The Internal Structure of Cities:

Koster, H.R.A.

2013

10 The gains of trains: The effect of station openings on house prices*

10.1 Introduction

Congestion problems on the road and environmental constraints are causing a renewed interest for public transport (Cervero, 2004). In the US as well as in Europe existing public transport systems are upgraded and new lines are built. Many large cities in the world, such as Portland, Montreal, San Franscisco, Vancouver and Copenhagen are developing policies to stimulate mixed and dense urban development around public transport nodes in order to reduce automobile dependency and increase the use of public transport. Also in the Netherlands such policies have been developed. This so-called *Transit-Oriented Development* concept is based on two critical assumptions. First, households and workers tend to frequently use railway services and therefore are willing to pay to reside close to stations. Second, households and firms prefer to reside in dense and mixed neighbourhoods. In this chapter we test the first assumption using data on the housing market.

There exists an empirical literature measuring the effects of public transport innovations. The economic benefits of these innovations are usually measured by reductions in travel time, direct user costs and accident costs, using stated-choice experiments. However, mainly because of wider economic benefits, these evaluation methods may lead to an underestimate of the benefits of improved accessibility (Gibbons and Machin, 2005).[222] Therefore, when the economic benefits are mainly captured by households that live close to improved transport infrastructure, hedonic price methods are preferred, because they aim to monetarise preferences of households for neighbourhood attributes. See also Bayer et al. (2007), who argue that for continuous housing and location attributes, such as distance to railway stations, hedonic price approaches, rather than discrete choice approaches, are a good way to estimate the mean willingness to pay (WTP).

A general conclusion in the literature is that households are willing to pay a premium to live close to public transport stations as proximity to stations implies shorter travel times (see Gatzlaff and Smith, 1993; Baum-Snow and Kahn, 2000; and the meta-analysis of Debrezion et al., 2007). However, most studies are cross-sectional, which probably leads to upward biased estimates, as the effects of stations are correlated with pleasing unobserved spatial factors (e.g. shopping malls). Recent contributions try to avoid the bias inherent to cross-sectional studies by using repeated sales prices (McMillen and McDonald, 2004; Grimes and Young, 2010) or employing a difference-in-difference methodology based on openings of stations (Gibbons and Machin, 2005).

In this chapter we investigate the impacts of railway proximity on house prices in Dutch cities between 1995 and 2007 based on a repeated sales sample. We focus on openings of *small* stations on *existing* tracks that are announced a short period (about two or three years) before the actual opening

* The authors gratefully acknowledge NVM for providing data. Participants of a number of NICIS-KEI seminars are thanked for useful comments.

[222] For example, improved transport access may reduce frictions in the labour market, increase the intensity of knowledge spillovers, and lower input costs. But also reductions in (local) environmental externalities could be a benefit of transport innovations.

of the station. Our chapter improves on the existing literature in the sense that we use an extensive *repeated sales dataset* (more than 20,000 properties), which allows us to examine house price differences at the level of the individual property. We therefore avoid any bias in the estimates that may arise when using cross-sectional or area-aggregated panel data.

This chapter proceeds as follows. In Section 10.2, we discuss the study's context and present the data. Section 10.3 considers the econometric methodology and presents and discusses the result. Section 10.4 reviews evidence on the effect of public transport accessibility in the previous chapters. Section 10.5 concludes.

10.2 Rail innovations and data

A. Rail innovations in the Netherlands

The rail network of the Netherlands is one of the densest in the world (Claessens et al., 1998). Until 1930 the network has been expanded very quickly. The Dutch railway network was exploited by a large number of different private companies. From 1930 onwards, competition by bus and car led to the closure of a large number of lines and stations (Figure 10.1). Since the World War II the network has been mainly upgraded in terms of efficiency and capacity. Today, the network consists of about 2,800 kilometres of track.[223] It is mainly used for passenger rail services and almost 1,000,000 people are transported each day.

We focus on the effects of new stations built on existing railroad tracks between 1995 and 2007. The new stations are often located in the suburbs of cities, but sometimes also are new stations not close to any other station. The new stations have comparable railway service levels and provide stops for local commuter trains only. Figure 10.2 presents a map of the cities and the new stations.

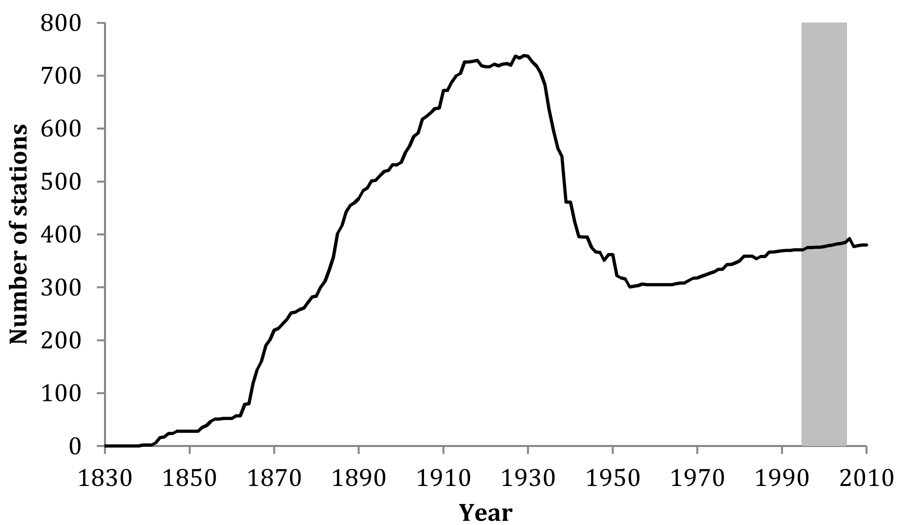

FIGURE 10.1 — NUMBER OF STATIONS BETWEEN 1830-2010
Note: The shaded area denotes the study period.

[223] Since the 1990s, no major network expansions have taken place, except for a high speed rail connection between Amsterdam and the Belgium border and a freight line between the port of Rotterdam and the Ruhr area in Germany.

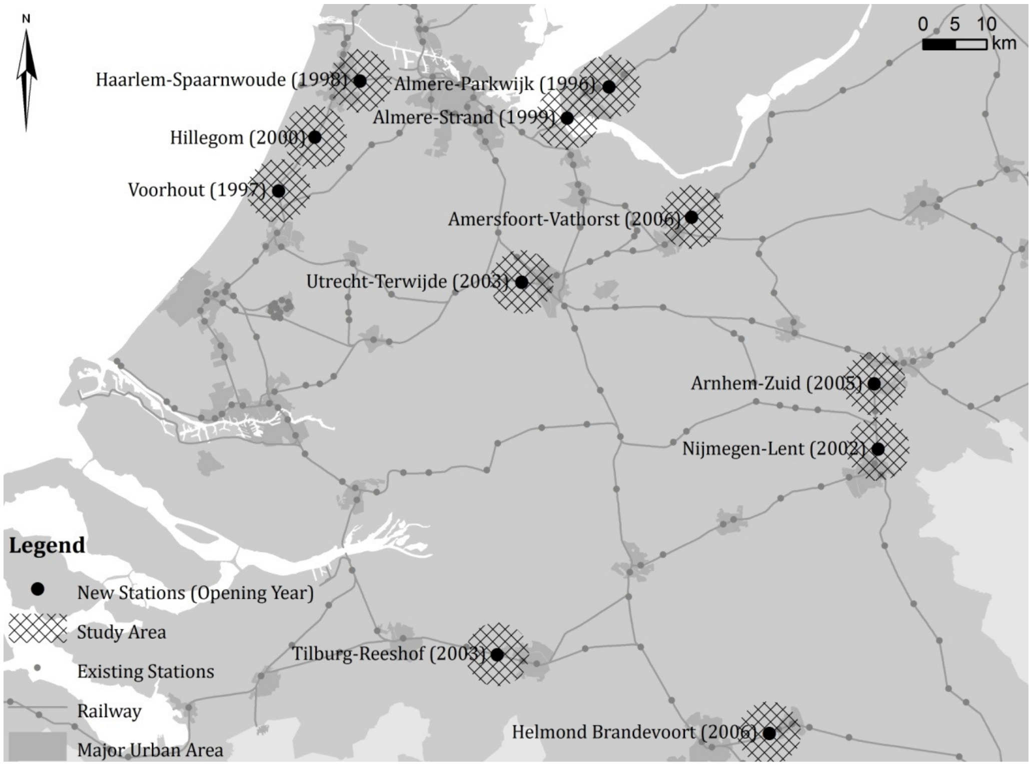

FIGURE 10.2 — MAP OF THE NETHERLANDS WITH SELECTED CITIES AND NEW STATIONS

B. Data

Our hedonic price analysis is based upon a house transactions dataset from the NVM (Dutch Association of Real Estate Agents). It contains information on about 80 percent of all transactions between 1995 en 2007. We focus on relevant properties, so we select all transactions within 5 kilometres of a new station. For 162,537 transactions, we know the transaction price, the exact address, and a wide range of house attributes such as size (in square meters), type of house, number of rooms and construction year.[224] For each property we calculate the distance to the nearest railway station using GIS-software. We also control for the population density in a neighbourhood.[225] For 1996, 1998, 2000 and 2002 we do not have information on population density, so for these years we impute the average values of the preceding and following year.

About 28 percent of the transactions are repeated sales, which we will use in our analysis.[226] The repeated sales dataset consists of 45,291 transactions of 21,959 residential properties. Most properties in this dataset (75 percent) are transacted twice. Table 10.1 presents summary statistics for price (in €)

[224] We exclude transactions with prices that are above € 1.5 million or below € 25,000 or a square meter price below € 250 or above € 5,000. Furthermore, we exclude transactions that refer to properties smaller than 25m^2 or larger than 300m^2. Old houses are sometimes demolished and replaced by new ones, so we delete observations which refer to properties that have changed type of house, or for which the size has changed more than 20m^2.

[225] Neighbourhoods are fairly small: the average distance to the centroid of a neighbourhood is only 286 meter.

[226] Appendix 10.A presents a table with means and standard deviations of the repeated sales sample. We also have compared the descriptives of the repeated sales sample to the full sample and it appears that the means and standard deviations are very similar.

TABLE 10.1 — DESCRIPTIVES OF REPEATED SALES SAMPLE BEFORE AND AFTER STATION OPENINGS

Price *(in €)*			Distance to Station *(in km)*		
Before	After	Annual Change (%)	Before	After	Change
149,977	195,565	0.080	1.892	1.730	-0.163
(71,460)	(87,901)	(0.066)	(1.391)	(1.108)	(0.800)

Notes: Standard deviations are between parentheses. The change in price is the compound annual growth rate.

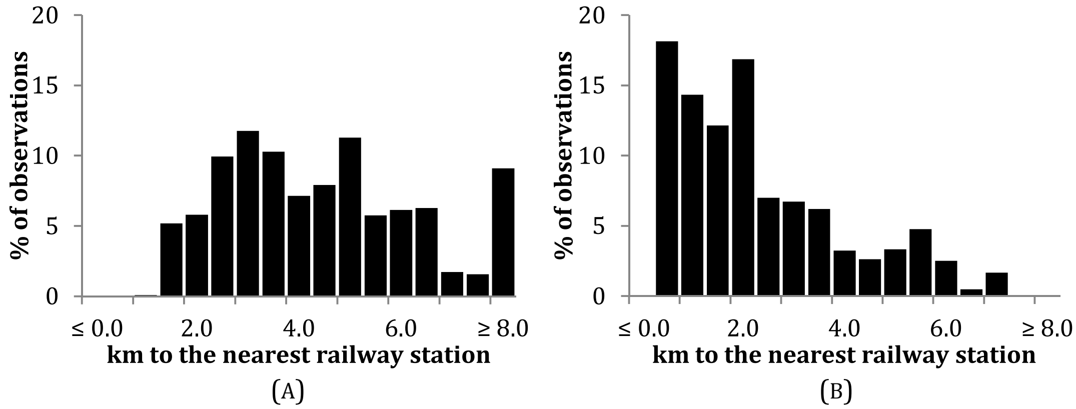

FIGURE 10.3 — DISTANCE BEFORE STATION OPENINGS AND DISTANCE REDUCTIONS TO RAILWAY STATIONS

and distance to station (in km) before and after a transaction, as well as their changes over time. The average yearly increase in prices is 8 percent. Only 9 percent of properties have experienced a reduction in station distance. Therefore, the average reduction in distance to station is low and only 185 meter. For a complete overview of all variables we refer to Table 10.A2 in Appendix 10.A.

Figure 10.3 presents histograms *only* for properties that experienced reductions in distance to stations. In Figure 10.3A we focus on the distance to the station before opening. About 80 percent is then located within 6 kilometres of a railway station. Figure 10.3B denotes the distance reductions due to station openings. Note that properties already located within one kilometre of a station, never experience a reduction in distance. This range of distances is economically meaningful because previous studies have shown that about 80 percent of trips to get to stations originate from locations within 5 kilometres of the station with an average of about 2 kilometres (Keijer and Rietveld, 2000). For these properties, the average distance decrease is 2 kilometre with a median of 1.67 kilometre. Other descriptive statistics can be found in Appendix 10.A.

10.3 Repeated sales models

A. Repeated sales and the hedonic price function

We assume that price p_{ht} of house h is a function of time-varying variables x_{ht} and z_{ht} where t denotes the year and v_h are time-invariant attributes. Then, the price function can be represented as follows:

$$\log(p_{ht}) = \alpha d_{ht} + \beta x_{ht} + \gamma v_h + \delta y_{ht} + \xi_h + \epsilon_{ht}, \quad (10.1)$$

where d denotes the distance to the nearest railway station and are α, β, γ and δ parameters to be estimated, y_t are year dummies, ξ_h and ϵ_{ht} denote respectively unobserved property-specific time-invariant and time-varying unobserved components. One way to deal with time-invariant unobserved attributes is to use postcode area fixed effects to estimate the effects of stations on property values. This essentially deals with all variation between areas (e.g. quality of housing stock) but assumes that the distribution of unobserved differences of sampled properties within postcodes does not vary over time, or at least is uncorrelated with changes in house prices. This assumption may not hold for two reasons. The first reason is that the economic value of certain unobserved house attributes may change (e.g. sufficient parking on street may be important for those without private parking space, but less so when the house is close to a station). The second reason is that sales turnover of houses with specific unobserved attributes (e.g. sufficient parking on street) may change over time (see Bajari et al., 2012). When this does not hold, repeated sales on the level of the individual property are preferred. We therefore use a difference-in-difference approach, which implies that we eliminate all time-constant variables:

(10.2) $$\log(p_{ht}) - \log(p_{hs}) = \alpha(d_{ht} - d_{hs}) + \beta(x_{ht} - x_{hs}) + \delta(y_{ht} - y_{hs}) + (\epsilon_{ht} - \epsilon_{hs}),$$

where s denotes the other transaction year in which the property is transacted. An alternative notation would be:

(10.3) $$\Delta \log(p_h) = \alpha\Delta d_h + \beta\Delta x_h + \delta\Delta y_h + \Delta\epsilon_h,$$

To be able to identify the causal effect of stations on properties we have to make the assumption that the assignment of new stations is independent from the arrival of other unobserved local (dis)amenities that are spatially correlated with distance to station (see Baum-Snow and Kahn, 2000; Gibbons and Machin, 2005; Redfearn, 2009b). We think this is a reasonable assumption for three main reasons. First, the new railway stations are built on existing tracks. These rail lines are there for a long time and for reasons that are uncorrelated with current (dis)amenities. Second, we do not expect that house prices influence the construction of new stations. As train stations in the Netherlands are used by rich as well as poor people, the planning institution does not have incentives to build stations in areas with on average high house prices (see Glaeser et al., 2008). Furthermore, the stations are not only built in growth-prone areas. It appears that properties that are within 1 km of a station experience nearly identical annual growth of house prices compared to properties that are located further away. Third, we use the *distance* to station as an indicator of the presence of a station instead of using dummy indicators whether a property is within an arbitrary range of the station. Besides that our approach seems more intuitive, as the impact of stations is continuously declining over space, it is also less likely that the *distance* to station variable is correlated with unobservable local price trends.

B. Results

Table 10.2 presents the results based on the repeated sales sample. Because errors $\Delta\epsilon_h$ are likely correlated across space, we cluster standard errors at the station area level. In Specification (1) we do not include Δx_h and Δy_h. The results then suggest that a kilometre decrease in distance to the nearest station increases house prices with 3.2 percent. However, price trends are likely correlated with changes in the distance to the nearest station. Specification (2) therefore includes year fixed effects Δy_h. It is now observed that distance to station coefficient has an unexpected sign and becomes statistically insignificant. This also holds if we include a range of control variables related to the house and include

TABLE 10.2 — REGRESSION RESULTS ON THE IMPACT OF STATIONS
(Dependent variable: the logarithm of price per square meter)

	(1) No control variables included	(2) Year fixed effects included	(3) Housing variables included	(4) Correct for anticipation effects
Distance to station	-0.032 (0.013)**	0.010 (0.008)	0.010 (0.007)	0.005 (0.016)
Population density *(log)*			-0.003 (0.015)	0.001 (0.016)
House size in m^2 *(log)*			0.056 (0.024)**	0.047 (0.025)*
Rooms			0.001 (0.001)	0.001 (0.001)
Garage			0.023 (0.012)*	0.016 (0.009)*
Garden			0.000 (0.004)	-0.003 (0.004)
Central heating			0.125 (0.012)***	0.129 (0.011)***
Listed			0.005 (0.024)	0.002 (0.024)
Transaction year dummies (14)	No	Yes	Yes	Yes
Number of observations	23,332	23,332	23,332	20,181
Within-R^2	0.010	0.723	0.747	0.746

Notes: In Specification (4) we exclude all transactions that are observed within 3 years before the station opening. Clustered standard errors (at the station-area level) are between parentheses.
*** Significant at the 0.01 level
** Significant at the 0.05 level
* Significant at the 0.10 level

population density in Specification (3). One may argue that anticipation effects lead to an underestimate of the effect of stations. As most stations were announced maximally three years before the opening, we have excluded all transactions that took place within 3 years before the station's opening. Specification (4), however, shows that the results are nearly identical: we cannot detect any statistically significant impact of station openings on house prices in the Netherlands.

So, in contrast to previous studies we cannot detect any effect of station openings on house prices. This may have several reasons. First, most new stations are relatively small and are located close to larger stations. These new stations may therefore have limited effects on travel time of passengers. Second, station openings may also imply negative externalities. Bowes and Ihlandfeldt (2002) mention increases in noise and the unsightliness of the station, especially if it includes a parking lot. Improved public transport access may also increase crime rates in station areas, because it is easier for criminals to access these areas. These negative effects may offset positive accessibility effects. Third, for only 2 percent of the trips, people use the train (Statistics Netherlands, 2012).[227] This low number implies that railway accessibility is of limited importance for a large share of the Dutch population, which may explain why improvements in railway accessibility do not command higher house prices.

C. Robustness

Because it is surprising that we do not find any positive effect of stations we have done several robustness checks. First, we have used alternative measures of railway accessibility, such as distance interval dummies and a distance weighted station potential, based on frequency.[228] Second, we have

[227] Trips by train are relatively long: the share of kilometres travelled by train is 8.3 percent.
[228] More specifically, d_{ht} is then defined as $d_{ht} = \max_r(\mathrm{e}^{-\theta k_{hr}} f_r)$, where r denotes a station, k_{hr} is the kilometre distance between h and r, θ is a distance decay parameter that is optimised within the model (as in Chapter 2) and

extended the sample to all transactions (about 2 million) in the Netherlands. The size of the latter sample should be large enough to detect small effects. Third, we have included postcode six-digit effects (PC6) rather than using a difference-in-difference equation (see equation 10.3). Fourth, we have excluded stations in suburbs because these stations usually experience substantial competition from other railway stations nearby. Fifth, we have included other neighbourhood variables, such as share ethnic minorities and average age, to control for potential sorting effects. All these regressions suggest that the effect of station openings on house prices is absent or even negative (but very small). Results of these robustness checks are available upon request.

10.4 Evidence from previous chapters

Although previous chapters did not put the identification of the effects of stations at centre stage, we have included an indicator of station proximity in all chapters. We therefore think that it is useful to summarise and investigate whether railway accessibility has an impact on commercial rents (Chapters 2-4), births of knowledge intensive business services (Chapter 5) and house prices (Chapter 6-10). Table 10.3 summarises the results for the preferred specification in each chapter.

It is shown that we could not detect any effect of stations on commercial property values and KIBS-births, suggesting that (changes in) railway accessibility does not increase profits of firms. This is not too surprising as the share of commutes by train is only 2 percent (Statistics Netherlands, 2012). This share is obviously higher near stations, but because the Netherlands is characterised by a dense railway network, the marginal effect of a new station is likely to be small. The effect of stations on house prices is found to be negative in Chapters 6, 7, 8. We note when station proximity is proxied by 'Distance to station', a positive coefficient indicates a negative effect. It also emphasised that in Chapter 6 and 7 we

TABLE 10.3 — EFFECT OF STATION PROXIMITY ON COMMERCIAL RENTS, KIBS-BIRTHS AND HOUSE PRICES

	Effect	Measurement	Identification
			PART1: FIRMS
Chapter 2	+	Station <250m	Existing stations , municipality fixed effects included
Chapter 3	+	Distance to station	New stations, PC6 fixed effects included
Chapter 4	±	Station <250m	Existing stations, PC4 fixed effects included
Chapter 5	+	Distance to station	Existing stations, municipality fixed effects included
			PART 2: HOUSEHOLDS
Chapter 6	–***	Distance to station	Existing stations, municipality fixed effects included
Chapter 7	–***	Station <250m	Existing stations, municipality fixed effects included
Chapter 8	–***	Station <250m	Existing stations, boundary fixed effects included
Chapter 9	+	Station <500m	Existing stations, only near bombing boundary
Chapter 10	–	Distance to station	New stations, difference-in-difference

Notes: We focus on the preferred specification in each chapter. When station proximity is proxied by 'Distance to station' a positive *coefficient* indicates a negative *effect*; + indicates a positive effect, – indicates a negative effect and ± reveals mixed evidence. We also denote the statistical significance of the effects:

*** Significant at the 0.01 level
** Significant at the 0.05 level
* Significant at the 0.10 level

f_r is the frequency measured as the number of trains per hour. So, this specification of railway accessibility allows for the possibility that households travel to a station that is further away if it has a high frequency (and likely more destinations to offer).

merely focus on the city region of Rotterdam. So, in this particular region, station areas are considered as less attractive (although the city centre of Rotterdam seems to be an exception, see Chapter 9). Chapter 8 focuses on conservation areas. The negative effect of stations is then explained by the fact that stations are often located close but not in conservation areas. In general, the evidence is in line with the analysis in this chapter: stations seem to have a negligible impact on rents and prices, likely because accessibility improvements are small and positive accessibility effects are offset by negative externalities caused by stations.

10.5 Conclusions

In the US, as well as in Europe, existing public transport network are expanded and upgraded. In the current chapter, we investigated the WTP to reside close to stations by employing hedonic price methods. Using an extensive repeated sales dataset over a period of 13 years, we were able to deal with potential biases inherent to previous studies, most importantly time-invariant heterogeneity. We could not detect any statistically significant impact of station openings on house prices. Furthermore, other studies that are part of this dissertation also did not detect a positive effect of proximity to stations. This makes it harder to justify substantial investments in the opening of new stations and questions the strong focus of contemporary planning concepts on new development around public transport nodes. However, it should be emphasised that potential positive welfare effects strongly depends on the existing road network and the accessibility of stations (e.g. by bicycle), implying that the benefits of public transport improvements are strongly dependent on the local situation. Also, improvements in public transport networks at least are expected to generate positive welfare effects for the current users of the network.

Appendix 10.A Descriptive statistics

TABLE 10.A1 — DESCRIPTIVE STATISTICS

	Mean	Std.Dev.	Min	Max
Price	175,448	84,801	25,865	1,475,000
Size in m^2 *(log)*	103.431	32.512	33	300
Rooms	3.891	1.185	1	16
Garage	0.030	0.170	0	1
Garden	0.807	0.395	0	1
Central Heating	0.903	0.296	0	1
Listed Building	0.003	0.053	0	1
Population density *(log)*	6809.225	3162.609	2500	22,410
Distance to station	1.796	1.245	0.055	10.733
Number of observations		45,291		
Number of properties		21,959		

TABLE 10.A2 — CHANGE OF VARIABLES BETWEEN TRANSACTIONS

	Before		After		Change	
	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
Price	149,977	71,460	195,565	87,901	45,587	54.385
Size in m^2 *(log)*	101.886	32.266	102.584	32.45	0.698	7.799
Rooms	3.823	1.160	3.872	1.208	0.049	0.656
Garage	0.032	0.175	0.026	0.159	-0.006	0.133
Garden	0.799	0.401	0.787	0.410	-0.012	0.274
Central heating	0.878	0.327	0.929	0.256	0.051	0.301
Listed building	0.002	0.048	0.003	0.056	0.001	0.054
Population density *(log)*	6,775.754	3156.195	6,825.363	3,058.467	49.608	1,009.989
Distance to station	1.892	1.391	1.730	1.108	-0.163	0.800