

Israel Diego (UMID: israeldi)

Stats 504

October 10th, 2019

U.S Household Family Types and Sizes

We aim to study the relationship between households of different types and sizes and population structure. We gain insight into changes in the household structures, aggregated by U.S census tracts, by conducting a principal components analysis (PCA) of household types and sizes from 1990 through 2010. We also determine the relationship between household types and population structure by comparing their principal component scores.

In order to study household structure for U.S. census tracts, we use a PCA approach. Each census tract consists of around 4,000 people with around 75,000 census tracts in the US at present. Household types are divided into family and non-family households, and each type is further subdivided into household size. For a household to be categorized as a family household, it needs to have at least two people living in the same household. Both family and non-family households can have household sizes up to seven people or more people. U.S. population structure is divided between men and women, and further subdivided into 18 age groups from less than 5 years old to more than 85 years old.

First, we note that since 1990 until present, there is a decreasing trend in the mean proportion of family households and an increasing trend in the mean proportion of non-family households. Thus, we predict that the gap between the number of family and non-family households will narrow in the following decades. We begin our analysis by looking at the principal components of household types by household size for 1990 through 2010.

The plots in Figure 1 reveal that the principal component loadings for the first and second layers are relatively similar for each decade for family households (blue) and non-family households (orange) . From the first layer of loadings, we see that two-person family household communities are close to the mean proportion of household over all census tracts. Additionally, family households with five to seven or more people load strongly negative, 25% lower than the mean proportion of U.S census tracts. On the other hand, the first layer loadings for one to two-person non-family households load strongly positive, roughly 20% higher than the mean

proportion of households over all census tracts. We also note that non-family households with five to seven or more people are close to the mean proportion.

Furthermore, we interpret results from the second layer of principal component loadings from Figure 1. For family households, we see a lower population of two to four-person households and a higher population of 6 to 7 or more person family households. In the case of non-family households, we note a higher population of two to three-person households. Thus 6 to 7 person family households and 2 to 3 person non-family households load the strongest in the positive direction, while 2 to 4-person family households load strongest in the negative direction.

Finally, we compare the principal component scores of U.S population structure and household types. We present the principal component scores that were most correlated with each other in Figure 2. First of which, we plot household layer 2 scores against population structure layer 1 scores. The scores have a pearson correlation coefficient of -0.67. If scores are negative for the second layer of household type and positive for the first layer of population structure, we predict that a lower population of two to four-person family households correspond to communities with young kids still at home and adults between the ages of 35-59, presumably the kids' parents. If we take a high score from layer two of house type and low score from layer one of population structure, we predict that higher population of two to three-person non-family households correspond to communities with young adults between the ages of 18-29. These results are intuitive since we are seeing a decline in small family households and an increase young adults cohabiting in the U.S.

In the second plot on Figure 2, we compare household type layer 1 scores against population structure layer 2 scores. The scores have a pearson correlation coefficient of 0.77. Given a low score on household layer 1 loadings and low score on population layer 2 loadings, we predict that five to seven or more-person family households correspond to a low population of young kids. On the other hand, high scores on both layers imply one to two-person non-family households correspond to more people in retirement. The latter result may be attributed to retiring baby-boomers living alone in their households or with a caretaker.

In conclusion, our principal component analysis predicts that the proportion of large families is declining and the proportion of small non-family households is increasing. When comparing household types and population structure, we see that a decrease of two to four-person family communities consist of families with young kids and their parents. In addition, we predict an increase in small non-family households consisting of young adults.

Figure 1

Component Loadings for Household Type by Household Size

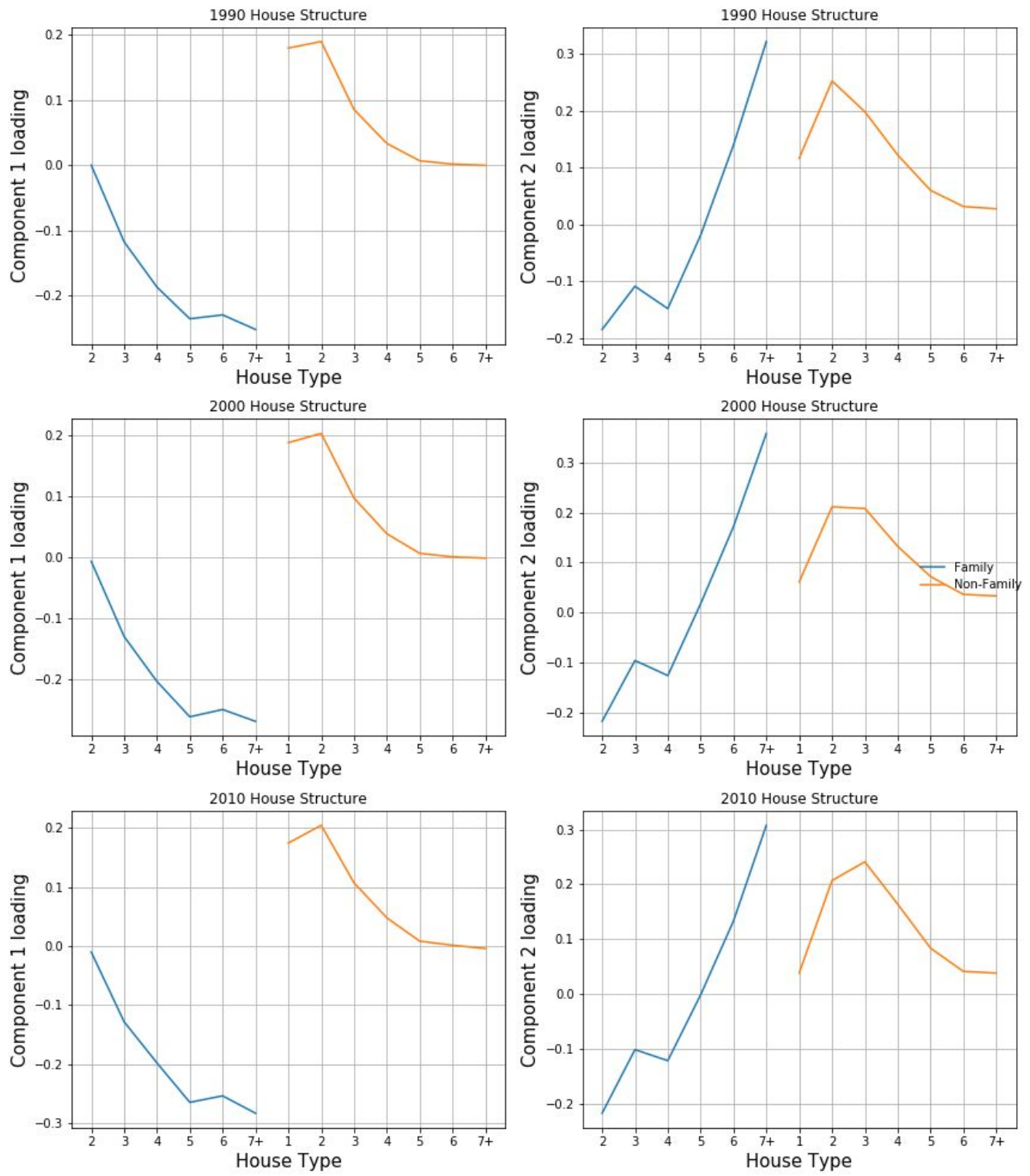


Figure 2

Principal Component Scores for Household Type and Population Structure

