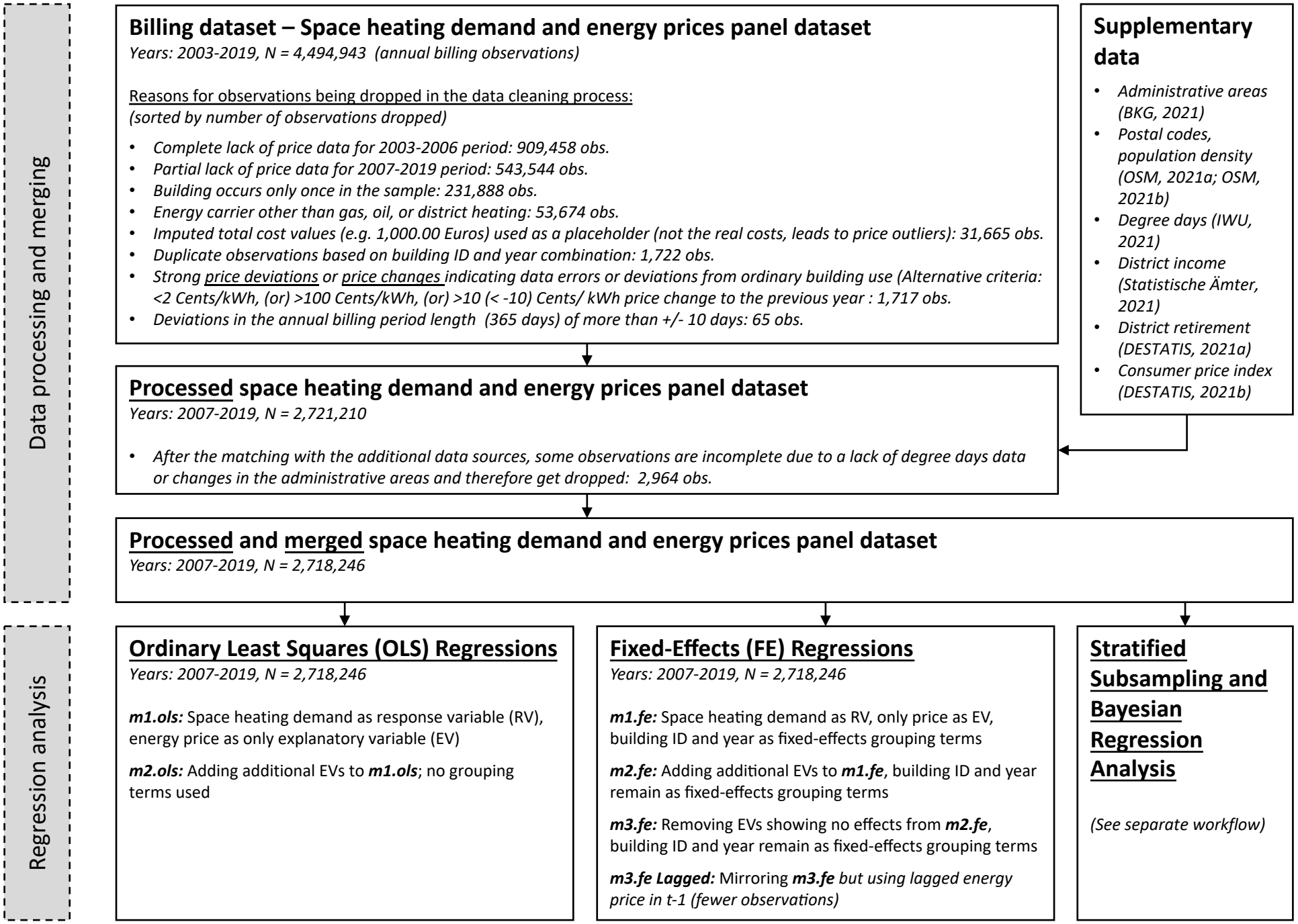
Data processing and merging
Billing dataset – Space heating demand and energy prices panel dataset
Years: 2003-2019, N = 4,494,943 (annual billing observations)
Reasons for observations being dropped in the data cleaning process:
(sorted by number of observations dropped)
• Complete lack of price data for 2003-2006 period: 909,458 obs.
• Partial lack of price data for 2007-2019 period: 543,544 obs.
• Building occurs only once in the sample: 231,888 obs.
• Energy carrier other than gas, oil, or district heating: 53,674 obs.
• Imputed total cost values (e.g. 1,000.00 Euros) used as a placeholder (not the real costs, leads to price outliers): 31,665 obs.
• Duplicate observations based on building ID and year combination: 1,722 obs.
• Strong price deviations or price changes indicating data errors or deviations from ordinary building use (Alternative criteria: <2 Cents/kWh, (or) >100 Cents/kWh, (or) >10 (< -10) Cents/ kWh price change to the previous year : 1,717 obs.
• Deviations in the annual billing period length (365 days) of more than +/- 10 days: 65 obs.
Supplementary data
• Administrative areas (BKG, 2021)
• Postal codes, population density (OSM, 2021a; OSM, 2021b)
• Degree days (IWU, 2021)
• District income (Statistische Ämter, 2021)
• District retirement (DESTATIS, 2021a)
• Consumer price index (DESTATIS, 2021b)
Processed space heating demand and energy prices panel dataset
Years: 2007-2019, N = 2,721,210
• After the matching with the additional data sources, some observations are incomplete due to a lack of degree days data or changes in the administrative areas and therefore get dropped: 2,964 obs.
Processed and merged space heating demand and energy prices panel dataset
Years: 2007-2019, N = 2,718,246
Regression analysis
Ordinary Least Squares (OLS) Regressions
Years: 2007-2019, N = 2,718,246
m1.ols: Space heating demand as response variable (RV), energy price as only explanatory variable (EV)
m2.ols: Adding additional EVs to m1.ols; no grouping terms used
Fixed-Effects (FE) Regressions
Years: 2007-2019, N = 2,718,246
m1.fe: Space heating demand as RV, only price as EV, building ID and year as fixed-effects grouping terms
m2.fe: Adding additional EVs to m1.fe, building ID and year remain as fixed-effects grouping terms
m3.fe: Removing EVs showing no effects from m2.fe, building ID and year remain as fixed-effects grouping terms
m3.fe Lagged: Mirroring m3.fe but using lagged energy price in t-1 (fewer observations)
Stratified Subsampling and Bayesian Regression Analysis
(See separate workflow)