The method used for data analyses is the OLS regression. The specification of our empirical model is:



The dependent variable is a dummy indicating whether the respondent is a smoker or not in the last month (if yes=1; otherwise=0). The explanatory variable $Family\_{it}$ is the number of children in a family. $x\_{it}$ represent other time-variation control variables，including the years of education, occupation and age difference between couples. We defined age difference as: husband’s age minus wife’s age. $θ\_{i}$ is family level fixed effect，$γ\_{t}$ is time effect， $ε\_{it}$ is error term.