustellenwerte in Berlin 1928*. Berlin: Emro.

Knoll, Katharina, Schularick, Moritz, & Steger, Thomas. (2014). No Price Like Home: Global House Prices, 1870 – 2012. *Working paper accessed via URL:* http://piketty.pse.ens.fr/files/Schularicketal2014.pdf*, last accessed February 17, 2016*.

Knoll, Katharina, Schularick, Moritz, & Steger, Thomas. (2017). No Price Like Home: Global House Prices, 1870-2012. *American Economic Review, 107*(2), 331-353.

Kuminoff, Nicolai V., Pope, Jaren C. 2014. Do "capitalization effects" for public goods reveal the public's willingness to pay? *International Economic Review*, 55(4), p.1227-1250.

Leggett, Christopher G., & Bockstael, Nancy E. (2000). Evidence of the Effects of Water Quality on Residential Land Prices. *Journal of Environmental Economics and Management, 39*(2), 121-144.

Leyden, F. (1933). *Gross-Berlin: Geographie einer Weltstadt*. Berlin: Gebr. Mann Verlag.

Linden, Leigh, & Rockoff, Jonah E. (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *American Economic Review, 98*(3), 1103-1127.

McDonald, John F., & Osuji, Clifford I. (1995). The effect of anticipated transportation improvement on residential land values. *Regional Science & Urban Economics, 25*(3), 261.

McMillen, Daniel P. (1996). One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach. *Journal of Urban Economics, 40*(1), 100-124.

McMillen, Daniel P., & McDonald, John F. (2004). Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983-1999. *Real Estate Economics, 32*(3), 463-486.

Mills, Edwin S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Centre. *American Economic Review, 57*(2), 197-210.

Minnesota Center for Fiscal Excellence. (2014). *50 State Property Tax Comparison Study 2013*. Camebridge, MA: Lincoln Institute of Land Policy.

Mohammad, Sara I., Graham, Daniel J., Melo, Patricia C., & Anderson, Richard J. (2013). A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice, 50*, 158-170.

Muth, R. (1969). *Cities and Housing*. Chicago: University of Chicago Press.

Natschka, W. (1971): Berlin und seine Wasserstrassen. Duncker & Humblot.

Navrud, Ståle. (2002). *The State-Of-The-Art on Economic Valuation of Noise*. Brussels: European Commission DG Environment.

Nelson, Arthur C. (1992). Effects of Elevated Heavy-Rail Transit Stations on House Prices with Respect to Neighborhood Income. *Transportation Research Record, 1359*, 127-132.

Nelson, Jon P. (1978). Residential choice, hedonic prices, and the demand for urban air quality. *Journal of Urban Economics, 5*(3), 357-369.

Nelson, Jon P. (2004). Meta-Analysis of Airport Noise and Hedonic Property Values: Problems and Prospects. *Journal of Transport Economics & Policy, 38*(1), 1-28.

Nelson, Jon P. (2008). Hedonic Methods in Housing Markets, Chapter Hedonic Property Value Studies of Transportation Noise: Aircraft and Road Traffic. Springer Verlag.

Osborne, Martin J., & Turner, Matthew A. (2010). Cost benefit analyses versus referenda. *Journal of Political Economy, 118*(1), 156-187.

Pope, Jaren C. (2008). Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics, 63*(2), 498-516.

Rahlf, Thomas. (2015). *Zeitreihendatensatz für Deutschland, 1834-2012*.

Reinhardt, Winfried. (2015). *Geschichte des Öffentlichen Personenverkehrs von den Anfängen bis 2014: Mobilität in Deutschland mit Eisenbahn, U-Bahn, Straßenbahn und Bus*

Wiesbaden: Springer Vieweg.

Rossi-Hansberg, Esteban, Sarte, Pierre-Daniel, & Owens, Raymond. (2010). Housing Externalities. *Journal of Political Economy, 118*(3), 485-535.

Schmiedeke, Carl Wilhelm. (1997). *Der Wagenpark der Berliner S-Bahn*. Hamburg: Lokrundschau-Verlag.

Senate Department for Urban Development and the Environment. (2013). *Strategic Noise Maps*. Berlin: Senate Department for Urban Development and the Environment.

Senatsverwaltung für Stadtentwicklung Berlin. (2006). *Urban and Environmental Information System*. Berlin.

Silverman, B. W. (1986). Density Estimation For Statistics and Data Analysis. *Monographs on Statistics and Applied Probability.*

Tyrväinen, Liisa, & Miettinen, Antti. (2000). Property Prices and Urban Forest Amenities. *Journal of Environmental Economics and Management, 39*(2), 205-223.

U.S. Department of Labor. (2006). *100 Years of U.S. Consumer Spending*. Washington, D.C.: U.S. Department of Labor,.

Voith, Richard. (1993). Changing Capitalization of CBD-Oriented Transportation Systems: Evidence from Philadelphia, 1970-1988. *Journal of Urban Economics, 33*(3), 361.

Wrigley, M., & Wyatt, P. (2001). *Transport Policy and Property Values*. Paper presented at the Royal Institution of Chartered Surveyors (RICS) 'Cutting Edge' Conference, University of the West of England.

Xu, Yangfei, Zhang, Qinghua, & Zheng, Siqi. (2015). The rising demand for subway after private driving restriction: Evidence from Beijing's housing market. *Regional Science and Urban Economics, 54*, 28-37.