Availability of Abitur and Non-Abitur Upper Secondary Schools and Housing Prices in NRW

Julian Schilling Benedikt Teuber Elena Preussner

03.02.2026

RUHR UNIVERSITÄT BOCHUM

RUB

Chair: Chair of empirical economics
Course: Data Analysis using R
Lecturer: Sven Werenbeck-Ueding

Table of contents I

Introduction

Theoretical Framework

Insights from relevant empirical literature

Empirical Design

Results

Policy Implications and further research

Appendix

Introduction

Research Question

Main question

To what extent does the local availability of upper secondary schools that offer an Abitur pathway, compared to other secondary schools, affect housing prices in North Rhine-Westphalia (NRW)?

- Goal: Estimating heterogeneity in the capitalization of secondary schools that offer the possibility to achieve higher educational outcomes.
- We assume that parents are willing to pay a premium for their homes, if a school lies within 3km at which their children can make an Abitur (Gymnasium or Gesamtschule)
- Therefore, we restrict our analysis on different types of secondary schools and exclude elementary schools and other specialised or private schools

Motivation and Institutional Context

- School availability and quality in general are important determinants of housing decisions and are capitalized into housing prices (..)
- However, existing evidence on capitalization differences across upper secondary school tracks is limited, while most hedonic price studies face endogeneity concerns.
- The **magnitude of capitalization effects may differ by school track**
- Germany's multi-track secondary school system provides a unique setting to study differential capitalization effects.
- In NRW the 2008/09 reform increased freedom in school-choice and weakens formal residence–school links as it may **increase behavioral selection** into high-quality school areas.

Key Insight:

When school choice is flexible, households with strong school preferences are more likely to relocate to access better schools — reinforcing the link between school quality and housing prices. (Bayer et al., 2007)

Theoretical Framework

Tiebout sorting

- The decision-making process of residents include the availability and quality of provided public goods and services within a municipality
- Under the assumption of perfect mobility, residents pick that community that exactly satisfies their preferences (Tiebout, 1956)
- If such a community or municipality is not feasible, a perfect substitute (if existent) is to be chosen

What do parents value?

- Educational quality is an important part of the set of considered public goods, because quality schooling is often decisive in later life-outcomes (e.g. labor market opportunities, gained income, health etc.)
- The way in which parents sort into the housing market directly influencing the level of residentual segregation (Bayer et al., 2007)
- Ongoing debate, which dimension of education is valued by parents (outputs or learning environments containing sociodemographic composition) (Machin, 2011)
- We want to test whether the availability of upper secondary schools is such a dimension

Capitalization mechanism

- Parents are willing to pay a premium for housing units nearby top-tier school networks (Jayantha & Lam, 2015)

This capitalizes into the housing market via two mechanisms (La, 2015):

- Wealthier households that care about school quality bid up prices within the walking zone of a school
- generation of spillovers through changes in neighborhood composition

Hypothesis

Based on the theoretical background, we expect the following results:

- Education is one of the most important public services (Zhang et al., 2020) and it is therefore reasonable to test the channel
- Educational opportunities play a role in parents' housing decisions
- Parents value the opportunity to achieve higher educational outcomes for their children (Hörnig & Schäfer, 2025) because of their decisiveness for later life outcomes
- Parents are therefore willing to pay a premium for houses near a secondary school offering the opportunity to obtain an *Abitur*

Formal Hypothesis

τ_{school}(Estimated treatment effect): $H_1 : \beta_1(D_i) > 0$

τ_{abitur}(Estimated treatment heterogeneity): $H_2 : \beta_2(D_i \times A_i) > 0$

Insights from relevant empirical literature

Literature Insights (excerpt)

- The empirical literature mainly focused on the capitalization effects of s chool quality
- UK: Strong capitalization of **primary school** performance into housing prices. (Gibbons & Machin, 2003)
- US: Stronger price effects from **middle and high school** quality than from elementary schools.(Sedgley et al., 2008)
- France: Secondary school quality capitalized more strongly in areas without **private school alternatives**. (Fack & Grenet, 2010)

Empirical Design

Basic Identification Assumptions

Theoretical Assumptions:

- Parents (Households) derive utility from the perceived quality of schooling (Brunner et al., 2012) available to their child
- Parents consider the availability of Abitur and non-Abitur upper secondary schools in their housing decisions.

Empirical Assumptions:

- Parents' preferences for their children's education are reflected in hedonic price regressions.
- Conditional on the controlling for both housing and neighborhood characteristics, the treatment assignment can considered to be random.
- Within the treatment area and the control area, the capitalization effect is constant.

Assumptions (II)

- Buildings inside and outside of the treatment zones share the same average housing and neighborhood characteristics.
- Offer prices are time-independently exceeding sale prices at a constant rate.
- Property prices can be interpreted as the willingness to pay for amenities because they are determined by relevant characteristics

Data and Variables

Housing Data:

- Geo-referenced listings of sales properties (RWI – Leibniz Institute for Economic Research & ImmobilienScout24, 2023)
- Variables: living area, site area, number of rooms, number of bathrooms, year of construction, cellar

School Data:

- Locations and types of schools (Grammar School, Comprehensive School, etc. based on grid cells) (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2023)

Data and Variables

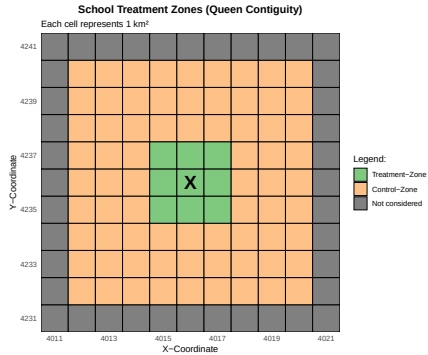
Regional Data:

- Includes information on neighborhood characteristics (e.g. income levels, migration rates, average age and the availability of infrastructure such as doctors or a supermarket) (Esri Deutschland GmbH, 2022; Geofabrik GmbH, 2026; Landesdatenbank NRW | Statistisches Landeamt | Landesbetrieb Information und Technik Nordrhein-Westfalen (IT.NRW), 2022; Statistisches Bundesamt (Destatis), 2022a, 2022b)

Quasi-experimental approach

We define spatial treatment zones to estimate the causal effect of upper secondary schools on housing prices using grid-cells:

- **Treatment-Zone:** Grid-cells **within 2 km** radius
- **Untreated zone:** Grid-cells **2–4 km** from any school
- Houses exposed to multiple secondary schools (double-treated) are excluded



Robustness check

As a robustness check, we exclude a **Buffer zone** of 1km around the treatment zone to prevent spillover contamination, therefore reducing the amount of control units

Identification Framework

Potential Outcome Model (POM): following Rubin, 1974

$$\ln(P_{ij}) = \begin{cases} \ln(P_{1ij}) & \text{if } D_i = 1, \\ \ln(P_{0ij}) & \text{if } D_i = 0. \end{cases}$$

where:

$\ln(P_{1ij})$: Price of the house, when it lies in the treatment-zone near to a secondary school.

$\ln(P_{0ij})$: Price of the house, when it would not have been 'exposed' to a secondary school (counterfactual).

Note
Since the counterfactual is not observed, we use the most similar house lying in the non-treated zone as a proxy for the counterfactual.

Matching Strategy

Identification Assumption (CIA):

$$\ln(P_{0ij}), \ln(P_{1ij}) \perp D \,|X_i$$

Estimation method: We use **matching** on observable covariates to compare treated and untreated buildings:

- Building characteristics (e.g., living area, number of rooms, building age,)
- Neighborhood characteristics (e.g., average income or age, share of immigrants, structural features)

After successful matching, the treatment effect is estimated through the following equation (Doko Tchatoka & Varvaris, 2021):

$$\begin{aligned}\tau_{\text{school}} &= \mathbb{E}\big[\ln(P_{1ij}) - \ln(P_{0ij}) \,\big|\, X_i\big] \\ &= \mathbb{E}\big[\ln(P_{1ij}) \,\big|\, X_i, D=1\big] - \mathbb{E}\big[\ln(P_{0ij}) \,\big|\, X_i, D=0\big]\end{aligned}$$

Econometric Model (OLS Specification) - First model

We estimate the following log-linear autoregressive hedonic regression for elementary and secondary schools each (Lu et al., 2023):

$$\log(\mathsf{P}_i) = \alpha + \beta_1 D_i + \mathbf{X}_i'\gamma + \mathbf{N}_i'\delta + \mathsf{FE}_{r(i)} + \varepsilon_i$$

Where:

- log(P_i): price of a building per m^2
- D_i: Is within treatment distance to a secondary school (=1) or not (=0)
- $\mathbf{X}_i$: set of building characteristics
- $\mathbf{N}_i$: set of neighborhood characteristics
- $\mathsf{FE}_{r(i)}$: Regional fixed effects to account for spatial effects at the municipality level
- ε_i: error term

Multiple treatment regime - Second model

We are especially interested in the price premium of an available school that offers academic track compared to other secondary schools (treatment heterogeneity).

The basic specifiaction is extended referring to a multiple treatment regime:

$$\log(\mathsf{P}_i) = \alpha + \beta_1 D_i + \beta_2 (D_i \times A_i) + \mathbf{X}_i' \gamma + \mathbf{N}_i' \delta + \mathsf{FE}_{r(i)} + \varepsilon_i$$

Where:

- A_i: Nearest school offers academic track (=1), otherwise (=0)
- All other variables are the same as for the main specification

Challenges and Limitations

- **Endogeneity:** better schools tend to be located in affluent neighborhoods and students endowed with those priviledged backgrounds generally achieve higher educational outcomes (Fack & Grenet, 2010)
- **Assumptions**: It may be unrealistic that the capitalization effects of upper secondary schools are uniform across the whole space (Wen et al., 2018)
- **School access rules:** In some regions, school choice or private alternatives may weaken capitalization effects.
- **Interpretation:** It is arguably that property prices can be interpreted as the willingness to pay for amenities (Jayantha & Lam, 2015) and therefore the difference between the groups as a premium for educational opportunities
- **Price validity:** Property prices from ImmoScout are *asking prices* — not actual transaction prices. This has implications on the interpretations of the results and has to be taken into account.
- **Data:** Incomplete sets of observed building ($\mathbf{X}_i$) and neighborhood ($\mathbf{N}_i$) characteristics

Results

Descriptives - Summary Statistics

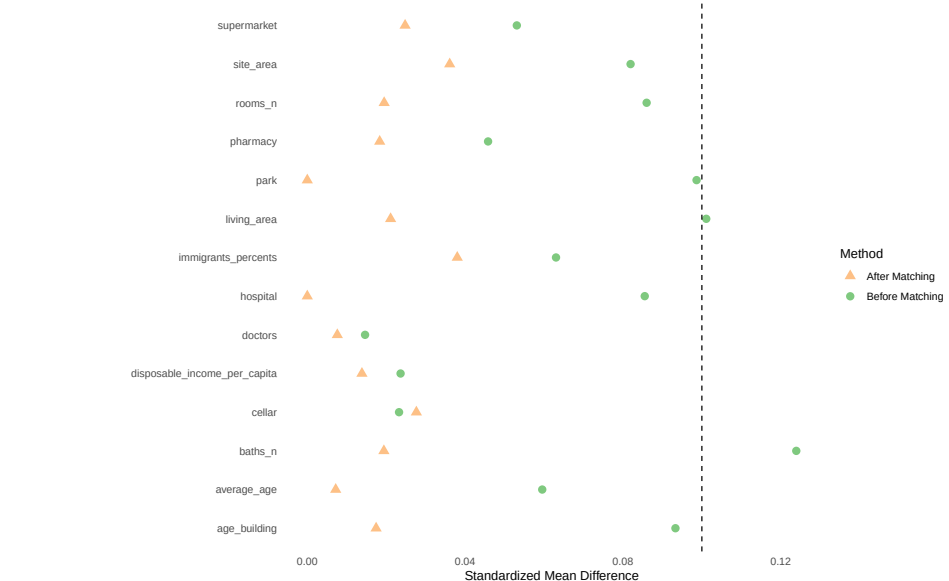
	Any School		Abitur		Non-Abitur		Control	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Price per sqm	3113.02	1194.97	3226.07	1218.31	2856.20	1098.12	3238.50	1314.23
Living Area	173.74	68.53	172.08	67.55	177.53	70.57	180.67	77.43
Site Area	591.08	395.79	577.79	395.65	621.25	394.57	623.50	474.98
Rooms	6.08	2.59	6.05	2.61	6.14	2.55	6.30	2.97
Baths	1.92	1.13	1.91	1.10	1.95	1.19	2.06	1.37
Building Age	37.91	39.82	37.32	39.12	39.25	41.35	41.63	42.23
Cellar	0.52	0.50	0.52	0.50	0.51	0.50	0.53	0.50

Descriptives - Balance Check

School nearby vs. Control

Variable	Treatment	Control	Std. Diff.	t-stat
living_area	173.744	180.673	-0.101	5.828
site_area	591.075	623.503	-0.082	4.633
rooms_n	6.080	6.303	-0.086	4.935
baths_n	1.919	2.060	-0.124	6.991
age_building	37.911	41.627	-0.093	5.478
cellar	0.516	0.528	-0.023	1.387
immigrants_percents	11.864	12.299	-0.063	3.468
average_age	44.649	44.465	0.060	-3.527
disposable_income_per_capita	26679.919	26600.951	0.024	-1.421
pharmacy	0.295	0.316	-0.046	2.719
supermarket	0.246	0.269	-0.053	3.143
hospital	0.003	0.007	-0.086	4.288
doctors	0.263	0.256	0.015	-0.875
park	0.001	0.005	-0.099	4.575

Balance Check



Main specification

~~Effect of Secondary School Proximity on House Prices~~

	Unmatched sample			Matched sample		
	(1)	(2)	(3)	(1)	(2)	(3)
School nearby	0.015	0.006	0.005	0.003	0.010	0.010
	(0.012)	(0.009)	(0.009)	(0.013)	(0.010)	(0.011)
Observations	18 703	18 703	18 703	9537	9537	9537
R^2	0.397	0.614	0.617	0.415	0.626	0.629
Building controls	-	Yes	Yes	-	Yes	Yes
Neighborhood controls	-	-	Yes	-	-	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Robust standard errors clustered at the municipality level in parentheses.

Heterogeneity Model

~~Effect of School with academic track proximity on House Prices~~

	Unmatched sample	Matched sample
	Full	Full
School nearby	-0.013	-0.009
	(0.015)	(0.015)
School nearby × Gymnasium nearby	0.026	0.027
	(0.016)	(0.017)
Observations	$18\,703$	9537
R^2	0.617	0.629
Building controls	Yes	Yes
Neighborhood controls	Yes	Yes
Region fixed effects	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Robust standard errors clustered at the municipality level in parentheses.

Robustness check

~~Effect on housing prices~~

	Unmatched sample		Matched sample	
	(1)	(2)	(1)	(2)
School nearby	0.008	−0.012	0.008	−0.012
	(0.016)	(0.024)	(0.015)	(0.025)
School nearby × Gymnasium nearby		0.028		0.027
		(0.026)		(0.027)
Observations	7991	7991	4723	4723
R^2	0.649	0.649	0.652	0.653
Building controls	Yes	Yes	Yes	Yes
Neighborhood controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Robust standard errors clustered at the municipality level in parentheses.

Policy Implications and further research

Appendix

Controls (1/2)

Table 1: Property characteristics

Variable	Description
living_area	Living area of the dwelling (m^2)
site_area	Site area of the dwelling (m^2)
rooms_n	Number of rooms
baths_n	Number of baths
age_building	Age of the building (years)
cellar	Dummy for presence/absence of a cellar

Controls (1/2)

Table 2: Neighborhood characteristics

Variable	Description
immigrants_percents	Share of immigrants in the 1 km² grid (%)
average_age	Average age of residents in the 1km² grid (years)
pharmacy	(=1) if the grid-cell contains at least 1 pharmacy, (=0) otherwise
hospital	(=1) if the grid-cell contains at least 1 hospital, (=0) otherwise
doctors	(=1) if the grid-cell contains at least 1 doctor's office, (=0) otherwise
park	(=1) if the grid-cell contains at least 1 park, (=0) otherwise

Literature I

- Bayer, P., Ferreira, F., & McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, *115*(4), 588–638. Retrieved December 2, 2025, from http://www.jstor.org/stable/10.1086/522381
- Brunner, E. J., Cho, S.-W., & Reback, R. (2012). Mobility, housing markets, and schools: Estimating the effects of inter-district choice programs. *Journal of Public Economics*, *96*(7), 604–614. https://doi.org/https://doi.org/10.1016/j.jpubeco.2012.04.002
- Doko Tchatoka, F., & Varvaris, V. (2021). Neighbourhood, school zoning and the housing market: Evidence from new south wales. *Journal of Housing Economics*, *54*(C). https://doi.org/10.1016/j.jhe.2021.101790
- Esri Deutschland GmbH. (2022). *Zensus 2022 – gitterzellen (feature layer)*. ArcGIS Online. Retrieved February 3, 2026, from https://www.arcgis.com/home/item.html?id=87285c05ac3b498a8cbf257c3c23a37d

Literature II

- Fack, G., & Grenet, J. (2010). When do better schools raise housing prices? evidence from paris public and private schools. *Journal of Public Economics*, *94*(1), 59–77. https://doi.org/https://doi.org/10.1016/j.jpubeco.2009.10.009
- Geofabrik GmbH. (2026). *Nordrhein-westfalen – openstreetmap-daten (download)*. Retrieved February 3, 2026, from https://download.geofabrik.de/europe/germany/nordrhein-westfalen.html
- Gibbons, S., & Machin, S. (2003). Valuing english primary schools. *Journal of Urban Economics*, *53*(2), 197–219. https://doi.org/https://doi.org/10.1016/S0094-1190(02)00516-8
- Hörnig, L., & Schäfer, M. (2025). The value of school choice opportunities [The European Association for Labour Economists (EALE) Conference 2024, Bergen, Norway, 5-7 September 2024]. *Labour Economics*, *95*, 102742. https://doi.org/https://doi.org/10.1016/j.labeco.2025.102742

Literature III

- Jayantha, W. M., & Lam, S. O. (2015). Capitalization of secondary school education into property values: A case study in hong kong. *Habitat International*, *50*, 12–22.
- La, V. (2015). Capitalization of school quality into housing prices: Evidence from boston public school district walk zones. *Economics Letters*, *134*(C), 102–106. https://doi.org/10.1016/j.econlet.2015.07.001
- Landesdatenbank NRW | Statistisches Landeamt | Landesbetrieb Information und Technik Nordrhein-Westfalen (IT.NRW). (2022). *Verfügbares einkommen der privaten haushalte – gemeinden, 2022*. Retrieved February 3, 2026, from https://www.landesdatenbank.nrw.de/ldbnrw/online
- Lu, Y., Shi, V., & Pettit, C. (2023). The impacts of public schools on housing prices of residential properties: A case study of greater sydney, australia. *ISPRS International Journal of Geo-Information*, *12*, 298. https://doi.org/10.3390/ijgi12070298

Literature IV

- Machin, S. (2011). Houses and schools: Valuation of school quality through the housing market. *Labour Economics*, *18*(6), 723–729. https://doi.org/10.1016/j.labeco.2011.05.005
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2023). *Open data portal*. Retrieved October 3, 2023, from https://www.schulministerium.nrw/open-data
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies.. *Journal of Educational Psychology*, *66*, 688–701. https://api.semanticscholar.org/CorpusID:52832751
- RWI – Leibniz Institute for Economic Research & ImmobilienScout24. (2023). *Rwi real estate data - campus file cross-section* (Version 1). https://doi.org/10.7807/immo:red:cross:v4

Literature V

- Sedgley, N. H., Williams, N. A., & Derrick, F. W. (2008). The effect of educational test scores on house prices in a model with spatial dependence. *Journal of Housing Economics*, *17*(2), 191–200. https://doi.org/https://doi.org/10.1016/j.jhe.2007.12.003
- Statistisches Bundesamt (Destatis). (2022a). *Zensus, datensatzbeschreibung zur tabelle „ausländeranteil in gitterzellen, 2022“*. Retrieved February 3, 2026, from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Zensus2022/_publikationen.html#1403950
- Statistisches Bundesamt (Destatis). (2022b). *Zensus, datensatzbeschreibung zur tabelle „durchschnittsalter in gitterzellen“*. Retrieved February 3, 2026, from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Zensus2022/_publikationen.html#1403950
- Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*(5), 416–424. Retrieved December 8, 2025, from http://www.jstor.org/stable/1826343

Literature VI

- Wen, H., Xiao, Y., Hui, E. C., & Zhang, L. (2018). Education quality, accessibility, and housing price: Does spatial heterogeneity exist in education capitalization? *Habitat International*, *78*, 68–82. https://doi.org/https://doi.org/10.1016/j.habitatint.2018.05.012
- Zhang, J., Li, H., Lin, J., Zheng, W., Li, H., & Chen, Z. (2020). Meta-analysis of the relationship between high quality basic education resources and housing prices. *Land Use Policy*, *99*, 104843. https://doi.org/https://doi.org/10.1016/j.landusepol.2020.104843