Happy at Home or Stuck in Place? The Determinants of Declining Local Residential Mobility

Lei Ma*

October 6, 2021

Abstract

The local residential move rate has declined by 50 percent over the past 35 years, but little is known about the underlying factors driving the long-run decline. This paper begins by examining the impact of demographic shifts on local moving trends. I find that changes in age, homeownership, race, income and marital status can explain approximately 25 percent of the overall decline. Then I explore two key mechanisms contributing to the long-term decline. Using an instrumental variable approach, I evaluate the causal impact of family life events and housing costs on local moves. The results indicate that the decrease in marriage rates accounts for at least 30 percent of the move rate decline, while rising rents have significantly lowered the mobility of renters in the last 15 years. These findings imply that the overall decline in local moves is primarily attributed to the diminishing benefits associated with relocations, but it is evident that rising housing costs are increasingly trapping renters in place.

*Department of Economics, Boston University. Email: leim@bu.edu.

1 Introduction

The geographic mobility of Americans has declined by over 50 percent over the past four decades. In 1981, about 18 percent of Americans moved to a different residence each year, but the fraction dropped to 9 percent in 2020. A large body of literature documents the decline in interstate moves and tests different theories to explain the long-run trend[1]. However, few studies focus on local residential moves although over 60 percent of all moves are within the same county or city. Similar to long-distance moves, the local residential move rate has also dropped by half in the past four decades.

This paper fills the gap in the literature by providing stylized facts about the trend of local residential moves and providing explanations for the declining trend. I find that population compositional changes explain at most 25 percent of the decline. Instead, delayed marriage explains at least 30 percent of the decline over the past thirty-five years and rising rent explains about 80 percent of the decline of renters over the past fifteen years.

The first part of the paper documents the trends of local move rates over the past thirty-five years. I use three measures of local move rate. The primary measure is the share of individuals that had moved within the same county in the Current Population Survey (CPS). I also use the share of individuals that had moved within the same public use micro area (PUMA) and metropolitan statistical area (MSA) in the American Community Survey (ACS). All measures show consistent patterns that the local move rate has declined by 50 percent.

Next, I explore the role of sociodemographic factors on local moves. I begin by presenting descriptive evidence of local move rates for different demographic groups. I find that the local move rate decreases for every population subgroup regardless of age, race, income, education, and housing tenure status, but people who are younger, have lower income, are white, and do not have a college degree contribute most to the aggregate decline. Then I test how changes in population composition contribute to the aggregate decline in local moves. I adjust the local move rate using population characteristics in a pooled OLS regression. I find that changes in age, homeownership, marital status, income, and race jointly can explain at most 25 percent of the entire decline. The descriptive evidence suggests that there exist broader factors that are responsible for the declining trend.

I also examine the correlation between local housing market conditions and local moves

[1] See Molloy et al. (2011) for a comprehensive review of the trend of internal migration during the past three decades in the United States. They use various data sources and find a reduction in migration of all distances across different sociodemographic groups. Many papers attempt to find plausible explanations for the declining trend of interstate migration. Many papers attribute the declining mobility to lower labor market returns in alternative locations, such as Molloy et al. (2014), Kaplan and Schulhofer-Wohl (2017), and Ganong and Shoag (2017). A more detailed literature review is at the end of this section.

given that over half of all local moves are due to housing-related reasons. I find that households living in cities with higher housing costs and lower vacancy rates are less likely to make local moves although the magnitudes are relatively small. It suggests that households take housing prices and search frictions into consideration when making local move decisions.

The second part of the paper aims to find plausible explanations for the declining trend in local moves. The decline in local moves can reflect either a decrease in the benefits of residential moves or an increase in the cost of moves. These two explanations have opposite implications on household welfare. If the decline in moves is mostly due to lower benefits, households are happy staying in their current residences. But if the higher cost of moves is the reason that deters households from making moves, households suffer welfare losses because they are locked in their current residence and could not easily adjust housing consumption.

I propose two theories that reflect these two mechanisms. First, households might face fewer moving shocks to make local moves. Moving shocks lower the utility of staying in the current house and incentivize households to move to alternative residences to improve housing match quality. One example of common moving shocks is marriage shocks, that is, when individuals switch from single to married.

I explore the long-term causal relationship between marriage shocks and local moves. The main challenge of the analysis is that there is no publicly available individual-level panel data at a large scale that allows me to observe changes in marital status. My solution is to construct pseudo-panel data from the cross-sectional CPS data. I group individuals into mutually exclusive groups based on their birth cohort, gender, race, and education, and then track changes in marital status between years at the group level. The assumption is that individuals within the same group have a high degree of homogeneity and share similar move patterns. The construction of marriage shocks contains measurement errors due to imperfect grouping, and the OLS estimates also suffer from omitted variable bias. To address the endogeneity issues, I use the variations of the potential marriage market size to instrument for marriage shocks. The intuition is that individuals are more likely to get married if they face a large pool of potential marriage partners. The IV estimates show that getting married is associated with an increase in the local move rate in the same year by about 15.5 to 20.6 percentage points. The back-of-envelope calculation suggests that the decline in marriage rate explains at least 30 percent of the entire decline in local move rate over the past thirty-five years.

The second theory is that the cost of moving might be increasing and thus discourages local moves. If households want to move to a better house to improve housing match quality but the new house would cost significantly more than the current house, households may not be able to afford the move due to financial constraints. Then they have to stay in their current house with unideal living conditions and suffer welfare losses.

To test the theory that higher alternative housing costs discourage local moves, I examine the causal relationship between local moves and housing cost differences before and after moves. I use the ACS data from 2005 to 2019 to conduct the analysis because the CPS contains no information about housing characteristics. The challenge is that I only observe the current housing cost for each household, but I need to know the housing cost in alternative residences to calculate housing cost differences. I use two methods to construct the counterfactual housing cost if the household were to make a different moving decision. The first method uses the distribution of housing costs of households with similar socioeconomic characteristics, and the second method estimates the marginal housing cost assuming the household were to increase the size of the house after the move. The OLS estimates suffer from measurement errors and omitted variable biases, so I employ an instrumental variable approach to account for potential endogeneity. I use house price growth to instrument for housing cost differences since any house price growth is more likely to be capitalized in the housing cost of recent movers. The magnitudes of IV estimates using different methods differ by quite a lot, but they are consistent with the qualitative conclusion that renters are more sensitive to housing cost differences while homeowners do not respond much to housing value differences. The results suggest that homeowners can hedge against future housing price growths because they can use home selling gains to purchase their next home, while renters do not accumulate home equity.

Related literature

This paper is directly related to the literature on geographic mobility. Most papers focus on the large decline in interstate and inter-city mobility over the past several decades (Molloy et al. (2011); Kaplan and Schulhofer-Wohl (2017)). Declining long-distance mobility is problematic because it reflects that some workers are not optimizing their location choices to maximize economic returns. At the macro level, lower mobility can cause the economy to be flexible because of misallocated labor. So many papers focus on the relationship between migration and labor market factors to explain the decline in interstate mobility (Molloy et al. (2014); Kaplan and Schulhofer-Wohl (2017); Notowidigdo (2020); Johnson and Kleiner (2020)). Another strand of literature finds that rising housing prices and tight housing supply in productive cities restrain population inflow (Ganong and Shoag (2017); Hsieh and Moretti (2019)). Some other papers focus on the impact of the ability to pay for new housing on residential mobility (Chan (2001); Ferreira et al. (2010); Di Maggio et al. (2019); Bleemer et al. (2021); Giannone et al. (2020)). For example, Chan (2001) and Ferreira et al. (2010) find that homeowners are less mobile when house prices fall during housing bust because they have low or negative equity to pay for new housing. Di Maggio et al. (2019) and Bleemer et al. (2021) find that higher student debt increases co-residence with parents and inhibits geographic mobility of young adults.

However, fewer studies focus on short-distance moves although short-distance moves consist of over half of all types of moves and they have significant implications on household welfare, neighborhood development, and housing market dynamics. At the household level, local residential moves are often the only way to accommodate changing housing preferences over the lifecycle. At the neighborhood level, local residential moves directly affect neighborhood population composition and can potentially change neighborhood characteristics (for example, Glaeser et al. (2018)). At the aggregate level, home selling and buying activities associated with local moves affect the volume of housing transactions and local housing price dynamics (Ngai and Tenreyro (2014); Ngai and Sheedy (2020)).

In this paper, I fill the literature gap by presenting empirical evidence of the long-run trend in local mobility and providing explanations for the decline. My paper is guided by the early literature on the moving process. Several early theoretical work assumes that move decisions are triggered by dissatisfaction with current housing characteristics from changing household needs, location amenities, or changing tolerance of dissatisfaction (Speare (1974); Wolpert (1965); Brown and Moore (1970); Ngai and Sheedy (2020)). I assume that individuals are hit by moving shocks that lower housing match quality and the decrease in housing match quality propels individuals to initiate moves. My empirical strategy is also inspired by early empirical work on the correlates of local move decisions. Rossi (1955); Speare Jr (1970); Brown and Kain (1972); Weinberg (1979) offers correlational evidence that intra-urban residential moves are accompanied by lifecycle changes, such as changes in family composition and marital status. I employ an instrumental variable method to recover the causal relationship between local residential moves and family life events.

This paper is organized as follows. Section 2 describes the data. Section 3 shows the descriptive patterns of local mobility. Section 4 presents empirical evidence of the causal relationship between life events and local moves. Section 5 presents empirical evidence of housing cost differences and local moves. Section 6 concludes.

2 Data

This paper uses two datasets, the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) from 1986 to 2020 and the American Community Survey (ACS) from 2005 to 2019. Both datasets are cross-sectional surveys over a large sample of the US population and include a rich set of socioeconomic controls at the individual and household levels. I restrict the sample to adults over age 18 not living in group quarters and also exclude immigrants.

Both the CPS and ACS record whether the individual had changed residence in the past year. The local move measure used in the CPS is move within the same county. However, county may not be the appropriate geographic unit to capture local moves because some cities consist of multiple counties and within county move rate underestimates the actual local move rate. So I also use the move rate within Public Use Micro Area (PUMA)[2] and metropolitan statistical area (MSA)[3] in the ACS to measure local moves. One limitation of both datasets is that they only provide an indicator of whether the respondent has moved, so I cannot observe the moving history or frequency of moves each year. The CPS provides a much longer time series, so I use the CPS when I examine the long-term trend in local move rate. The ACS contains information on housing characteristics, such as housing value, rent, and number of bedrooms, so I use the ACS to investigate the relationship between housing factors and local moves.

3 Empirical patterns of local mobility

This section presents some basic facts about local residential moves over the past three decades and then describes the roles of sociodemographic variables and housing market conditions on local moves. I use the descriptive evidence as guidance to find plausible hypotheses to explain the long-term decline.

3.1 Basic facts about local moves

Fact 1: Local moves constitute the majority of all moves

Figure 1 shows the annual move rate by move distance from 1986 to 2020 using the CPS. The green line shows the share of population who have moved within the same county, which is the measure for local move rate in the CPS. The orange and purple lines show two measures of long-distance moves, within-state moves and across-state moves. Figure 1 shows that the local move rate is about three times higher than long-distance move rates and all types of moves have experienced different degrees of declines over the past thirty-five years. Although local moves consist of a significant share of all residential moves, it has received much less

[2]PUMAs are statistical geographic areas of at least 100,000 residents defined by the Census. An easy but crude way to understand it is that it represents groups of counties for small counties.

[3]Move within MSA is not a variable directly recorded in the ACS. I construct it using the MSA of residence in the previous and current years. Individuals not living in MSAs are excluded from the sample.

attention than long-distance moves. In this paper, I focus on finding explanations for the declining trend in local residential mobility.

Fact 2: The local move rate has declined by 50 percent over the past three decades

To take a deeper look at the evolution of local move rate, Figure 2 uses three measures of local moves to track changes over time. The primary measure of local move rate is the within county move rate using the CPS in the green line. The share of population who have moved within the same county has dropped from around 11 percent in 1986 to 5.2 percent in 2020. To address the underestimation concern of within county moves and cross-validate the declining pattern, I also show the move rate within PUMA and MSA from the ACS in the orange and purple lines respectively. The move rates within PUMA and MSA are higher than within county move rates because PUMAs and MSAs cover multiple counties. Although all three measures do not perfectly capture local moves, they are mostly parallel to each other and have declined by similar magnitudes (in percentage terms). It gives me the confidence to claim that the local move rate has been declining by about 50 percent in the past three decades.

Fact 3: Homeowners and renter households exhibit different time series patterns after 2000

Since homeowners and renters have different considerations when making move decisions, I document the local move rate by housing tenure status in Figure 3. First, renters are four times more likely to make local moves compared to homeowners and account for a larger share of aggregate local moves. This is consistent with the intuition that it is much cheaper for renters to make local moves relative to homeowners because they face lower financial and non-financial costs when changing residences. Second, the local move rates of both homeowners and renters have declined by about 50 percent between 1986 and 2020. It suggests that there exist some common forces driving the decline in local moves regardless of housing tenure status. However, the time series patterns do not perfectly align for homeowners and renters after 2000. For example, the local move rate of homeowners declines sharply between 2000 and 2010, but not renters. This is probably because homeowners are more subject to housing market fluctuations during this period. In particular, many homeowners have low or negative equity during the housing crisis, which makes home selling and buying more costly (Ferreira et al. (2010)). On the other hand, renters experience large declines in local move rates after 2010, which might be correlated with the rising rents over the past ten years. The different time series patterns after 2000 suggest that housing market conditions could be one determinant of local move decisions.

Fact 4: Housing-related reasons are the primary reason for local moves

Local residential moves typically do not involve changes in local labor market, but almost always lead to changes in housing characteristics. Figure 4 decomposes within county moves by reason for moving and housing tenure status. At least half of within county moves are due to housing-related reasons[4] regardless of housing tenure status, while family-related reasons constitute another 30 percent of within county moves[5]. It implies that the goal of many local moves is to improve housing match quality. When households are hit by moving shocks, such as changes in family composition, income shocks, or natural disasters, the current housing no longer meets the households' housing preferences, so they have the desire to move to a different house to improve the housing match quality. Figure 4 also shows that housing-related moves have declined the most compared to other types of moves. It emphasizes the important role of housing-related reasons when exploring the mechanisms driving the long-term decline.

3.2 Role of sociodemographic characteristics on local moves

In this subsection, I explore the role of sociodemographic characteristics on local moves. To begin with, I present descriptive evidence on how local move rates vary for different sociodemographic groups.

Figure 5 plots within county move rate by age, household income, race, and education over the past three decades. Panel (a) uses a binscatter plot to show the local move rate by age during three time periods, 1990-2000, 2000-2010, and 2010-2020[6]. It shows that the local move rate has declined for people of all ages over the past thirty years, but people below age 35 experience the largest decline in local move rate. Panel (b) shows the local move rate by income using a similar binscatter plot. Similar to panel (a), individuals at all income levels experience declines in local moves at different degrees, but individuals with annual household income below $50,000 experience the largest decline. The pattern by income might be correlated with the pattern by age because people tend to earn lower

[4]Housing-related reasons include want new or better housing, want to become homeowners, want better neighborhood, want cheaper housing, and other unspecified housing-related reasons.

[5]One thing to bear in mind is that the CPS only allows the respondent to choose one response for this question although the move might be due to multiple causes. It is possible that some family-related moves are also housing-related moves.

[6]I follow the guideline of Cattaneo et al. (2019) to graph the binscatters. I use 20 quantile-based bins of the main variable. I also add housing tenure status as the control variable.

income when they are young and experience large income growth as they age. Panel (c) and Panel (d) show the time series of local move rate by race and education attainment. All individuals regardless of their race and education have experienced declines in within county move rate at similar magnitudes over the three decades. It implies that any racial- and education-related factors probably are not the main drivers of the decline.

Figure 5 shows that the decline in local residential mobility is not exclusive to a few specific groups, but instead a universal experience for people regardless of their age, income, race, and education. However, it is not immediately obvious which group contributes most to the aggregate decline. Figure 6 shows the contribution to the aggregate local move rate by different population subgroups.

Panel (a) divides the total within county move rate into eight age groups. Individuals younger than age 35 are the major participants of local moves and contribute most to the aggregate decline over the past thirty-five years. Although other age groups also experience declines in local move rate, their contribution to the aggregate time trend is relatively constant. Panel (b) shows that individuals with annual household income less than $50,000 consist of over 70 percent of all local moves and they also contribute almost all the aggregate decline in local moves. Panel (c) shows that the majority of local moves are made by the White population and they contribute most to the aggregate decline. Panel (d) shows that individuals without a college degree are responsible for over 50 percent of all local moves in the 1980s and drive the long-term decline, while the contribution of individuals with a college degree remains roughly the same. Figure 5 implies that any significant factors driving the long-run decline should affect the younger, low-income, and less educated individuals more.

To examine the extent to which the sociodemographic factors can explain the long-term decline in local moves, I adjust the local move rate by the sociodemographic trends in the population using a pooled regression:

$$Move_{it} = \alpha + \beta X_{it} + \delta_t + \varepsilon_{it},$$

where $Move_{it}$ is an indicator variable of whether individual i have moved within county in year t, α is the constant term, X_{it} represents a vector of controls, and δ_t are year fixed effects. Figure 7 plots the year fixed effects δ_t using different controls and normalizes the year fixed effect in 1986 to be zero. The year fixed effects are equivalent to residual annual average of local move rate unexplained by the control variables relative to the base year.

The green line shows the year fixed effects without any controls. It shows that the within county move rate has declined by about 6 percentage points between 1986 and 2020, consistent with the absolute local move rate change shown in Figure 2. Then I add various controls to quantify how much population compositional changes can explain the long-run decline in within county moves. Adjusting for changes in housing tenure, age, marital status, income, and race gives flatter lines, indicating a slightly smaller decline in within county moves, but they jointly can explain at most one-fourth of the entire decline in local moves. The results are consistent with the fact that all population subgroups have experienced different degrees of declines in local moves over the years (shown in Figure 5 and Figure 6). It implies that there exist factors other than shifts in sociodemographic trends driving the long-term decline in local moves.

3.3 Role of housing market conditions on local moves

Given that a large share of local moves are housing-related moves (shown in Figure 4), I explore how housing costs and housing vacancies affect local moves using the ACS from 2005 and 2019. Figure 8 shows the binscatter plots of within PUMA move rate on median housing costs and housing vacancies at the PUMA-year level for homeowners and renters separately. I control for age, household income, marriage rate, college share, and metropolitan status when constructing the binscatter plots.

Panel (a) and (b) show that there is an inverse relationship between median monthly housing cost and within PUMA move rate for both homeowners and renters. For example, renter households living in PUMAs and years with monthly median rent over \$1000 are about 10 percent less likely to make local moves compared to areas with median monthly rent at \$500. Panel (c) and (d) show that households are less likely to move in a tight housing market. When the housing vacancy rate is low, there are fewer houses available on the market and thus making housing searches more difficult. These two empirical patterns provide suggestive evidence that higher housing costs and tight housing markets discourage local moves.

4 Marriage shocks and local mobility

In this section, I test the hypothesis that fewer moving shocks contribute to the long-run decline in local moves. In particular, I examine the causal relationship of marriage shocks on the probability of making local residential moves in the same year. I define marriage shock as the event when an individual switches from single to married. In other words, the variable of interest is the change in marital status, not just the level.

To show some motivating evidence, Panel (a) of Figure 9 plots the average marriage rate over the past four decades by age group. The average marriage rate has dropped from 63.34

percent to 52.86 percent from 1980 to 2020. Younger individuals below age 40 experience the largest drop in marriage rate, about two times larger than the average decrease. However, the marriage rate is not equivalent to the frequency of marriage shocks because the marriage rate measures the stock of people who are married and it does not necessarily reflect the share of people who experience changes in marital status. Panel (b) shows the marriage rate over the lifecycle by birth cohort. The difference in marriage rates between two ages can be interpreted as the share of population experiencing marriage shocks although it is not entirely accurate because it also captures divorce shocks. Nevertheless, it shows that recent cohorts delay marriage until later years and are less likely to experience marriage shocks in early adult life.

The descriptive evidence suggests that the decline in marriage shocks for younger individuals might explain the lower move rate over the years. If changes in marriage shocks are the dominating force of the decline in local moves, it implies that individuals voluntarily choose not to move due to a lower benefit and thus incurs no welfare losses. I test this mechanism formally in the following section.

4.1 Construct pseudo-panel data

To examine how marriage shocks affect local moves, ideally I would regress the indicator of local moves on changes in marital status at the individual level:

$$Move_{ict} = \beta \cdot MarriageShock_{ict} + \delta X_{ict} + \gamma W_{ct} + \varepsilon_{ict},$$

where $Move_{ict}$ is an indicator variable measuring whether an individual makes a local move or not, $MarriageShock_{ict}$ is an indicator variable of whether the individual experiences a change in marital status between $t-1$ and t or not, X_{ict} is vector of individual characteristics, and W_{ct} represent place-based factors that affect local moves. The main coefficient of interest is on β, which measures the propensity of local moves following the marriage shock in the same year. The construction of $MarriageShock_{ict}$ requires knowledge about the marital status in the current and previous periods. However, the cross-sectional CPS data only report the level of marital status in the current period although the change is what is relevant. To circumvent the issue, I construct pseudo-panel data from the CPS by grouping individuals with similar characteristics together and track changes at the group level.

The construction of pseudo-panel data from cross-sectional data has been explored in Deaton (1985); Moffitt (1993); Verbeek (2008). The idea of the pseudo-panel data is to group similar individuals into mutually exclusive cells and track changes for each group over time. The assumption is that households within the same group share similar move patterns and characteristics, so the group average is representative of the group population. The validity of the pseudo-panel approach largely depends on the grouping method. Any grouping method needs to make sure that each individual belongs to only one group for all periods, so any grouping variables need to be time-invariant. Examples of time-invariant variables include birth year, race, gender, and education for adults who have completed formal education. Moreover, individuals within the same group should have a high degree of homogeneity to minimize sampling errors.

My primary grouping method consists of 176 household groups defined by birth cohort, gender, race, and education. I use 11 birth cohorts using five-year intervals from 1940 to 1994, 2 genders of male and female, 4 racial groups of Black, White, Asian, and other, and 2 education levels of with and without college. All household groups are constructed at the state and year level. I exclude individuals under age 22 in my sample because they may not have finished formal education. Table 1 reports the summary statistics of the number of observations and standard deviation within each household group. The standard deviations are not all close to zero but are acceptable. The average cell size is about 31 and the maximum cell size is less than 300. On one hand, finer cells defined by more grouping variables can increase within-group homogeneity, but the cell size will become even smaller and potentially captures more noises. On the other hand, coarse cells may group heterogeneous individuals together and enlarge sampling errors. To address the sampling issue, I run robustness checks using alternative grouping variables in the Appendix.

After constructing the pseudo-panel data, I calculate the share of married individuals within each household group g in city c and year t, then I use changes in the variable to proxy for marriage shocks. I estimate the relationship between marriage shocks and local move rate at the group level:

$$Move_{gct} = \beta \cdot MarriageShock_{gct} + \delta X_{gct} + \gamma W_{ct} + \alpha_g + \varepsilon_{gct},$$

where $Move_{gct}$ is the average local move rate for household group g in city c and year t, $MarriageShock_{gct}$ measures the changes in the share of individuals who are married in each group g between two consecutive years $MarriageShock_{gct} = Married_{gct} - Married_{gct-1}$, X_{gct} and W_{ct} are two sets of controls at the group and state level, and α_g is an time invariant group specific fixed effect.

There are two main identification threats to the OLS estimates. First, the group-level variables contain two types of measurement errors. The group mean may fail to capture the true mean if the group contains few observations. In this case, the group mean can be driven by outliers instead of being representative of the whole group. Given that the average sample size of household groups is only 31 (shown in Table 1), the sampling error can be big.

To alleviate some of the concerns, I exclude small cells that have less than 10 observations and I also experiment with other alternative grouping variables to run robustness checks in Appendix. The second type of measurement error is that changes at the group level only capture the net change in marriage rate. The problem is that marriages can be dissolved, so the net change in marriage rate captures both marriage shocks and divorce shocks. For example, if the marriage rate and divorce rate both increase by 10 percent, the group-level change in marriage rate is zero on net although 10 percent of individuals in the group experience marriage shocks. So I also use the change in the share of individuals who are never married $\Delta NeverMarried_{gct}$ to measure marriage shocks since the change from never married to married is not reversible.

The second threat to identification is omitted variable bias. One example of omitted variables is moving and transaction costs. A higher moving and transaction cost is associated with a lower propensity to move and newly married couples typically face higher moving and transaction costs. First, newly married individuals are more likely to become homeowners than single individuals, and homeownership typically is associated with much higher transaction costs, such as realtor's fee, search costs, and mortgage-related costs. Second, they need to pay for additional negotiation and coordination costs on top of regular moving costs they need to pay if they were single. Thus, the unobserved moving and transaction costs cause the OLS coefficient to underestimate the true coefficient.

4.2 Instrument

I use an instrumental variable approach to address the endogeneity issues of $MarriageShock_{gct}$. The instrument needs to satisfy both the relevance and exclusion restriction conditions. First, the instrument needs to be correlated with $MarriageShock_{gct}$. Suppose $MarriageShock_{gct}$ is positive, the instrument should explain why there is an increase in newly married individuals in group g. Second, the instrument should only affect household local move decisions through these lifecycle shocks. In other words, the instrument should be uncorrelated with the structural error term ε_{gct} and omitted variables.

The instrument uses variations in the size of local potential marriage market to identify the causal relationship between marriage shocks and local moves. Specifically, I match each group g, which is defined by gender, birth cohort of five-year intervals, education, and race, to group k that is of the same birth cohort and education but is of the opposite gender. Then I measure the share of not married individuals in group k to estimate the size of marriage market for group g. Since the endogenous variable is whether more individuals in group g in city c experience marriage shocks in year t, I use the lagged share of single individuals in group k living in c in previous years $t-j$, where $j = \{1, 2, 3, 4\}$. The underlying assumptions are that individuals are more likely to find good matches in thicker marriage market and only the local marriage market size during the search process matters for marriage decisions.

The instrument exploits three mechanisms about marriage. First, it takes two individuals of opposite genders to form a heterosexual marital relationship. Second, the assortative matching theory suggests that individuals tend to marry people who share similar backgrounds. Third, a thicker marriage market enables individuals to find good matches quickly and thus make them more likely to get married. Here is a simple example to demonstrate the intuition of the instrument. It is more likely for a college-educated woman born between 1985-1989 to get married in 2010 if a large share of college-educated men of similar ages in the same city is not married before 2010.

The first stage equation is:

$$MarriageShock_{gct} = \rho \cdot Single_{k,c,t-j} + \pi X_{gct} + \theta W_{ct} + \alpha_g + e_{gct},$$

where groups g and k include individuals with the same education and in the same birth cohort but of the opposite gender, and $Single_{k,c,t-j}$ measures the share of individuals who are not married in group k in c in year $t-j$, $j = \{1, 2, 3, 4\}$. Figure 10 shows the binscatter plots of the first stage of the instrument $Single_{k,c,t-1}$ using two measures of marriage shocks. Panel (a) uses changes in marriage rate to proxy for marriage shocks and Panel (b) uses changes in the share of never married individuals. All binscatter plots control for group-level mean age, income, homeownership rate, changes in the number of children, income growth, and state fixed effects. The binscatter plots provide evidence that the relevance condition is satisfied. Panel (a) displays a positive relationship between changes in marriage rate and a large local marriage market. Panel (b) shows a negative correlation between changes in share of never married individuals and local marriage market size. It is consistent with the finding in Panel (a) because a decrease in the share of never married individuals indicates more individuals in the group get married this year. Figure A.1 and Figure A.2 show the first stage binscatter plots of instruments $Single_{k,c,t-j}$, where $j \in \{1, 2, 3, 4\}$. The correlation persists although it becomes noisier as the number of lagged years increases. Moreover, the joint F-statistics of the first stage are all over 30, implying that it does not suffer from a weak instrument problem. In summary, the evidence shows that marriage shocks and potential marriage market size are strongly correlated.

The instrument also satisfies the exclusion restriction if the variation in potential marriage market size is independent of the unobserved factors that affect local moves. Whether the individual faces a large or small marriage market in previous year has no meaningful implications on whether the individual dislikes her home or not, nor does it have any direct impact on the moving costs. It is reasonable to claim that the local mobility decisions are not directly affected by marriage market size except through its impact on the probability of marriage shocks.

4.3 Results

Table 2 reports the OLS and IV results on marriage shocks. I use two measures to proxy marriage shocks. Panel A reports the result using the net change in marriage rate $\Delta Married_{gct}$, and Panel B reports the result using the change in the share of never married $\Delta NeverMarried_{gct}$.

I start my discussion from the results in Panel A. The OLS coefficient shown in column (1) is only 0.004. It indicates that the relationship between marriage shocks and the propensity to make local moves in the same year is positive but close to zero. To correct the downward bias of the OLS coefficient, columns (2) to (4) report the IV estimates using different sets of instruments. The joint F-statistics in the first stage are all larger than 40, suggesting the instrument is not weak. All IV coefficients are much larger than the OLS coefficients and are statistically significant. The IV coefficients using different lagged measures of single rate of the opposite gender are similar in magnitudes, ranging from 0.145 to 0.17. It suggests that a 10 percentage points increase in marriage rate is associated with 1.45 to 1.7 percentage points increase in local move rate.

The problem with the change in marriage rate is that it is a noisy measure of true marriage shocks because it also captures divorce shocks. Therefore, I also use the change in the share of individuals who are never married to proxy marriage shocks. The advantage of this measure is that the change from never married to married is unidirectional and free from divorce shocks. Different from changes in marriage rate, a decrease in the share of never married indicates marriage shocks, so I expect the coefficients on $\Delta NeverMarried_{gct}$ to have the opposite signs to the coefficients on $\Delta Married_{gct}$.

The results in Panel B are consistent with the expectations. The OLS estimate in column (1) is negative and small in magnitude, indicating an attenuation bias of the OLS coefficient. The IV estimates in columns (2) to (4) correct the bias and all give statistically significant coefficients that are negative and much larger in magnitudes. Moreover, the magnitudes of the IV coefficients are similar to the IV coefficients in Panel A. The estimates suggest that a 10 percentage points decrease in the share of never married is associated with 1.71 to 2.06 percentage points increase in local move rate. The results in Panel B support the qualitative conclusion from Panel A although the quantitative magnitudes are slightly larger in Panel

B.

As discussed in the Section 4.1, the validity of the pseudo-panel approach depends on the grouping method. Table A.1 runs robustness checks using alternative group definitions. Columns (1) to (4) report the IV results using birth cohort, gender, and race at the state and year level to group individuals. Columns (5) to (8) uses birth cohort, gender, and education to define groups. These two sets of grouping relax grouping criteria and increases the average cell size, but the measurement error problem can be exacerbated if individuals within each group become less homogeneous and the group means capture more noises. Consistent with the expectation, Columns (1) to (4) has the largest average cell size and the IV estimates are much larger compared to the preferred estimates in Table 2. The average cell size in columns (5) to (8) is slightly larger than the primary grouping method, and the IV estimates are of similar magnitudes although they have lower power. Nevertheless, all the coefficients in Table A.1 have the same signs as the preferred estimates in Table 2. It confirms that marriage shock is a key determinant of local move decisions.

To put it in perspective, I conduct back-of-envelope calculations to estimate how much of the decline in local move rate can be explained by changes in marriage rate. The IV estimates are only valid for the compliant population with the instrument, that is, those who will only get married if facing a large local marriage market defined by the same birth cohort and education. I recognize that it is a strong assumption. Although most people marry someone with similar education and ages, there exist many married couples with large age differences and different education. Also, some newly married couples have long-distance relationships before marriage, which is not captured by the local marriage market size.

My back-of-envelope calculation assumes that everyone is a complier and these violations do not occur. It also assumes that changes in marriage rate equal to changes in marriage shocks. The marriage rate of individuals under age 40 has dropped by about 16.46 percentage points between 1986 and 2020, and the within county move rate has dropped by 7.51 percentage points. Using estimates from Panel A, it translates into 2.39 to 2.80 percentage points decrease in local move rate and accounts for about 32 to 37 percent of the entire decline in local move rate. Similarly, the share of never married individuals has increased by 21.46 percentage points from 36.45 percent in 1986 to 57.91 percent in 2020. The estimates from Panel B suggest that the declining marriage shocks cause the local move rate to decrease by about 3.67 to 4.42 percentage points and explain 48 to 58 percent of the entire decline in within county moves. Given that the extrapolation replies on many strong assumptions, I use the lower bound of the back-of-envelope calculations to quantify the magnitude of the causal impact. The conservative estimate says that the fewer marriage shocks since 1986 contributes to at least 30 percent of the aggregate decline in local move rates.

5 Housing cost and local mobility

In this section, I explore how housing market conditions affect local moves. In particular, I measure the causal impact of marginal housing cost difference on local moves. The null hypothesis is that households are financially constrained to make the move if the house that improves the housing match quality costs significantly more than the current house. Another way to understand it is that the new house the household can afford offers little quality improvement, so there is little need to make the move. The main variable of interest is the difference in the marginal housing cost before and after the move after adjusting for the housing quality difference.

I use the ACS to conduct the analysis on housing costs because the CPS contains no information about housing characteristics. The ACS has a much larger sample size but it is only available since 2005, so I only evaluate the causal impact of rising housing costs over the past 15 years.

5.1 Estimation

Since I only observe the current housing cost of households, I use two methods to estimate the counterfactual housing cost if the household were to make a different move decision. Then I take the difference between the observed actual housing cost and the counterfactual housing cost as the marginal housing cost difference. I estimate the counterfactual cost for renters and homeowners separately.

The first method uses the distribution of housing costs of households with similar demographic and socioeconomic characteristics. The assumption is that similar households have similar housing preferences and pay similar housing costs. The drawback of the approach is that it is agnostic to changes in housing characteristics before and after moves. I divide households into groups defined by age, income, and family structure. I use six age groups, 18-25, 25-34, 35-44, 45-54, 55-64, 65+, five income groups, <\$25k, \$25-50k, \$50-75k, \$75-100k, >\$100k, and two family structures, married and not married. There are 120 groups in total. For each household, I use the housing cost of other households within the same household group but who make a different moving decision to estimate the counterfactual housing cost. To be more specific, to estimate the counterfactual housing cost for a stayer household if they were to move, I use the average housing cost of mover households in the same household group or take a random draw from the distribution, and vice versa for movers.

The second method takes account of changes in housing characteristics. It estimates the additional housing cost the household needs to pay for an additional bedroom or room. More specifically, I match each household to other households in the same household group, then I use the average housing cost of those households who live in houses with one additional bedroom or room but who have made a different moving decision to estimate the counterfactual housing cost of the household if they were to upgrade their home. The underlying assumption is that households would increase the size of home upon moving. One interpretation is that the variable represents the marginal housing cost of housing upgrade per move.

After predicting the counterfactual housing cost, I construct the housing cost difference as

$$MP_{ict} = \begin{cases} \ln \hat{p}_{move(i)ct} - \ln p_{ict} & \text{if } move_{ict} = 0 \\ \ln p_{ict} - \ln \hat{p}_{stay(i)ct} & \text{if } move_{ict} = 1 \end{cases},$$

where $\ln p_{ict}$ is the log of current housing cost of household i in city c and year t, and $\ln \hat{p}_{move(i)ct}$ and $\ln \hat{p}_{stay(i)ct}$ are the estimated counterfactual housing cost if the household were to make a different moving decision. The housing cost difference measures the percentage change in housing cost before and after moves.

Since the CPS does not include any information about housing cost and housing characteristics, I use the ACS to conduct the analysis. The baseline OLS specification is

$$Move_{ict} = \beta_0 + \beta_1 MP_{ict} + \beta_2 X_{ict} + \beta_3 Z_{ct} + \gamma_c + \gamma_t + \gamma_{g(i)} + u_{ict},$$

where $Move_{ict}$ is an indicator of whether household i has moved within MSA c in year t, MP_{ict} is the estimated housing cost difference, X_{ict} is a set of household characteristics, including household income, household head age, gender, education, marital status, family size, presence of children, and Z_{ct} is a set of city controls, including housing vacancy rate, population, change in population, and average housing cost burden.

The OLS specification suffers from endogeneity. First, the estimated housing cost difference is likely to contain measurement errors because the estimation process does not capture all the relevant housing characteristics and household preferences. The estimates can also include a lot of noises if some household groups contain few observations. Omitted variable bias is another concern. There are many variables relevant to household local move decisions that I do not observe in my data, including household match quality with the current and alternative residences, housing preferences, transaction cost, and the number of houses available in relevant housing market segments. Some of these variables might be correlated with the endogenous variable and create biased OLS estimates. For example, I do not observe changes in family characteristics, which is shown to increase the likelihood of local moves in Section 4. Family life events are also associated with higher housing cost differences because they are typically accompanied by housing upgrades, so the true coefficient would be larger in magnitude. Reverse causality is also possible. If there are a lot of local moving activities going on, the local housing demand surge will push up housing prices.

I use an instrumental variable approach to address the potential endogeneity. I use the group-city-level housing price growth $\Delta \ln HP_{g(i),c,t}$ to instrument for MP_{ict}. To construct the instrument, I estimate the housing price growth or rent growth in city c between $t-1$ and t for the group $g(i)$ the household i belongs to. The instrument uses variations in housing price growths across cities, years, and housing market segments to identify the causal relationship between housing cost difference and local move rate.

The instrument satisfies the relevance condition because it generates differential variations in the marginal price of housing upgrades. Here is a simple example to provide intuition. Consider two cities with different house cost growth, San Francisco and in Cleveland. On average, San Francisco has had higher house price growth than Cleveland. The source of housing price growth is mostly due to housing demand and land values instead of changes in housing supply and housing quality. It would be more expensive to pay for house upgrades in San Francisco than Cleveland. The housing price growth is much more likely to be fully capitalized in the actual new housing cost of recent movers than stayers.

On the other hand, the instrument satisfies the exclusion restriction if changes in housing cost growth are orthogonal to unobserved determinants of household local move decisions. As discussed above, the main omitted variable is family life events. It is plausible that whether an individual chooses to get married or have children has limited correlation with local housing cost growth.

5.2 Results

I conduct the analysis for homeowners and renters separately. Table 3 reports the results of the first method that uses the distribution of housing cost. Panel A reports the OLS and IV estimates for renters. All IV coefficients are statistically significant negative and have much larger magnitudes than the OLS estimates. Moreover, the IV estimates using the group average and random draw have similar magnitudes, ranging from -0.532 to -0.673. They indicate that a 10 percent increase in housing cost difference is associated with about 5 to 6 percentage points decline in local move rate. Panel B reports the results for homeowners using housing value differences. The OLS coefficients are very small and close to zero. The IV coefficients are unstable without and with fixed effects and even flip signs. I interpret these results as evidence that homeowners are less affected by house price differences if they are affected in any way. Housing not only serves a consumption purpose for homeowners but is also a crucial non-financial asset. Homeowners can extract home equity from selling their previous house and pay for the new house to cover the price difference. So homeowners can hedge against housing cost growth induced by moves, while renters do not have this option. On the other hand, house value is not equivalent to the actual housing cost of homeowners because they have financing options to choose their monthly payments, and different mortgage terms can result in different housing costs for the same house.

Table 4 reports the results of the second method that uses the marginal price of house upgrades. All coefficients are negative and most are statistically significant, but the magnitudes are smaller than the first method in Table 3. Similarly, renters are more sensitive to the difference in housing costs compared to homeowners. For renters, the estimates suggest that a 10 percent increase in the difference between current and alternative rent is associated with about a 1.8 percentage point decrease in local move rate. For homeowners, the OLS and IV coefficients are all negative and statistically significant, but the magnitudes are very small and close to zero. The results suggest that rent growth makes it more costly for renters to make local moves, but homeowners are largely immune to housing value differences between houses after 2010.

All methods yield a consistent qualitative conclusion that renters are less likely to make local moves if outside rent grows substantially, while homeowners can hedge against housing price growths. However, the magnitudes of IV coefficients differ by quite a lot. I use the estimates from the second method to conduct a conservative back-of-envelope calculation. The average rent has increased about 16 percent during the past 15 years. It translates into 2.8 percentage points decline in local move rate, accounting for about 80 percent of the entire decline in local move rate for renters from 2005 to 2019. Again, the back-of-envelope should be taken with caution because the estimate relies on many assumptions. First, it assumes that every move is an upgrade move that adds one more room or bedrooms. It is probably true for younger households, but not for older households who are more likely to choose downsize. Second, the estimate is only valid for households whose housing cost difference follows the same trajectory of the average housing price growth at the group-city-year level. Nevertheless, the back-of-envelope calculation suggests that renters are likely to suffer welfare losses from rising rent because they are stuck in place.

6 Conclusion

In this paper, I first document the evolution of local move rate using various measures in the past 35 years. I find that every population subgroup has experienced different degrees of declines in the propensity of local moves, and demographic shifts can explain at most 25 percent of the entire decline.

Next, I test two theories to explain the decline. The first theory examines the long-run causal relationship between the decline in marriage shocks and the decline in local move rate. I find that individuals are much more likely to make within county moves following marriage shocks in the same year, but individuals are less likely to experience marriage shocks in recent years conditional on age. The back-of-envelope calculation suggests that the decline in marriage shocks since 1986 can explain at least 30 percent of the decline in within county move rate. It implies that these households voluntarily choose not to move because there are fewer benefits associated with local moves.

The second theory examines how marginal housing cost before and after move affects within MSA moves in the past 15 years. I find that the local move decisions of renters are very sensitive to rent growth, while homeowners are largely immune to housing price differences. The back-of-envelope calculation suggests that about 80 percent of the decline in local move rate of renters are attributed to higher rent. It implies that many renters are stuck in their current residences because they cannot afford housing upgrades.

References

Bleemer, Zachary, Meta Brown, Donghoon Lee, Katherine Strair, and Wilbert Van der Klaauw, "Echoes of rising tuition in studentsâ borrowing, educational attainment, and homeownership in post-recession America," *Journal of Urban Economics*, 2021, *122*, 103298.

Brown, H James and John F Kain, "Moving behavior of San Francisco households," *The Detroit prototype of the NBER urban simulation model*, 1972, *2.*

Brown, Lawrence A and Eric G Moore, "The intra-urban migration process: a perspective," *Geografiska Annaler: Series B, Human Geography*, 1970, *52* (1), 1–13.

Cattaneo, Matias D, Richard K Crump, Max H Farrell, and Yingjie Feng, "On binscatter," *arXiv preprint arXiv:1902.09608*, 2019.

Chan, Sewin, "Spatial lock-in: Do falling house prices constrain residential mobility?," *Journal of urban Economics*, 2001, *49* (3), 567–586.

Deaton, Angus, "Panel data from time series of cross-sections," *Journal of econometrics*, 1985, *30* (1-2), 109–126.

Ferreira, Fernando, Joseph Gyourko, and Joseph Tracy, "Housing busts and household mobility," *Journal of urban Economics*, 2010, *68* (1), 34–45.

Ganong, Peter and Daniel Shoag, "Why has regional income convergence in the US declined?," *Journal of Urban Economics*, 2017, *102*, 76–90.

Giannone, Elisa, Qi Li, Nuno Paixao, and Xinle Pang, "Unpacking Moving," Technical Report, Working Paper 2020.

Glaeser, Edward L, Hyunjin Kim, and Michael Luca, "Nowcasting gentrification: using yelp data to quantify neighborhood change," in "AEA Papers and Proceedings," Vol. 108 2018, pp. 77–82.

Hsieh, Chang-Tai and Enrico Moretti, "Housing constraints and spatial misallocation," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 1–39.

Johnson, Janna E and Morris M Kleiner, "Is occupational licensing a barrier to interstate migration?," *American Economic Journal: Economic Policy*, 2020, *12* (3), 347–73.

Jr, Alden Speare, "Home ownership, life cycle stage, and residential mobility," *Demography*, 1970, *7* (4), 449–458.

Kaplan, Greg and Sam Schulhofer-Wohl, "Understanding the long-run decline in interstate migration," *International Economic Review*, 2017, *58* (1), 57–94.

Maggio, Marco Di, Ankit Kalda, and Vincent Yao, "Second chance: Life without student debt," Technical Report, National Bureau of Economic Research 2019.

Moffitt, Robert, "Identification and estimation of dynamic models with a time series of repeated cross-sections," *Journal of Econometrics*, 1993, *59* (1-2), 99–123.

Molloy, Raven, Christopher L Smith, and Abigail K Wozniak, "Declining migration within the US: The role of the labor market," Technical Report, National Bureau of Economic Research 2014.

_ , _ , and Abigail Wozniak, "Internal migration in the United States," *Journal of Economic perspectives*, 2011, *25* (3), 173–96.

Ngai, L Rachel and Kevin D Sheedy, "The decision to move house and aggregate housing-market dynamics," *Journal of the European Economic Association*, 2020, *18* (5), 2487–2531.

_ and Silvana Tenreyro, "Hot and cold seasons in the housing market," *American Economic Review*, 2014, *104* (12), 3991–4026.

Notowidigdo, Matthew J, "The incidence of local labor demand shocks," *Journal of Labor Economics*, 2020, *38* (3), 687–725.

Rossi, Peter Henry, *Why families move: A study in the social psychology of urban residential mobility*, Free Press, 1955.

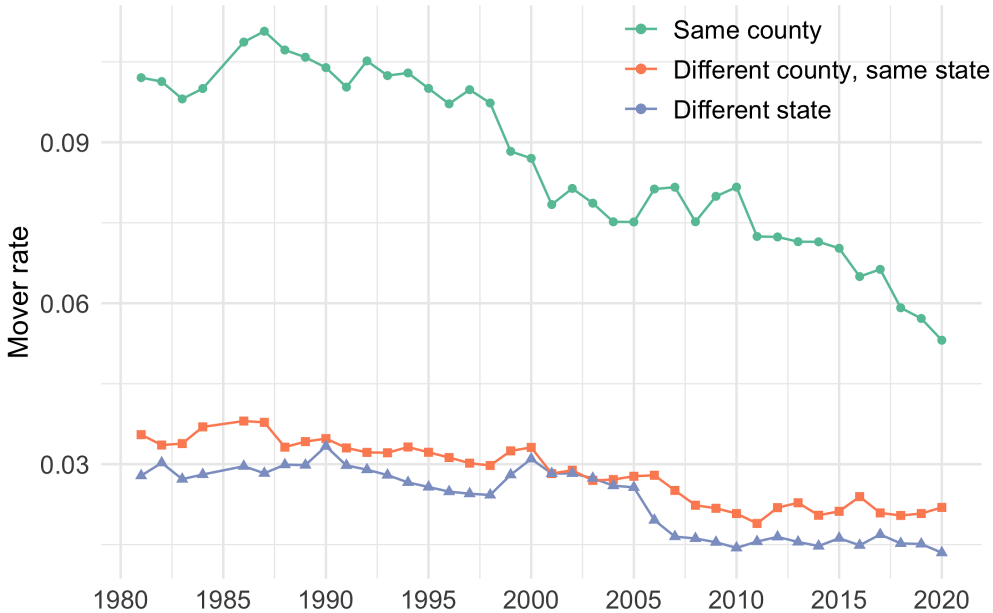
Speare, Alden, "Residential satisfaction as an intervening variable in residential mobility," *Demography*, 1974, *11* (2), 173–188.

Verbeek, Marno, "Pseudo-panels and repeated cross-sections," in "The econometrics of panel data," Springer, 2008, pp. 369–383.

Weinberg, Daniel H, "The determinants of intra-urban household mobility," *Regional Science and Urban Economics*, 1979, *9* (2-3), 219–246.

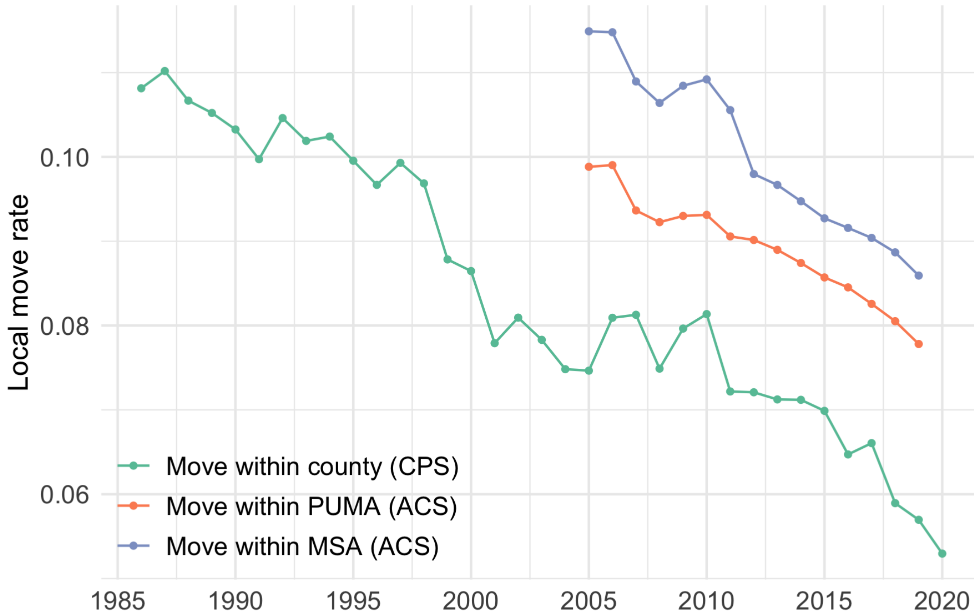
Wolpert, Julian, "Behavioral aspects of the decision to migrate," in "Papers of the Regional Science Association," Vol. 15 Springer 1965, pp. 159–169.

Figure 1: Move rate by distance of move



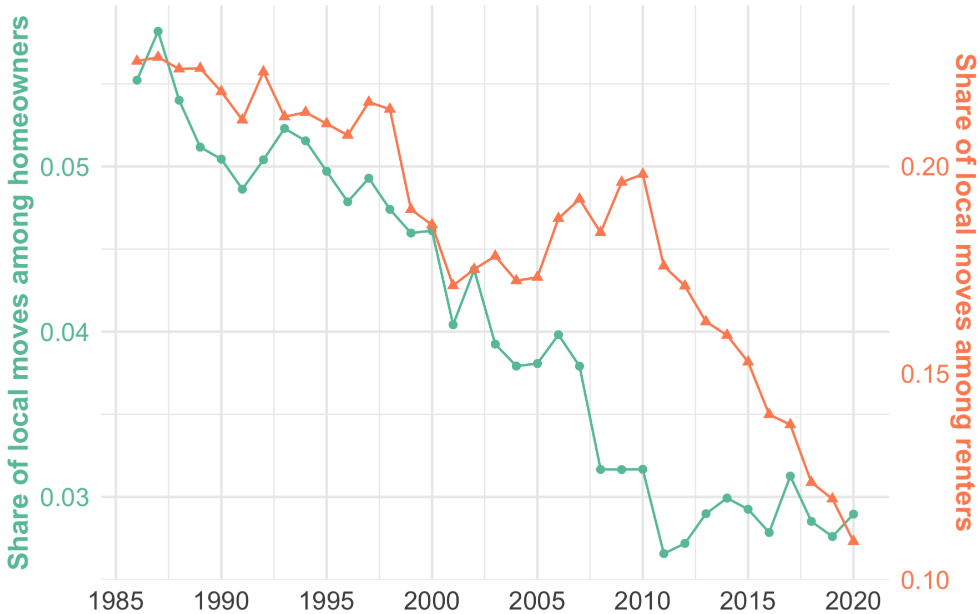
Note: This figure shows annual move rate by distance of move using the CPS from 1986 to 2020. The move rate is calculated by the number of adult movers divided by the total number of adults.

Figure 2: Local move rate in the U.S.



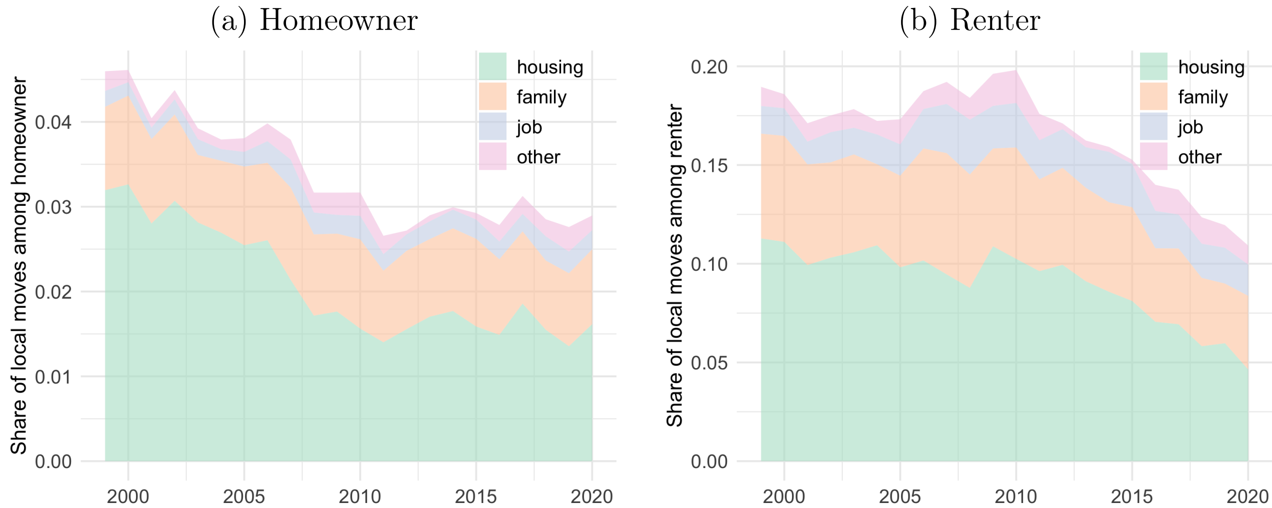
Note: The figure shows the local move rates using three different measures. The within county move rate is from the IPUMS CPS and the within PUMA move rate is from the IPUMS USA. The within MSA move rate is constructed using the MSAs in current and previous residences from the ACS.

Figure 3: Within county move rate by housing tenure status



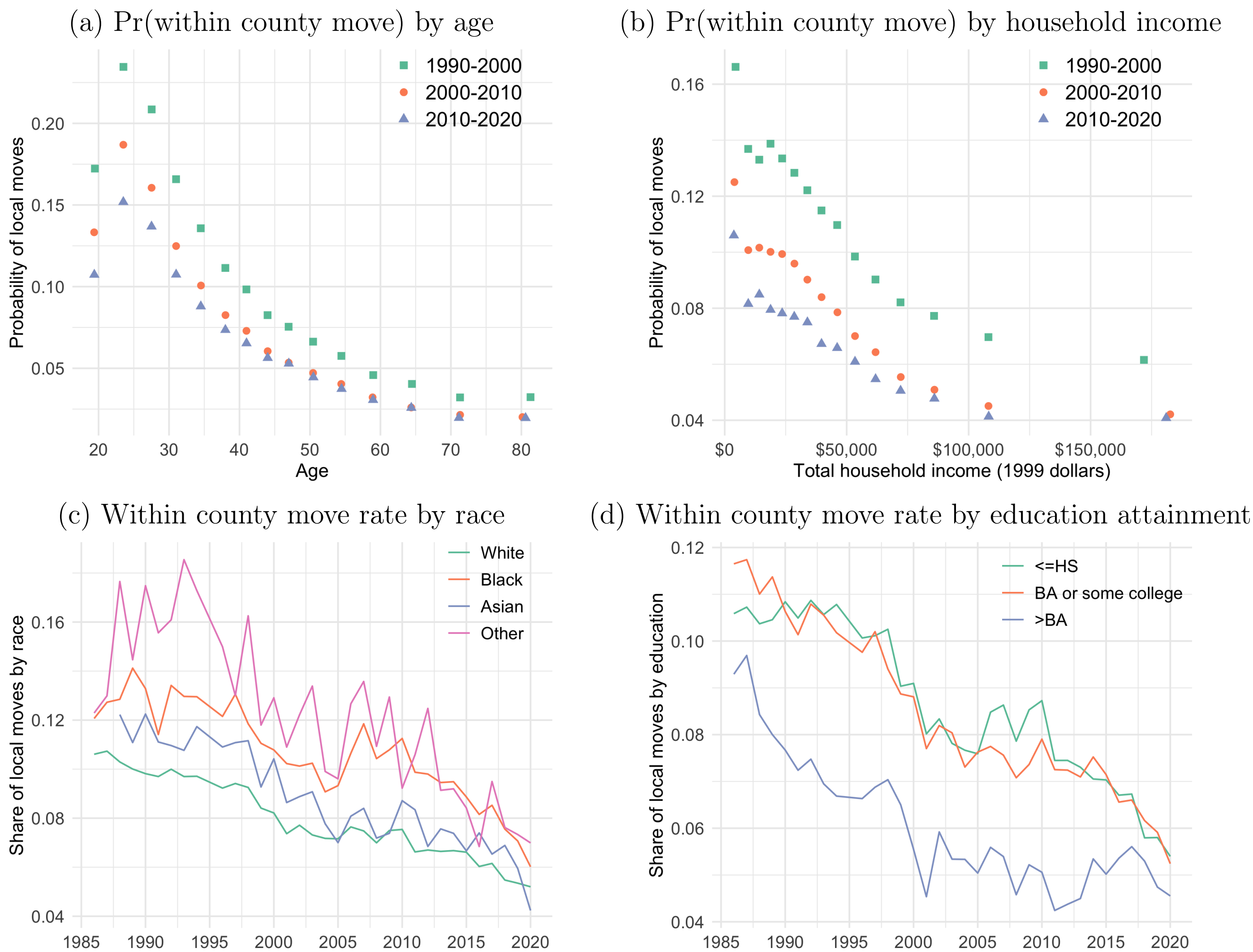
Note: This figure shows within county move rate for homeowners and renters separately using the CPS. The move rate of homeowners is shown in the green line and corresponds to the left y-axis, and the move rate of renters is shown in the orange line and corresponds to the right y-axis.

Figure 4: Within county moves by housing tenure and reason for moving



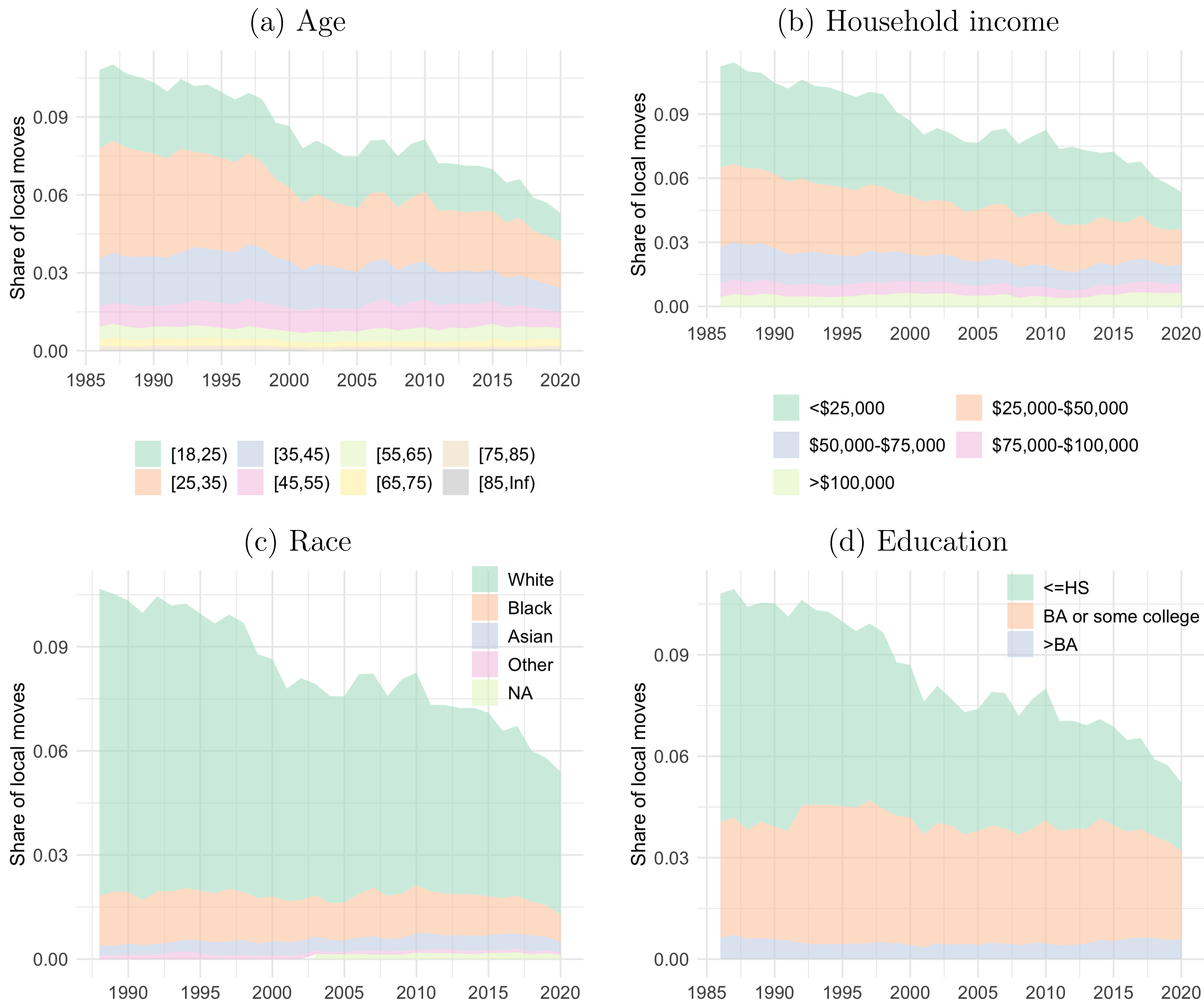
Note: This figure decomposes within county moves by reason for moving and housing tenure using the CPS. The CPS starts to survey this question from 1999, so data before 1999 is not available.

Figure 5: Within county move rate by population subgroup



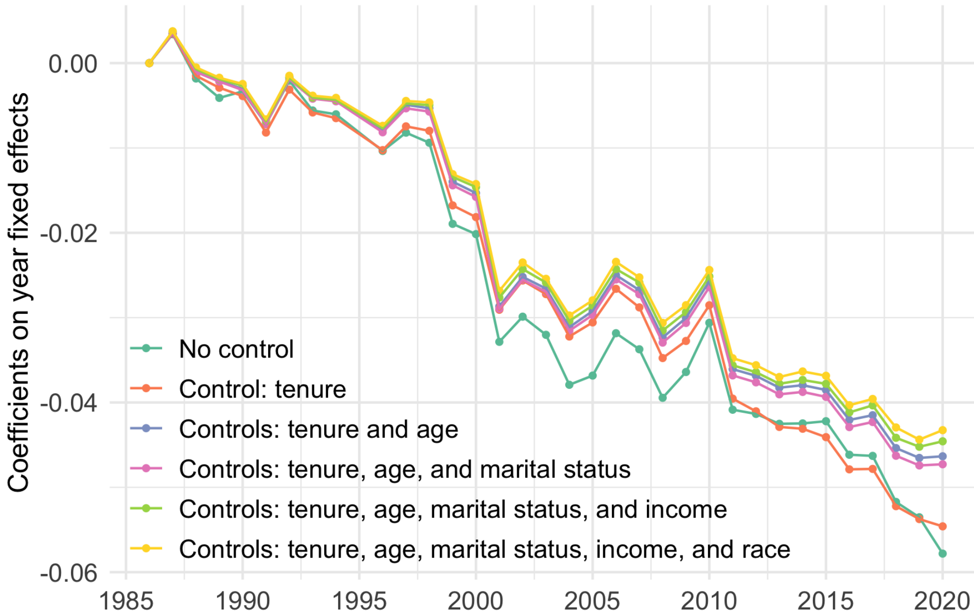
Note: The four panels use the CPS to document the local move rate for different sociodemographic groups. The first two panels use binscatter plots to decompose local move rate by age and income over three time periods, 1990-2000, 2000-2010, and 2010-2020. The last two panels show the time series of local move rate by race and education from 1986 to 2020.

Figure 6: Contribution of within county moves by population subgroups



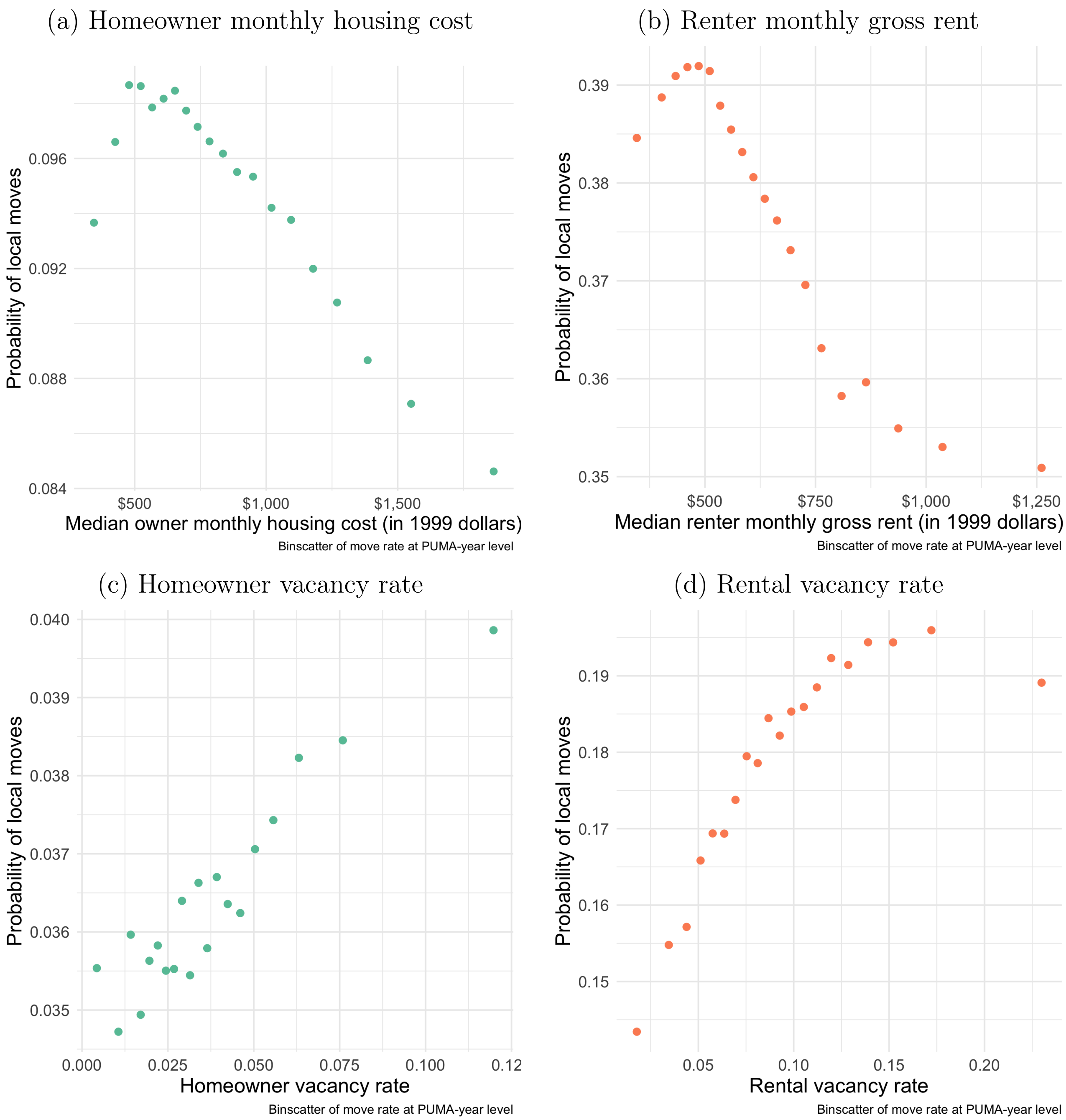
Note: The four panels use the CPS to show the contribution to the aggregate decline in local move rate by different sociodemographic groups.

Figure 7: Time series of year fixed effects δ_t



Note: This figure plots the year fixed effects from the pooled OLS regression using the CPS. It regresses within county move on a constant term, year fixed effects, and different control variables.

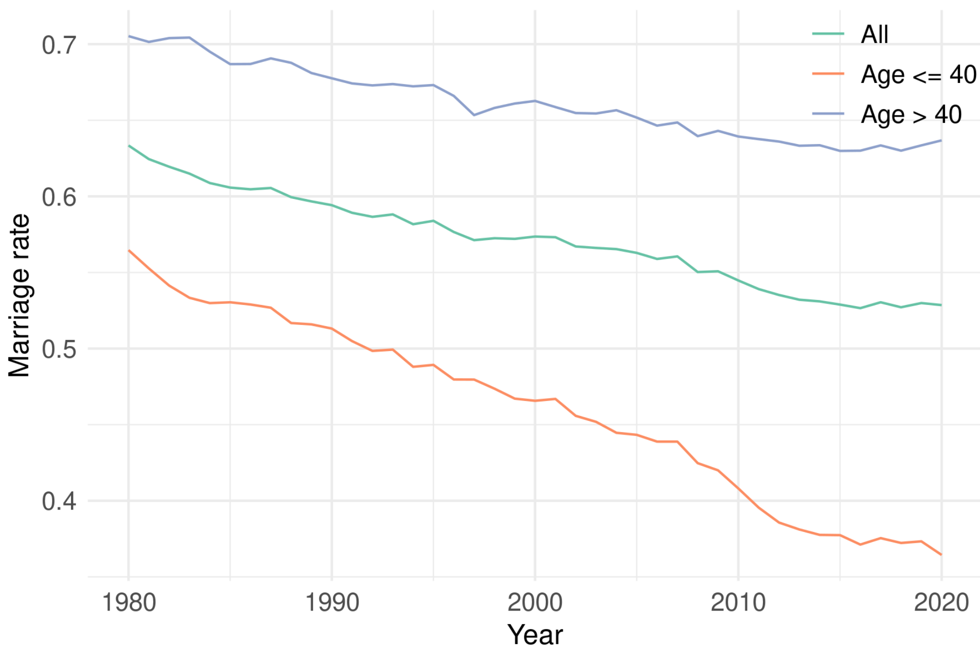
Figure 8: Within PUMA moves and housing market conditions



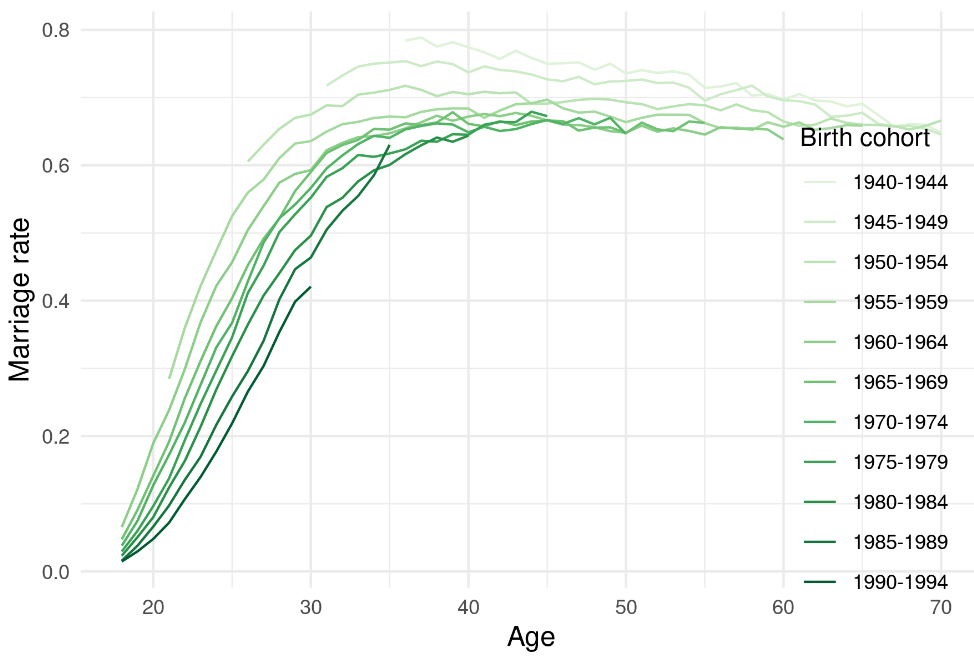
Note: This figure shows the relationship between within PUMA move and housing market conditions using the ACS from 2005 and 2019. Each dot in the binscatter plots represent data at the PUMA-year level. The variables measuring housing market conditions include monthly housing cost and housing vacancies. The owner monthly housing cost includes mortgage payment, property taxes, and other housing expenses. The renter monthly housing cost includes rent, utility bills, and other housing expenses. The vacancy rates are constructed for homeowners and renters separately. All plots include average age, college share, mean household income, marriage rate, and metropolitan status at the PUMA-year level as controls.

Figure 9: Marriage rate

(a) Marriage rate by age group and year

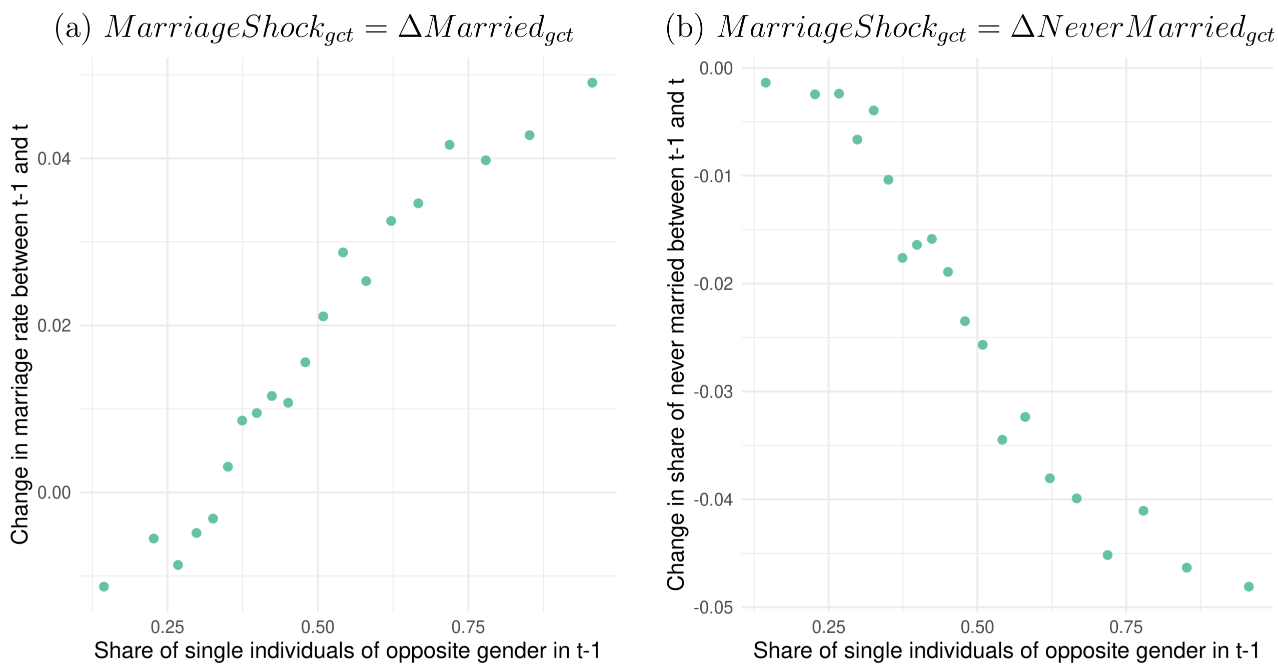


(b) Marriage rate over lifecycle by birth cohort



Note: This figure constructs the average marriage rate and marriage rate conditional on age for different birth cohorts from the CPS. The sample includes adults over age 18 not living in group quarters.

Figure 10: First stage of instrument, $MarriageShock_{gct}$ on $Single_{k,c,t-1}$



Note: This figure shows the first stage relationship between the instruments $Single_{k,c,t-1}$ and the endogenous variable $MarriageShock_{gct}$ after controlling for group-level mean age, income, education, changes in the number of children, income growth, homeownership rate, and state fixed effects. Panel (a) uses changes in marriage rate to proxy for marriage shock and Panel (b) uses changes in never married rate to proxy for marriage shock.

Table 1: Cell size and variability of individual groups

Group	Birth cohort, gender, race, education		
	$\bar{N}_{gct}$	$\min(N_{gct})$	$\max(N_{gct})$
Cell size	31	10	248
Cell variability	$S\bar{D}_{gct}$	$\min(SD_{gct})$	$\max(SD_{gct})$
Moved within county	0.24	0	0.50
Age	1.34	0	1.95
Never married	0.31	0	0.50
Married	0.44	0	0.50
Divorced	0.31	0	0.50

Note: This table shows the sizes of household group sand standard deviations of key variables within household groups in the pseudo-panel data.

Table 2: Life events and within county moves

Dependent variable: moved within county					
	OLS	IV: $Single_{kct-j}$			
	(1)	(2)	(3)	(4)	(5)
		$j = 1$	$j = \{1, 2\}$	$j = \{1, 2, 3\}$	$j = \{1, 2, 3, 4\}$
Panel A: $MarriageShock_{gct} = Married_{gct} - Married_{gct-1}$					
$\Delta Married_{gct}$	0.004	0.170**	0.164**	0.145**	0.155**
	(0.003)	(0.077)	(0.074)	(0.068)	(0.067)
Controls	Yes	Yes	Yes	Yes	Yes
HH group, year, state FEs	Yes	Yes	Yes	Yes	Yes
Observations	67,485	67,466	64,132	60,467	56,399
R^2	0.422	0.388	0.390	0.396	0.367
First stage F-stat		49.48	48.09	46.02	47.15
Panel B: $MarriageShock_{gct} = NeverMarried_{gct} - NeverMarried_{gct-1}$					
$\Delta NeverMarried_{gct}$	−0.010***	−0.206**	−0.196**	−0.171**	−0.183**
	(0.003)	(0.093)	(0.088)	(0.080)	(0.078)
Controls	Yes	Yes	Yes	Yes	Yes
HH group, year, state FEs	Yes	Yes	Yes	Yes	Yes
Observations	67,485	67,466	64,132	60,467	56,399
R^2	0.422	0.380	0.384	0.393	0.363
First stage F-stat		37.62	36.68	34.84	35.8

Note: This table displays the OLS and IV results of regression of within county move rate on marriage shocks at the individual group level. Marriage shocks are proxied by changes in the share of married and share of never married individuals within each group g. Individual groups g are defined by eleven birth cohorts, two education levels, two genders, and four races. Each group g living in c and year t is matched to group k that is of the same birth cohort and education but is of the opposite gender. Column (1) reports the OLS results and columns (2)-(5) report the IV results using different lagged measures of the share of single individuals of the opposite gender in group k in previous years $t - j$. The control variables include changes in the number of children between two years, changes in income, age, homeownership rate, income, education, age distribution at the state level, state-level housing price index, and state-level unemployment rate. The sample excludes people younger than age 22 to exclude those who have not finished formal education. Standard errors are clustered at the state level and are reported in parentheses. Levels of significance: *p<0.1; **p<0.05; ***p<0.01.

Table 3: Within MSA move and housing cost difference using distribution of housing cost

	Dependent variable: moved within MSA							
	Panel A: Renters							
	(a) Use group average				(b) Use random draw			
	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) OLS	(6) OLS	(7) IV	(8) IV
Rent diff	-0.077^{***}	-0.075^{***}	-0.637^{**}	-0.532^{**}	0.030^{***}	0.030^{***}	-0.673^{*}	-0.558^{*}
	(0.003)	(0.003)	(0.277)	(0.221)	(0.003)	(0.003)	(0.372)	(0.294)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FEs		Yes		Yes		Yes		Yes
Observations	3,163,254		3,163,116		3,163,254		3,163,116	
	Panel B: Homeowners							
	(a) Use group average				(b) Use random draw			
	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) OLS	(6) OLS	(7) IV	(8) IV
House value diff	-0.023^{***}	-0.023^{***}	-0.674	-0.019	0.003^{***}	0.003^{***}	0.494	-0.012
	(0.001)	(0.001)	(0.731)	(0.016)	(0.0005)	(0.0005)	(0.391)	(0.017)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FEs		Yes		Yes		Yes		Yes
Observations								

Note: This table displays the OLS and IV results of household-level likelihood of local move on housing cost difference estimated using the distribution of housing cost. The analysis use the ACS from 2005 to 2019. The control variables include household head age, gender, marital status, number of children, family size, income, city-level housing vacancy rate, city-level housing cost burden, city-level population and population change. The fixed effects include household group, city, and year fixed effects. All standard errors are clustered at the city level. Levels of significant: *p<0.1; **p<0.05; ***p<0.01.

Table 4: Within MSA move and housing cost difference using marginal cost of housing upgrade

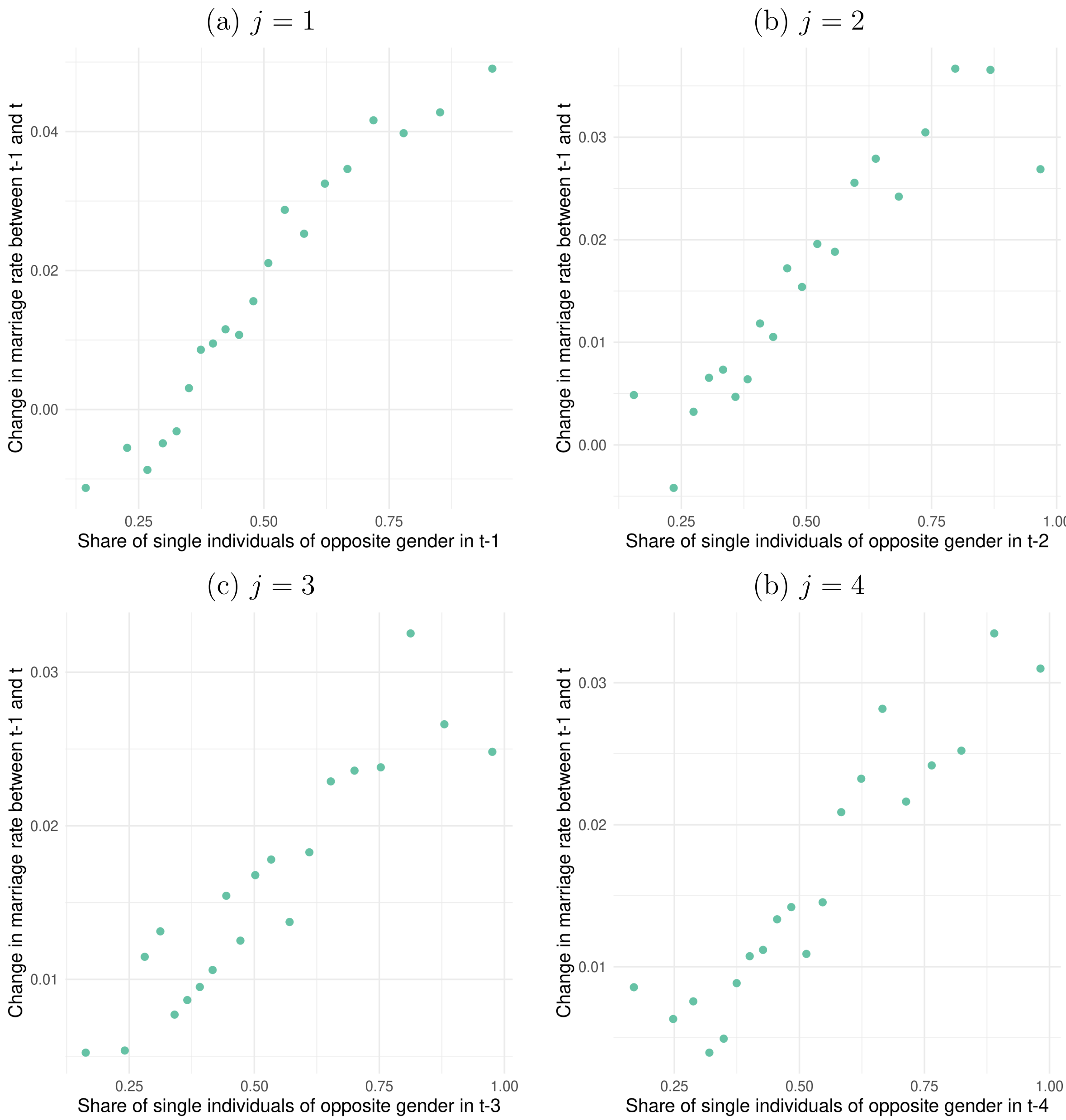
Dependent variable: moved within MSA

	(a) Marginal cost to get one more bedroom				(b) Marginal cost to get one more room			
Panel A: Renters								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IV	IV	OLS	OLS	IV	IV
Rent diff	-0.099^{***}	-0.097^{***}	-0.182^{***}	-0.179^{***}	-0.045^{***}	-0.044^{***}	-0.187^{***}	-0.184^{***}
	(0.014)	(0.014)	(0.017)	(0.017)	(0.007)	(0.006)	(0.015)	(0.014)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FEs		Yes		Yes		Yes		Yes
Observations	2,624,339		2,621,646		2,596,669		2,589,191	
Panel B: Homeowners								
	(a) Marginal cost to get one more bedroom				(b) Marginal cost to get one more room			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	IV	IV	OLS	OLS	IV	IV
House value diff	-0.037^{***}	-0.036^{***}	-0.015^{***}	-0.002	-0.018^{***}	-0.018^{***}	-0.022^{***}	-0.012^{***}
	(0.002)	(0.003)	(0.004)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FEs		Yes		Yes		Yes		Yes
Observations	4,937,629		4,934,459		5,071,059		5,066,231	

Note: This table displays the OLS and IV results of household-level likelihood of local move on housing cost difference using the marginal price of house upgrades. The analysis use the ACS from 2005 to 2019. The control variables include household head age, gender, marital status, number of children, family size, income, city-level housing vacancy rate, city-level housing cost burden, city-level population and population change. The fixed effects include household group, city, and year fixed effects. All standard errors are clustered at the city level. Levels of significant: *p<0.1; **p<0.05; ***p<0.01.

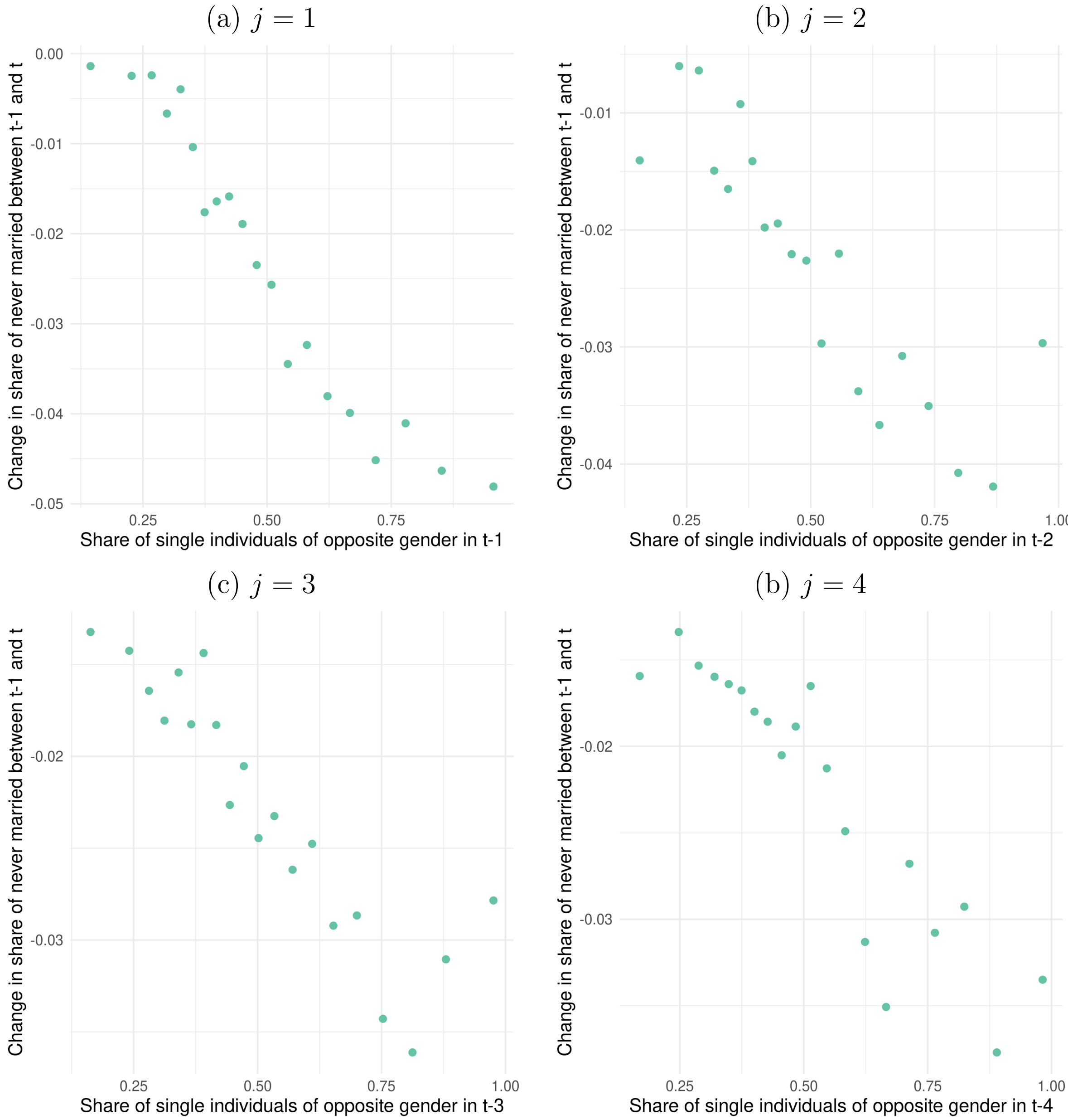
Appendix figures and tables

Figure A.1: First stage of instrument, $\Delta Married_{gct}$ on $Single_{k,c,t-j}$



Note: This figure shows the first stage relationship between the instruments $Single_{kct-j}$ and the endogenous variable $MarriageShock_{gct} = Married_{gct} - Married_{gct-1}$, where group g and k include individuals with the same education and in the same birth cohort but of the opposite gender and $j = \{1, 2, 3, 4\}$.

Figure A.2: First stage of instrument, $\Delta NeverMarried_{gct}$ on $Single_{k,c,t-j}$



Note: This figure shows the first stage relationship between the instruments $Single_{kct-j}$ and the endogenous variable $MarriageShock_{gct} = NeverMarried_{gct} - NeverMarried_{gct-1}$, where group g and k include individuals with the same education and in the same birth cohort but of the opposite gender and $j = \{1, 2, 3, 4\}$.

Table A.1: Within County Moves and Marriage Shocks, Robustness Checks Using Alternative Groups

	Dependent variable: moved within county. IV: $Single_{kct-j}$							
	(1) $j=1$	(2) $j=\{1,2\}$	(3) $j=\{1,2,3\}$	(4) $j=\{1,2,3,4\}$	(5) $j=1$	(6) $j=\{1,2\}$	(7) $j=\{1,2,3\}$	(8) $j=\{1,2,3,4\}$
$\Delta Married_{gct}$	0.437***	0.362***	0.233**	0.234***	0.173**	0.153*	0.111	0.118*
	(0.150)	(0.127)	(0.101)	(0.082)	(0.082)	(0.077)	(0.069)	(0.067)
$\Delta NeverMarried_{gct}$	−0.559***	−0.449***	−0.273**	−0.269***	−0.213**	−0.185*	−0.132	−0.142*
	(0.208)	(0.163)	(0.123)	(0.097)	(0.102)	(0.093)	(0.083)	(0.079)
Grouping variables	Birth cohort, gender, and race				Birth cohort, gender, and education			
Number of groups	88				44			
Average cell size	64				39			

Note: This table displays the IV results of regression of within county move rate on marriage shocks using alternative grouping variables. The first four columns use birth cohort, gender, and race to group individuals. The last four columns use birth cohort, gender, and education to group individuals. Levels of significance: *p<0.1; **p<0.05; ***p<0.01.

Table A.2: Within County Moves and Life Events, Heterogeneity Analysis

	Dependent variable: moved within county. Instrument $Single_{kct-j}, j \in \{1,2,3,4\}$					
	(1) Homeowner	(2) Renter	(3) Female	(4) Male	(5) No college	(6) College
$\Delta Married_{gct}$	0.040	0.407***	0.172**	0.097	0.241**	0.206***
	(0.056)	$(0,116)$	(0.078)	(0.081)	(0.113)	(0.074)
$\Delta NeverMarried_{gct}$	−0.050	−0.418***	−0.205**	−0.114	−0.290***	−0.237***
	(0.061)	(0.140)	(0.095)	(0.092)	(0.131)	(0.085)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
State & year FEs	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table displays the IV results of regression of within county move rate on marriage shocks for different population subgroups. Levels of significance: *p<0.1; **p<0.05; ***p<0.01.