Subsampling

Cleaned and matched space heating demand and energy prices panel dataset

Years: 2007-2019, N = 2,718,246 (annual billing observations)

- Create a random subsample stratified by energy carrier group
- N = 400 buildings per energy carrier group (Gas, Oil, District heating)

Stratified random subsample of space heating demand and energy prices panel dataset

Years: 2007-2019, N = 8,031 (annual billing observations)

Bayesian regression analysis, Heterogeneity analysis

Bayesian Regression Analysis (Re-create models, based on subsample)

Years: 2007-2019, N = 8,031

Single level structure models:

- ***b.1:*** Space heat demand as response variable (RV), energy price as only explanatory variable (EV) (approach mirrors ***m1.ols***).
- ***b.2:*** Adding additional EVs to ***b.1***; no grouping terms used (approach mirrors ***m2.ols***).

Multilevel structure models:

- ***bm.1:*** Space heat demand as RV, only price as EV, building ID and year as multilevel grouping terms with varying intercepts (partial pooling)
- ***bm.2:*** Adding additional EVs to ***bm.1***, population density and heating surface EVs not included since they lead to model divergence, building ID and year remain as multilevel grouping terms with varying intercepts (partial pooling)
- ***bm.2 Lagged:*** Model mirroring ***bm.2*** but using ln(price) from t-1

Bayesian Regression Analysis (Heterogeneity in price responsiveness)

Years: 2007-2019, N = 8,031

***Step 1: Graphical investigation of* heterogeneity**

- Scatterplots of ln(price) against ln(demand).
- Grouped by third variable to see if heterogeneity can be detected (EVs considered: Year, state, energy carrier group, degree days, heating surface, district per capita household income, and regional retirement share).
- Scatterplots indicate no group-level heterogeneity for most variables; heterogeneity found for energy carrier group (most pronounced), heating surface and district per capita household income

Step 2: Regression with interaction term

- Construct model (***bm.3***) with interaction between ln(price) and energy carrier group; models with heating surface and district per capita household income show either no relevant difference or lead to model divergence
- ***bm.3:*** Model mirroring ***bm.2*** which is extended by an interaction term between ln(price) and energy carrier group.

Sub-Sampling

Years: 2009-2019, N = 2,718,240 (annual billing observations)

- Create a random stratified sub-sample by energy carrier group
- N = 200 buildings per energy carrier group (gas, oil, district heating)
- Additional criterium: Must contain renovation information from energy performance certificates

Certificates (EPC) data

Years: 2008-2019,
N = 57,156 (individual buildings)

Stratified random sub-sample of space heat demand and energy prices panel dataset

Years: 2009-2019, N = 4,410

Bayesian Regression Analysis (Focus on Robustness and Heterogeneity)

Bayesian Regression Analysis

Years: 2009-2019, N = 4,410

- *Re-create models used for the full sample for the sub-sample and compare results*

Robustness (account for potential biases)

Years: 2009-2019, N = 4,410

Using regressions:

- Omitted variables: Add information on renovation activities to account for demand side shifts unobserved in the full sample scenario
- Reverse causality: Run model with a lag of ln_price (t-1)

Heterogeneity in price responsiveness

Years: 2009-2019, N = 4,410

Using graphs:

- Scatterplots of ln(price) against ln(demand)
- Grouped by third variable to see if heterogeneity can be detected (Investigated: year, state, energy carrier group, degree days, heating surface, regional income, and regional retirement share)
- Plots indicate no group-level heterogeneity for all variables except for energy carrier group

Using regressions:

- Run model with interaction between ln(price) and energy carrier