Federal Reserve Bank of New York
Staff Reports

Population Aging, Migration Spillovers, and the Decline in Interstate Migration

Fatih Karahan
Serena Rhee

Staff Report No. 699
November 2014
Revised April 2017

Population Aging, Migration Spillovers, and the Decline in Interstate Migration
Fatih Karahan and Serena Rhee
Federal Reserve Bank of New York Staff Reports, no. 699
November 2014; revised April 2017
JEL classification: D83, J11, J24, J61, R12, R23

Abstract

We investigate the role of the aging of the U.S. population in the decline in interstate migration since the mid-1980s. Using an instrumental variables strategy on cross-state data, we show that an aging workforce causes the migration rates of all age groups in a state to drop. This demonstrates that the effect of aging on migration includes indirect effects that go beyond the direct effect of raising the workforce share of groups with lower migration rates. We then develop an island model in which firms can hire workers either locally or from other locations, and show that an aging population leads firms to recruit more locally. This improves the local job prospects of all workers, which causes migration rates of all age groups to fall, consistent with our empirical findings. Our quantitative analysis suggests that this channel accounts for around half the migration decline, substantially more than what would be predicted solely by the direct effect.

Key words: interstate migration, labor mobility, population aging

---

Karahan: Federal Reserve Bank of New York (e-mail: fatih.karahan@ny.frb.org). Rhee: University of Hawaii at Manoa (e-mail: rhees@hawaii.edu). The authors thank Harold Cole, Morris Davis, Jan Eeckhout, Cecilia Fieler, Nezih Guner, Soojin Kim, Dirk Krueger, Alex Ludwig, Guido Menzio, Ryan Michaels, Guillermo Ordoñez, Serdar Ozkan, Ayşegül Şahin, Martin Schneider, Sam Schulhofer-Wohl, Robert Shimer, Chris Smith, Ahmet Ali Taşkin, and seminar and conference participants at numerous institutions. The authors also thank Darius Li, Samuel Kapon, and Rachel Schuh for providing superb research assistance and Rob Shimer for providing historical data on birthrates across U.S. states. This paper was written in part while Rhee was visiting the RAND Corporation. Rhee thanks the National Institutes of Health for their support (under award number 2T32AG000244). The views expressed in this paper are those of the authors and do not necessarily reflect the position of the National Institutes of Health, the Federal Reserve Bank of New York, or the Federal Reserve System.

1 Introduction

The rate of interstate migration in the United States has fallen steadily from 3 percent in the mid-1980s to less than 1.5 percent in 2010. The majority of interstate moves are job-related, which has prompted some analysts to express concern that the declining migration rate might harm the labor market.[1] We investigate the reasons behind the interstate migration decline, since its effect on the labor market depends critically on what is driving it. This paper's focus is on the impact of the aging population on interstate migration. Migration rates decline sharply over the life cycle: The migration rate of workers below age 40 is about twice as large as that of workers older than 40. From the 1980s to 2010, the share of individuals older than 40 in the working-age population increased from 45 percent to almost 60 percent. Thus, the aging of the population would seem to be a logical explanation for lower mobility.[2]

However, empirical analysis shows that the direct effect of an aging population does not appear to be the primary explanation for falling interstate migration. In section 2.4, in a simple analysis of the shift in the shares of different age groups, we find that the direct effect of an aging population accounts for only 20 percent of the migration decline. Kaplan and Schulhofer-Wohl (2013) and Molloy et al. (2013) evaluate the role of demographic changes, including age, education, and household structure. These studies also find direct demographic effects to be too small to explain the bulk of the decline. Instead, a declining trend common to all demographic subgroups accounts for most of it.

At first glance, this pattern suggests that aging of the population is not quantitatively important and that instead one should look for common factors affecting the migration decisions of the entire labor force regardless of age.[3] This paper takes another view. We argue that aging can affect migration patterns of all workers through equilibrium effects. For that reason, simple shift-share analyses are not suitable to capture the full range of the effects of aging. In particular, we show that an increase in the share of middle-aged workers

[1] A large fraction of interstate migrants report having moved for a new job, job search, or other job-related reasons. For example, about 50 percent of all interstate moves during the 2000s were job-related (March CPS; authors' calculations). The idea that geographical mobility might be important for the labor market goes back to at least Blanchard and Katz (1992), who showed that labor mobility is an important adjustment mechanism of local labor markets to adverse shocks. Oswald (1996) hypothesized that some of the differences in unemployment across countries can be attributed to differences in geographic mobility. Several papers have found migration to be an important determinant of individual labor market outcomes (e.g. Kennan and Walker (2011) and Gemici (2011)).

[2] Throughout the paper, we restrict attention to workers between the ages of 25 and 59. We label workers older than 40 as middle aged.

[3] Following this insight, Kaplan and Schulhofer-Wohl (2013) argue that the development of information technology and the decrease in the geographic specificity of occupations are responsible for lower migration rates. Molloy et al. (2013) propose a decline in labor turnover as a possible explanation.

in the labor force lowers the equilibrium migration rate for all workers, an effect we call *migration spillovers*. These spillovers explain the lower migration rates of workers in all age groups.

We exploit cross-state variation in population aging and migration rates to estimate the indirect effects of aging on migration patterns. Following Shimer (2001), we instrument a state's age composition using lagged cumulative birthrates. This instrument relies on the exclusion restriction that, conditional on state and time fixed effects, any economic conditions that shifted fertility rates in the past do not affect current migration decisions except through their effects on current age composition.

Without indirect effects, we would expect age composition to have no impact on the average migration rates of different age groups. However, we observe empirically that the aging of the population does have large indirect effects. Specifically, our baseline specification suggests that the average migration rate of a state declines by around 30 percent if the share of middle-aged individuals increases 10 percent. Using alternative specifications, we find the cross-state elasticity to be at least -2. Investigating a unit of analysis smaller than states, we find negative elasticities, though with lower precision because of a weaker instrument at this level of aggregation.

In the second part of the paper, we explore the economic forces that drive the effects of population aging. Our hypothesis is that an aging local population leads firms to change their recruiting methods and hire more from the local labor force. To test this theory quantitatively, we develop an equilibrium search model consisting of many locations that differ in their attractiveness to workers of different ages. At each location, there are workers of various ages with differing moving costs and job destruction rates. Workers can look for jobs in the local market, where they meet local firms, or they can search globally for jobs with firms in any location. Similarly, firms in a given location can look for workers only in the local market or they can recruit in the global market.[4] Older workers have high moving costs. Thus, it is profitable for firms in the local market to hire older workers because these workers have fewer options outside the local market and are more profitable. Consequently, aging of the population in a local market causes local firms to recruit more heavily in that market. This raises the local job finding rate, in turn lowering the mobility of all workers in that local labor market.

We calibrate a version of the model with 50 locations by targeting several labor market and migration-related moments during the 1980s. Although we do not target state-level

[4]Therefore, local labor markets are distinct from the global market as they allow local residents and firms target their search efforts. Indeed, Oyer and Schaefer (2012) document that, even in a highly specialized labor market for lawyers, geographic proximity explains nearly one-third of the sorting patterns between law firms and lawyers.

data, the model reproduces remarkably well the negative cross-state relationship between population aging and migration. As the expected profits from hiring within a local market rise, firms find it increasingly advantageous to recruit locally rather than globally. As a result, workers find local jobs at a higher rate and have less need to move, regardless of their age.

A unique prediction of our hypothesis is regarding where firms hire workers. The model predicts that the share of in-state hires increases in response to population aging. Analyzing data from the Survey Income and Program Participation (SIPP), we find that this share has increased from 89.8% in 1985 to 91.4% in 2013. Consistent with the data, our model predicts an increase of 2 percentage points. We further test the model by comparing the cross-state elasticity of in-state hires to that in the data. The elasticity in the model falls within the confidence interval of the empirical estimate.

We use the model as a measurement device to gauge the contribution of population aging to declining migration rates. Keeping all parameters of the model constant, we change only the age composition to mimic the U.S. population in each year. The model explains around 50 percent of the decline (0.7 percentage points) until the early 2000s, and little of the decline in later years.[5] According to the model, about 70 percent of the decline comes from the equilibrium effect. Just 30 percent reflects the direct effect of an aging population. Consistent with the data, our model generates large declines in migration rates for workers of *all* ages through the indirect effect. Thus, our results suggest that accounting for migration spillovers is important in explaining the effects of an aging population.

Finally, we use our model to assess the implications of lower geographic mobility for aggregate unemployment. Our explanation for the long-run decline in migration suggests that lower geographic mobility is not a cause for concern regarding the labor market. We find that the large decline in migration implies a drop in aggregate unemployment. The upward pressure on the unemployment rate caused by the limited search opportunities of older workers is largely offset by the equilibrium effect that increases the rate at which workers of all ages find jobs.

Related literature Our paper is most closely related to Kaplan and Schulhofer-Wohl (2013) and Molloy et al. (2013), who also study the decline in interstate migration over the same period. While the explanations in these papers differ, both show that migration has been declining for various groups of workers, accounting for a large portion of the aggregate decline. Our contribution is to show that population aging is an important factor in this

[5] There are two ways to interpret the failure of the model in generating declining migration after 2000. First, the decline in this period might be attributable to the development in information technology as explained in Kaplan and Schulhofer-Wohl (2013). Second, Hyatt et al. (2016) show that other datasets do not point to a decline in the interstate migration rate after 2000.

common component.

Guler and Taskin (2012) evaluate the role of the rise in female labor force participation and the resulting rise in the share of dual-income households. We differ from their work most notably because we explain the decline in migration rates within demographic groups, which accounts for most of the decline in the data. Coate and Mangum (2017) document that locations experiencing a larger fall in migration also experienced faster growth of income dispersion, which they interpret as increasing the returns from local job search and reducing incentives to move.

Geographic mobility is not the only measure of labor reallocation that has been trending down in the U.S. An emerging literature examines declines in other measures of reallocation. Davis et al. (2010) study the fall in the flow rate into unemployment from the 1980s to the mid-1990s. Fujita (2011) studies the decline in labor turnover. Decker et al. (2014), Haltiwanger et al. (2012), Hyatt and Spletzer (2013) and Hyatt and Spletzer (2015) document declines in gross worker and job flows. These papers all find evidence that a significant share of the declines they investigate are due to changes within demographic groups. Thinking through the lens of a shift-share framework, this suggests that demographic changes, such as population aging, cannot be the primary explanation. While we study geographic mobility across local labor markets, our paper opens the possibility that demographic shifts might have shaped U.S. worker reallocation more broadly over the past several decades through indirect effects.[6]

Using a similar empirical methodology to our work, Shimer (2001) argues that population aging in a state causes an increase in that state's unemployment rate. This finding seems to contradict our model, which predicts slight declines in local unemployment. However, Foote (2007) shows that when more recent data is added to the analysis, the estimated effect in Shimer (2001) switches sign and becomes consistent with our model.[7]

The rest of the paper is organized as follows. Section 2 documents the stylized facts on the decline in interstate migration and presents the cross-state analysis. Section 3 presents the quantitative model, while section 4 discusses our calibration and results. Section 5 concludes.

[6]In a similar vein, Karahan et al. (2017) show that shifts in population growth can explain about 30% of the decline in new business formation through equilibrium effects.

[7]To be more precise, depending on how one deals with correlation patterns in the residuals, the estimates may not rule out the null of no effect on unemployment.

2 Empirical analysis

This section investigates whether the age composition in a labor market affects individual migration probabilities.

2.1 Empirical strategy

The main empirical analysis for testing and measuring such effects relies on regional differences in the timing and extent of population aging. The empirical specification is given below:

$$m_{ijt} = \alpha_j + \beta_t + \gamma\omega_{jt} + \delta X_{it} + \epsilon_{it}, \tag{1}$$

where m_{ijt} is an interstate migration indicator that takes a value 1 if individual i leaves location j in year t. α_j and β_t are a full set of state and time dummies, respectively. In some specifications, we allow for location-specific time trends. We control for a vector of individual observables, X_{it}, that vary across specifications, but always include a full set of age dummies. In the baseline specification X_{it} includes a college dummy, a gender dummy, and a dummy for race (white). These controls along with the fully nonparametric age controls ensure that we compare similar people who reside in different locations. ω_{jt} captures differences in the age composition in j. We summarize the age composition with the fraction of the working age population (25–59) older than 40.

The age composition in a local labor market is likely endogenous with respect to migration decisions. This could arise due to various reasons, for example, if shifts in demand for labor simultaneously affect migration decisions and demographics by attracting young workers. To identify the causal effect of population aging, it is critical that we isolate the component that is orthogonal to labor demand conditions.[8] We follow Shimer (2001) and exploit the variation in age composition induced by (cumulative) lagged birthrates. More specifically, our instrument in location j and year t, denoted by z_{jt}, is given by

$$z_{jt} = \sum_{k=25}^{39} br_{j,t-k},$$

where $br_{j,t}$ is the birthrate in location j in year t. The sum is taken across all years between $t-39$ and $t-25$. Identification of causal estimates rely on the exclusion restriction that, conditional on state and time effects, any lagged economic conditions that shifted fertility rates in the past do not affect current migration decisions,except through their effects on

[8]Note that any permanent unobserved differences are controlled for by state fixed effects. Similarly, aggregate shocks are absorbed in year fixed effects. The remaining threat to identification is due to shifts within locations relative to the aggregate.

current age composition.

2.2 Data

To implement this strategy empirically, we should decide on the definition of a location. Given our goal of measuring equilibrium effects, a location should proxy a local labor market. An additional restriction is on measurement: We should observe moves across locations and have demographic information on birthrates and age composition. Given the availability of migration data across states over a relatively long period, our baseline analysis reported in section 2.4.1, uses state-level data. As states may not always correspond to unique local labor markets, we repeat our analysis at a finer level using data from the American Community Survey (ACS). Section 2.4.2 presents the results using this dataset.

We now provide a brief explanation of the various data sources used in the analysis. See Appendix A for more details.

Migration information

State-level analysis draws on data from the March Current Population Survey (CPS). To focus on migration that is not motivated by changes in schooling (in particular, college attendance and graduation) or retirement, we restrict our sample in the main analysis to civilians between the ages of 25 and 59. Our data is for the period 1986–2013 and exclude people who report a move from Alaska, Hawaii or from outside the U.S.[9] After 1996, we exclude observations with imputed migration data to avoid complications arising from changes in CPS imputation procedures.[10] CPS respondents are asked in March about their migration behavior since the previous March. Since most of this reference period is in the previous calendar year, we assume that the March CPS measures moves that took place in the previous calendar year.[11] These restrictions leave us with a sample of 2,003,437 observations.[12] Appendix A.2 contains further details and some summary statistics about the sample.

To estimate migration elasticities in a more granular way, we turn to the ACS, which allows us to measure migration at a finer geographical level than U.S. states. While Public

[9] As we explain below, the omission of Alaska and Hawaii is due to the lack of information on birthrates in the earlier years of the sample. Nevertheless, this omission does not affect our qualitative conclusions, in the sense that keeping these states in the sample and focusing on a shorter time series does not change the results.

[10] Specifically, we exclude observations based on the following flags for imputation: qmigrat1, qmigrat1g, qmigst1a and qmigst1b. See Kaplan and Schulhofer-Wohl (2012) for a detailed explanation of the issues around imputation in the CPS.

[11] Therefore, when we match the CPS sample to demographic information, we lag the CPS data by one year. The results do not depend on this particular assumption.

[12] Our CPS sample starts in 1986 due to data availability. For years 1982–84, we were unable to construct our instrument due to the lack of Statistical Abstracts of the U.S. corresponding to those years. In 1985, the variable used to measure interstate migration (migsta1) is not available.

Use Microdata Areas (PUMA) are the geographical unit used for the individual records in the ACS PUMS, the migration variables are associated with Migration PUMAs (MIGPUMAs), which are constructed from one or multiple PUMAs. Prior to 2012, the ACS used PUMAs based on the Census 2000, whereas Census 2010-based PUMAs were introduced to the ACS starting in 2012. We restrict the ACS data to the years between 2005 and 2011. Similar to the CPS, we restrict the sample to 25–59 year old civilians. See Appendix A.3 for further details.

Demographic information

Our empirical strategy requires us to have local-level data on age composition and birthrates. For the state-level analysis, we obtain annual population estimates for each state by 5-year age groups from the Census Bureau.[13] These data allow us to compute the share of middle-age population, ω_{jt}, for each state.

To obtain exogenous variation in middle-age share, we instrument ω_{jt} with cumulative lagged birthrates. Birthrates are measured in births per thousand residents and are available in the various Statistical Abstracts of the United States.[14] Our measure that predicts ω_{jt} is the sum of birthrates in a state over the years $t-39$ through $t-25$.

For the analysis using the ACS sample, we obtain demographic data at the county level and then aggregate them to MIGPUMA level to estimate MIGPUMA level migration elasticities. This involves a two-step process. First, using a county–PUMA crosswalk, county-level birthrates and population shares are aggregated to PUMA level. Second, using a PUMA–MIGPUMA crosswalk, these variables are aggregated to MIGPUMA level. We explain this process in Appendix A.3.

2.3 Motivating aggregate facts

We report here some aggregate facts that motivate this paper. Figure 1 plots the evolution of interstate (gross) migration rates from the March CPS (blue solid line), along with its long-run trend estimated using an HP filter with a smoothing parameter of 100 (red dashed line). This figure points to a 50 percent decline over three decades with little variation in business cycle frequencies.

A natural candidate for explaining the migration trend is the aging of the population over the past 30 years. To display the substantial aging of the U.S. population, we compute the

[13]The age groups that are used in our analysis are 25–29, 30–34, ..., 54-59. All population files are downloaded from `http://www.census.gov/popest/data/historical/index.html`.

[14]We are grateful to Rob Shimer for providing us with his data constructed from the Statistical Abstracts for the period 1940–91. Data are not available for the years 1941, 1942, and 1943. Data are also not available for Hawaii and Alaska prior to 1960. We drop these states entirely from the analysis. This omission does not affect the results in any meaningful way.

share of the working-age population (25–59) older than 40.[15] Figure 2 shows that this share increased from around 45 percent during the mid-1980s to about 60 percent by early 2000s and then stabilized. The large migration differences across age groups are well documented. In fact, in our data, individuals between 25 and 29 are almost four times more likely to move to a different state than those between 55 and 59.

Mechanically, the migration rate is a weighted average of age-specific migration rates. Thus, demographic changes alter the weights and the migration rate. To evaluate the direct effect of this compositional change, we conduct a simple accounting exercise. The migration rate in year t, m_t, can be expressed as:

$$m_t \equiv \sum_{i=1}^{N} s_{i,t} \times m_{i,t},$$

where $s_{i,t}$ and $m_{i,t}$ are group-specific population shares and migration rates, respectively. We consider 7 age groups: 25–29; 30–34; 35–39; 40–44; 45–49; 50–54; 55–59. Fixing the migration rate of each group to its 1980 level, we construct a counterfactual migration rate by changing only the shares of age groups:

$$\hat{m}_t = \sum_{i} s_{i,t} \times \overline{m}_{i,t}.$$

Under this formulation, any change in the migration rate over time, $\Delta\hat{m}_t$, is driven by the change in the share of each age group; that is,

$$\Delta\hat{m}_t = \sum_{i} \Delta s_{i,t} \times \overline{m}_{i,t}.$$

The red line in figure 3 plots the counterfactual migration rate if the within-group rates remained constant. It shows that the direct effect of population aging accounts for about 3 percentage points, or 20% of the decline. The corollary is that declines in "within-group" rates account for most of the decline. Figure 4 plots the migration rates for individuals in various age groups, confirming that migration has declined for all age groups. A successful explanation of the migration trend must be consistent with this. Our paper argues that the group-specific migration rates $m_{i,t}$ are functions of the age composition $\{s_{i,t}\}$. We show theoretically and empirically that population aging can generate a broad-based decline in migration rates.

[15]This cutoff is motivated by the fact that migration rates decline sharply until around 40 and are more or less constant throughout the rest of working life.

2.4 Measuring equilibrium effects of population aging on migration

We turn to our main empirical analysis and evaluate the role of population aging in the decline in interstate migration. We begin our analysis using state-level data.

2.4.1 State-level analysis

As explained in section 2.1, the estimation of the effect of population aging in (1) relies on cross-state variation in the age composition, which we instrument with cumulative lagged birthrates. Table 1 reports several "first-stage" regressions of the age composition on the instrument. Column (1) includes state and year fixed effects, runs the specification in levels, and estimates a semi-elasticity of the middle-age share with respect to lagged birthrates of -0.0463. The instrument, along with the fixed effect, explains slightly more than 95% of this variation. To see how much of this explanatory power is due to the instrument alone, we regress the middle-age share on the instrument after residualizing both of these variables by regressing them separately on state and time dummies. The results are reported in column (2). Of the variation in the middle-age share not explained by the fixed effects, the instrument explains about 14%. Columns (3) and (4) repeat the same analysis in a log specification. The takeaway is the same: Birthrates in a state have a predictive power for future age composition in that state, with an elasticity of -0.35. Figure 5 provides a scatter plot of the residual middle-age share on residual instrument.

The results of the first-stage analysis give us confidence that we can estimate the effects of population aging with sufficient precision. In estimating (1), we use the instrumental variable probit estimator as our benchmark, but we also consider linear probability models. The vector of controls (X_{it}) includes a full set of age dummies, a dummy for college degree, a dummy for white and a dummy for female. In each specification, observations are weighted using sampling weights and, unless otherwise stated, standard errors are clustered around state.[16]

To illustrate the economic significance of the estimated effects for various specifications, we report two other statistics. The first is the effect of population aging on migration by 2010. We compute the difference between the estimated model's predicted migration for 2010 and a prediction for 2010 that assigns to each U.S. state its average age composition over the period 1990–1992 (inclusive). In calculating the latter, α_j, β_t, and X_{it} are held the same as in the baseline prediction. Only the age composition in 2010 is modified for each state j to $\hat{\omega}_{j2010} = \frac{1}{3}\sum_{t=1990}^{1992} w_{jt}$. The second quantity we compute is the elasticity of the migration rate with respect to the middle-age share in 2010. To compute this elasticity, we apply a 10% decline to the middle-age share of each state in 2010, calculate the aggregate migration

[16] We use "wtsupp" to weigh the observations.

rate implied by the model, and divide the percent change in the aggregate migration rate by 0.01.

Table 2 presents the results. The first column shows the results for our baseline specification. We estimate a negative and statistically significant effect of the middle-age share on individual migration propensities. According to the point estimate, population aging can explain about 1.3 percentage points of the decline in migration between early 1990s and 2010. The 90% confidence around this effect is between 0.3 and 1.4 percentage points. The point estimate corresponds to a migration elasticity around -3. The rest of Table 2 investigates the robustness of our main finding. Column (2) estimates (1) as a linear probability model using 2 stage least squares. Note that in this case we estimate a much larger elasticity, but with much lower precision.[17] In column (3), we employ a two-way cluster that allows the standard errors to be correlated across time and space, and find that the market-level effects are still significant. One generic concern in regressions such as (1) is whether state fixed effects can flexibly control for the heterogeneity in the data. To investigate whether our results are driven by the choice of the panel model, in column (4) we allow states to have heterogeneous linear time trends.[18] Allowing for heterogeneous trends dampens the estimated effects somewhat, but the effects are still economically relevant: Under this specification, population aging can explain about 0.8 percentage points of the decline, with a 90% confidence interval between 0.1 and 1.5 percentage points. The elasticity corresponding to the point estimate is -2.3.

2.4.2 Analysis at lower levels of aggregation

While state-level analysis points toward large market-level effects of population aging on migration, there is one potentially important caveat: Most U.S. states are geographically large and do not closely correspondwith unique local labor markets. We now repeat a similar analysis using ACS data, which allows us to measure moves across MIGPUMAs, which are smaller geographic units. One additional advantage of focusing on MIGPUMAs is that they provide many more observations of locations than states.[19] In principle, this larger sample should allow us to test the presence of market-level effects with more power. However, there are two important drawbacks. First, for reasons explained in section 2.2, we must focus our analysis on a shorter period, 2005 to 2011. Second, the instrument is probablyweaker when using smaller units of analysis, because people are more likely to move out of their birth location in 25 years if that location is measured at a very fine level. Therefore, one should

[17]The coefficient is significant at 10 percent.

[18]Estimation of probit models with endogenous regressors involve nonlinear solvers and do not easily converge if there are many dummy variables. Therefore, we opted to estimate the heterogeneous trends specification using a linear probability model.

[19]In our final dataset there are 1028 MIGPUMAs.

expect that predicting the age composition with lagged birthrates is more difficult at the MIGPUMA level than at the state level.Whether the analysis using MIGPUMAs has more power for testing market-level effects is an empirical question. Table 3 provides the "first-stage" results for MIGPUMAs analogous to Table 1. Note that while the lagged birthrate is significant in all specifications, assuring the validity of the instrument, the contribution of the instrument to the goodness of fit is much smaller compared with the state-level data, 0.028 in the log-log specification.

Table 4 shows the results based on the ACS sample. Column (1) reports the results for the baseline specification, column (2) estimates a linear probability model, and column (3) uses a two-way cluster to deal with potential correlation in standard errors across both space and time by employing a two-way cluster. Importantly, we estimate a much larger elasticity of migration with respect to the age composition of the population. There are two ways to interpret the larger elasticity at the MIGPUMA level. First, state-level results may be understating true elasticity because states do not correspond with unique local labor markets. Second, the ACS data or the variation across MIGPUMAs may not possess enough power to measure the effect precisely. The standard deviation suggests the latter is the likely explanation. Nevertheless, all specifications point to an effect that is statistically significant at 10 percent.

2.5 Possible explanations for the cross-sectional facts

In this section, we examine several potential mechanisms through which population aging might reduce the migration rate of individuals. We then turn to our preferred explanation arising from equilibrium effects.

One potential explanation is based on the change in relative prices in response to age composition. Such a change would occur if workers of different ages are not perfect substitutes (see, for example, Borjas 2003; and Wasmer 2003). Suppose that the share of middle-aged workers increases, say, due to fewer births 25 years ago. In that case, the scarcity of young workers would push up their relative wages and cause them to move out at lower rates. This channel operates through a decline in young worker migration rates.[20] We test this hypothesis by interacting the share of middle-age population with a dummy for being older than 40.[21] Column (2) of Table 5 shows the results.[22] While the effect on young worker migration rate is negative and sizable, the effect on older workers is larger. This finding runs counter to the implications of the hypothesis discussed above.

[20]In fact, the change in relative prices should have the opposite effect on older workers.

[21]To instrument for this variable, we include the interaction of our instrument with this dummy as an excluded variable.

[22]Column (1) reproduces the baseline specification for convenience.

Another possible mechanism is based on informal care arrangements between parents and their children. If middle-aged workers move less frequently, the mobility of their children may be reduced if these children expect their parents' help to insure against labor market shocks (as documented in Kaplan 2012) or if the children provide informal care to their parents (Groneck and Krehl 2014). If our facts are driven by intergenerational relationships, such as these informal care arrangements, then an aging population should also reduce the mobility of the elderly. Column (3) of Table 5 investigates this possibility, by estimating market-level effects on a sample of 60–85 year old individuals. We find no evidence for a statistically significant effect on individuals above 60. If anything, we estimate a slightly positive effect on this population.

One conclusion we draw from the analysis in Table 5 is that the effect we document disproportionately affects people in the labor force. We now consider our theory of migration spillovers, which emphasizes a general equilibrium effect of population aging that operates through the labor market.

3 The Model

In this section, we develop an island model of the labor market in the spirit of Lucas and Prescott (1974). We use the model to argue that aging of the population affects the recruiting strategy of firms through a composition externality in a way that reduces the migration rate of all individuals in the economy.

3.1 The environment

The economy consists of a finite number of "islands," indexed by $i = 1, 2, \cdots, N$. Time is continuous. The locations are identical in terms of their productivity, which is normalized at one.

There is a unit measure of infinitely lived, risk-neutral, heterogeneous workers across these locations. Workers are endowed with one unit of indivisible labor and discount future at rate r. There are $j = 1, \cdots, J$ types of workers, with type j making up ω^j measure of total population. In order to move, a type-j worker pays a moving cost c, which is i.i.d. over time and follows a type-specific distribution $G^j(c)$. Types also differ along two other dimensions: First, a worker of type j derives a flow utility of ϵ_i^j when she resides in location i.[23] Second, as we explain below, workers of type j separate from their employer at rate δ^j.

The economy is also populated by a continuum of homogeneous firms which have access to a linear production technology that turns one unit of labor into y units of output. To

[23] Without this preference heterogeneity, all locations will have an identical composition of workers. We use these parameters in our quantitative analysis to generate exogenous variation in the age composition across locations so as to test the model's cross-sectional implications.

form a match and produce output, workers and firms must carry out a time-consuming search process. There are two modes of search: A firm can recruit workers *locally* from the local population or *globally* from any other location. More specifically, each location has a local market accessible only to local workers and firms. Positions advertised in the local market are only available to local residents. By contrast, a job-post in a global market is visible to all workers.

Regardless of the mode of search, firms pay κ to create a vacancy. Matches are determined randomly by a constant-returns-to-scale matching function $m(v,u)$. Denoting $\theta = v/u$ as the vacancy-to-unemployment ratio (or the market tightness), the job-finding rate for workers and the contact rate for firms can be written as $p(\theta) \equiv m(v,u)/u$ and $q(\theta) \equiv m(v,u)/v$, respectively. We let v_i denote the measure of vacancies in the local market as v_i and $u_i = \sum_j u_i^j$ denote the measure of unemployed workers that search in this market. In a global market, the relevant measure for vacancies is $v_g = \sum_{i=1}^N v_{gi}$, where v_{gi} is the measure of vacancies of firms in location i that are advertised in the global market. Similarly, the unemployed in the global market is given by $u_g = \sum_{i=1}^N u_i$.[24]

Upon a successful job search, workers and firms negotiate a wage contract via Nash bargaining and start an employment spell. We assume the match dissolves exogenously at type-specific rate δ^j.

In our setup, types correspond to age groups.[25] Our assumptions on heterogeneity mean that workers at different ages face different moving costs, have different separation rates into unemployment, and prefer to live in different locations.

3.2 Value functions

We now turn to the decision problems faced by workers and firms.

Workers Let U_i^j and $W_i^j(w)$ denote the lifetime utility of a worker of type j in location i, who is unemployed and employed, respectively. Then, the following equations characterize these values:

$$rW_i^j(w) = w + \epsilon_i^j + \delta^j \left(U_i^j - W_i^j(w)\right) \tag{2}$$

$$\begin{aligned} rU_i^j &= b + \epsilon_i^j + \left(p_{il} + p_g \frac{v_{ig}}{v_g}\right)\left(W_i^j\left(w_{il}^j\right) - U_i^j\right) \\ &\quad + p_g \sum_{k \neq i} \frac{v_{kg}}{v_g} \mathbb{E} \max\left\{0, W_{kg}^j\left(w_{kg}^j(c)\right) - U_i^j - c\right\}, \end{aligned} \tag{3}$$

[24]For workers, there are no costs to searching in any market. Thus, all unemployed workers search in their local market as well as in the global market. Therefore, the relevant unemployment when defining the global market tightness is the measure of unemployed workers across all locations.

[25]Thus, instead of modeling the life cycle of workers, we model age as a fixed type. This shortcut allows us to obtain analytical expressions and gain useful insight into the analysis of equilibrium effects.

where w_{il}^{j} is the wage in the local market of i and $w_{kg}^{j}(c)$ is the going wage in the global market paid by a firm in location k. Equation (2) states that a worker employed at wage w keeps receiving this wage until the match dissolves at rate δ^{j}. As stated in equation (3), unemployed workers receive a flow utility of leisure b, which includes unemployment benefits, home production and utility from leisure activities. Workers obtain the location- and type-specific flow utility, ϵ_{i}^{j}, and search for jobs in two markets: the local market of location i and the global market. Meeting with a local firm happens in the local market at rate p_{il} and in the global market at rate $p_g v_{ig}/v_g$. In the latter expression, p_g is the rate at which any meeting occurs in the global market and v_{ig}/v_g is the probability that the meeting is with a local firm. Because no moving cost has to be paid, a meeting with a local firm in both markets turns into a job. Unemployed workers meet nonlocal prospective employers in the global market with probability $p_g(1 - v_{ig}/v_g)$. In this case, they draw a moving cost c from the type-specific moving cost distribution and determine if a job is mutually beneficial.[26] Upon negotiating the wage, the firm subsidizes workers' relocation expenses.

It is easy to show that the decision to accept a distant job offer and move is characterized by a cutoff rule. We let c_{ik}^{j} denote the cutoff moving cost such that workers in location i who meet a firm in location k move if and only if $c < c_{ik}^{j}$.

Firms Let $J^{j}(w)$ denote the value of a firm matched with a type-j worker at wage w, and let V_{il} and V_{ig} denote the values of creating a vacancy in the local and global markets, respectively. The firm is the residual claimant of the output, thus the value of operating a firm matched with a type-j worker is given by

$$rJ^{j}(w) \;=\; y - w - \delta^{j} J^{j}(w).$$

[26]Note that we assume moving cost is revealed to the worker upon meeting a firm in a different location. The i.i.d. assumption simplifies the analysis of the model and has no meaningful effect on our results. In fact, an earlier version of the paper modeled moving cost as a permanent trait of the worker and reached similar conclusions.

The value to a firm of creating a vacancy in the local and global market is given in (4) and (5), respectively.

$$
\begin{aligned}
rV_{il} &= -\kappa + q_l \sum_{j=1}^{J} \frac{u_i^j}{u_i} \left[J^j \left(w_{il}^j \right) - V_{il} \right] && (4) \\
rV_{ig} &= -\kappa + q_g \Bigg[\sum_{j=1}^{J} \frac{u_i^j}{u} \left\{ J^j \left(w_{il}^j \right) - V_{ig} \right\} && (5) \\
&\quad + \sum_{k \neq i} \sum_{j=1}^{J} \frac{u_k^j}{u} \mathbb{E}^j \max \left\{ 0, J^j \left(w_{kg}^j (c) \right) - V_{ig} \right\} \Bigg],
\end{aligned}
$$

where $u = \sum_i u_i$ is aggregate unemployment in the economy. Equation (4) shows that the expected profit associated with creating a vacancy has two components. The first is the probability of meeting a worker in this market, q_l. The second is the composition of workers in the economy, u_i^j / u_i, which matters because the profitability of a worker to the firm depends on the worker's type.

Similarly, equation (5) defines the value of creating a job in the global market. In this market, a firm meets a local worker at rate $q_g u_i^j / u$, and a worker from a different location k at rate $q_g u_k^j / u$. Importantly, in the latter event, the firm and worker pair decide if it is profitable to form a match and agree on a wage. Thus, one important difference between the local and the global market is that some meetings in the global market may not turn into matches if the moving cost is high enough, unlike the local market, where all meetings turn into matches.

3.3 Wage determination

We assume that wages are set through Nash bargaining and let η denote workers' bargaining power. This assumption implies that the worker obtains η-share of the surplus generated by the match and the firm captures the rest.

Since the outside option for workers may depend on the mode of job search, there are two types of surplus functions to be solved for: The surplus of a local firm--worker pair and the surplus of a firm matched with a nonlocal worker. For a match formed in the local market i with a type-j worker, the match surplus S_i^j is given as

$$
S_i^j \equiv J_i^j + W_i^j - U_i^j.
$$

When a firm in location i meets a type-j worker in a location k, the match surplus depends on the worker's moving costs c. Let $S_{ki}^j (c)$ denote the joint surplus of a match of a firm in location i with a type-j worker in a different location k. This surplus is given by

$$S_{ki}^{j}(c) \equiv J_i^j + W_i^j - U_k^j - c.$$

In appendix B.1, we derive a system of equations that characterizes the surplus functions in terms of the model's fundamentals. Those equations are also useful in solving the model.

3.4 Equilibrium

We now proceed to define steady state equilibrium for this environment. To that end, we first complete the model by describing the law of motion in the economy and discuss the entry decision of firms.

Law of motion and the steady state The following equations define the laws of motion for the measures of employed and unemployed workers across locations and types:

$$\begin{aligned} \dot{u}_i^j &= -\left(p_{il} + p_g\left[\frac{v_i}{v} + \sum_{k\neq i}\frac{v_k}{v}G^j\left(c_{ik}^j\right)\right]\right)u_i^j + \delta^j e_i^j \\ \dot{e}_i^j &= -\delta^j e_i^j + \left(p_{il} + p_g\right)u_i^j + p_g\sum_{k\neq i}G\left(c_{ki}^j\right)u_k^j \\ 1 &= \sum_{i=1}^{N}\left(u_i^j + e_i^j\right). \end{aligned} \tag{6}$$

We study the steady state of this environment and thus impose $\dot{u}_i^j = \dot{e}_i^j = 0$.

Free-entry conditions We assume that there is a continuum of firms that can create jobs in any market. This implies that jobs are created up to the point that firms make zero ex-ante profits. Equations (7) and (8) show the relevant free entry conditions in local and global markets, respectively:

$$\kappa = (1-\eta)\,q_{il}\sum_{j=1}^{J}\frac{u_i^j}{u_i}S_i^j \tag{7}$$

$$\kappa = (1-\eta)q_g\left\{\sum_{j=1}^{J}\frac{u_i^j}{u}S_i^j + \sum_{k\neq i}\sum_{j=1}^{J}\frac{u_k^j}{u}\mathbb{E}^j\max\left(0, S_{ki}^j(c)\right)\right\}. \tag{8}$$

We note again that all meetings in the local market turn into an employment spell, but some of the nonlocal meetings in the global market do not.

Equilibrium Definition We are now ready to define equilibrium for this environment.

Definition 1. Steady-state equilibrium

A steady-state equilibrium consists of cutoff values for moving, $\left\{c_{ik}^j\right\}$, wages in local and global markets, w^j_{il} and $w^j_{kg}(c)$, measures of unemployed by type and location, $\{u^j_i\}$, market tightness $\{\theta_{il}\}$ and θ_g, and vacancy shares of firms in the global market by location, $\{v_{ig}/v_g\}$, such that

1. Cutoff values $\{c^j_{ik}\}$ solve workers' migration problem in (2) and (3).
2. Wages in the local market, w^j_{il}, and the global market, $w^j_{kg}(c)$, solve the Nash bargaining problem.
3. Measures of employed and unemployed by location and type, e^j_i and u^j_i, satisfy the laws of motion in steady-state given in (6).
4. Free-entry conditions hold in all markets so that equations (7) and (8) are satisfied.

3.5 Composition externalities and migration spillovers

We are interested in the effect of population aging on migration. In particular, we argue that changes in the age structure of the population, modeled as shifts in $\{\omega^j\}$ toward older age groups, trigger a general equilibrium effect through the labor market, reducing the migration rate of all individuals in the economy.

We now use a stylized version of the model to illustrate the main forces behind this effect. We consider a special case in which individuals have identical preferences across locations so that demographic composition of each location is the same in equilibrium. We also assume a common separation rate δ across age groups. Denoting the surplus of a type-j worker-firm pair in the local and global market as S^j_i and S^j_g, respectively, we derive the following expressions:[27]

$$S^j_l = \frac{y - b - \eta p_g \frac{N-1}{N} \mathbb{E}^j \max\left\{0, S^j_g(c)\right\}}{r + \delta + \eta\left(p_l + \frac{p_g}{N}\right)} \tag{9}$$

$$S^j_g(c) = S^j_l - c. \tag{10}$$

Equations (9) and (10) make clear the economics of age differences in the model. First, firms that meet workers in the local market take into account that workers have the option to look for a job in the global market. This option value decreases with moving costs, as shown by the term $\mathbb{E}^j \max\left\{0, S^j_g(c)\right\}$, and increases the local surplus. Thus, in the local market, meetings with older workers, who have higher average moving costs, generate a higher surplus. In the global market, these relationships switch signs because firms have to subsidize workers' moving costs. It can be shown that the surplus in the global market is *lower* for older workers.

[27] The derivations are shown in Appendix B.3.

Now, consider two economies that only differ in the age composition of their population. The one with the older population has a labor market in which recruiting workers locally is more profitablebecause workers have a lower option of searching in the global market. The free-entry condition in equation (7) dictates that firms in the economy with an older population recruit more intensively in the local market, thereby lowering q_l and increasing p_l, the rate at which workers in any location find local jobs. By contrast, equation (8) dictates that firms shy away from hiring workers in the global market, lowering p_q, the rate at which workers meeting firms require a move. As a consequence of firms' increased recruiting in the local market, all workers regardless of age find local jobs at a faster rate and global jobs at a lower rate. The equilibrium in the older-population economy thus dictates a lower migration rate for all workers. This effect is driven by a composition externality which we label migration spillovers.

The model makes two predictions: First, the migration rate of all workers will decline in a labor market with an aging population, as documented in section 4. Second, as population ages, a larger share of new hires will be local residents who do not require moves. Thus, the decline in mobility in this environment reflects the changing recruiting practices of firms in response to changing demographics.

In what follows, we test these predictions of the theory with data and use the model to measure the aggregate effects of an aging U.S. population on the interstate migration rate.

4 Quantitative analysis

Our model studies the effect of compositional changes in the population on migration.[28] We now calibrate the model and test it by comparing its cross-sectional predictions with the data. Finally, the calibrated model is used to evaluate the role of an aging population in declining migration rates.

4.1 Calibration

We calibrate a version of the model with identical locations.[29] Each type of worker corresponds to a specific age group in the data. We use the model to match a number of targets related to mobility and labor markets, as explained below. Appendix B.2 presents

[28] It is worth emphasizing that the model is general and can be used to study the implications of changes in the U.S. population other than the aging population. Some examples are the rise in the share of dual-income households and changes in the homeownership rate. We focus in this paper on the aging of the population, because (1) the magnitude of demographic change is large; (2) the timing lines up well with the trend in migration; and (3) population aging is plausibly exogenous to migration and the labor market.

[29] This means that we are turning off location preference parameters, ϵ_i^j, by setting them to zero. The nonsymmetric version, a perturbation of the symmetric model, is used later to evaluate the model by studying its cross-sectional implications.

the details of the computational algorithm used to solve for a steady-state equilibrium.

Calibration strategy The calibration proceeds in two steps. First, we calibrate some parameters externally. The second step calibrates the remaining parameters internally by using an exactly identified simulated method of moments (SMM) and targeting moments from the 1980s. It is important to emphasize that we do not target the cross-location relationship between the age composition and the migration rate documented in section 2.4.1. Instead, we test the model's performance along this dimension after calibration.

Functional forms The matching function is Cobb-Douglas. The contact rate functions for both local and global markets are given by,

$$p(\theta) = \nu\theta^{1-\gamma}, \quad q(\theta) = \nu\theta^{\gamma}.$$

The parameter γ governs the elasticity of the matching function and ν is the matching efficiency.

Parameters set outside the model We set N, the number of locations, to 50.[30] We focus on the seven age groups between the ages of 25 and 60 described in section 2 ($J = 7$). The share of each age group in the population is computed from the March CPS using 1981 data. To calibrate the job destruction rates by age group, δ^j, we follow Shimer (2012) and compute the continuous time analog of the separation probabilities by age group for the entire sample period. We then take the average over the 12 months of 1981.[31] The time discount rate r is set to match a quarterly discount rate of 1 percent. This requires setting $r = 0.0033$. The bargaining parameter η is set to 0.5. The flow utility of unemployment is taken from Hall and Milgrom (2008) and set to 0.71.[32] Finally, to set the matching function parameters ν and γ, we follow Diamond and Sahin (2014) and construct a measure of the monthly job-finding rate for the period 1977–85 as the sum of E–E, U–E and N–E transition rates from the CPS. This series is then regressed in logs on a constant and the market tightness series constructed in Barnichon (2010). We obtain an estimate of 0.77 for ν and 0.25 for γ. Parameters calibrated outside the model are summarized in table 6.

Parameters calibrated with the simulated method of moments Eight parameters remain to be estimated: the vacancy posting cost, κ, and the moving cost distribution for each age group. We assume that the moving cost for each age group is distributed exponentially

[30] This choice is not without loss of generality and affects the cross-sectional elasticites that we use to test the model. Aging of population in one location affects other locations through the global market. The quantitative magnitude of this force depends on how big that location is in the economy. Since our goal is to explain the cross-state elasticities reported in section 2.4.1, we let the model have 50 locations.

[31] These data were constructed by Robert Shimer. For additional details, please see Shimer (2012) and his web page http://sites.google.com/site/robertshimer/research/flows.

[32] We provide the robustness of our results to an alternative choice of b in Appendix B.4.

with mean μ_i for $i = 1, \cdots, J$.[33] We now describe our targets and how they are computed in the model and the data.

Targeted moments and their model counterparts We follow Shimer (2012) to compute the continuous time job finding rate for the unemployed. This procedure gives us an estimate for each month, and we target their average in 1981. To compute the model counterpart of this measure, we first compute the job finding rate for each age group, which is given by

$$f^j = p_l + \frac{p_g}{N} + p_g \frac{N-1}{N} G^j\left(c^j\right),$$

where the first two terms sum to the rate at which the worker gets a local job and the last term accounts for the rate at which the worker gets a job elsewhere.[34] The average job-finding rate is then defined as

$$f = \frac{1}{u} \sum_{j=1}^{J} u^j f^j.$$

Finally, we target the trend component of interstate migration rates for each age group in 1981. Taking out the cyclical component allows us to address the problem that recession of the early 1980s could have put migration off its trend. To extract the trend component, $trend_t$, we apply a Hodrick-Prescott filter on the aggregate migration series with a smoothing parameter of 100. We compute the trend component of the age-specific migration rates, $trend_{jt}$, by scaling the raw numbers ($\hat{m}_{jt}$) by the ratio of the aggregate trend to aggregate migration in 1981 ($\hat{m}_{t=1981}$):

$$trend_{jt} = \hat{m}_{jt} \frac{trend_{1981}}{\hat{m}_{1981}}.$$

This approach assumes that the trend component is the same across all age groups. We target the values in 1981, $trend_{j,1981}$, for $j = 1, \cdots, J$.

The model counterpart of the annual migration rate is computed as the fraction of workers who change their location at least once over the course of a year. This is achieved in two steps. First, we compute the monthly migration rates separately for each age group, location, and employment status. Then, we obtain the aggregate migration rate by weighting age-specific migration rates with the respective population shares by first computing this object separately for each age group, location, and employment status, and then obtaining the average by weighting with the respective population shares. Note that even though only

[33] We choose this distribution as it takes only positive values and leaves us with only one parameter to be estimated per age group.

[34] The term $\frac{p_g}{N}$ reflects the fact that, on average, out of N contacts that occur in the global market, 1 is with a firm in the same location.

unemployed workers are allowed to move in the model at any time, employed workers may move over the course of a year if they have an unemployment spell in between. We assume that an unemployed worker meets with at most one job in a month and we use monthly migration rates to compute the annual migration rate, taking into account time aggregation as follows.[35] Let $m^j_{e,i,t}$ and $m^j_{u,i,t}$ denote the t-period migration probabilities for an employed and unemployed worker of type j residing in i, respectively. In other words, these objects measure the probability that a worker will move at least once in the next t periods. We compute $m^j_{e,i,12}$ and $m^j_{u,i,12}$, since they correspond to annual migration rates. Equations (11)–(14) define these objects recursively:

$$\begin{aligned} m^j_{e,i,1} &= 0 && (11)\\ m^j_{u,i,1} &= P^j_{ig} && (12)\\ m^j_{e,i,t+1} &= (1-\Lambda)\, m^j_{e,i,t} + \Lambda m^j_{u,i,t} && (13)\\ m^j_{u,i,t+1} &= P^j_{ig} + \left(1-P^j_{ig}\right) P_{il} m^j_{e,i,t} + \left(1-P^j_{ig}\right)(1-P_{il})\, m_{u,i,t}, && (14) \end{aligned}$$

where P^j_{ig} denotes the monthly probability of finding a job in another location, P_{il} denotes the monthly probability of a local job offer and Λ denotes the monthly separation probability. These are given by the following expressions:

$$\begin{aligned} P^j_{ig} &= 1-\exp\left\{-p_g \sum_{k\neq i} \frac{v_{kg}}{v_g} G^j\left(c^j_{ik}\right)\right\}, \quad \text{for } k \neq i. \\ P_{il} &= 1-\exp\left\{-\left(p_{il}+\frac{v_{ig}}{g}p_g\right)\right\} \\ \Lambda &= 1-\exp\left\{-\lambda\right\}. \end{aligned}$$

The estimation minimizes the equally weighted sum of squared percent deviations of model moments from their targets. Table 7 summarizes the moments used in the estimation and provides the fit of the model to the targeted moments. Table 8 summarizes the estimated parameters. The model does an excellent job of matching the calibration targets. Recall that migration decreases with age in the data. This monotonicity might lead to the conjecture that moving costs increase with age.[36] However, this is not necessarily the case in this model.

---

[35] While this assumption is not necessary to compute migration rates in a continuous time environment, we find that it is the fastest approach numerically and easiest to implement. One can approximate the continuous time version arbitrarily well by shortening the length of a "period." We found that our approach provides a good approximation to the continuous-time counterpart.

[36] Kennan and Walker (2011) use a structural model to estimate moving costs and find them to be large, much larger than our estimates. The difference is likely driven by the search frictions in our model, which makes it difficult for workers to move even in the absence of explicit costs. For a group of workers that face

For example, while individuals between 45 and 49 have a lower migration rate than those between 40 and 44, the latter group's average moving cost is higher.[37]

4.2 Testing the model on cross-sectional data

The model's primary goal is to quantify the effect of population aging in the U.S. on interstate migration, taking into account equilibrium effects. First, we validate the key forces in the model by comparing the model's cross-location implications with the data.[38]

There is no heterogeneity across locations in the calibrated model. Locations have identical productivity and the preference values ϵ_i^j are all set to zero. We perturb these preference parameters to generate variation in age composition across locations. Specifically, we set the youngest age-group's location preference for island 1 as $\epsilon_{i=1}^{j=1} = \epsilon$, where ϵ is uniformly distributed over the interval $[0, M]$. We also set the preference of group members toward the other islands as $\epsilon_i^{j=1} = -\epsilon$ for $i \geq 2$. Hence, young workers derive a positive flow utility from living in location $i = 1$ and a disutility of the same magnitude when they live in other locations. Thus, the equilibrium corresponding to these prefererences features a larger share of young workers in location 1 relative to other locations.[39] Preferences of the other age groups are unchanged. We draw 100 such perturbations and compute the relevant elasticities for each economy as explained below. We report the average elasticity across simulations.[40]

Age composition and migration rates Let m_i denote the outmigration rates of workers in location i, and $m_{\leq 39,i}$ and $m_{\geq 40,i}$ denote the same for workers aged 25–39 and 40–59, respectively. Similarly, let ξ denote the elasticity of migration with respect to the share of middle-age workers $share_{\geq 40,i}$, and $\xi_{\leq 39}$ and $\xi_{\geq 40}$ denote the elasticities for young and

no moving cost, interstate migration would be less than 4 percent. In contrast, the model in Kennan and Walker (2011) features no additional friction so that their estimate of moving cost is a composite of various costs and frictions.

[37] If the cutoff costs were the same across types, then trivially the mobility rate would be decreasing in the mean of the moving cost distribution. However, it is easy to show that the cutoff cost is increasing in the mean in a way that the relationship between the mean and the migration rate is not necessarily monotonic.

[38] Recall that the calibration did not target any data about the cross section of U.S. states

[39] While there are 50 locations in the model, solving for an equilibrium in this setup requires solving all equilibrium objects for only two locations: Location 1 and the rest. Despite the fact that this perturbation reduces the model effectively to a two-location model, the number of locations is important for the cross-location differences in migration rates as it influences the magnitude of the general equilibrium effects. For example, if there were only two locations in the model, a change in the population composition in one location would exert a big effect on equilibrium objects of the other location through the global market. Specifying the number of locations to the number of states in the U.S. disciplines the magnitude of the general equilibrium effect.

[40] Increasing the number of repetitions does not have any material effect on the estimated elasticities.

middle-aged workers. These statistics are calculated in the model as[41]

$$\begin{aligned}
\xi &= \frac{\log(m_1) - \log(m_2)}{\log(share_{\geq 40,1}) - \log(share_{\geq 40,2})} \\
\xi_{\leq 40} &= \frac{\log(m_{\leq 39,1}) - \log(m_{\leq 39,2})}{\log(share_{\geq 40,1}) - \log(share_{\geq 40,2})} \\
\xi_{\geq 40} &= \frac{\log(m_{\geq 40,1}) - \log(m_{\geq 40,2})}{\log(share_{\geq 40,1}) - \log(share_{\geq 40,2})}.
\end{aligned}$$

Table 9 reports the results, which are compared with the IV estimates from the probit model in Table 2. Comparing the point estimates, we find that although our model generates a sizable negative elasticity for migration (-1.36), it understates the estimated effect in the data (-3.09). Furthermore, consistent with the empirical finding, the migration rate of older workers is more elastic (-1.54) than that of the young (-1.00). While the model elasticities lie below the point estimates, they are within 95% confidence intervals. We conclude that the model agrees with the data on the cross-sectional elasticity of migration.

Age composition and the fraction of local hires Our second testconcerns the positive relationship between the share of middle-age workers and hiring of state residents. This feature of the data is especially important because it speaks directly to the model's main mechanism. Using data from the Survey of Income and Program Participation (SIPP), we compute the fraction of local hires among total hires for each state-year combination.[42] The number of total hires is defined as everyone in the state who reports being unemployed one month prior to the survey month but is employed at the time of the survey. Local hires are defined as those among the total hires who did not move across states over this period. In the data, we compute the elasticity of the share of local hires with respect to the share of population older than 40 to be 0.38, significant at 10 percent.

The model counterpart of the share of local hires, lh_i, is computed by simply dividing the measure of workers who are hired locally to all hires in that location. The associated cross-sectional elasticity, ς, is then given by

$$\varsigma = \frac{\log(lh_1) - \log(lh_2)}{\log(share_{\geq 40,1}) - \log(share_{\geq 40,2})}.$$

Table 10 reports the resulting elasticity, comparing it with the empirical estimate.[43] We find

[41]Note that we use location 1 and location 2 to compute these elasticities. Location 2 refers to any location that has unperturbed preference parameters as explained above.

[42]Further details about the SIPP sample are given in appendix A.2.

[43]Note that the sample size of the SIPP is relatively small for this exercise. Using a one-month period to calculate the share of local hires results in many observations being either 0 or 1. As the zeros are dropped in estimating the elasticity, table 10 uses 781 observations to estimate the elasticity. We also estimate the

the magnitude of the cross-sectional correlation computed in the model to be smaller than that in the data, but within a 90% confidence interval.

Plausibility of the mechanism The model's basic mechanism is based on the idea that jobs can be filled by either a local search or a broad-based search. The interaction between these two markets is what gives rise to the spillovers. It is plausible that for jobs with mundane tasks, firms may not engage in global search because it may be costlier and offer no clear advantage to recruiting in a local labor market. Thus, such spillovers may be weaker or nonexistent for these segments of the labor market. We now investigate this possibility by allowing the effects of population aging to be heterogeneous across occupations. More specifically, following Autor and Dorn (2013), we group occupations into four broad categories based on whether the occupation's typical tasks are routine or nonroutine, and whether they require manual or cognitive skills. Column (1) in Table 11, extends the baseline specification in column (1) of Table 2 to include these occupation dummies as controls. The sample size drops to around 1.5 million since we only assign an occupation category to people who are employed and whose occupation can be matched. Nevertheless, we estimate a precise effect significant at 5 percent that corresponds to a slightly larger elasticity than the baseline.

Next, we interact these categorical dummies with middle-age shares and instrument these with the interaction of the same dummies, leaving out the dummy for manual routine occupations. Therefore, the coefficient of -0.2 in the first row should be interpreted as the effect of population aging on individuals in these occupations. Consequently, all the interaction terms measure the additional effect on workers in other occupations. The effect is larger in cognitive and nonroutine occupations relative to manual and routine occupations.

4.3 Effects of population aging on interstate migration

To isolate the role of the aging population, we compute the model-generated aggregate migration rate for each year between 1981 and 2013. Figure 6 plots equilibrium migration rates in the model's steady state equilibrium corresponding to each year's demographic composition, comparing them with the data. The model generates a 0.7 percentage point decline in interstate migration, which is about 47 percent of the 1.5 percentage point decline in the data. The model is exceptionally successful in explaining declining mobility until the early 2000s, but explains little of the decline after that. This unexplained portion could be attributable to recent changes in information technology as described in Kaplan and

relationship using the level of the share as opposed to its log and find a coefficient of 0.3923 with a p-value of 0.09. It is reassuring that the qualitative results are not an artifact of dropping the zero observations. Using a longer period of, say, three or six months to compute the local hires share, we also find a positive elasticity.

Schulhofer-Wohl (2013). An alternative interpretation is based on Hyatt et al. (2016), who argue that the interstate migration rate did not fall by much after 2000. More specifically, they document that a substantial fraction of CPS respondents, who appear having moved in administrative Census data, do not report their move in the survey. Importantly, this fraction has been increasing since 2000. Through the lens of this finding, our model is quite successful in explaining the evolution of the interstate migration rate over a long horizon.

As shown in the accounting exercise in section 2.4, the direct effect of population aging accounts for a 0.2 percentage point decline in interstate migration. The remaining 0.5 percentage point decline in the model reflects equilibrium effects, the migration spillovers. This large equilibrium effect can best be understood by focusing on the changes in age-specific migration rates: Figure 8 plots the life cycle profiles of migration in 1980 and 2000 from the model and compares them with the data. Our model is able to generate large declines in the migration rates of *all* age groups through a change in the age composition. In fact, the model explains more than 30% of the decline in the migration rates of each age group.[44]

To illustrate the mechanism behind this decline in mobility, we plot the contact rates of unemployed workers in the local market, p_l, and in the global market, $N/\left(N-1\right)p_g$, over time. Figure 7 shows that, while the local job finding rate increases, workers' job-finding rates in other locations decreases as population ages. The combination of these two changes results in lower mobility rates.

We conclude that accounting for equilibrium effects in the labor market is important for properly assessing the role of population aging. Studies quantifying only the direct effect understate the total effect.

4.4 Understanding the mechanism

As Figure 7 shows, population aging in the model works through a change in the recruiting mode of firms. A direct implication of this mechanism is an increase over time in the share of local hires, that is, newly employed workers inside the state. Figure 9 plots the model counterpart of this share and compares it with the data.[45] In the data, this share increased steadily until early 2000s and has leveled around 97 percent. Similar to the data, our model predicts this share goes up until the mid-2000s and then remains flat. While the model overstates the share's level, the data support the theoretical implication that this share has risen over time, supporting our hypothesis.

[44] There have been other studies to explain this across-the-board decline in migration. The unexplained portion of the within-group decline can be attributable to the mechanisms explained in Kaplan and Schulhofer-Wohl (2013) and Molloy et al. (2013). Kaplan and Schulhofer-Wohl (2013) find that the decline in the geographic-specificity of occupations and changes in information technology can account for at least one-third of the decline.

[45] We plot the 5-year moving average from the data.

4.5 The decline in migration and aggregate unemployment

In this section, we explore the implications of lower migration rates for the labor market. One popular theory is that a decline in migration might indicate that workers are less able to take distant jobs, which in turnmay cause unemployment to rise.[46] Here we study the implications of our estimated model for aggregate unemployment.

Figure 10 plots the time series for aggregate unemployment in the model. As Figure 6 illustrates, population aging causes lower mobility for all workers. Since mobility is an important means of finding jobs in the model, population aging might lead to increased unemployment. However, despite the large fall in migration, we find that unemployment *falls* by about 0.3 percentage points in response to the aging population.

To better understand this seemingly counter intuitive result, it is worth while reviewing the cause of the migration decline. As section 3.5 explains, migration decreases because firms post more jobs locally and fewer jobs in other locations. Thus, workers do not move as frequently because they are better able to find jobs locally. Migration drops because the increase in local job-finding rates more than offsets the negative effect from the change in age composition, raising the overall job finding rate of all workers.

5 Conclusions

This paper investigates the long-run decline in interstate migration in the United States over the last several decades. It demonstrates that compositional changesdo not only affect migration directly, but also indirectly, an effect that we label *migration spillovers*. As the local population ages, local jobs become easier to find, and the migration rates of *all* workers declines in equilibrium. We find strong evidence for this mechanism in the cross-section of U.S. states. Our quantitative analysis suggests that population aging explains a substantial part of lower mobility and that the equilibrium effect accounts for most of this decrease.

The spillovers discussed this paper are more general and can be applied to any compositional change, such as the rise in dual-income households or changes in homeownership rates. A similar spillover may exist in response to changes in frictions faced by a group of workers. In particular, a large literature examines the mobility effects of frictions arising in the housing market. Most of the empirical literature in this field compares the outcomes of homeowners with those of renters. These studies implicitly assume that renters are not affected by such frictions. By contrast, our theory implies that these housing market imper-

[46] This idea goes back to Blanchard and Katz (1992), who showed that labor mobility is an important adjustment mechanism of local labor markets to adverse shocks. Oswald (1996) hypothesized that some of the differences in unemployment across countries can be attributed to differences in geographic mobility. More recent work such as Borjas (2003), Cadena and Kovak (2016), Karahan and Rhee (2013), Nenov (2015), Sterk (2015), and Yoon (ming) focus on the role of labor mobility for various outcomes in local labor markets.

fections could affect renter migration rates through spillovers. Thus, existing studies that don't take into account potential spillovers might yield biased results.

Various measures of worker reallocation show declining trends, including job-to-job transitions, separations into unemployment, job creation and destruction, and excess worker reallocation. In each of these cases, a substantial share of the decline occurs within demographic groups. Consequently, a simple shift-share analysis would suggest that population aging is not a major cause of these declines. However, our paper points to another conclusion: that population aging might have affected the labor market over the past several decades through the equilibrium effects in the labor markets. Further work is needed to fully uncover the labor market implications of population aging.

References

Aaronson, D. and J. Davis (2011). How much has house lock affected labour mobility and the unemployment rate? *Chicago Fed Letter No. 290, September.*

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the u.s. labor market. *American Economic Review 103(5)*, 1553–1597.

Barnichon, R. (2010). Building a composite help-wanted index. *Economics Letters 109*, 175–178.

Blanchard, O. and L. Katz (1992). Regional evolutions. *Brookings Papers on Economic Activity*.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics 118(4)*, 1335–74.

Cadena, B. and B. Kovak (2016). Immigrants equilibrate local labor markets: Evidence from the great recession. *American Economic Journal: Applied Economics 8(1)*, 257–290.

Coate, P. and K. Mangum (2017). Declining labor mobility and the role of income dispersion.

Davis, S., J. Faberman, J. Haltiwanger, R. Jarmin, and J. Miranda (2010). Business volatility, job destruction and unemployment. *American Economic Journal: Macroeconomics 2(2)*, 259–87.

Decker, R., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2014). The role of entrepreneurship in us job creation and economic dynamism. *The Journal of Economic Perspectives 28*(3), 3–24.

Diamond, P. A. and A. Sahin (2014). The beveridge curve revisited. *Working Paper*.

Foote, C. (2007). Space and time in macroeconomic panel data: Young workers and state-level unemployment revisited. *FRB Boston Working Papers Series No. 07-10*.

Fujita, S. (2011). Declining labor turnover and turnover. *Federal Reserve Bank of Philadelphia Working Paper 11-44*.

Gemici, A. (2011). Family migration and labor market outcomes. *Working paper*.

Groneck, M. and F. Krehl (2014). The exchange motive and informal long-term care: Evidence from the hrs exit-interviews. *mimeo*.

Guler, B. and A. A. Taskin (2012). Homeownership and unemployment: The effect of market size. *Working Paper*.

Hall, R. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American Economic Review 98*, 1653–1674.

Haltiwanger, J., H. Hyatt, E. McEntarfer, and L. Sousa (2012). Business dynamics statistics briefing: Job creation, worker churning, and wages at young businesses. *Kauffman Foundation Statistical Brief*.

Hyatt, H. R., E. McEntarfer, K. Ueda, and A. Zhang (2016). Interstate migration and employer-to-employer transitions in the u.s.: New evidence from administrative records data. *US Census Bureau Center for Economic Studies* (Paper No. CES-WP-16-44).

Hyatt, H. R. and J. R. Spletzer (2013). The recent decline in employment dynamics. *IZA Journal of Labor Economics 2*.

Hyatt, H. R. and J. R. Spletzer (2015). The recent decline of single quarter jobs. *IZA DIscussion Paper* (No. 8805).

Kaplan, G. (2012). Moving back home: Insurance against labor market risk. *Journal of Political Economy 120(3)*, 446–512.

Kaplan, G. and S. Schulhofer-Wohl (2012). Interstate migration has fallen less than you think: Consequences of hot deck imputation in the current population survey. *Demography 49(3)*, 1061–1074.

Kaplan, G. and S. Schulhofer-Wohl (2013). Understanding the long-run decline in interstate migration. *Federal Reserve Bank of Minneapolis Working Paper 697*.

Karahan, F., B. W. Pugsley, and A. Şahin (2017). Demographic origins of startup deficit.

Karahan, F. and S. Rhee (2013). Geographic reallocation and unemployment during the great recession: The role of the housing bust. *Federal Reserve Bank of New York Staff Reports 601.*

Kennan, J. and J. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79(1)*, 211–251.

King, M., S. Ruggles, J. T. Alexander, S. Flood, K. Genadek, M. B. Schroeder, B. Trampe, and R. Vick (2010). Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Lucas, R. and E. Prescott (1974). Equilibrium search and unemployment. *Journal of Economic Theory 7*, 188–209.

Molloy, R., C. L. Smith, and A. Wozniak (2013, April). Declining migration within the us: The role of the labor market. *Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C. 27.*

Nenov, P. (2015). Regional reallocation and housing markets in a model of frictional migration. *Review of Economic Dynamics 18*(4), 863–880.

Oswald, A. (1996). A conjecture on the explanation for high unemployment in the industrialized nations: Part 1. *Warwick Economic Research Paper 475*.

Oyer, P. and S. Schaefer (2012, December). Firm/employee matching: An industry study of american lawyers. (18620).

Shimer, R. (2001). The impact of young workers on the aggregate labor market. *Quarterly Journal of Economics Vol. 116, No: 3 (Aug., 2001)*, 969–1007.

Shimer, R. (2012, April). Reassassing the ins and outs of unemployment. *Review of Economic Dynamics 15(2)*, 127–148.

Sterk, V. (2015, September). Home equity, mobility and macroeconomic fluctuations. *Journal of Monetary Economics 74(C)*, 16–32.

Wasmer, E. (2003). Between-group competition in the labor market and the rising returns to skill. us and france 1964-2000. *Working Paper*.

Yoon, C. (forthcoming). Estimating a dynamic spatial equilibrium model to evaluate the welfare iimplication of regional adjustment processes: The decline of the rust belt.

6 Tables and Figures

Table 4 – Market-level effects of population aging: ACS Sample

	(1) Baseline	(2) Linear	(3) Twoway
Middle-age share	-6.945*	-0.790*	-0.790**
	(4.016)	(0.477)	(0.388)
College	0.0772***	0.00868***	0.00868***
	(0.00556)	(0.000569)	(0.000918)
White	0.0956***	0.00995***	0.00995***
	(0.00247)	(0.000329)	(0.000675)
Female	-0.0242**	-0.00223*	-0.00223
	(0.0122)	(0.00134)	(0.00173)
Age Dummies	Yes	Yes	Yes
Cluster	MIGPUMA	MIGPUMA	MIGPUMA & Year
Elasticity	-7.306	-6.461	-6.461
Observations	8,759,910	8,759,910	8,759,910

Notes: This table reports estimates of the specification in 1 using the ACS sample. Column (1) estimates a probit model, whereas columns (2) and (3) estimate linear probability models. All specifications include a college dummy, a dummy for white, a dummy for gender, and a full set of dummies for age, MIGPUMA, and year. The dependent variable of interest is the (log) share of middle-age individuals in the working age population. Standard errors are clustered around MIGPUMAs in (1) and (2), and (3) implements a two-way cluster around MIGPUMA and year. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

Table 1 – Results of the first stage analysis at state level

	(1) Share mid-age Raw	(2) Share mid-age Residuals	(3) Share mid-age, log Raw	(4) Share mid-age, log Residuals
Instrument	-0.0463***			
	(0.00315)			
Instrument (residual)		-0.0463***		
		(0.00306)		
Instrument (in log)			-0.347***	
			(0.0183)	
Instrument (in log, resid.)				-0.347***
				(0.0178)
State dummies	Yes	No	Yes	No
Year dummies	Yes	No	Yes	No
Observations	1372	1372	1372	1372
R^2	0.955	0.143	0.963	0.218

Notes: This table reports the results of the first stage associated with the instrument. The instrument is the sum of state-level birthrates lagged by 25 to 39 years. The dependent variable is the share of middle-age individuals in the working age population in (1), the residual of middle-age share in (2), the logarithm of the middle-age share in (3), and the residual of the logarithm in (4). Standard errors are clustered around state. See text for further details. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

Table 2 – Market-level effects of population aging: CPS sample

	(1) Baseline	(2) Linear	(3) Twoway	(4) Het. trends
Middle-age share	-2.219**	-0.206*	-0.206*	-0.0665*
	(0.967)	(0.112)	(0.122)	(0.0353)
College	0.234***	0.0137***	0.0137***	0.0137***
	(0.0170)	(0.00116)	(0.00206)	(0.00115)
White	-0.00129	0.000348	0.000348	0.000382
	(0.0224)	(0.00118)	(0.00137)	(0.00120)
Female	-0.0138***	-0.000771***	-0.000771***	-0.000777***
	(0.00327)	(0.000187)	(0.000186)	(0.000186)
Age Dummies	Yes	Yes	Yes	Yes
Cluster	State	State	State & Year	State
Effect on migration in 2010	-0.013	-0.025	-0.025	-0.008
90% CI	(-0.014, -0.003)	(-0.047, -0.003)	(-0.049, -0.0005)	(-0.015, -0.001)
Elasticity	-3.085	-5.919	-5.919	-2.314
Observations	2,003,386	2,003,386	2,003,386	2,003,386

Notes: This table reports estimates of the specification in 1. Column (1) estimates a probit model, whereas columns (2)–(4) estimate linear probability models. All specifications include a college dummy, a dummy for white, a dummy for gender, and a full set of dummies for age, state, and year. (4) also includes a state-specific linear trend. The dependent variable of interest is the (log) share of middle-age individuals in the working age population. Standard errors are clustered around state in (1), (2), and (4), and (3) implements a two-way cluster. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

Table 3 – Results of the first stage analysis at MIGPUMA level

	(1) Share mid-age	(2) Share mid-age Residuals	(3) Share mid-age	(4) Share mid-age Residuals
Instrument	-0.0281***			
	(0.00469)			
Instrument (residual)		-0.0281***		
		(0.00431)		
Instrument (in log)			-0.0292***	
			(0.00229)	
Instrument (in log, resid.)				-0.0292***
				(0.00211)
MIGPUMA dummies	Yes	No	Yes	No
Year dummies	Yes	No	Yes	No
Observations	6,575	6,575	6,575	6,575
R^2	0.975	0.006	0.976	0.028

Notes: This table reports the results of the first stage associated with the instrument at the MIGPUMA level. The instrument is the sum of state-level birthrates lagged by 25 to 39 years. The dependent variable is the share of middle-age individuals in the working age population in (1), the residual of middle-age share in (2), the logarithm of the middle-age share in (3), and the residual of the logarithm in (4). Standard errors are clustered around state. See text for further details. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

Table 5 – Heterogeneity by age groups

	(1) Baseline	(2) Heterogeneity by age	(3) Older sample (60–85)
Middle-age share	-2.219**	-1.906**	0.433
	(0.967)	(0.951)	(2.257)
Middle-age share × (age>40)		-0.634***	
		(0.139)	
Gender/Race/College Dummies	Yes	Yes	Yes
Age/Year Dummies	Yes	Yes	Yes
Cluster	State	State	State
Elasticity	-3.085	-2.648	0.720
Observations	2,003,386	2,003,386	623,991

Notes: This table investigates the heterogeneity in the effects of population aging on migration. All specifications include a college dummy, a dummy for white, a dummy for gender, and a full set of dummies for age, state, and year. The dependent variable of interest is the (log) share of middle-age individuals in the working age population. Column (1) reproduces our benchmark probit model, column (2) investigates if the effect is different on individuals older than 40, and column (3) estimates the effect on individuals between 60 and 85. Standard errors are clustered around state. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

Table 6 – Externally calibrated parameters

Parameter		Value
Time discount rate, r		0.0033
Value of leisure, b		0.71
# Locations, N		50
Workers' bargaining power, η		0.50
Matching efficiency, ν		0.77
Elasticity of the matching function, γ		0.25
Population share by age group	25–29	20.15%
	30–34	18.17%
	35–39	14.42%
	40–44	12.02%
	45–49	11.33%
	50–54	11.99%
	55–59	11.91%
Separation rate by age group	25–29	.0425
	30–34	.0310
	35–39	.0250
	40–44	.0210
	45–49	.0192
	50–54	.0176
	55–59	.0158

Table 7 – Matching the calibration targets

Moment		Data	Model
Average job-finding rate		0.4160	0.4160
Annual migration rate by age group	25–29	5.26%	5.26%
	30–34	3.82%	3.82%
	35–39	2.96%	2.96%
	40–44	1.99%	1.99%
	45–49	1.98%	1.98%
	50–54	1.42%	1.42%
	55–59	1.38%	1.38%

Note: Table 7 shows the fit of the model on targeted moments of the data.

TABLE 8 – Internally calibrated parameters

Parameter		Value
Vacancy posting cost, κ		0.4418
Mean of the moving-cost distribution by age group, μ	25–29	0.4363
	30–34	0.5215
	35–39	0.6261
	40–44	1.0565
	45–49	0.9205
	50–54	1.4253
	55–59	1.2658

TABLE 9 – Elasticity of migration: Model vs. Data

	Aggregate	25–39	40–59
Model	-1.360	-1.004	-1.542
Data (IV probit)	-3.085	-2.648	-3.948
95% CI	[-7.704,-0.294]	[-7.027,-0.038]	[-6.983,-0.053]

Notes: Table 9 shows the cross-sectional elasticity of outmigration computed from the model (first row) and compares it to the IV probit estimate (second row) and the 95

TABLE 10 – Elasticity of local hires: Model vs. Data

	Data	Model
Elasticity of the share of local hires	0.3821*	0.1093
w.r.t. share of population > 40	(0.231)	

Notes: Table 10 shows the IV estimate (first column) of the cross-sectional elasticity of local hires and compares it to the model counterpart (second column). The dependent variable in the first column is the share of local hires in a month at the state level. Standard errors are in parentheses.

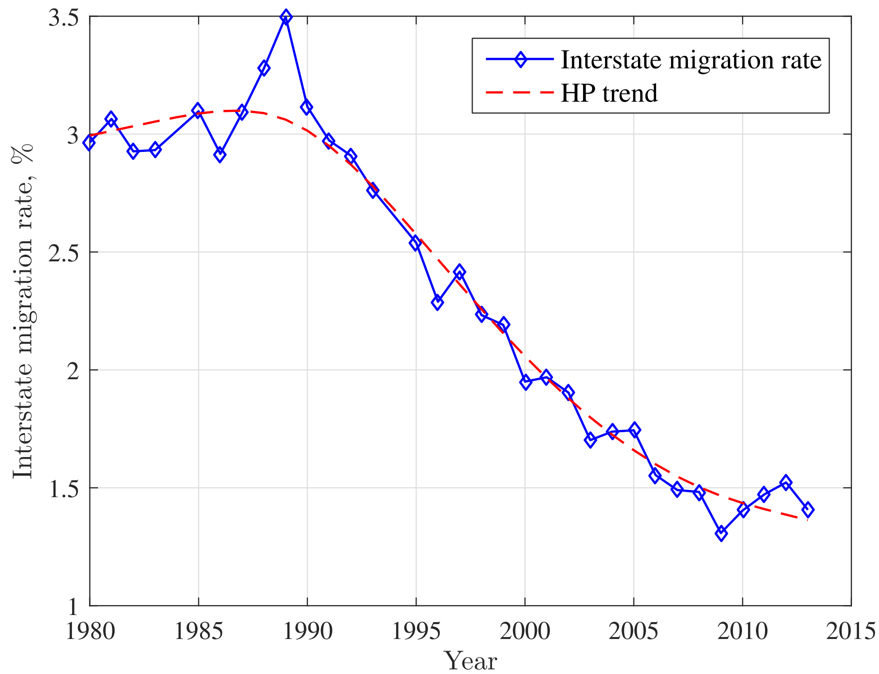
Table 11 – Market-level effects of population aging: CPS sample

	(1) Occ. dummies	(2) Interactions
Middle-age share	-2.407**	-0.200*
	(1.110)	(0.121)
manual nonroutine	-0.0124	-0.00967**
	(0.0112)	(0.00427)
cognitive routine	-0.0141	-0.0105**
	(0.00905)	(0.00457)
manual rout.	-0.110***	-0.0299***
	(0.0141)	(0.00802)
cognitive nonroutine × share		-0.0381***
		(0.0140)
manual nonroutine × share		-0.0320***
		(0.0118)
cognitive routine × share		-0.0287**
		(0.0125)
Age Dummies	Yes	Yes
Cluster	State	State
Observations	1,506,234	1,506,234

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

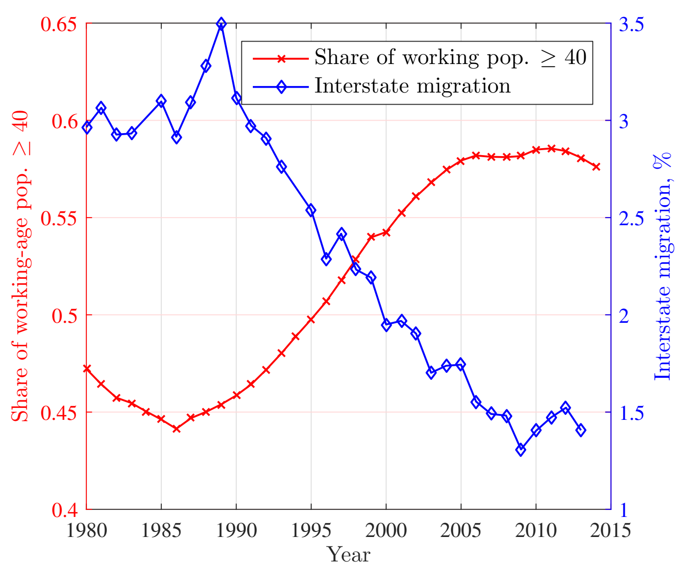
Notes: This table investigates how the effects of population aging differ across occupations. All specifications include a college dummy, a dummy for white, a dummy for gender, dummies for three occupation groups (see text for details), and full sets of dummies for age, state, and year. Column (2) includes interactions of these occupation dummies with the share of middle-age population. The left-out interaction term is for manual routine occupations, so that the coefficients on the interactions measure the additional effect of population aging relative to individuals in manual routine occupations. Standard errors are clustered around state. *, **, *** denotes statistical significance at the 10%, 5%, 1% level, respectively.

FIGURE 1 – Interstate migration rate



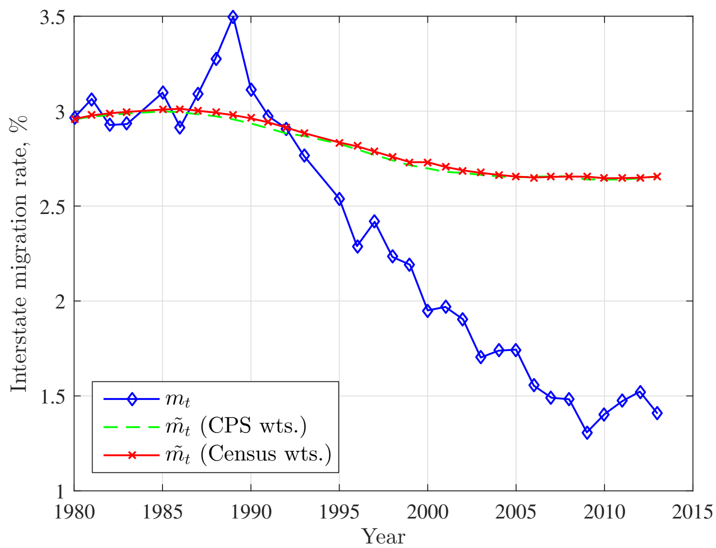
Notes: The blue line shows the fraction of individuals that move across states in a given year and the red dashed line shows its trend, obtained with an HP filter with a smoothing parameter of 100. Source: March CPS; authors' calculations. See text for details.

FIGURE 2 – Population aging and interstate migration



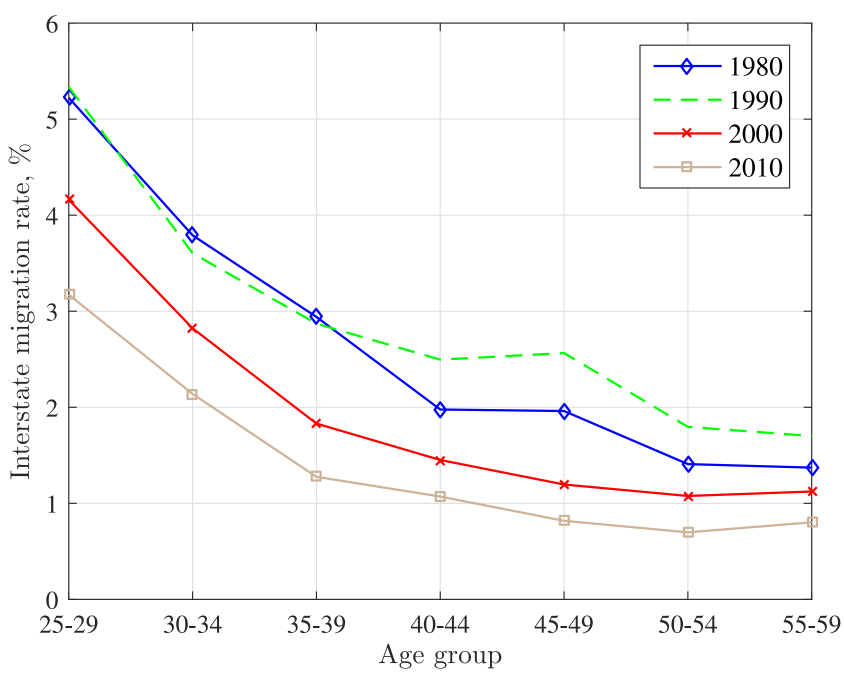
Notes: The blue line shows the interstate migration (right axis) and the red dots show the share of the working age population (25–59) older than 40 (left axis). Source: March CPS; U.S. Census Bureau; authors' calculations.

Figure 3 – Direct effect of population aging on migration



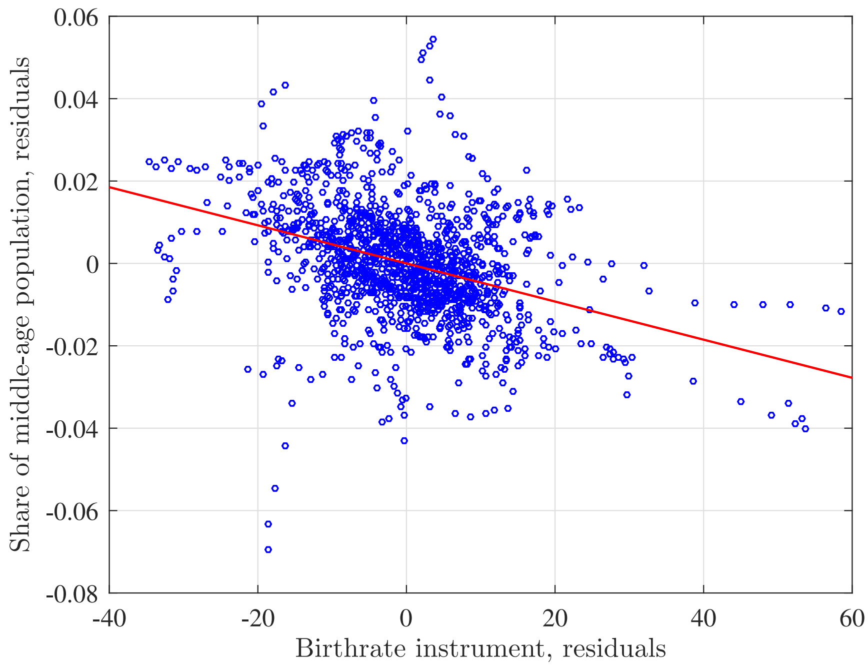
Notes: The blue line shows the interstate migration rate. The red and green lines show counterfactual migration rates, calculated by holding constant the life cycle profile of migration to its 1980 level and letting the age composition of the population vary as it does in the decennial census and CPS, respectively. See text for details.

Figure 4 – Age-specific migration rates over time



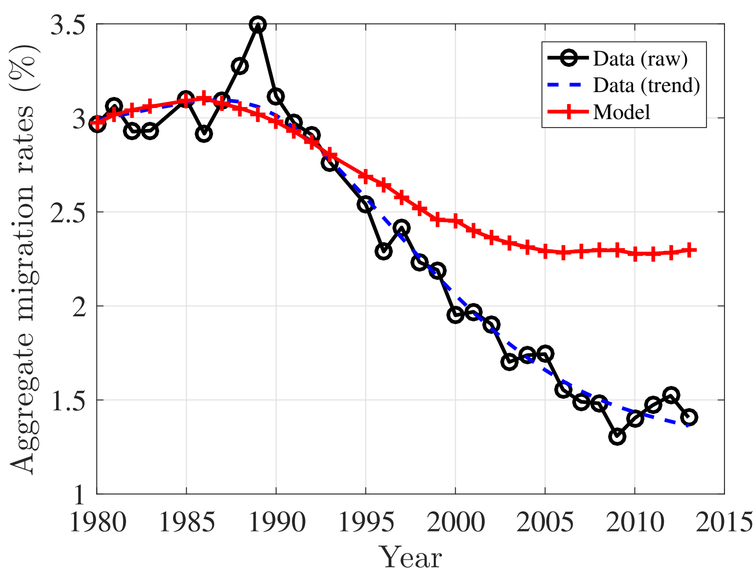
Notes: This figure plots the life cycle profile of migration in various years. Source: March CPS and authors' calculations.

Figure 5 – Lagged cumulative birthrates and age composition (residuals)



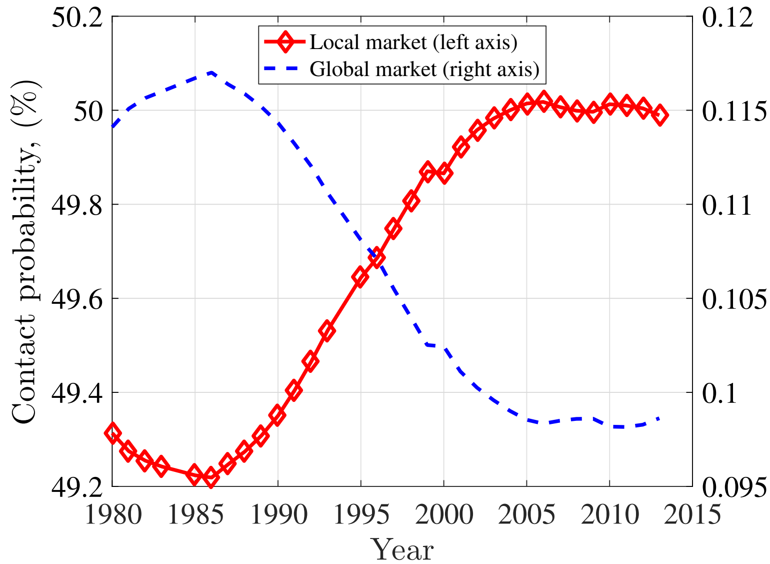
Notes: This figure plots the share of middle-age workers, residualized by regressing on state and time fixed effects, against cumulative lagged birthrates, residualized in a similar fashion. The red line plots a linear fit to the scatter plot.

Figure 6 – Aging population and the decline in migration: Data vs. Model



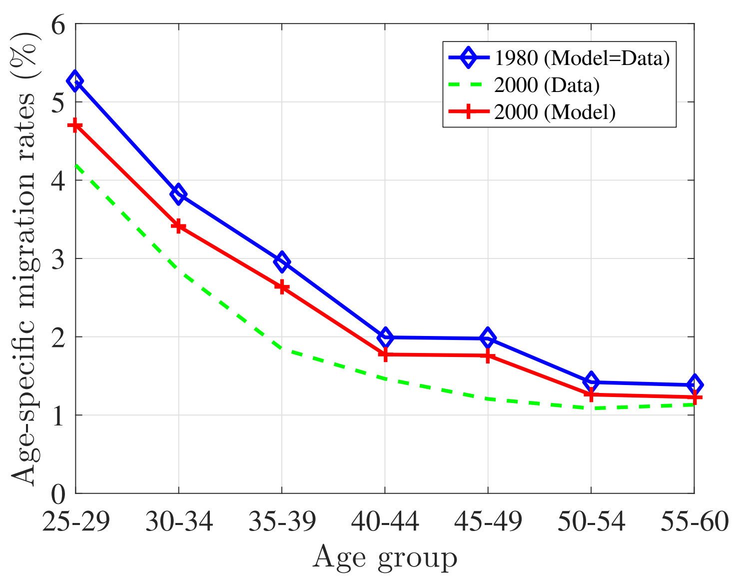
Notes: Figure 6 plots the model-implied interstate migration rates from 1981 to 2013 and compares them to the data. Annual migration in the model is computed as the fraction of all population who move at least once in a 12-month period. The trend component of the series is obtained with an HP filter with a scaling parameter of 100.

FIGURE 7 – Contact rates for unemployed workers in local and global markets



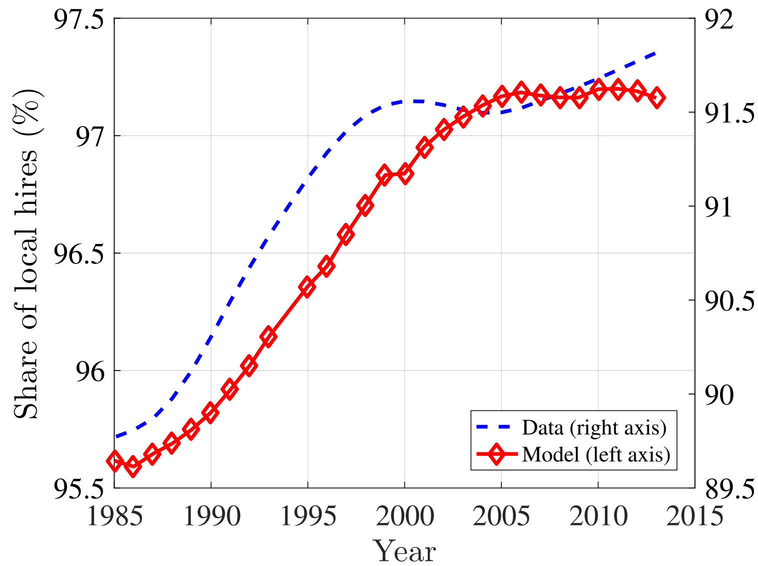
Notes: This figure plots the rate at which unemployed workers contact with potential employers in the local and global markets. The red lines shows the contact probability in the local market (left axis) and the blue dashed line shows the contact probability in the global market (right axis).

FIGURE 8 – Quantifying the importance of migration spillovers



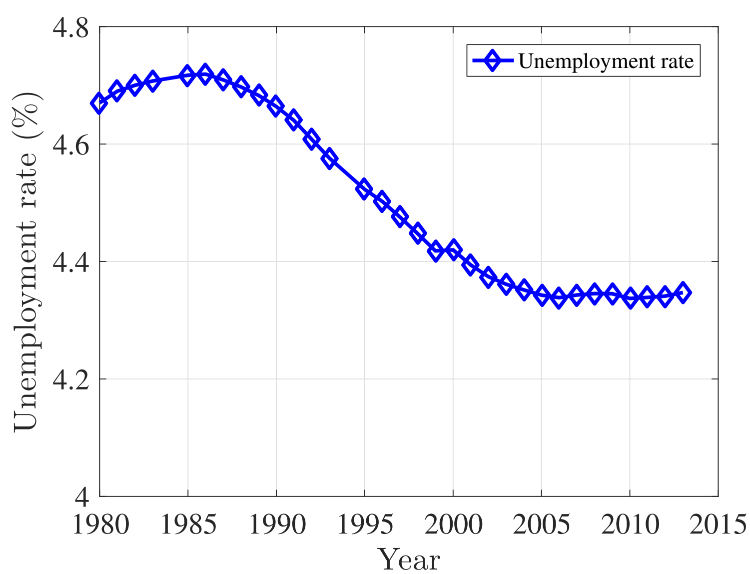
Notes: Figure 8 shows the life-cycle profile of migration in various years. The x-axis shows the seven age groups used in our estimation. The y-axis is the migration rate in percentage points. The blue dashed line is the profile in 1980 (in the data and in the model), and the green dashed line is for 2000. The series in red is the migration rates in the model for year 2000. The model values correspond to the steady state equilibrium associated with the age composition of the working-age population in each year.

Figure 9 – Share of hires from other locations



Notes: Figure 9 shows the model counterpart of the share of hires from outside the firm's own state (left axis) and compares it to its empirical counterpart (right axis). The empirical counterpart is the 5-year moving average of the raw series computed from SIPP.

Figure 10 – Declining mobility and aggregate unemployment



Notes: Figure 10 plots the series for the aggregate unemployment rate in the model. The unemployment rate corresponds to the steady state associated with the age composition of the working-age population in each year.

A Online Appendix: Data

A.1 Demographics

We obtain annual population estimates by 5-year age groups in each state from the Census for the period 1980–2013.[47] These data allow us to compute the share of a given age group for each state-year cell. Following Shimer (2001), we obtain exogenous variation in age composition in a state by instrumenting with cumulative lagged birthrates. These are measured in births per thousand residents and are available in the various Statistical Abstracts of the United States.[48] Data are not available for 1941, 1942, and 1943. Therefore, we start using birthrates from 1944 onward. Focusing on the population share of 25–40 year old implies that our data on lagged birthrates starts in 1984.

TABLE A.1 – Share of young workers (25 - 39)

	1985 - 89		2008 - 12
West Virginia	51.8%	New Hampshire	35.2%
Arkansas	52.2%	Maine	35.3%
New Jersey	52.6%	Vermont	35.5%
Colorado	58.3%	Texas	44.9%
DC	59.4%	Utah	51.0%
Utah	59.8%	DC	53.2%

TABLE A.2 – Cumulative lagged birthrates (25 to 39 years)

	1985 - 89		2008 - 12
New York	341.68	Connecticut	207.52
Rhode Island	342.46	Rhode Island	209.92
New Jersey	348.86	Massachusetts	211.36
Mississippi	470.14	New Mexico	312.58
Utah	490.96	Idaho	316.30
New Mexico	542.14	Utah	415.90

[47] The age groups that are used in our analysis are 25–29, 30–34, ..., 54–59. All population files are downloaded from http://www.census.gov/popest/data/historical/index.html.

[48] We are grateful to Rob Shimer for providing us with his data constructed from the Statistical Abstracts for the period 1940–91. Data are unavailable for Hawaii and Alaska prior to 1960. We drop these states entirely from the analysis. This omission does not affect the results in any meaningful way.

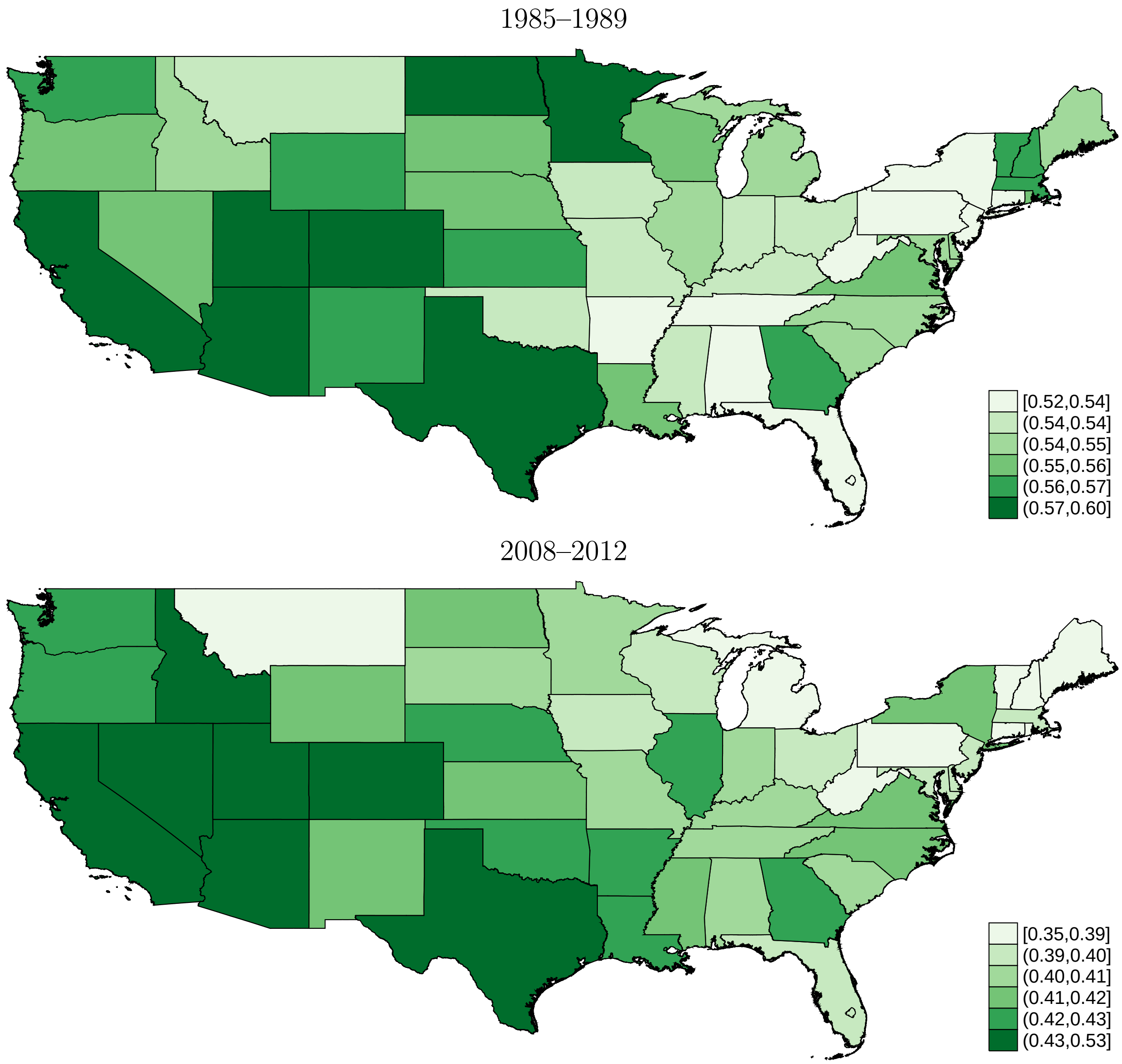


FIGURE A.1 – Regional variation in age composition

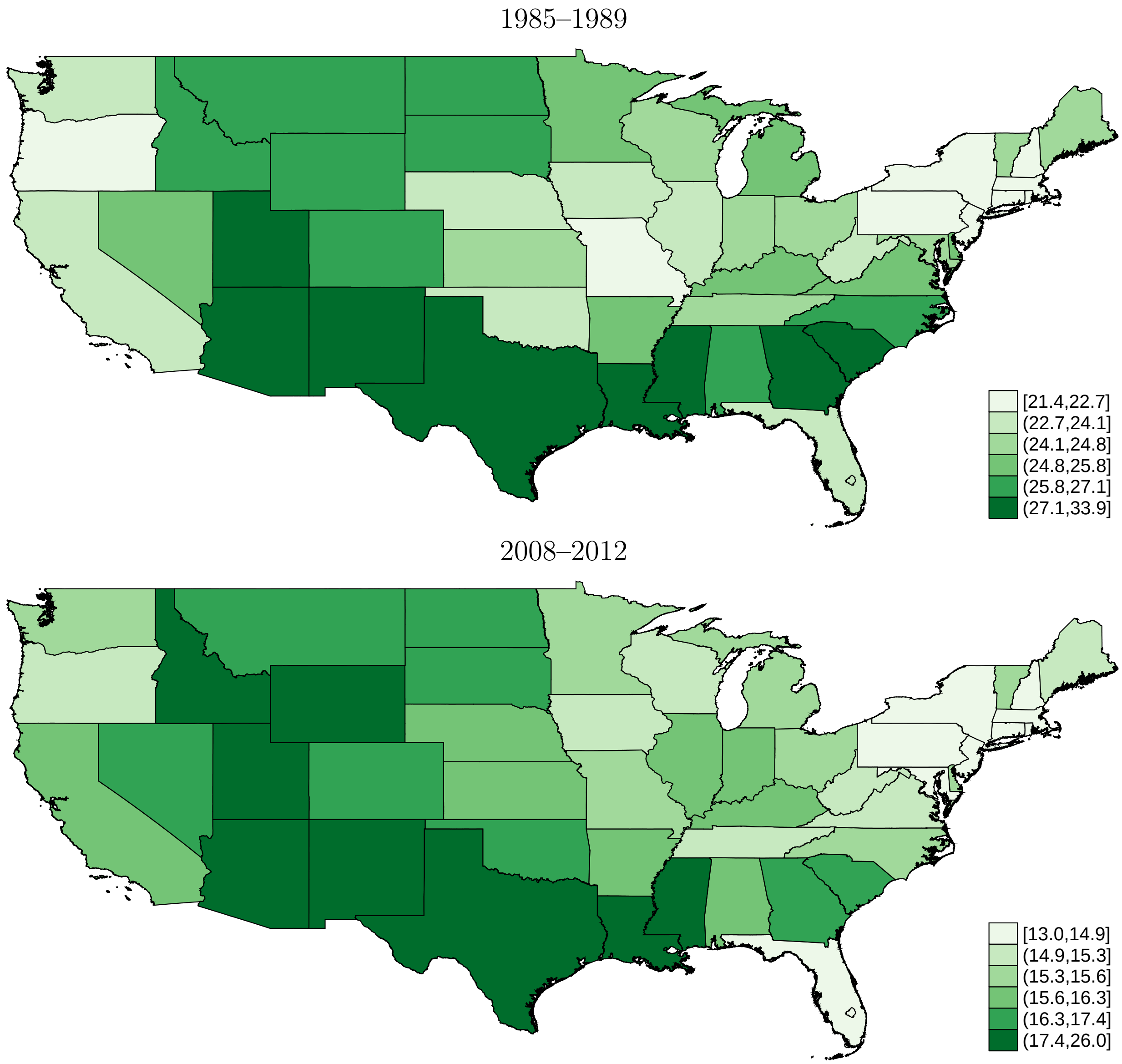


FIGURE A.2 – Regional variation in fertility rates

A.2 March CPS

Our analysis focuses on the period after 1980. Throughout the paper, we consider two age groups: young workers (25–39) and middle-aged workers (40–59).

Migration rates are computed using micro data from the Annual Social and Economic Supplement to the Current Population Survey (March CPS). In order to focus on migration that is not motivated by changes in schooling (in particular, college attendance and graduation) or retirement, we restrict the sample to nonmilitary/civilian individuals who are between the ages of 25 and 59 at the time of the survey. March CPS data are obtained from the Integrated Public Use Micro data Series (King et al. (2010)).[49] After 1996, we exclude

[49]The data can be obtained here: https://cps.ipums.org/cps/.

observations with imputed migration data to avoid complications arising due to changes in CPS imputation procedures.[50]

A.3 American Community Survey

A.3.1 Construction of ACS Sample

Since the American Community Survey Public Use Microdata Sample (ACS PUMS) allows to measure migration at a more granular geographical level than U.S. states, we proceed to construct to an ACS PUMS sample on which to test our model. While Public Use Microdata Areas (PUMS) are the geographical unit used for the individual records in the ACS PUMS, the migration variables are associated with Migration PUMAs (MIGPUMAs), which are constructed from one or multiple PUMAs. Prior to 2012, the ACS PUMS used PUMAs based on the Census 2000, whereas Census 2010-based PUMAs were introduced to the ACS PUMS starting in 2012 with no simple crosswalk to show the relationship with Census 2000-based PUMAs. Since MIGPUMAs are built from PUMAs, the MIGPUMAs based on both censuses are also not readily reconciliable.[51] As a result, we restrict the ACS PUMS data to the years between 2005 and 2011 inclusive to avoid complications surrounding irreconciliable PUMA and MIGPUMA geographies based on two different censuses.

To operationalize the test on the ACS sample, we need to obtain data on fertility rates and age composition of population at the MIGPUMA-level. However, such information is available at the county level and we need to come up with a method to aggregate counties to MIGPUMAs. To do this, we first map counties to PUMAs and then map PUMAs to MIGPUMAs. Let us now turn to the details of these procedures.

A.3.2 Counties to PUMAs

First, we obtain a crosswalk between counties and 2000 PUMAs from the Population Studies Center at the University of Michigan.[52] This dataset includes two variables that quantify the relationship between counties and PUMAs. The first variable, `p2cnty`, shows the percent of the PUMA population that is within a given county. The second variable, `pofcnty`, is the percent of the county population that is within a given PUMA. Thus, having both variables assigned to 1 would indicate that the county and PUMA are identical. If `p2cnty` $= 1$, while the other variable is less than 1, then the PUMA is entirely encapsulated within the county. Similarly, if `pofcnty` $= 1$, while the other variable is less than 1, then the county is located completely within a PUMA. The county and PUMA intersect in the

[50] See Kaplan and Schulhofer-Wohl (2012) for a detailed explanation.

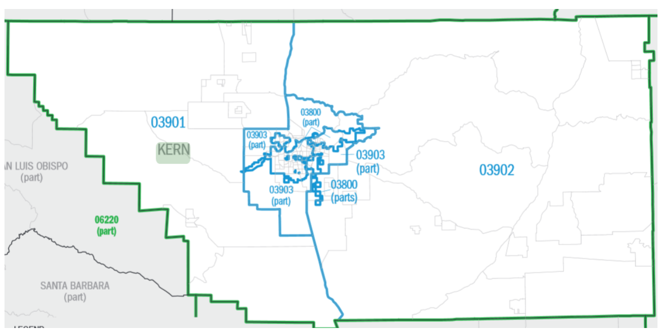
[51] For a further discussion of PUMA and MIGPUMA geographies based on both censuses, please see the IPUMS webpage on this topic: https://usa.ipums.org/usa/volii/10migpuma.shtml.

[52] Data can be obtained at http://www.psc.isr.umich.edu/dis/workshop/references/gis/puma2county.txt. Associated documentation can be found at this link.

case that both variables are less than 1. We show a few maps to illustrate various cases of the relationship between counties and PUMAs.

Figure A.3 shows an example of multiple PUMAs within a county.[53] In this case, PUMAs 03800, 03901, 03902, and 03903 compose Kern County, California, encompassing roughly 37, 19, 22, and 21 percent of the county population, respectively. Since the data that we plan to merge into the ACS PUMS data is at the county level, we have to assign the same county-level data to all four PUMAs. This requires the assumption that the county characteristics in the data are reflected uniformly throughout the entire county. If the central part of Kern County covering PUMAs 03800 and 03903 are drastically different from the outer regions encompassing PUMAs 03901 and 03902, then this indiscriminate assignment of Kern County charateristics to all four PUMAs would be inaccurate. However, since the county-level data that we plan to merge into the ACS PUMS data is not further disaggregated into a geographical level that can reconcile potential county-level heterogeneity, we have no recourse other than making the assumption of uniformity within counties.

Figure A.3 – Example of PUMAs within County



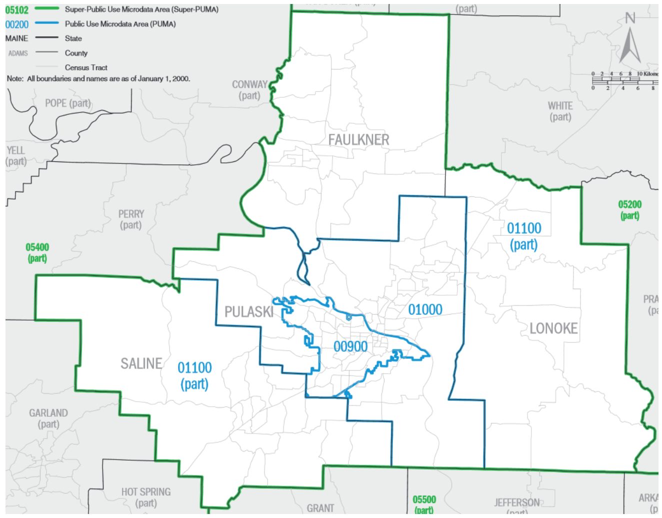
Note: Figure A.3 shows a map of PUMAs 03800, 03901, 03902, and 03903, which compose Kern County, CA. All four PUMAs are assigned the same county-level characteristics under the assumption of uniformity within counties.

Figure A.4 shows an example of multiple counties within a PUMA.[54] In this case, Faulkner, Lonoke, and Saline counties compose PUMA 01100, despite the non-contiguity of the counties. Since Faulkner, Lonoke, and Saline counties account for roughly 39, 24, and 38 percent of the PUMA population, respectively, we can safely use these as weights for aggregating county-level data to the PUMA level.

[53] The figure is a partial screenshot taken of page 32 of the California Census 2000 PUMA maps, which are found at http://www2.census.gov/geo/maps/puma/puma2k/ca_puma5.pdf.

[54] The figure is a partial screenshot taken of page 4 of the Arkansas Census 2010 PUMA maps, which are found at http://www2.census.gov/geo/maps/puma/puma2k/ar_puma5.pdf.

FIGURE A.4 – Example of Counties within PUMA

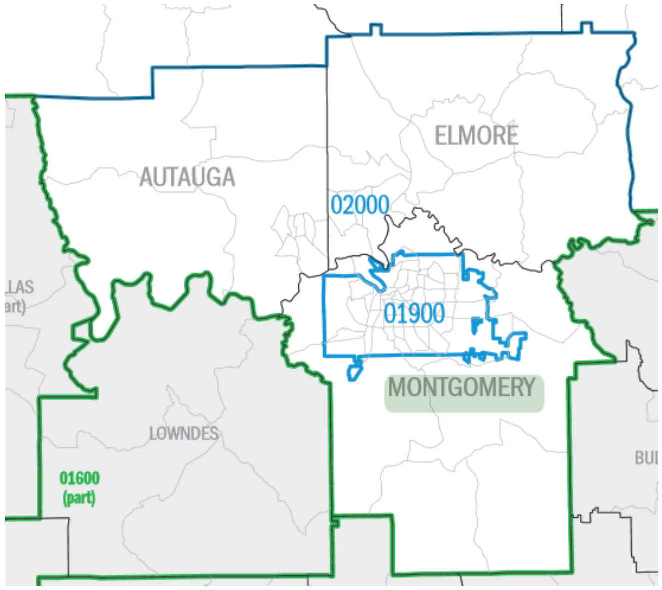


Note: The figure shows a map of Faulkner, Lonoke, and Saline counties, which compose PUMA 01100, which is not continguous. The requirement for PUMAs to be constructed from continugous counties and/or census tracts was introduced for the 2010 PUMAs, not the 2000 PUMAs. For further discussion of PUMA geography, please see the documentation from the Missouri Census Data Center located at http://mcdc.missouri.edu/allabout/geo_pumas.shtml.

Figure A.5 shows an example of the intersection between a PUMA and county.[55] PUMA 01900 is encapsulated within PUMA 2000. In this case, 90 percent of the Montgomery County population is within PUMA 01900, whereas the other 10 percent reside in PUMA 2000. Since we are aggregating from the county to PUMA, we are ultimately more interested in the percentage allocation of the PUMA population across counties (`p2cnty`). The 10 percent of Montgomery County residing in PUMA 02000 represents 17 percent of the PUMA 02000 population, whereas the other 83 percent of PUMA 02000 reside in Elmore and Autauga counties. Thus, Montgomery County is given a weight of 0.17 in the aggregation of county-level variables to PUMA 02000. Since PUMA 01990 is completely within Montgomery County, we consider both entities to have the same population characteristics. Both cases require the assumption that the population characteristics are uniform across the entire county, since the county-level variables are not disaggregated into county parts by PUMA level. If that were the case, the county-PUMA crosswalk would be extraneous. This uniform assumption would breakdown if the Montgomery County population within PUMA 01900 is drastically different from that within PUMA 02000.

[55]The figure is a partial screenshot taken of page 6 of the Alabama Census 2000 PUMA maps, which are found at http://www2.census.gov/geo/maps/puma/puma2k/al_puma5.pdf.

FIGURE A.5 – Example of County and PUMA Overlap



Note: Figure A.5 shows a map of PUMAs 01900 and 02000 in Alabama. While PUMA 01900 is within Montgomery County, the rest of the county is part of PUMA 02000.

While the breakdown of the county uniformity assumption could certainly render certain PUMA aggregates inaccurate, the bulk of the county-PUMA crosswalks do not require this assumption. Table A.3 shows a tabulation of various county-PUMA crosswalk types:

TABLE A.3 – County-PUMA Crosswalk Types

	Crosswalk		PUMA-Level Auxiliary Data	
County-PUMA Type	Frequency	Percent	Frequency	Percent
1	179	4.13	1,253	8.64
2	2,650	61.16	3,479	23.99
3	1,255	28.96	8,785	60.57
4	29	0.67	56	0.39
5	220	5.08	931	6.42

Note: This table shows the breakdown of county-PUMA types for two datasets: 1) the county-PUMA crosswalk and 2) PUMA-level auxiliary data from 2005 to 2011 (birthrate and population share data that is eventually merged into the ACS PUMS data after further aggregation to the MIGPUMA level). The county-PUMA types are as follows: 1 = county and PUMA are identical 2 = county is entirely in PUMA 3 = PUMA is entirely in county 4 = PUMA is partially in a county, where at least half of the county population is in the PUMA 5 = PUMA is partially in a county, where less than half of the county population is tn the PUMA

Since the county-PUMA crosswalk is neither a one-to-many nor a many-to-one mapping, we cannot tabulate the county-PUMA type uniquely by counties or PUMAs. Thus, we simply tabulate the county-PUMA types using the entire county-PUMA crosswalk. For the PUMA-level auxiliary data, which is the birthrate and population share data that is eventually merged in to the ACS PUMS data after further aggregation to the MIGPUMA level, the county-PUMA types are tabulated uniquely by PUMAs, since the aggregation process from the county to PUMA level does in fact leave us with unique PUMAs. Thus, in this aggregation process, we assign the maximum of the five county-PUMA types for each PUMA for cases in which a PUMA is assigned to multiple county-PUMA types. For the description of the five county-PUMA types, please see the note below Table A.3. Note that we purposely assigned the county-PUMA types in the order of increasing need for assumptions. For instance, Type 1 (county and PUMA are identical) and Type 2 (county is entirely in a PUMA) requires no assumptions for the aggregation, since the former is an identity, while the latter requires a simple weighted average. On the other hand, Type 3 (PUMA is entirely in a county) requires the assumption of county uniformity when we assign the county characteristics to the PUMA encapsulated within one part of the county. Types 4 and 5, where are cases of county-PUMA intersection, also require this assumption, but could break down even more easily given that only part of the county population resides within the PUMA. As a result, the assignment of the maximum county-PUMA type for the PUMA-level auxiliary dataset is a more conservative tabulation. However, note that the smaller percentage of Type 2 PUMAs in the PUMA-level auxiliary data tabulation compared to that in the crosswalk is solely due to the collapse of multiple counties into one PUMA entry in the dataset after the aggregation step. The crosswalk is not aggregated at the PUMA level, so it retains more Type 2 county-PUMA relationships. In fact, prior to the aggregation step of the PUMA-level auxiliary dataset, the percentage distribution of the five county-PUMA types are identical to that of the crosswalk. More importantly, Type 4 and 5 county-PUMA relationships, being fairly rare in the crosswalk at less than 6 percent of all county-PUMA matches, appears very minimally in the PUMA-level auxiliary dataset.

A.3.3 PUMAs to MIGPUMAs

Second, we obtain a crosswalk between 2000 PUMAs and 2000 MIGPUMAs from the IPUMS website.[56] Unlike the county-PUMA crosswalks, there are no intersections between PUMAs and MIGPUMAs. There either exists a one-to-one mapping between PUMAs and MIGPUMAs or a composition of several PUMAs as one MIGPUMA. Table A.4 shows the tabulation of PUMAs in various PUMA-MIGPUMA compositions, which is defined as the number of PUMAs within a given MIGPUMA. Over 37 percent of the PUMAs are mapped as the MIGPUMA itself, whereas nearly 10 percent of the PUMAs in the crosswalk share their MIGPUMA boundaries with another PUMA. Since the PUMA-MIGPUMA crosswalk

[56] Data can be found at https://usa.ipums.org/usa/volii/00migpuma.shtml.

is a many-to-one mapping, the tabulations by unique MIGPUMAs would be very different. For instance, whereas 115 PUMAs are in MIGPUMAs comprised of five PUMAs (5.55 percent of all PUMAs), only 23 MIGPUMAs are in this type of mapping (2.24 percent of all MIGPUMAs).

Table A.4 – Number of PUMAs within MIGPUMAs

	Crosswalk (by PUMA)		Crosswalk (by MIGPUMA)		ACS PUMS Dataset	
Number	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	771	37.23	771	75.00	8,859,284	42.36
2	204	9.85	102	9.92	2,102,423	10.05
3	105	5.07	35	3.40	1,040,262	4.97
4	112	5.41	28	2.72	1,011,406	4.84
5	115	5.55	23	2.24	1,035,446	4.95
6 - 15	508	24.54	59	0.06	4,472,183	21.38
$>$ 15	256	12.37	10	0.01	2,394,537	11.45

Note: The first set of tabulations uses the 2000 PUMA-MIGPUMA crosswalk, which has MIGPUMAs repeated in multiple entries if multiple PUMAs are within the MIGPUMAs. The second set is tabulated by unique MIGPUMAs in the crosswalk. The third set shows the tabulation of MIGPUMA types in the ACS PUMS dataset applied to our model.

Having constructed the crosswalks, we proceed by aggregating birthrate data from the county to PUMA level. The county-level birthrate data is based on work done by Jean Roth at NBER.[57] The prepared dataset contains birth rates, defined as births per 1,000 people, for each county from 1940 to 2005. For each year and county, we take the average of the birth rates from 20 to 34, 25 to 34, 20 to 39, and 25 to 39 years ago. We also merge in the county-PUMA crosswalk by year. In reality, we merge the county-level birth data (master dataset) with the county-PUMA crosswalk by year and county FIPS as a one-to-many merge in Stata. This requires the county-PUMA crosswalk to be replicated for each year appearing in the master dataset. Since the master dataset spans the years 2005 to 2011, the crosswalk is replicated seven times with an added `year` column in order to make the one-to-many merge by year and county FIPS. After the merge, we collapse the dataset by year and PUMA to aggregate the four county-level average birth rates to the PUMA level by computing the weighted mean of the county-level average birth rates with the variable `p2cnty` as the analytical weight. The birthrate dataset, now aggregated to the PUMA level, is saved in preparation for a final merge with the PUMA-MIGPUMA crosswalk.

In order to compute the main covariates, working-age population shares, we must first compile county-level population data. We first obtain county-level population estimates from

[57] Natality data can be found at http://www.nber.org/data/vital-statistics-natality-data.html.

2010 to 2014 from the Census.[58] We retain the observations for which the variable `year` is between 3 and 7 inclusive, as this spans the years 2010 to 2014. Note that we choose to use the July 2010 population estimates rather than the actual census figures for 2010 simply for the sake of consistency with the 2011 to 2014 data. Intercensal county-level estimates prior to 2010 are found in a separate Census dataset.[59] Since this dataset has a further disaggregation by gender, which is not needed for our purposes, we drop observations for which the variable `sex` is either 1 or 2, retaining only observations that refer to the total county population. As this dataset is in wide form, we reshape the dataset so that each record refers to a particular age group (variable `agegrp`) in a county by year. We then drop observations for which the year is 2010, since we will use the 2010 data from the 2010 to 2014 dataset. After reconciling the variable names in both datasets, we append them into a single dataset with county population data from 2000 to 2014. In preparation for the merge with the county-PUMA crosswalk, we retain only records between the years 2005 and 2011 inclusive.

We then proceed to calculate the four population shares with two choices for the working-age population (ages 20 to 59 and ages 25 to 59) crossed with two choices for the numerator (ages 35 to 59 and ages 40 to 59). Since the data is in long form, we first create indicators for total population (`agegrp`=0), ages 20 to 59 (`agegrp` from 5 to 12), ages 25 to 59 (`agegrp` from 6 to 12), ages 30 to 59 (`agegrp` from 8 to 12), and ages 40 to 59 (`agegrp` from 9 to 12). Within each county and year, we sum up the population column using the four age range indicators to calculate the total population for the four age ranges by county and year, allowing us to then calculate the four population shares. Since the calculated population shares are identical within each county and year, we keep only the observations for the total population in preparation for the merge with the county-PUMA crosswalk. Similar to the merge of the county-level birth data with the county-PUMA crosswalk, we also have to merge the county-level population share data with the county-PUMA crosswalk replicated for each year from 2005 to 2011. Following the one-to-many merge of the county-level population share data with the replicated county-PUMA crosswalk by year and county FIPS, we collapse the dataset by year and PUMA to calculate the weighted average of the four population shares with `p2cnty` as the analytical weight. Having aggregated the

[58] Data can be obtained at http://www.census.gov/popest/data/counties/asrh/2014/CC-EST2014-ALLDATA.html by clicking on the "All States" link at the bottom of the page. Associated documentation is found at http://www.census.gov/popest/data/counties/asrh/2014/files/CC-EST2014-ALLDATA.pdf.

[59] Data can be downloaded from http://www.census.gov/popest/data/intercensal/county/files/CO-EST00INT-AGESEX-5YR.csv with the associated documentation located at http://www.census.gov/popest/data/intercensal/county/files/CO-EST00INT-AGESEX-5YR.pdf. The main Census webpage on county intercensal estimates from 2000 to 2010 can be found at http://www.census.gov/popest/data/intercensal/county/county2010.html.

population share dataset to the PUMA level, we save it in preparation for the merge with the PUMA-MIGPUMA crosswalk.

In this following consolidation step, we merge the PUMA-level birthrate and population share datasets with the PUMA-MIGPUMA crosswalk. Similar to the previous merges with the county-PUMA crosswalk, we have to first replicate the PUMA-MIGPUMA crosswalk for the years from 2005 to 2011 and add a `year` column to indicate the year. After doing a one-to-one merge of the PUMA-level birthrate and population share datasets by year, state, and PUMA, we merge this consolidated PUMA-level dataset with the replicated PUMA-MIGPUMA crosswalk by year, state, and PUMA. We then collapse this dataset by year, state, and MIGPUMA to calculate the weighted average of the mean birthrates and population shares with the PUMA population as the analytical weight. The PUMA population is in the dataset as a result of the earlier merges of the county-PUMA crosswalk. Since the variable `migpuma` is a unique geographic identifier with the two-digit state FIPS concatenated with the five-digit PUMA code, we rename this as `migpuma1_code`. This prevents the confusion with the ACS PUMS variable `migpuma`, which is the five-digit PUMA code without the state FIPS.

A.3.4 Preparing the Final Sample

In the final step, we apply some final data cleaning and constraints to the ACS PUMS data before the merge with the MIGPUMA-level birthrate and population share data. To emphasize that the MIGPUMA code in the ACS PUMS data refers to the person's place of residence a year ago, we rename `migpuma` to `migpuma1`. We drop all records for which `migpuma1` is either 1 (did not live in the U.S. or Puerto Rico a year ago) or 2 (lived in Puerto Rico a year ago and now lives in the U.S.), since we do not have any birthrate and population share data for areas outside of the continental U.S. We create unique PUMA and MIGPUMA codes by appending the corresponding state FIPS with the five-digit `puma` and `migpuma` variables, respectively. Note that `st` is the state FIPS for `puma`, whereas `migsp`, which we renamed as `migsp1`, is the state FIPS corresponding to `migpuma1`. For consistency of nomenclature, the unique PUMA and MIGPUMA codes are named `puma_code` and `migpuma1_code`, respectively. Following Hurricane Katrina, the Louisiana PUMAs 01801, 01802, and 01905, were consolidated into PUMA 77777. Unfortunately, these three original PUMAs lie within two different MIGPUMAs, making it difficult to manually map PUMA 77777 to a particular MIGPUMA. For the sake of simplicity, we drop the ACS PUMS records for which `puma_code` is 2277777. We then proceed to do a many-to-one merge of the ACS PUMS data with the PUMA-MIGPUMA crosswalk by `puma_code` to obtain the MIGPUMA of each person's current residence.

In the ACS PUMS data, records with missing values for `migpuma1_code` could be the result of simply not moving within the last year, which is indicated with the variable `mig` having a value of 1. Thus, we replace the missing values of `migpuma1_code` with `migpuma_code` for all cases in which `mig` has a value of 1. This ensures that the next step of establishing the migration indicator will not be skewed towards inter-MIGPUMA migration simply due to the missing values for `migpuma1_code`. We create the migration indicator `migmove` by assigning 1 to cases for which `migpuma_code` and `migpuma1_code` are different. Lastly, we merge the ACS PUMS data with the consolidated MIGPUMA-level birthrate and population share dataset by year and `migpuma1_code` to obtain MIGPUMA-level birthrate and population share data associated with each person's MIGPUMA location from a year ago. Only two MIGPUMAS from the ACS PUMS data (3601490 and 4000690) did not have any corresponding birthrate and population share data merged into the master dataset. This final merged dataset is applied to our model.

A.4 SIPP Data

SIPP is a large representative sample of households interviewed every four months (a "wave") for two to four years. The first panel begins in 1984, and a new cohort is added around the time when the previous cohort exits. The latest wave that we use was started in 2008, and contains data for years 2008–2013. We have around 10.4 million individual-month observations between 1984 and 2013. Migration information can be constructed in all but the first wave of each panel. Table A.5 presents some summary statistics of our sample. When constructing aggregate or statewide measures, we use the individual weights provided with the survey. As explained in Aaronson and Davis (2011), SIPP is useful for studying migration behavior because it tracks households when they move to different addresses and because it contains various demographic information.[60]

TABLE A.5 – Summary Statistics for the SIPP Sample

Variable	Statistic
# Individuals	10,376,325
Married (%)	66.3
Holding a college degree (%)	25.8
In the labor force (%)	80.6
Employed (%)	77.2
Age	41.4

Note: This table shows some summary statistics of the SIPP sample that is used in the paper. Prior to 1996, we impute college attainment by years of schooling. After 1996, we observe the conferral of the degree. A person is counted as employed if they report being continuously employed for a month. A person is counted in the labor force if he is either employed or reports having looked for a job for at least one week.

[60]Data can be downloaded from http://thedataweb.rm.census.gov/ftp/sipp_ftp.html.

B Online Appendix: Model

This section describes the details of the computation of the model. In section B.1, we derive the system of equations that are used to solve for the surplus functions and the cost cutoffs for mobility decisions, c_{ik}^{j}. Section B.2 describes the algorithm used to solve for a steady-state equilibrium of the model, and section B.3 provides the derivations for the expressions used in section 3.5.

B.1 Surplus Functions

There are two types of surpluses that we need to solve for. Recall that S_i^j denotes the surplus associated with a firm-worker match in the local market, where the worker is of type j. Similarly, $S_{ik}^j(c)$ denotes the surplus in the global market associated with a firm-worker match, where the firm is located in k and the worker is of type j and located in i. These objects are defined as follows:

$$\begin{aligned} S_i^j &\equiv J_i^j + W_i^j - U_i^j \\ S_{ik}^j(c) &\equiv J_k^j + W_k^j - U_i^j - c. \end{aligned}$$

We first start with the local surplus:

$$\begin{aligned} rS_i^j &= y - b - \delta^j S_i^j - \eta\left(p_{il} + \frac{v_{ig}}{v_g}p_g\right)S_i^j \\ &\quad - p_g \sum_{k\neq i}\frac{v_{kg}}{v_g}\mathbb{E}\max\left\{0, W_k^j - U_i^j - c\right\} \\ \left\{r + \delta^j + \eta\left(p_{il} + \frac{v_{ig}}{v_g}p_g\right)\right\}S_i^j &= y - b - \eta p_g \sum_{k\neq i}\frac{v_{kg}}{v_g}\mathbb{E}\max\left\{0, S_{ik}^j(c)\right\} \\ \Rightarrow S_i^j &= \frac{y - b - \eta p_g \sum_{k\neq i}\frac{v_{kg}}{v_g}\mathbb{E}\max\left\{0, S_{ik}^j(c)\right\}}{r + \delta^j + \eta\left(p_{il} + \frac{v_{ig}}{v_g}p_g\right)}. \qquad (15) \end{aligned}$$

This final expression is achieved by first substituting in the expressions for the value functions, noticing that the worker always gets η-share of the relevant surplus (local or global) and then solving for S_i^j. The expectation is with respect to the moving cost distribution of type j worker, G^j. Turning to the global surplus function, we proceed similarly to obtain:

$$\begin{aligned}
rS_{ik}^{j}(c) &= y-b+\Delta_{ki}^{j}-rc-\delta^{j}S_{k}^{j}-\eta\left(p_{il}+\frac{v_{ig}}{v_g}p_g\right)S_i^j-\eta p_g\sum_{v\neq i}\frac{v_{vg}}{v_g}\mathbb{E}\max\left\{0,S_{iv}^{j}(c)\right\}\\
S_{ik}^{j}(c) &= -c+\frac{1}{r}\Bigg\{y-b+\Delta_{ki}^{j}-\delta^{j}S_{k}^{j}-\eta\left(p_{il}+\frac{v_{ig}}{v_g}p_g\right)S_i^j\\
&\qquad -\eta p_g\sum_{v\neq i}\frac{v_{vg}}{v_g}\mathbb{E}\max\left\{0,S_{iv}^{j}(c)\right\}\Bigg\},
\end{aligned} \tag{16}$$

where $\Delta_{ki}^{j}=\epsilon_k^j-\epsilon_i^j$ is the relative preference of the worker for k over location i.

Let c_{ik}^{j} denote the threshold for moving to location k from i for a worker of type j; i.e., $S_{ik}^{j}\left(c_{ik}^{j}\right)=0$. Note that S_{ik}^{j} is linear in c with a slope of -1. Let $S_{ik}^{j}=a_{ik}^{j}-c$. Then, $a_{ik}^{j}=c_{ik}^{j}$, so that $S_{ik}^{j}(c)=c_{ik}^{j}-c$. Substituting this into the expressions (15) and (16), we arrive at the following expressions:

$$S_i^j = \frac{y-b-\eta p_g\sum_{v\neq i}\frac{v_{vg}}{v_g}\int^{c_{iv}^{j}}\left(c_{iv}^{j}-c\right)dG^j(c)}{r+\delta^j+\eta\left(p_{il}+\frac{v_{ig}}{v_g}p_g\right)}. \tag{17}$$

$$\begin{aligned}
S_{ik}^{j} &= -c+\frac{1}{r}\Bigg\{y-b+\Delta_{ki}^{j}-\delta^{j}S_{k}^{j}-\eta\left(p_{il}+\frac{v_{ig}}{v_g}p_g\right)S_i^j\\
&\qquad -\eta p_g\sum_{v\neq i}\frac{v_{vg}}{v_g}\int^{c_{iv}^{j}}\left(c_{iv}^{j}-c\right)dG^j(c)\Bigg\}\\
c_{ik}^{j} &= \frac{1}{r}\Bigg\{y-b+\Delta_{ki}^{j}-\delta^{j}S_{k}^{j}-\eta\left(p_{il}+\frac{v_{ig}}{v_g}p_g\right)S_i^j\\
&\qquad -\eta p_g\sum_{v\neq i}\frac{v_{vg}}{v_g}\int^{c_{iv}^{j}}\left(c_{iv}^{j}-c\right)dG^j(c)\Bigg\}.
\end{aligned} \tag{18}$$

Note that equations (17) and (18) define the cost cutoffs. One the cutoffs are obtained, the value of the relevant surplus can be evaluated using (15) and (16).

B.2 Overview of the Computational Algorithm

This section describes the details of the computation used in this paper.

1. Start with an initial guess of the market tightnesses, $\{\theta_{il,0}\},\theta_{d,0}$ and vacancy shares, $\left\{\frac{v_{ig}}{v_g}\right\}_0$.

2. For each guess of $\{\theta_{il,n}\},\theta_{d,n}$ and $\left\{\frac{v_{ig}}{v_g}\right\}_n$ in iteration n:

   (a) For each worker type, solve the system of equations given by (17) and (18) to solve the cost cutoffs, c_{ik}^j.

(b) Use equations (15) and (16) to obtain the surplus functions, S_{il}^j and S_{ik}^j.

(c) Use the law of motions to solve for the steady-state values of employment and unemployment by type and location, e_i^j and u_i^j.

(d) Compute the deviation from the free-entry conditions in (7) and (8).

(e) If the average of squared percentage deviation is less than the tolerance level, $\varphi = 10^{-6}$, stop. Otherwise, update the guess for $\{\theta_{il,n}\}, \theta_{d,n}$ and $\left\{\frac{v_{ig}}{v_g}\right\}_n$ and move to iteration $n+1$.

B.3 Composition Externalities – An Example

Under the assumptions of identical job separation rates ($\delta^j \equiv \delta$) and location preferences, the population and the composition of each island is the same. Therefore, in global search, the probability of meeting a firm in a location k is equal to $\frac{v_{kg}}{v_g} = \frac{1}{N}$ for $k = 1, \cdots, N$. Using this expression, we simplify the surplus in the local labor market ((17)) of a match with a type-j worker in a location i as follows:

$$\begin{aligned} S_i^j &= \frac{y - b - \eta p_g \sum_{v \neq i} \frac{1}{N} \int^{c_{iv}^j} \left(c_{iv}^j - c\right) dG^j(c)}{r + \delta + \eta \left(p_{il} + \frac{p_g}{N}\right)} \\ &= \frac{y - b - \eta p_g \frac{N-1}{N} \int^{c^j} \left(c^j - c\right) dG^j(c)}{r + \delta + \eta \left(p_l + \frac{p_g}{N}\right)}, \end{aligned} \tag{19}$$

which is independent from the location of the worker. Thus we can denote the surplus $S_i^j \equiv S_l^j$, where l stands for the surplus from the local labor market.

Similarly, for a match between a firm in k and a worker in i in the global market can be rewritten as

$$\begin{aligned} S_{ik}^j(c) &= -c + \frac{1}{r}\Bigg\{ y - b + \Delta_{ki}^j - \delta^j S_k^j - \eta \left(p_{il} + \frac{v_i}{v_g} p_g\right) S_l^j \\ &\qquad - \eta p_g \sum_{v \neq i} \frac{v_{vg}}{v_g} \int^{c_{iv}^j} \left(c_{iv}^j - c\right) dG^j(c) \Bigg\} \\ &= -c + \frac{1}{r}\left\{ y - b - \delta S_l^j - \eta \left(p_l + \frac{p_g}{N}\right) S_l^j - \eta p_g \frac{N-1}{N} \int^{c^j} \left(c^j - c\right) dG^j(c) \right\} \\ &= -c + \frac{1}{r}\left[y - b - \left\{\delta + \eta \left(p_l + \frac{p_g}{N}\right)\right\} S_l^j - \eta p_g \frac{N-1}{N} \int^{c^j} \left(c^j - c\right) dG^j(c) \right] \end{aligned}$$

Substituting the expression for S_l^j in equation ((19)),

$$\begin{aligned}
r\left(S_{ik}^j + c\right)(c) &= \left\{ y - b - \eta p_g \frac{N-1}{N} \int^{c^j} \left(c^j - c\right) dG^j(c) \right\} \times \frac{r}{r + \delta + \eta\left(p + \frac{p_g}{N}\right)} \\
S_{ik}^j(c) &= -c + \frac{y - b - \eta p_g \frac{N-1}{N} \int^{c^j} \left(c^j - c\right) dG^j(c)}{r + \delta + \eta\left(p + \frac{p_g}{N}\right)} \qquad (20) \\
&= -c + S_l^j.
\end{aligned}$$

Therefore, we confirm that the match surplus in a meeting in the global market is also independent of the location. We define the above equation ((20)) as $S_{ik}^j(c) \equiv S_g^j(c)$.

B.4 Online Appendix: Robustness analysis

In this section, we investigate the robustness of our findings with respect to the choice of the value of leisure. Specifically, we set this value to 0.95, recalibrate our model to the same targets, and use it to quantify the effect of population aging on interstate migration.

The value of parameters calibrated outside the model are the same as in the baseline, except that we now use a higher value of leisure, $b = 0.95$. We calibrate the rest of parameters internally using the same moments used in Section 4. We report the model's fit to the targeted moments in table A.6. The estimated parameters are summarized in table A.7.

TABLE A.6 – Matching the calibration targets ($b = 0.95$)

Moment		Data	Model
Average job-finding rate		0.4160	0.4160
Annual migration rate by age group	25–29	5.26%	5.26%
	30–34	3.82%	3.82%
	35–39	2.96%	2.96%
	40–44	1.99%	1.99%
	45–49	1.98%	1.98%
	50–54	1.42%	1.42%
	55–59	1.38%	1.38%

Note: Table A.6 shows the model's fit on targeted moments of the data.

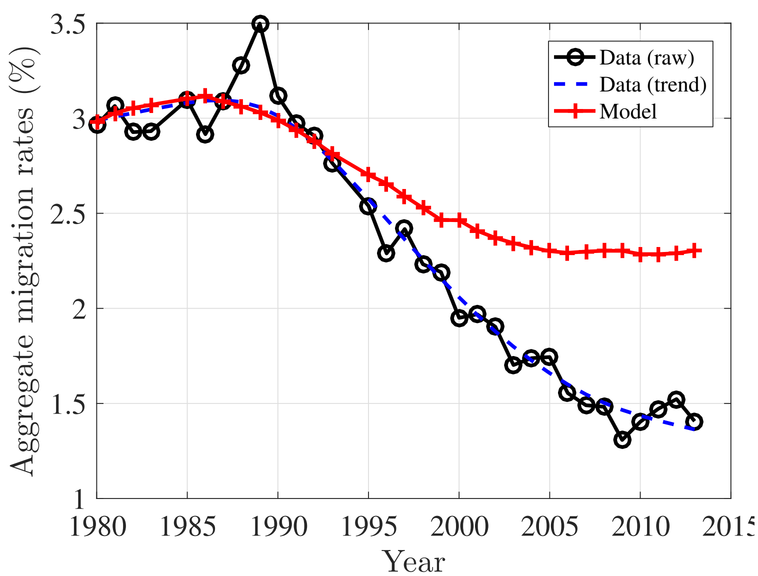
Table A.7 – Internally calibrated parameters ($b = 0.95$)

Parameter		Value
Vacancy posting cost, κ		0.0762
Mean of the moving-cost distribution by age group, μ	25–29	0.0752
	30–34	0.0899
	35–39	0.1080
	40–44	0.1822
	45–49	0.1587
	50–54	0.2457
	55–59	0.2182

Note: Table A.7 reports the values of parameters calibrated through SMM.

Using the estimated parameters, we redo our quantitative analysis and evaluate the role of aging population on migration rates. First, we compute the time series for the aggregate migration rates by changing the share of age groups corresponding to the demographic changes of the United States. Figure A.6 illustrates the model generated aggregate migration rates. Similar to the baseline, this calibration generates large effects of population aging on migration.

Figure A.6 – Aging population and the decline in migration: Data vs. Model ($b = 0.95$)



Note: Figure A.6 plots the model-implied interstate migration rates and compares them to the data. Annual migration in the model is computed as the fraction of all population who move at least once in a 12-month period. The trend component of the series is obtained with an HP filter with a scaling parameter of 100.