# The Effects of School Proximity and the School Social Index on Housing Prices

4th February 2026

Yanyi Ji    Daniel Tobien    Luc Wichtmann

# Agenda

**1. Motivation, Existing Literature & Research Question**

*Why school proximity and social indices matter.*

**2. Data & Methodology**

*Data used and our empirical model.*

**3. Results & Main Findings**

*Presentation and interpretation of our findings.*

**4. Limitations & Conclusion**

*Limitations of our work and concluding thoughts.*

# Motivation, Existing Literature & Research Question

# Motivation

- **Personal Interest**: Analyzing the real estate market's reaction to various factors is particularly interesting from an economic standpoint, as it offers insights into market dynamics. Additionally, the topic is personally intriguing, adding an extra layer of motivation to explore it further.
- **Academic and Practical Relevance**: Housing markets are influenced by various social and economic factors, and school quality is also cited as a underestimated determinant of property values, especially in research (Seo and Simons 2009).

# Motivation

Understanding the relationship between school distance, school quality and property prices can provide valuable insights for researchers and individuals interested in the real estate market.

# Related Literature Background

Several papers have examined the effect of proximity to schools on housing prices:

- **Empirical Strategies**: While hedonic models are the dominant approach to estimating housing price effects, alternative empirical strategies have also been proposed (see Black and Machin 2011)
- **Findings**: Empirical evidence suggests that proximity to schools positively affects housing prices (Rosiers, Lagana, and Theriault 2001; Huang and Hess 2018)

# Related Literature Background

- **Other Factors**: Related work also documents capitalization effects of school reputation (Chin and Foong 2006), differences between public and private schools (Sah, Conroy, and Narwold 2016) and school quality (Edusei, Espey, and Lin 2007; Metz 2015)
- **Environment of studies**: Much of the literature focuses on settings with binding school catchment areas and measures school quality primarily through test scores rather than demographic composition.

# Research Question

We study a setting *without binding school catchment areas* (NRW, Germany) and assess whether both *school proximity* and a *demographically based school social index* are capitalized into housing prices.

# Research Question

**Research Question:** How do school proximity and school social indices affect the prices of detached houses, and do these effects differ between primary and secondary schools in North Rhine–Westphalia?

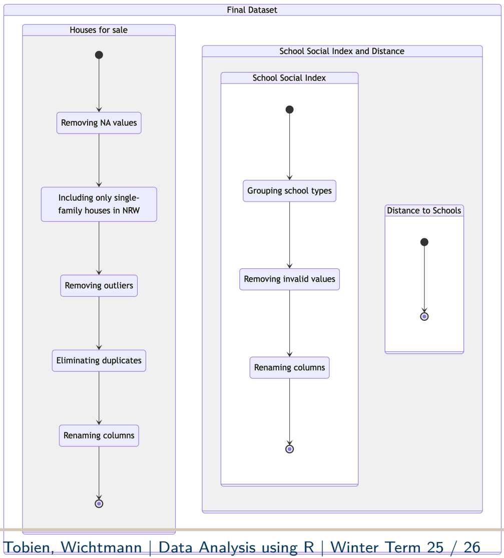
# Data & Methodology

# Data

- **Housing Data**: Cross-section dataset for houses for sale published by the Research Data Center Ruhr (FDZ Ruhr) (RWI and ImmobilienScout24 2023).
- **School Data**: School social index dataset from North Rhine-Westphalias Ministry of Schools and Education (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen 2024).

Data used from both datasets refer to the year 2022.

# Pre-Processing of Data



# Empirical Framework

We estimate regression-based hedonic models to

- examine how school proximity and school characteristics are capitalized into housing prices and
- whether the impact of school proximity on housing prices differs by the school's social index

# Empirical Framework: Hedonic Price Model

$$\log(P_i) = \beta_0 + \beta_1 \mathsf{Dist}_i + X_i'\gamma + \varepsilon_i$$

- $P_i$ denotes the listing price of house $i$
- $\mathsf{Dist}_i$ measures distance to the nearest school
- $X_i$ is a vector of housing characteristics
- $\gamma$ is a vector of coefficients corresponding to $X_i$
- $\varepsilon_i$ captures unobserved factors

# Empirical Framework: Hedonic Price Model + Interaction

$$\log(P_i) = \beta_0 + \beta_1(\text{Index}_i \times \text{Dist}_i) + X_i'\gamma + \varepsilon_i$$

- $P_i$ denotes the listing price of house $i$
- $\text{Index}_i$ is the school social index
- $\text{Dist}_i$ measures distance to the nearest school
- $X_i$ is a vector of housing characteristics
- $\varepsilon_i$ captures unobserved factors

# Empirical Framework: Final Model

$$\log(P_i) = \beta_0 + \beta_1 \text{Dist}_i + \beta_2 \text{Index}_i \\ + \beta_3 (\text{Dist}_i \times \text{Index}_i) + X_i' \gamma + \mu_{\ell(i)} + \varepsilon_i$$

- $P_i$ denotes the listing price of house $i$
- $\text{Dist}_i$ measures distance to the nearest school
- $X_i$ is a vector of housing characteristics
- $\text{Index}_i$ is the school social index
- $mu_{\ell(i)}$ are municipality fixed effects
- $\varepsilon_i$ captures unobserved factors

# Results and Main Findings

# Distance to Schools and House Prices: Hypotheses & Regression Strategy

**H1: House prices are higher closer to schools** *(negative coefficients on distance variables)*

---

**Empirical framework**

- **Outcome:** $\ln(\text{Price})$
- **Key variables:** distance to nearest **primary** school; distance to nearest **secondary** school
- **Strategy:** estimate **continuous**, **quadratic**, and **binned** distance gradients

**Baseline OLS controls**

*Standard housing attributes*

**Distance designs**

*Alternative specifications*

# Regression Specifications

## 1. Naive (Unconditional)

$$\ln(P) = \alpha + \beta_1 D^{prim} + \beta_2 D^{sec} + \varepsilon$$

- **Goal:** Establish raw spatial correlation.
- **Risk:** High Omitted Variable Bias.

## 2. Baseline (Preferred)

$$\ln(P) = \alpha + \ \mathbf{D} + \ \mathbf{X} + \varepsilon$$

- **Controls ($\mathbf{X}$):** Area, Plot, Rooms, Age.
- **Hypothesis:** $\beta < 0$ (Negative semi-elasticity).

## 3. Non-Linearity (Polynomial)

$$\ln(P) = \text{Base} \cdots + \delta(D)^2 + \ldots$$

- **Logic:** Test for convex/concave decay.
- **Test:** $\delta \neq 0$ implies varying MWTP.

## 4&5. Robustness (Single-Type)

$$\ln(P) = \alpha + \beta_k D^k + \ \mathbf{X} + \varepsilon$$

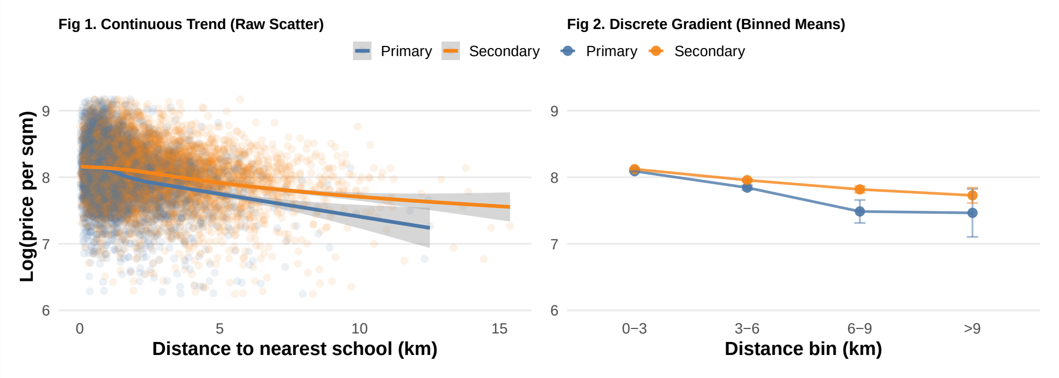
- **Method:** Estimate Primary/Secondary separately.
- **Check:** Stability against multicollinearity.

# Descriptive Evidence: Price-Distance Gradient

**Negative Slope (H1)**
Clear decay in prices as distance increases across specifications.

**Non-Linearity**
Steeper drop in the first 3-6 km, flattening afterwards.

**Heterogeneity**
Primary schools show a visibly steeper gradient.



# Main Results: Continuous Model

| | Naive (Both) | Baseline (Both) | Polynomial (Both) | Baseline (Primary) | Baseline (Secondary) |
|---|---|---|---|---|---|
| Distance to primary school (km) | −0.070*** | −0.071*** | −0.107*** | −0.103*** | |
| Distance to primary school (km)^2 | | | 0.007*** | | |
| Distance to secondary school (km) | −0.031*** | −0.034*** | −0.023*** | | −0.057*** |
| Distance to secondary school (km)^2 | | | −0.001 | | |
| Log living area | | 0.880*** | 0.879*** | 0.886*** | 0.888*** |
| Log plot area | | 0.053*** | 0.054*** | 0.046*** | 0.042*** |
| Rooms | | −0.012*** | −0.012*** | −0.012*** | −0.012*** |
| House age | | −0.002*** | −0.002*** | −0.002*** | −0.002*** |
| Intercept | 13.327*** | 8.677*** | 8.688*** | 8.649*** | 8.661*** |
| Num.Obs. | 5794 | 5794 | 5794 | 5794 | 5794 |
| R2 | 0.060 | 0.436 | 0.438 | 0.425 | 0.419 |
| R2 Adj. | 0.060 | 0.436 | 0.437 | 0.425 | 0.418 |

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

# Main Results: Binned Model

| | Naive (Both) | Baseline (Both) | Baseline (Primary) | Baseline (Secondary) |
|---|---|---|---|---|
| (Intercept) | 13.226*** | 8.594*** | 8.571*** | 8.600*** |
| dist_primary_bin3-6 | −0.147*** | −0.146*** | −0.268*** | |
| dist_primary_bin6-9 | −0.361*** | −0.369*** | −0.577*** | |
| dist_primary_bin>9 | −0.833*** | −0.538*** | −0.793*** | |
| dist_secondary_bin3-6 | −0.142*** | −0.146*** | | −0.179*** |
| dist_secondary_bin6-9 | −0.241*** | −0.256*** | | −0.331*** |
| dist_secondary_bin>9 | −0.209*** | −0.294*** | | −0.435*** |
| log_area | | 0.887*** | 0.899*** | 0.891*** |
| log_plot_area | | 0.042*** | 0.031*** | 0.038*** |
| zimmeranzahl | | −0.011*** | −0.012*** | −0.011*** |
| house_age | | −0.002*** | −0.002*** | −0.002*** |
| Num.Obs. | 5794 | 5794 | 5794 | 5794 |
| R2 | 0.046 | 0.420 | 0.401 | 0.413 |
| R2 Adj. | 0.045 | 0.419 | 0.400 | 0.412 |

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

# Compare: Continuous vs Binned

| | Baseline (Continuous) | Baseline (Binned) |
|---|---|---|
| Distance to primary school (km) | $-0.071$*** | |
| Distance to secondary school (km) | $-0.034$*** | |
| Primary: 3–6 km (ref: 0–3) | | $-0.146$*** |
| Primary: 6–9 km (ref: 0–3) | | $-0.369$*** |
| Primary: $>$9 km (ref: 0–3) | | $-0.538$*** |
| Secondary: 3–6 km (ref: 0–3) | | $-0.146$*** |
| Secondary: 6–9 km (ref: 0–3) | | $-0.256$*** |
| Secondary: $>$9 km (ref: 0–3) | | $-0.294$*** |
| Num.Obs. | 5794 | 5794 |
| R2 | 0.436 | 0.420 |
| R2 Adj. | 0.436 | 0.419 |

*Note:*

Binned coefficients are relative to the reference group: 0-3 km. Controls included but omitted from display.

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

# Model comparison & robustness (main models)

Continuous fits slightly better;

binned is slightly more robust to influential points.

### Model comparison

## Table 4. Model comparison

In-sample fit and 5-fold out-of-sample performance

| Model | N | $R^2$ | Adj. $R^2$ | AIC | BIC | CV RMSE | CV MAE | CV $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Baseline (Continuous) | 5794 | 0.436 | 0.436 | 4973.8 | 5027.1 | 0.371 | 0.288 | 0.434 |
| Baseline (Binned) | 5794 | 0.420 | 0.419 | 5143.6 | 5223.6 | 0.377 | 0.293 | 0.416 |

**In-sample:** higher $R^2$/Adj. $R^2$ is better; lower AIC/BIC is better.

**Out-of-sample (5-fold CV):** lower RMSE/MAE is better; higher $R^2$ is better.

# Model comparison & robustness (main models)

Continuous fits slightly better;

binned is slightly more robust to influential points.

**Robustness checks**

## Table 5. Robustness checks (main models)

Diagnostics summary (heteroskedasticity, multicollinearity, influence)

| Model | BP p-value | Max VIF | Influential (D>4/n) | Max Cook's D |
|---|---|---|---|---|
| **Good Reference** | <0.001 | 27.57 | 315 | 0.033 |
| **Average Reference** | <0.001 | 109.13 | 315 | 0.033 |
| **Bad Reference** | <0.001 | 578.70 | 315 | 0.033 |

Inference uses **HC1 robust SE**.

Lower BP p-values indicate heteroskedasticity; higher VIF indicates collinearity; larger Cook's D

# Economic Mechanisms & Interpretations

## 1. Capitalization

Why are houses closer to schools more expensive?

- **Commuting Costs:** Time saved is directly capitalized into property values
- **Residential Sorting:** Families cluster here, increasing local demand

## 2. Primary vs Secondary

Why is the effect stronger for primary schools?

- **Age Constraints:** Young children need

## 3. Non-Linearity

Why does the gradient flatten beyond 3km?

- **Walkability Premium:** Only exists within $<$3km distance
- **Substitution:** Public transit replaces driving at longer distances

## 4. Limitation

What is missing from this analysis?

- **Distance $\neq$ Quality:** Proximity captures accessibility, not desirability
- **Next Step:** Incorporate **Social Index**

# Distance to Schools and House Prices: Hypothesis

**H1: House prices are higher related to schools**
*Expected: Negative coefficients on distance variables*

## Empirical Framework

- **Outcome**: Log house price
- **Key variables**: Distance to nearest *primary* school, distance to nearest *secondary* school
- **Modeling Strategy**: Continuous, quadratic, and binned specifications

# Distance to Schools and House Prices: Regression Strategy

**Baseline OLS Controls**

**Standardizing housing attributes**

- Living area
- Plot area
- Number of rooms
- House age

**Two Distance Designs**

**Alternative Specifications**

1. Continuous
   - quadratic ($km^2$)
2. Binned
   - 0–3 km, 3–6 km, 6–9 km, $>$9 km

# Regression Specifications: Continuous Distance

## 1. Naive Specification (Unconditional)

*Raw price gradient without controls*

$$\ln(P) = \alpha + \beta_1 D^{prim} + \beta_2 D^{sec} + \varepsilon$$

- **Goal:** Establish raw spatial correlation.
- **Risk:** High Omitted Variable Bias (OVB).

# Regression Specifications: Continuous Distance

## 2. Baseline Model (Preferred)

*Conditional on attributes* ($\mathbf{X}$)

$$\ln(P) = \alpha + \ \mathbf{D} + \ \mathbf{X} + \varepsilon$$

- $\mathbf{X}$: Area, Plot size, Rooms, Age.
- **Assumption:** $E[\varepsilon|\mathbf{D}, \mathbf{X}] = 0$.
- **Hypothesis:** $\beta < 0$ (Negative semi-elasticity).

# Regression Specifications: Continuous Distance

### 3. Polynomial (Non-Linearity)

*Testing for convex/concave decay*

$$\ln(P) = \text{Baseline} \cdots + \delta_1 (D^{prim})^2 + \ldots$$

- **Test:** $\delta \neq 0$ implies marginal willingness to pay varies with distance.

# Regression Specifications: Continuous Distance

### 4 & 5. Single-Type Robustness

*Addressing multicollinearity*

$$\ln(P) = \alpha + \beta_k D^k + \mathbf{X} + \varepsilon$$

- **Method:** Estimate $D^{prim}$ and $D^{sec}$ separately.
- **Check:** Ensure $\hat{\beta}$ signs remain stable when isolated.

# Descriptive Evidence: Price-Distance Gradient

**Negative Slope (H1)**

Clear decay in prices as distance increases across both specifications.

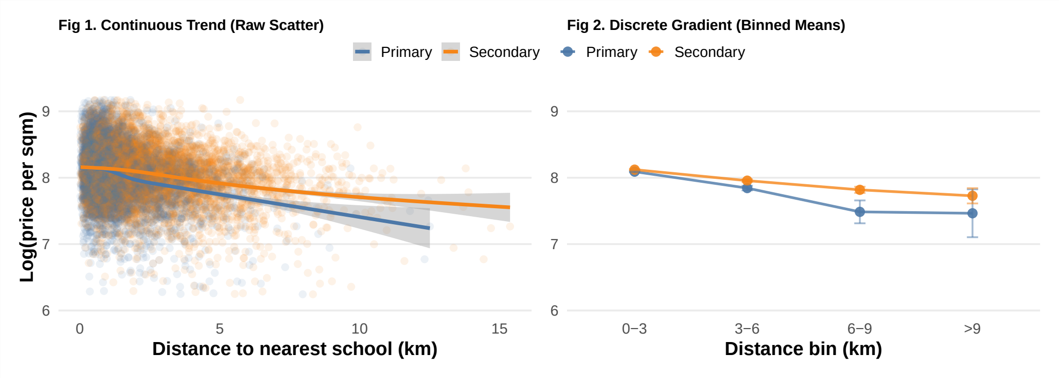
**Non-Linearity**

Steeper drop in the first 3-6 km, flattening afterwards.

**School Heterogeneity**

Primary schools (Blue) show a visibly steeper gradient than Secondary schools.

# Descriptive Evidence: Price-Distance Gradient



# Main Results: Continuous vs Binned

## Continuous Model

| | Naive (Both) | Baseline (Both) | Polynomial (Both) | Baseli |
|---|---|---|---|---|
| Distance to primary school (km) | -0.070*** | -0.071*** | -0.107*** | -0 |
| Distance to primary school (km)^2 | | | 0.007*** | |
| Distance to secondary school (km) | -0.031*** | -0.034*** | -0.023*** | |
| Distance to secondary school (km)^2 | | | -0.001 | |
| Log living area | | 0.880*** | 0.879*** | 0 |
| Log plot area | | 0.053*** | 0.054*** | 0 |
| Rooms | | -0.012*** | -0.012*** | -0 |
| House age | | -0.002*** | -0.002*** | -0 |
| Intercept | 13.327*** | 8.677*** | 8.688*** | 8 |
| Num.Obs. | 5794 | 5794 | 5794 | |
| R2 | 0.060 | 0.436 | 0.438 | |

# Main Results: Continuous vs Binned

## Binned Model

| | Naive (Both) | Baseline (Both) | Baseline (Primary) | Baseline (Secondary) |
|---|---|---|---|---|
| (Intercept) | 13.226*** | 8.594*** | 8.571*** | 8.600*** |
| dist_primary_bin3-6 | -0.147*** | -0.146*** | -0.268*** | |
| dist_primary_bin6-9 | -0.361*** | -0.369*** | -0.577*** | |
| dist_primary_bin>9 | -0.833*** | -0.538*** | -0.793*** | |
| dist_secondary_bin3-6 | -0.142*** | -0.146*** | | -0.179*** |
| dist_secondary_bin6-9 | -0.241*** | -0.256*** | | -0.331*** |
| dist_secondary_bin>9 | -0.209*** | -0.294*** | | -0.435*** |
| log_area | | 0.887*** | 0.899*** | 0.891*** |
| log_plot_area | | 0.042*** | 0.031*** | 0.038*** |
| zimmeranzahl | | -0.011*** | -0.012*** | -0.011*** |
| house_age | | -0.002*** | -0.002*** | -0.002*** |

# Main Results: Continuous vs Binned

## Compare: Continuous vs Binned

| | Baseline (Continuous) | Baseline (Binned) |
|---|---|---|
| Distance to primary school (km) | -0.071*** | |
| Distance to secondary school (km) | -0.034*** | |
| Primary: 3–6 km (ref: 0–3) | | -0.146*** |
| Primary: 6–9 km (ref: 0–3) | | -0.369*** |
| Primary: >9 km (ref: 0–3) | | -0.538*** |
| Secondary: 3–6 km (ref: 0–3) | | -0.146*** |
| Secondary: 6–9 km (ref: 0–3) | | -0.256*** |
| Secondary: >9 km (ref: 0–3) | | -0.294*** |
| Num.Obs. | 5794 | 5794 |
| R2 | 0.436 | 0.420 |
| R2 Adj. | 0.436 | 0.419 |

# Model comparison & robustness (main models)

Continuous fits slightly better;

binned is slightly more robust to influential points.

**Model comparison**

## Table 4. Model comparison

In-sample fit and 5-fold out-of-sample performance

| Model | N | $R^2$ | Adj. $R^2$ | AIC | BIC | CV RMSE | CV MAE | CV $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Baseline (Continuous) | 5794 | 0.436 | 0.436 | 4973.8 | 5027.1 | 0.371 | 0.288 | 0.434 |
| Baseline (Binned) | 5794 | 0.420 | 0.419 | 5143.6 | 5223.6 | 0.377 | 0.293 | 0.416 |

**In-sample:** higher $R^2$/Adj. $R^2$ is better; lower AIC/BIC is better.

**Out-of-sample (5-fold CV):** lower RMSE/MAE is better; higher $R^2$ is better.

# Model comparison & robustness (main models)

## Robustness checks

### Table 5. Robustness checks (main models)

Diagnostics summary (heteroskedasticity, multicollinearity, influence)

| Model | BP p-value | Max VIF | Influential (D>4/n) | Max Cook's D |
|---|---|---|---|---|
| **Good Reference** | <0.001 | 27.57 | 315 | 0.033 |
| **Average Reference** | <0.001 | 109.13 | 315 | 0.033 |
| **Bad Reference** | <0.001 | 578.70 | 315 | 0.033 |

Inference uses **HC1 robust SE**.

Lower BP p-values indicate heteroskedasticity; higher VIF indicates collinearity; larger Cook's D indicates influential points.

# Interpreting the Results: Economic Mechanisms

## 1. Capitalization

**Why are houses closer to schools more expensive?**

- **Commuting Costs & Convenience**
  - Time saved on daily school runs is directly capitalized into housing values.
  - Particularly critical for dual-income households with tight schedules.
- **Residential Sorting**
  - Families with children self-select into these areas.
  - This clustering increases local demand and price pressure compared to non-school areas.

# Interpreting the Results: Economic Mechanisms

## 2. Primary vs Secondary

**Why is the effect stronger for primary schools?**

- **Age-Related Mobility Constraints**
  - Young children cannot commute independently; parents *must* escort them.
  - This makes proximity a "non-negotiable" constraint for primary school parents.
- **Neighborhood Definition**
  - Primary schools act as the center of the local "micro-neighborhood."
  - Secondary schools serve larger regions, and older students use public transport, reducing the premium on immediate proximity.

# Interpreting the Results: Economic Mechanisms

## 3. Non-Linearity

**Why does the price gradient flatten out?**

- **Diminishing Returns (Threshold Effects)**
  - The "Walkability Premium" exists only within the first few kilometers.
  - Once driving is required, the marginal difference between 5km and 8km is negligible.
- **Substitution Effects**
  - At larger distances, alternative transport modes (school buses, public transit) replace parental driving, flattening the cost curve.

# Interpreting the Results: Economic Mechanisms

## 4. Limitation

**What is missing from a distance-only perspective?**

- **Distance $\neq$ Desirability**
  - Proximity captures *accessibility*, but not *quality*.
  - Living next to a poorly-performing school may actually be a disamenity (noise, traffic) without the educational benefit.
- **The "Social" Omitted Variable**
  - We need to control for **Social Index/Composition** to distinguish between "access to any school" and "access to a *good* school."

# Exploring the Social Index Impact on House Prices

**H2: Social index moderates the effect of school proximity on house prices**
*Expected: Better schools leading to stronger price premiums*

- **Key Finding from Distance Analysis:** We've already established that *households are willing to pay a premium* for proximity to schools
- **Next Question:** Now, we investigate *whether this price premium is driven by school quality*, measured through the *Social Index*

# What is the Social Index?

**Definition**: The Social Index is a measure that shows the socio-economic and demographic background of a schools students.

**Components of the social index are**:

- Parental Education
- Socio-economic Status of the Family
- Migration Background
- Single-Parent Household
- Unemployment in the Family
- Housing Situation

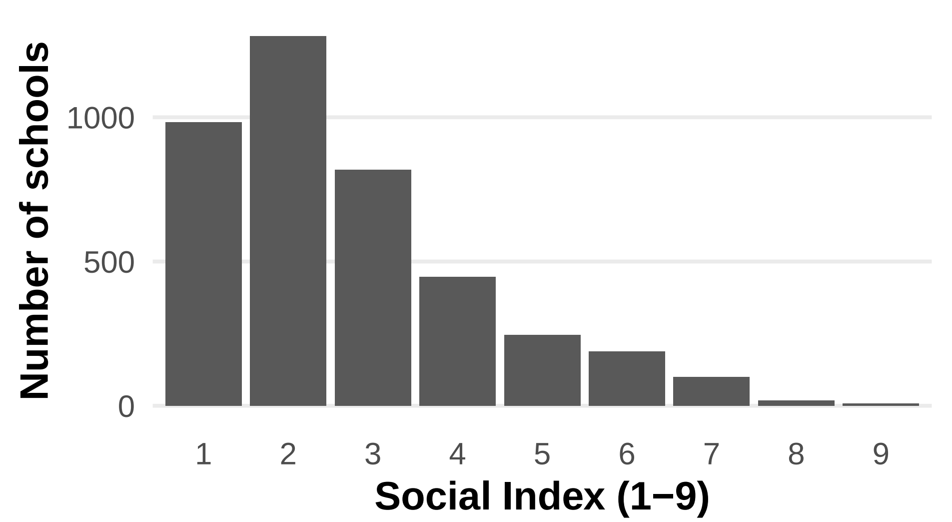
–> It provides insights into the **social environment** that might influence housing prices in that area.

# Categorization of the Social Index

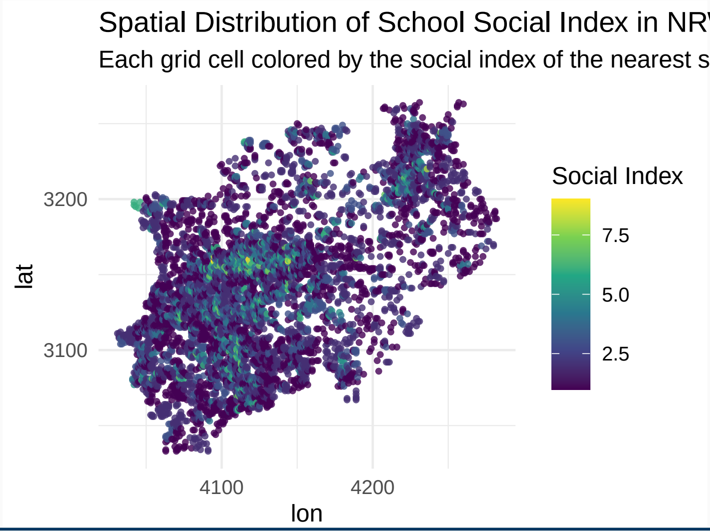
**Social Index Classification:**

- **Good Schools**: Social Index 1-2
- **Average Schools**: Social Index 3-4
- **Bad Schools**: Social Index 5-8
- **Skewed Distribution**: The Social Index classes are heavily skewed, with many schools falling into classes 1, 2 & 3. We excluded class 9 for our analysis because of too less observations

# Categorization of the Social Index: Bar Chart



# Spatial Distribution of School Social Index in NRW



# Hedonic Regression Model for Social Index

**Model Specification:**

# Social Index: Regression Results

**Table 1:** Table 1. OLS regressions (log house price): Continuous distance specifications

| | Naive (Both) | Baseline (Both) | Polynomial (Both) | Baseline (Primary) | Baseline (Secondary) |
|---|---|---|---|---|---|
| Distance to primary school (km) | −0.070*** | −0.071*** | −0.107*** | −0.103*** | |
| Distance to primary school (km)^2 | | | 0.007*** | | |
| Distance to secondary school (km) | −0.031*** | −0.034*** | −0.023*** | | −0.057*** |
| Distance to secondary school (km)^2 | | | −0.001 | | |
| Log living area | | 0.880*** | 0.879*** | 0.886*** | 0.888*** |
| Log plot area | | 0.053*** | 0.054*** | 0.046*** | 0.042*** |
| Rooms | | −0.012*** | −0.012*** | −0.012*** | −0.012*** |
| House age | | −0.002*** | −0.002*** | −0.002*** | −0.002*** |
| Intercept | 13.327*** | 8.677*** | 8.688*** | 8.649*** | 8.661*** |
| Num.Obs. | 5794 | 5794 | 5794 | 5794 | 5794 |
| R2 | 0.060 | 0.436 | 0.438 | 0.425 | 0.419 |
| R2 Adj. | 0.060 | 0.436 | 0.437 | 0.425 | 0.418 |

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

# Social Index: Interpretaition of the Results

**Primary School Proximity**

- *Distance to primary school* is still *significant* and negatively affects house prices.
- Interaction Effect: The effect is *stronger for average schools* (Estimate = -0.041, $p = 0.009$), suggesting that the price drop is greater when the school quality is average.
- *No significant effect* for bad schools ($p = 0.2996$).

# Social Index: Interpretaition of the Results

## Secondary School Proximity

- *Distance to secondary school* is also still *significant* and negatively impacts house prices.
- The effect is *stronger for average schools* (Estimate = 0.0307, p = 0.0096), meaning proximity to secondary schools is more impactful in areas with average school quality.
- *No significant effect* for bad schools (p = 0.7414).

# Limitations and Concluding Thoughts

# Limitations

- **Direct Distance**: The distance is not measured from the houses to the nearest school directly but from the school to the center of the raster cell.
- **Fragmented Grid**: We do not have data for every grid cell of NRW but only a fragmented set of grids containing houses from the available datasets.
- **Listing prices**: The prices we have are not transaction but listing prices.

# References I


Black, Sandra E., and Stephen Machin. 2011. "Housing Valuations of School Performance." *Handbook of the Economics of Education* 3: 485–519. https://doi.org/10.1016/B978-0-444-53429-3.00010-7.

Chin, Hoong Chor, and Kok Wai Foong. 2006. "Influence of School Accessibility on Housing Values." *Journal of Urban Planning and Development* 132 (3): 120–29. https://doi.org/10.1061/(ASCE)0733-9488(2006)132:3(120).

Edusei, Kwame Owusu-, Molly Espey, and Huiyan Lin. 2007. "Does Close Count? School Proximity, School Quality, and Residential Property Values." *Journal of Agricultural and Applied Economics* 39 (1): 211–21. https://doi.org/10.1017/S1074070800022859.

# References II


Huang, Peng, and Timothy Hess. 2018. "Impact of Distance to School on Housing Price: Evidence from a Quantile Regression." *The Empirical Economics Letters* 17 (April).

Metz, Neil. 2015. "Effect of Distance to Schooling on Home Prices." *Review of Regional Studies* 45: 151–71. https://doi.org/10.52324/001c.8060.

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 2024. "The School Social Index in North Rhine–Westphalia." https://www.schulministerium.nrw/schulsozialindex.

Rosiers, Francois Des, Antonio Lagana, and Marius Theriault. 2001. "Size and Proximity Effects of Primary Schools on Surrounding House Values." *Journal of Property Research* 18 (2): 149–68. https://doi.org/10.1080/09599910110039905.

# References III


RWI, and ImmobilienScout24. 2023. "RWI Real Estate Data – Campus File Cross-Section." RWI – Leibniz Institute for Economic Research. https://doi.org/10.7807/immo:red:cross:v4.

Sah, Vivek, Stephen J. Conroy, and Andrew Narwold. 2016. "Estimating School Proximity Effects on Housing Prices: The Importance of Robust Spatial Controls in Hedonic Estimations." *The Journal of Real Estate Finance and Economics* 53 (1): 50–76. https://doi.org/10.1007/s11146-015-9520-5.

Seo, Youngme, and Robert Simons. 2009. "The Effect of School Quality on Residential Sales Price." *Journal of Real Estate Research* 31 (3): 307–28.