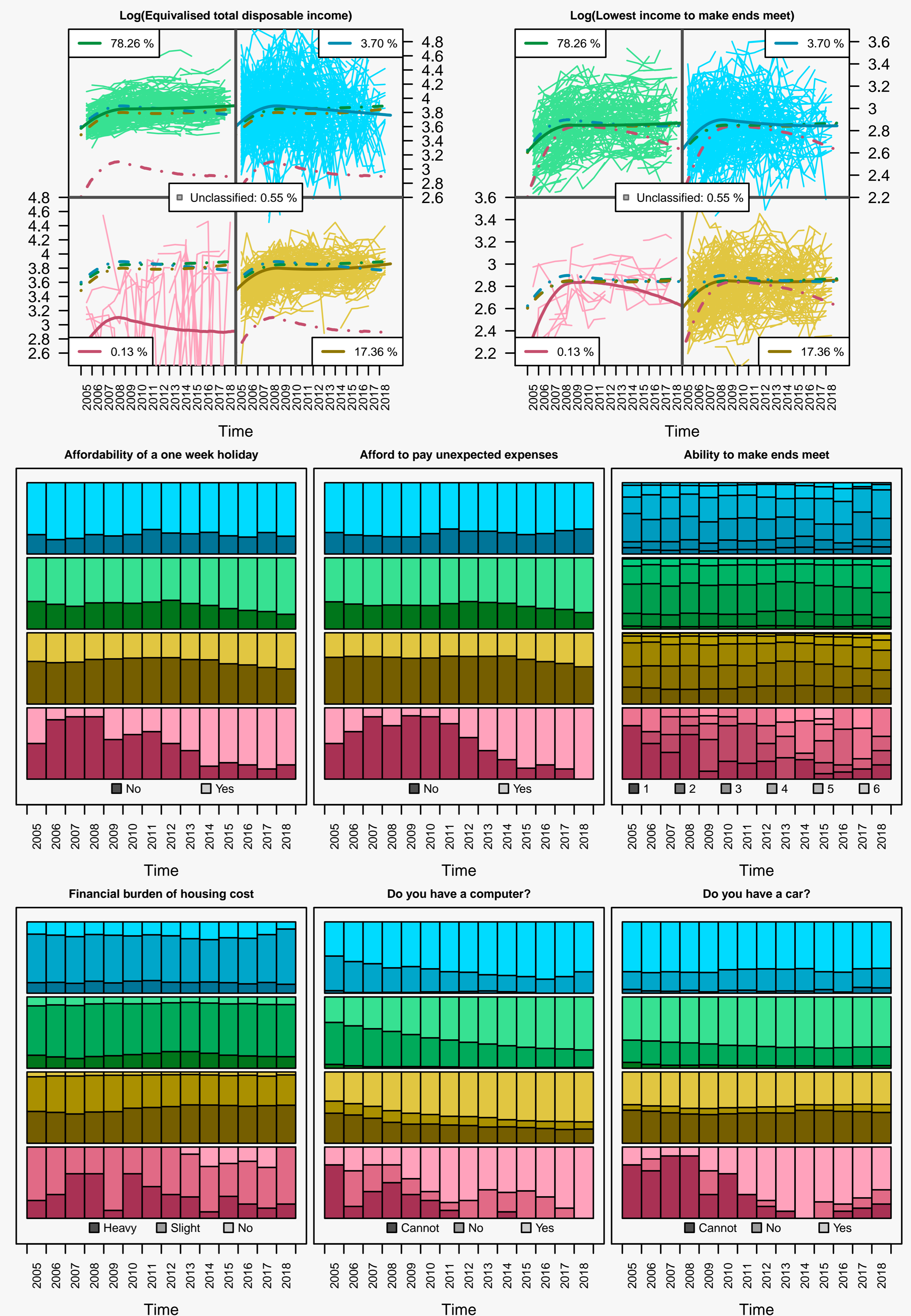


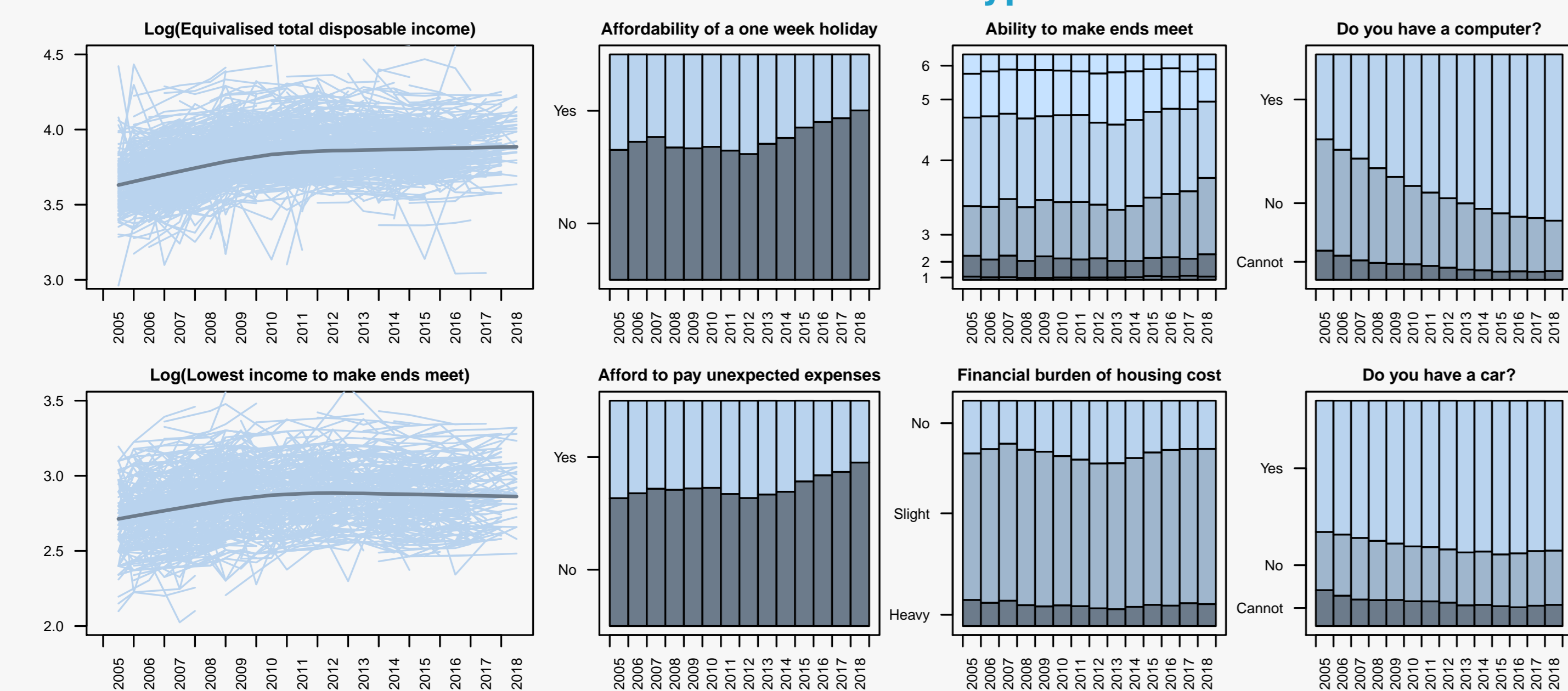
EU-SILC dataset

- EU-SILC = European Union - Statistics on Income and Living Conditions
- Longitudinal multidimensional data on income, poverty, social exclusion and living conditions measured on private households
- Annually gathered data via questionnaires targeted both on households and its members
- Available data: $n = 23\,360$ households from the Czech Republic (years 2005 – 2018)
- Outcomes
 - Numeric outcomes (modelled on logarithmic scale)
 - HX090 – Equivalised total disposable income [EUR/year]
 - HS130 – Lowest income to make ends meet [EUR/month]
 - Binary outcomes (Yes / No)
 - HS040 – Affordability of a one week holiday
 - HS060 – Afford to pay unexpected expenses
 - Ordinal outcomes (self-evaluation by the respondent)
 - HS120 – Ability to make ends meet
 - HS140 – Financial burden of the total housing cost
 - Categorical outcomes (Yes / No – cannot afford / No – other reason)
 - HS090 – Do you have a computer?
 - HS110 – Do you have a car?
- Explanatory variables:
 - time (quadratic spline parametrization)
 - level of urbanization (rural, town, city, Prague)
 - equivalised household size
 - the highest ISCED level attained within the household
 - presence of a baby, a student, ...
- Research goals
 - To discover unobserved heterogeneity in modelled outcomes
 - To identify hidden groups of similar longitudinal evolution
 - To partition households into these groups to determine the level of social-economic status

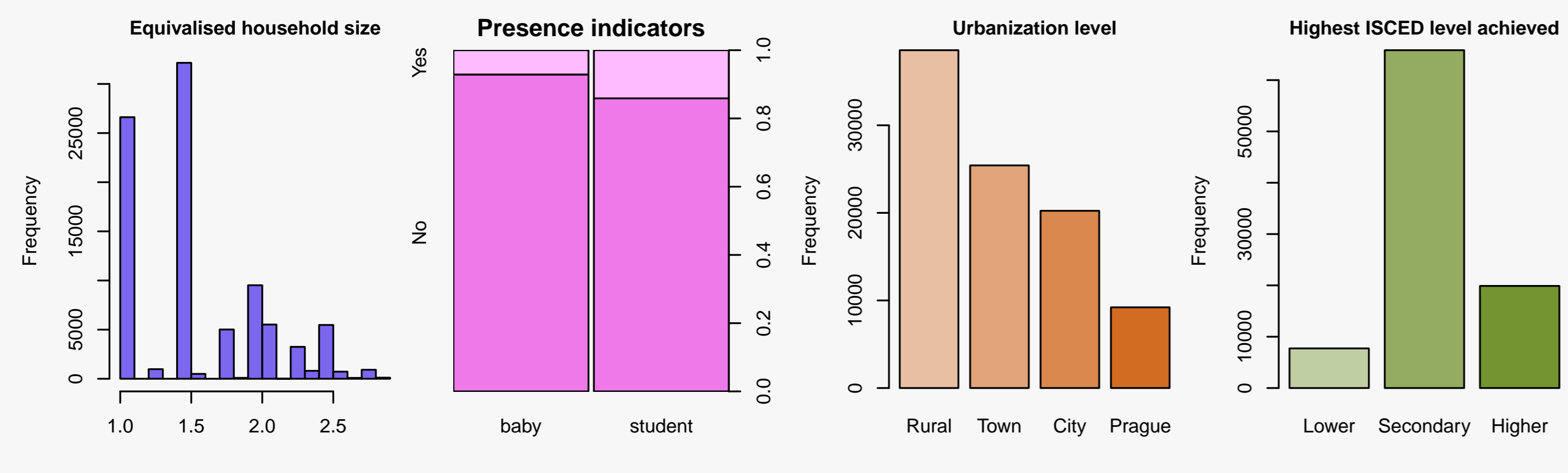
Classified households into $G = 4$ groups



Evolution of the modelled outcomes of different type in time



Explanatory variables



Methodology - GLMM-based clustering of longitudinal mixed-type data

- Random-effects models to capture specifics of each individual household
- Related outcomes joined through a general covariance matrix of combined vector of random effects
- Numeric outcomes - Classical normal linear mixed-effects model (LME)
- Binary outcomes - Logistic regression with random effects
- Ordinal outcomes - Ordinal logistic regression via cumulative probabilities
- Categorical outcomes - Multinomial logistic regression with random effects
- A mixture of these models - groups differ in selected parameters
 - spline parametrization of time evolution, within-household variance of numeric outcomes
- For more details see (Vávra and Komárek, 2022)

Effects of other covariates

