Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 19-49
November 2019
https://doi.org/10.21799/frbp.wp.2019.49

Fast Locations and Slowing Labor Mobility

Patrick Coate
National Council on Compensation Insurance

Kyle Mangum
Federal Reserve Bank of Philadelphia Research Department

ISSN: 1962-5361

Fast Locations and Slowing Labor Mobility

Patrick Coate* Kyle Mangum†

November 2019

Abstract

Declining internal migration in the United States is driven by increasing home attachment in locations with initially high rates of population turnover. These "fast" locations were the population growth destinations of the 20th century, where home attachments were low, but have increased as regional population growth has converged. Using a novel measure of attachment, this paper estimates a structural model of migration that distinguishes moving frictions from home utility. Simulations quantify candidate explanations of the decline. Rising home attachment accounts for most of the decline not attributable to population aging, and its effect is consistent with the observed spatial pattern.

Keywords: declining internal migration, labor mobility, home attachment, rootedness, local ties, conditional choice probability estimation

JEL codes: J61, R23, R11, C50

*National Council on Compensation Insurance. Contact: pecoate@gmail.com.

†Federal Reserve Bank of Philadelphia. Contact: kyle.mangum@phil.frb.org. We thank Kelly Bishop, Matthias Brachert, Chris Clapp, Jeff Lin, Pia Orrenius, Chris Palmer, and Danny Tannenbaum for helpful discussions at various stages of this research, as well as numerous seminar and conference participants for constructive comments. We are grateful to Solomon Tarlin for tireless research assistance.

1 Introduction

Internal migration rates in the U.S. have steadily trended downward in recent decades (Molloy et al. (2011), Kaplan and Schulhofer-Wohl (2017)). This observation is alarming for policymakers because migration is considered a primary labor market adjustment mechanism. The decline is pervasive across demographic strata, and explanations have proved elusive. Americans have typically been regarded a mobile population (Moretti (2012)), pioneers always in search of better opportunities, and there is rising concern that America has "lost its mojo" (Thompson (2016)).

This paper finds declining internal migration in the U.S. is primarily due to the increasing importance of home attachment. Migration propensity depends strongly on home attachment, or, in other words, preference for one's place of origin, and these attachments have increased on average because regional population growth has converged. Over the 20th century, the U.S. population expanded across the continent, and Sunbelt locations of the West and South grew explosively. New cities, populated by transplants, had high rates of gross out-migration because of weak attachments–hence we deem these "fast locations." In more recent decades, the population growth rates across regions have converged, and fast locations are increasingly populated by natives with high degrees of home attachment instead of weakly attached transplants. Consequently, migration out of these places has declined. Because fast locations make up an outsized share of total migrants, the national average has fallen. Thus, the decline is the result of not just individual demographics, but spatial demographics: where people are residing and from where they came.

We demonstrate empirically the role of home attachment in the national decline of labor mobility, and that it uniquely fits the spatial pattern of the decline. We begin with a descriptive analysis of migration behavior and the geographic history of population growth and home attachment. First, we show evidence that home attachment matters for migration. At all ages and skill levels, Americans living in their birthplaces are significantly less likely to migrate than transplants from other places. Moreover, those living away from home are significantly more likely to return there, indicating a *preference* for home, not selection on unobserved moving costs, drives the reluctance to leave.[1] Moreover, the *intensity* of one's home attachment predicts differential migration rates even among natives. More "rooted" natives, measured as those born to native parents, are less likely to leave (and more likely to return) than the unrooted. Notions of home attachment are not altogether new (see the literature review below), but in this paper, we demonstrate that spatial heterogeneity in home attachment largely explains the cross sectional and dynamic patterns in out-migration flows.

To understand the spatial heterogeneity, we document key facts related to local migration rates and population history. We show that local labor market in- and out-migration rates are

[1] Indeed, it is the high likelihood of returning home that drives the identification of home preference in our model, as described below.

highly correlated, and cities vary substantially in their degree of population turnover. This pattern is useful because it gives rise to our primary motivating fact: It is the "fast," high turnover locations whose out-mobility has slowed, while slower mobility cities have declined little or not at all. The distinguishing feature of fast locations is their relatively recent population growth; they are predominantly in West and South regions of the U.S., the centers of population growth in the 20th century. Knowing that home attachment is central to an individual's migration decision, we show how the history of population growth has affected their degree of regional nativity, directly and over time as successive generations grow up in these newly populated regions. We close our descriptive analysis showing the relationship between turnover and other local labor market attributes. Fast cities have more disperse (i.e., unequal) income distributions, but other major features like population size, mean income, education levels, or age composition do not explain the cross sectional facts or the spatial heterogeneity in the decline.

With these facts as foundation, we develop and estimate a structural model of migration. The model allows us to jointly account for multiple factors that affect migration propensity and to quantify the importance of home attachment. In the model, agents differ by location of birth, location of residence, age, education, and place in the income distribution. Migration is costly, with costs varying by distance. Home locations offer their natives utility premia, with the size of the premium dependent on the intensity of home attachment at time of birth. To measure the intensity of home attachment (i.e., "roots") our empirical model, we use a state-cohort matching method leveraging harmonized historical census data. Using state and year of birth, we derive the probability that an individual's birth state is the same as his or her parents. We show the measure is meaningfully predictive of migration propensity and destination choice probability among natives but–as a placebo check–not among regional transplants.

We use a dynamic discrete choice model in order to incorporate multiple dimensions of individual and spatial heterogeneity and the associated variance in opportunity sets. A nested formulation accounts for the asymmetry of home preference in move-in and move-out elasticities. The dynamic specification accounts for agents' heterogeneous continuation values, so that we can reliably estimate from a cross section the flow utility parameters free of the bias of expected future utility flows, even when pooling data from different cohorts and origins. We derive a simple linear estimator leveraging the properties of finite dependence and conditional choice probability estimation, which is, to our knowledge, the first application of these techniques in a nested logit model.[2]

The model delivers estimates of parameters governing utility from residing in one's birthplace, moving costs by age and education group and by distance, and a composite of local net income and amenities. The identifying variation comes from differences in move rates across distance, by type, and by at-home status; in particular, differences between birthplaces and cohorts in the

[2]Section 1.1 reviews related work on dynamic discrete choice model estimation.

depth of roots helps identify the intensity of home attachment. We estimate the model on cross-sectional data from the American Community Survey using method of moments on the derived estimating equations. We then simulate the model using estimated parameters and population group weights derived from the same data. We show the model fits the data well on the degree to which agents in their birthplace move relative to those not at home without resorting to assigning additional move costs to home status. The model can also generate heterogeneity in move rates across local labor markets, mainly through differences in home attachment.

The key validation of the model, which derived its estimates of primitives from cross sectional data, is its prediction of mobility rates over time outside the estimation period. Specifically, we project migration rates in previous time periods, holding fixed the primitive parameters and varying the economy's attributes, such as population sizes by age and education group, at-home status and birthplace, depth of roots, and income opportunities. The model simulation projects a migration decline in line with the actual time series and consistent with its spatial heterogeneity, with fast locations declining the most and slow locations the least. Note that the model accomplishes this without changing any primitive parameters (such as move cost or the utility premium of home attachment) but only through changing environmental factors like the incidence and intensity of home attachment in the population.

Several factors contribute to the aggregate decline, and the model permits a decomposition of the sources to quantify the contribution of each. Demographic factors matter for the aggregate, but explain only a fraction of the decline and cannot rationalize the spatial pattern. Population aging is a relevant factor, but the aging effect is mitigated by increasing educational rates and relatively larger population growth in faster locations. Rising home attachment, however, can explain the majority of the decline and fits the spatial pattern. From the data, the rising rates of nativity in fast locations (and in some of these, increases in the intensity of attachment) cause the model to predict a steady decline in mobility out of these places. Because these make up an outsized share of migrants, the rise in home attachment accounts for a majority (roughly two-thirds) of the national decline. Changes in income opportunities, mostly in fast locations, explain the balance and also account for some outlier cases. We conclude that the model can accurately depict the changes affecting migration rates over the past several decades across the geography of the U.S.

The quantitative model indicates home attachment is empirically important in the migration decline, but it is limited to the observed spatial distribution of population and birthplace source.[3] To gain further intuition for the mechanism, we use a stylized version of the model to study the role of home attachment in migration behavior in an economy facing shocks to location attractiveness. This simulation exercise provides a laboratory environment that allows us to trace

[3]Put another way, home attachment is exogenous to the individual in the estimated model, but in the economy, spatial distributions of population evolve endogenously in a path-dependent manner.

the endogenous time profile of home attachment and its impact on migration propensity under a controlled set of location shocks. The stylized model features home attachment that varies according to each location's history of population ("roots") in comparison to environments of fixed home attachment or none at all. Amenity shocks generate population reallocation that eventually arrives at a new steady state, but the transition is prolonged by home attachment, particularly when the intensity of attachment is evolving with the path of population. Moreover, endogenous home attachment generates a persistent impulse to gross migration rates that does not resolve until long after the population reallocation has settled. We suggest that this profile fits the experience of the U.S. economy in the 20th century, which was characterized by continental expansion and elevated migration rates followed by a decline.

Synthesizing the descriptive evidence and its interpretation through the model, we see a nuanced connection of population change and migration. When economists think about migration, they typically presume the spatial equilibrium paradigm–migration as an equilibrating force, with households leaving bad places for better ones. This is accurate in a sense, but like many other labor market flows (see, e.g., Davis et al. (2012)), the gross flows far exceed the resulting net changes, and the degree of place attachment (such as home) largely determines the ratio of gross to net. Our analysis indicates that past population changes affect this degree of place attachment, and thus migration flows and population changes are jointly causal in an autocorrelated process. This perspective is important for interpretation and evaluation. Concerns about the decline in gross mobility have presumed that it will limit needed population adjustments. In this paper, we show such a concern is the "tail wagging the dog," as it is actually the long run population convergence that has driven the gross mobility decline.

We conclude with an evaluation of the migration decline in light of these findings. Rising home attachment is, in principle, a friction to the labor market in that it reduces elasticities to local economic shocks, a point we can easily make within our model. This is potentially a threat to labor market efficiency–the specter of "lost mojo." But we do not actually see direct impacts on the ability of the economy to reequilibrate. We emphasize that the decline is not even observed in relatively bad labor markets; to the contrary, it is the growing local labor markets that are declining in out-migration. Moreover, trends in annual net population reallocation are smaller than and predated by the long run trends in population growth.

The fundamental question for welfare is whether the migration decline represents an increase in barriers or a decrease in the economic incentives to relocating. Are frictions increasing and preventing moves that households would otherwise like to make, or do households no longer have to incur the costs of relocating to reach their optimal location? Home attachment falls most nearly under declining incentives to relocate. Proximity to home, friends, and family is an idiosyncratic nontraded good, a form of horizontal location quality, and increasing fractions of households are finding it optimal to stay in place.

1.1 Related Literature

There are several existing studies documenting the decline in geographic labor mobility. Fischer (2002) notes that in the U.S., migration (i.e., moving one's local labor market) peaked around the 1970s and 1980s, while residential mobility (moving house within a labor market) had been steadily trending downwards for much longer. After Fischer, much of the initial work that emerged during and soon after the Great Recession (Molloy et al. (2011), Cooke (2011), Cooke (2013), Kaplan and Schulhofer-Wohl (2017)) emphasized the secular nature of the decline, finding compositional and cyclical explanations insufficient in magnitude and scope, although somewhat important in their own rights.[4] The secular trend was provocative and puzzling. How could the "death of distance" coincide with geographic sclerosis? As labor mobility is thought to be one of the primary shock-adjustment mechanisms for regions[5] and individuals,[6] a natural concern arose that low mobility will result in spatial mismatch and lower aggregate productivity.[7]

Several studies have offered explanations, although none has addressed the spatial pattern we emphasize. Cooke (2013) associated coincident trends like the rise in dual-earner households and improvements in information technology that rendered migration unnecessary. Kaplan and Schulhofer-Wohl (2017) argued that advances in travel and information technology have improved the signal-to-noise ratio in household-location matches, making migration more efficiently targeted and consequently less frequent. Several studies suggest that changes in the labor market have altered migration incentives or opportunities.[8] Kaplan and Schulhofer-Wohl (2017) also argue (in a second component to their explanation) that returns to occupations have become more similar across space, causing migration to be less necessary. Partridge et al. (2012) asked whether location differences are more sufficiently priced so that quantities will not adjust, although ultimately they find reduced sensitivity of flows to spatial differences.[9]

One strand of this literature explicitly ties geographic mobility to job mobility, since most geographic moves also involve job changes (Molloy et al. (2014), Molloy et al. (2017)). This places the study of migration in the larger literature on declining labor market dynamism,

[4]To our knowledge, the only study to parse the migration changes across space was Frey (2009), although Frey focused on the cyclical dynamics of net migration in the early 2000s instead of the long run secular trend in gross migration in focus in our analysis.

[5]See Blanchard and Katz (1992), Bound and Holzer (2000), Carrington (1996), Zabel (2012), Hornbeck (2012).

[6]See Topel (1986) or Kennan and Walker (2011).

[7]It is also worth noting that a substantial literature is devoted to understanding why labor mobility is slow or stagnant and not always in the expected direction (see, e.g., Sjaastad (1962), Lkhagvasuren (2012) Yagan (2019) Notowidigdo (2011) Autor et al. (2013)).

[8]As two examples, Karahan and Rhee (2014) suggest that an aging population could have general equilibrium effects on migration by causing firms to recruit workers (of all ages) locally instead of nationally. Hood (2013) suggests that labor market shocks are becoming more similar across space.

[9]Partridge et al. (2012) studied net population movements, not the gross migration rates in focus in our paper, and did not directly incorporate the history of population change as we do. Nevertheless, we reach a conclusion they suggest, that the economy is moving to a "new long run spatial equilibrium" of sorts. Our illustrative model in Section 6 makes this point explicit.

which has found reductions in job mobility, flows in labor market status, firm growth rates, and entrepreneurship.[10] Our paper stakes out a distinctly geographic position and makes no direct connection to other forms of labor market dynamism. However, we hope that our findings, or our technique of leveraging local labor market heterogeneity to study broader trends, is informative for that literature as well.[11]

In emphasizing geography, our paper takes a broad perspective on the forces influencing migration decisions. The economics literature on migration has progressed from studying purely pecuniary incentives[12] to incorporating non tradable amenities,[13] idiosyncratic preferences, and move costs.[14] This includes recent work relating migration to home attachments[15] and social capital,[16] which also have substantial literatures in the population sciences outside economics.[17] Our study is the first to show how long run population dynamics affect local attachments.

Finally, we offer the first quantitative dynamic spatial model studying declining migration. Models of migration are properly understood as dynamic decision problems (Sjaastad (1962), Topel (1986), Kennan and Walker (2011)), but in order to explain the spatial pattern, we need more geographic and demographic detail than a stylized model can offer. To focus on the estimation of key primitives, our approach utilizes a structural, partial equilibrium model of location choice, with a nested formulation inspired by Monras (2018). We contribute to the estimation of dynamic choice models using conditional choice probability (CCP) estimation.[18] We derive for our model a linear method of moments estimator that is highly tractable despite a large type space. This is, to our knowledge, the first implementation of CCP estimation on aggregated choice data,[19] and the first application via a nested logit model.[20] We discuss the model estimation in detail in section 4 and Appendix C.

[10]See, for example, Molloy et al. (2016), Davis et al. (2012), Davis and Haltiwanger (2014), Decker et al. (2014), Decker et al. (2016), Hyatt and Spletzer (2013), Hyatt and Spletzer (2017).

[11]Some of these studies (Molloy et al. (2016), Molloy et al. (2017), and Decker et al. (2014)) describe differences across states in rates of job mobility and changes thereof with population inflows. The patterns do not align too closely with our findings, although we focus on outflows and there are asymmetries in flows that may be relevant.

[12]See Greenwood (1975) and Greenwood (1985) for reviews.

[13]For a seminal paper on amenities and migration, see Graves and Linneman (1979).

[14]See, for example, Kennan and Walker (2011), Bayer et al. (2009), Moretti (2011) Coen-Pirani (2010), Lkhagvasuren (2012), or Diamond (2016).

[15]See, for example, Dahl and Sorenson (2010), Kennan and Walker (2011), Coate (2014), Zabek (2018)

[16]See Carrington et al. (1996), Glaeser et al. (2002), David et al. (2010), Kan (2007), Alesina et al. (2015), Falck et al. (2012), Hotchkiss and Rupasingha (2018)

[17]For some examples, see Dawkins (2006), Michielin et al. (2008), Mulder and Malmberg (2014), Belot and Ermisch (2009), Clark and Lisowski (2019).

[18]For seminal work one methodologies, see Hotz and Miller (1993) and Arcidiacono and Miller (2011). For recent applications, see Bishop (2008), Ma (2019), and Davis et al. (2017).

[19]Though not explicitly characterized as CCP estimation, Artuc et al. (2010) use algebraic manipulation of a choice value function to derive an estimating equation, which bears resemblance to our approach, although their identification method is different.

[20]The results of Arcidiacono and Miller (2011) apply to any generalized extreme value distribution, and they specifically use nested logit as an example, but we are not aware of any applications of outside of conditional logit. We show that with some additional algebra, one can still reap the computational benefits of finite dependence.

2 Empirical Facts

We first introduce a set of novel empirical facts that help motivate, formulate, and preview the results of the model.

2.1 Data Overview

This study relies on an assemblage of data from several different publicly available sources. We briefly describe them here and leave details to appendices. Our migration data principally come from two sources, the American Community Survey (ACS, obtained from Ruggles et al. (2019)) for 2005 to 2017 and the migration flows tables from the U.S. Treasury's Internal Revenue Service (IRS) for 1991 to 2016. The ACS reports the respondent's current and one-year ago Public Use Microdata Area (PUMA) of residence, from which we can elicit migration probability (move or not) and direction (origin-destination pairs). The IRS infers migration events from changes in the address on individual tax returns in two successive years, publishing the total county-to-county flows in each year, as well as the total stayers in, inflows to, and outflows from individual counties. The IRS data underwent a change in method in the 2011-2012 tax year that resulted in noticeable differences in the sample represented. We present the data for the period 2012-2016 but only rely on the consistent sample of 1990-2011.

The ACS and IRS data are complementary. The ACS provides information on the respondent's personal characteristics (including birthplace), but unfortunately it has not been collected long enough to infer trends in migration flows.[21] The IRS data provide rich spatial detail at a longer horizon but not individual or household characteristics. We use the IRS for documenting trends and their differences over space but the ACS for analyses by person type and home status, and for the estimation of our quantitative model.

We also employ two sets of historical census data to describe the evolution of places. County level population counts are used to show population growth trends. Census microdata, containing place of residence and place of birth, are used to construct measures of home attachment. We utilize decennial census data from 1880 to 2010.

A fundamental issue in a study of migration is the definition of location. In this paper, we introduce a unit of analysis we term a local labor market (LLM), which is a close cousin to a commuting zone (CZ). Both fully partition the U.S. and delineate metropolitan agglomerations. Most of our analysis focuses on urban areas. We derived LLMs from CZs and modified when necessary to achieve constant boundaries or closer definitions of current metro areas. We can map PUMAs and counties into LLMs for each year of data.[22] Migration is defined as exiting

[21] Observation of a time trend in the ACS is further complicated by the cyclical features dominating the available years.

[22] The mapping files for each decade are available to other researchers; see our websites for more details.

one LLM for another, while moves within counties or PUMAs, or among counties or PUMAs of the same LLM, are non-migration events.

We use LLMs instead of U.S. states in order to more closely correspond to a local labor market. Different cities within states have different rates of mobility, offer heterogeneous labor market opportunities and amenities, and may have materially different population compositions.[23] However, we are confined to use the census measure of birthplace, the respondent's state of birth. We allow residents of LLMs that cross state boundaries to be "at home" if they were born in any state comprising the LLM.

Further details on data construction are discussed in Appendix E, and LLM definitions and summary statistics are contained in Appendix F.

2.2 The Importance of Home Attachment in the Migration Decision

A critical fact to establish at the outset is that "home" occupies a special status in the choice set; that is, home offers a utility premium not available elsewhere. There is a substantial literature in the social sciences documenting the importance of home, broadly defined, in determining migration decisions, so while this idea is not new, we show that the measures of home available in the ACS and census are predictive of migration propensity in the expected way.

Table 1 uses ACS data to report annual mobility rates by age and education in total and disaggregated by birthplace status. Some well-known patterns appear: The young are more mobile than the old, and college educated are more mobile than noncollege, especially in youth. But among all categories, there are major differences by birthplace status: Those living away from their birthplace are an order of magnitude more likely to migrate than those at home. The foreign born more closely resemble those living at home, although among the college educated, mobility rates for the foreign born are somewhat closer to the away-from-home rate.

It is important for interpretation to understand whether the difference in move rates by at-home status is due to an actual utility-enhancing component–a *preference* for home–producing strong attachment to the place, or if the gap between columns 2 and 3 is due to selection on unobserved move costs among those who have never left their initial places. In terms of a location decision model, the distinction will clarify whether an evolving spatial distribution of population affects the observed distribution of move costs or the value of opportunities in the choice set. If home is valuable, then it will be chosen more frequently even when not already living there.[24]

In the remainder of Table 1 we examine rate of moving home. Column 5 reports the likelihood of returning home: the conditional choice probability to moving into one's birth state when

[23] California, a state that looms large in our analysis, is a prime example. Los Angeles, San Francisco, and Bakersfield, for instance, have population compositions and incomes substantially different from one another and have declined in mobility at different rates.

[24] As at least prima facie evidence, in odds ratio models of move/stay and destination choice conditional on moving, the coefficient on birthplace is an order of magnitude larger for inflows than for (lack of) outflows.

Table 1: Move Rates by Age, Education, and At-Home Status

Education/ Age	Move-Out Rate (%) Total 1	In Birthplace 2	From Other US 3	Foreign-Born 4	Move-Home Rates: Conditional Choice Probability (%) Actual 5	Synthetic 6	Log Odds Ratio, Homers to Others 7
Noncollege							
20s	5.79	4.41	12.01	4.12	35.80	4.71	3.42
30s	3.52	2.62	6.89	2.66	29.97	4.48	2.84
40s	2.35	1.68	4.40	1.77	25.63	4.18	2.86
50s	1.91	1.29	3.41	1.53	23.45	3.86	2.84
College							
20s	10.02	7.33	15.02	9.89	31.32	4.84	2.69
30s	5.07	3.09	7.57	5.68	25.77	4.45	2.03
40s	2.56	1.52	3.77	2.70	18.65	4.08	1.86
50s	2.20	1.41	3.16	2.03	17.80	3.73	2.12

NOTES: The table reports mobility rates by birth place; i.e., whether the LLM is in one's birth state or not. In column 7, "Homers" is shorthand for a native born of an LLM living outside the LLM. (Source: ACS data.)

living away from it and migrating somewhere. Roughly one-fifth to one-third of moves are such returns. For comparison, column 6 reports the synthetic probability of moving home if choosing destinations at the probability of the general population of all migrants.[25] These are all under five percent, indicating that the high frequency movements home are not a coincidence of market size, but because of special status in the mover's choice set.

One alternative possibility to consider is that if migration networks are relatively local and any mover does not migrate far from her initial location, then perhaps the rate of choosing home is not a purposeful return but a coincidence of home being nearby. In Column 7, we condition on a worker's origin and take the log odds ratio of "homers," natives from a destination, relative to nonnatives. Column 7 reports the median of this ratio to be on the order of 2 to 3.5, meaning natives are 5 to 27 *times* more likely to choose their home location than nonnatives migrating from the same origin. Clearly the move-home propensity is not a coincidence of geography.

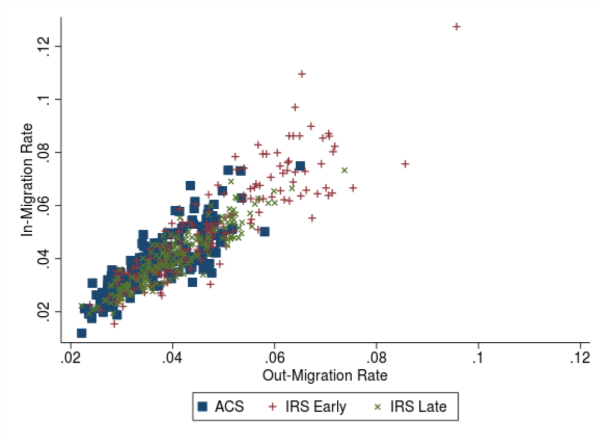
These patterns provide evidence of the existence of a home attachment via a utility premium. This explanation is important to bear in mind as we review the spatial pattern of the mobility decline and the evolution of population that preceded it.

2.3 Fast Locations Drive the Migration Decline

The decline in migration has been documented by several studies using different datasets and definitions of migration, and we do not attempt here a full rehashing of all components. Our aim is to document a fact new to the literature–that the migration decline has centered on the highly mobile cities. Afterwards, we describe some key features of these mobile locations, including the potential role of home attachment.

[25] In other words, we are adjusting the probability for the relative sizes of the LLMs. For example, because of its size, there are mechanically more New Yorkers living about the country, and also New York is a poplar destination.

Figure 1: In and Out-Migration by Local Labor Market



NOTES: The figures plot in- versus out-mobility rates for U.S. LLMs. IRS Early refers to IRS data for the period 1991-1993, IRS Late is that for 2009-2011, and ACS pools over 2005-2017. (Source: IRS and ACS data.)

As an initial matter, we first clarify the language of "fast" vis-a-vis "slow" locations. Gross migration rates are correlated–locations with high degrees of inflow also exhibit high degrees of outflow–and there is a large variance across places in the degree of turnover.[26] The use of "fast," for example, is a categorization strictly on the basis of population turnover (although we will describe other associated features of fast and slow places). Figure 1 displays scatterplots of out-mobility to in-mobility rates for 183 urban LLMs using the ACS data and the IRS data for 1991-1993 ("early") and 2009-2011 ("late"). The strong positive relationship is evident, with correlation coefficients near 0.9. Moreover, one can show (though we omit for brevity) that the rates are highly persistent year over year, and the inflow/outflow correlations maintain when splitting the population into subgroups such as age and education.[27] There is a large dispersion in turnover rates, with fast places turning over two to four times the number of residents per year as slow places.

The fast/slow distinction is important for introducing our primary new empirical fact: The decline in mobility is occurring primarily among fast LLMs. The plot on the left of Figure 2 compares the change in mobility rates in the IRS (from early to late) with the initial (early) mobility of the LLM.[28] The bubbles are proportional to LLM population size. The plot shows a clear negative correlation–the size of the decline is strongly related to the LLM's initial mobility level for markets of various sizes.

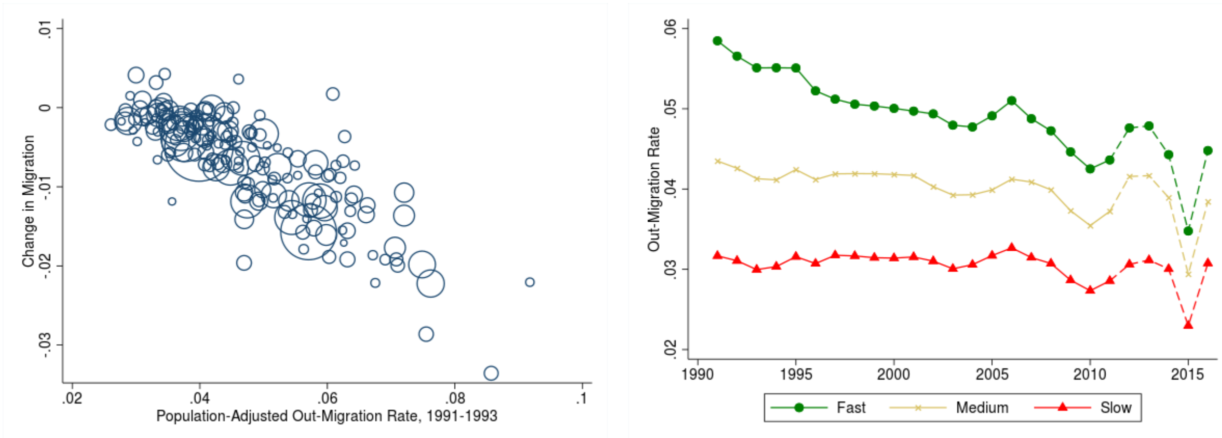
To see the complete time series, the righthand plot in the figure shows the annual out-

[26] This fact is not new; see seminal work by (Ravenstein (1885), Sjaastad (1962), Miller (1973), to more recently, Coen-Pirani (2010) and Mangum (2016)).

[27] Coen-Pirani (2010) reports a high degree of correlation of in- and outflows even when using fine subgrouping cells, including occupational categories, at the geography of U.S. state.

[28] Here we use a market size-adjusted mobility index. The index is simply the residual from a regression of average outflows on population size, which is inversely related. The raw data show a similar pattern, but this index is more precise.

Figure 2: Changes in Out-Migration Rate Over Time



NOTES: The left figure shows the change in metro level out-mobility compared to initial mobility rates; bubble size is proportional to metro population. The right figure shows the annual out-migration rates by initial mobility tercile. The data become dashed series 2012 and following to reflect a break in the method the IRS used to define move rates. (Source: IRS data).

migration rates for LLMs split into terciles by their mobility rate–fast, medium, and slow. The most mobile third of cities show a strong downward trend, dropping from about 5.7 percent to 4.6 percent from 1990 to 2011 (a 21 log point decline). The change for the middle third was much smaller, declining from about 4.3 to 4.0 percent (a seven log point decline). The least mobile third showed essentially no decline.[29]

We have also used the geographic detail in the IRS data to parse the migration flows network of origins to destinations, allowing us to check whether the slowing out-mobility is associated with lower inflows to particular destinations. See Appendix A for more detail. The networks of higher and lower mobility cities are not identical, and virtually everywhere, the plurality of migration is within region. However, the trends out of each origin local labor market are remarkably similar to all types of destinations: to near and distant locations, to small or large LLMs, to common destinations and the infrequently visited. Thus, the declines are clearly a general slowing from the origins.

Understanding the decline in outflows from fast locations is then essential for studying the national mobility decline. A simple accounting exercise helps fix ideas. By definition, the fast tercile of locations make up one-third of population but a greater share of out-migrants, about 44 percent of them in the early 1990s. Projecting how many migrants there would be in 2010 if all cities moved at 1990 levels, and taking the difference from actual as the number of "lost migrants," indicates that this one-third of places is responsible for 64 percent of the national decline.[30] In contrast, slow locations account for only 13 percent of the decline.

The spatial pattern invites two natural questions: What is different between fast and slow

[29]The differences in LLM-category trends are statistically significant by standard measures.

[30]As a particularly notable example, the cities of California make up 31 percent of the lost migrants, and Southern California alone - Los Angeles, Riverside/San Bernardino, and San Diego - makes up 18 percent.

locations? And then, what is changing?

2.4 Fast Locations: Centers of Continental Population Expansion

Which locations are fast? There is a strong regional component to current migration speed. Table 2 lists the location of fast, medium, and slow LLMs by region of the U.S. defined so that each region comprises one-fifth of the current U.S. population (see Appendix F for details). These closely correspond to the standard four census regions, but with states of the Rocky Mountains, Central Plains, and Southwest cut out as a fifth region denoted the "Frontier."[31] It is clear that western and southern states dominate the fast locations. For instance, 87 percent of the West's population lives in a fast LLM, but there are only two fast LLMs in the Northeast and Midwest combined.[32]

Table 2: Regional Location of Fast, Medium, and Slow LLMs

	Number of LLMs				Share of Regional Population			
Region	Fast	Medium	Slow	Total	Fast	Medium	Slow	Omitted
Northeast	2	8	18	28	0.09	0.22	0.63	0.07
Midwest	0	22	29	51	0.00	0.19	0.55	0.26
Southeast	16	14	10	40	0.24	0.32	0.14	0.30
Frontier	16	15	4	35	0.31	0.32	0.04	0.33
West	27	2	0	29	0.87	0.06	0.00	0.06

NOTES: See Figure F1 for regional definitions. The omitted category collects select excluded LLMs (military and college towns, as described in the main text) and rural areas unclassified as LLMs. (Source: IRS and census data.)

What is different about fast places? In short, these are LLMs and regions with a recent history of high population growth. The lefthand plot of Figure 3 plots decadal population growth rates from 1880 to 2010 for LLMs of each mobility tercile. Fast locations grew markedly over the last century. Cities such as Los Angeles, Phoenix, and Las Vegas burgeoned from small outpost towns with just a few thousand residents at the start of the 20th century to major urban areas at its end. It is especially apparent that population growth in fast cities peaked in the post war period and declined sharply thereafter, though they still are growing somewhat faster than medium and slow places.

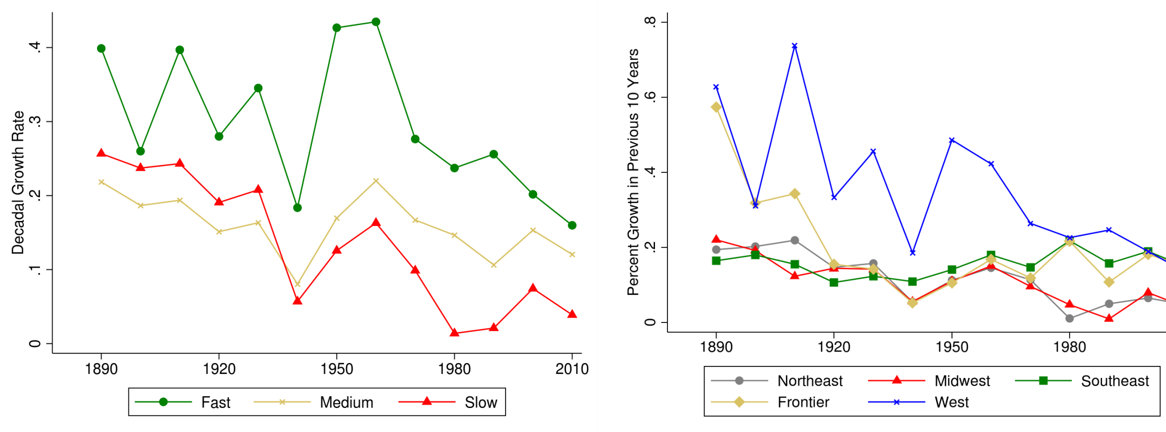
As the examples suggest, there was a major regional component to the growth trends. The 20th century was a time of continental expansion.[33] The righthand plot of Figure 3 shows population growth by region. For much of the 20th century, population growth in the Frontier and especially the West drastically outpaced the national as a whole, as these regions transformed from sparsely populated desert to urban growth engines. The Midwest and Northeast lagged

[31] Using this regional characterization, we can compare the spatial pattern of the decline in alternative datasets containing other geographic definitions. In Appendix A, we report migration by as measured by the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC, or commonly, the March CPS).

[32] These are Washington DC and Manchester, New Hampshire.

[33] See Chinitz (1986) for a discussion of American regional transformation.

Figure 3: Population Accumulation Over Time, by Region and by LLM Mobility Rates



NOTES: The figure reports decadal population growth rates by LLM speed category (left) and region (right). Definitions are detailed in Appendix F. (Source: Census county population estimates, harmonized over time by Manson et al. (2018).)

throughout. The South emerged as a population destination in the latter half of the century, and the Frontier region retains a relatively high rate of growth.[34]

Expansion to these new regions was relatively sudden (by historical standards), as new technologies made developable areas that were previously too remote or difficult to inhabit at a large scale. The reasons for the growth of these areas are varied, but a common theme is water technology. In the West, urban and rural areas developed simultaneously, as rivers were harvested for urban populations, irrigation for agriculture, and hydroelectric power for industry, including defense industries during World War II (Luckingham (1984), Reisner (1993)).[35] A notable example (with a colorful history) is the development of Los Angeles following the completion of the aqueduct in 1913 and its increasing use for urban water delivery in the 1920s. On the opposite coast (with an equally colorful history), the technology of water control aided Florida's development, as storm water was captured and swamps were drained (Barnett (2008)). Advances in transport technology and climate control made accessible new, desirable parts of the American continent (Ullman (1954), Trippett (1979), Luckingham (1984), Arsenault (1984), Glaeser and Tobio (2008)). Railroads connected the population centers of the east to the west and the Florida peninsula (Wiggins (1995)). Later in the 20th century, air conditioning played an important role. Besides making hot summers tolerable, air conditioning enabled the construction of high-density residential structures and large-scale industrial production. Overall, the technological trends meant the 20th century was a particular phase of history characterized by an opening of the American continent to urbanization at a scale not previously experienced–and one that had converged by century's end.

[34]Texas is contained in the Frontier region and is responsible for a large fraction of recent population growth.
[35]This contrasts with the slower rural-to-urban development of the wetter eastern U.S.

2.5 The Evolution of At-Home Status Across Space

The history of population growth had implications for the birthplace source composition of cities. Figure 4 uses census microdata on place of birth and place of residence to report the proportion of residents in each set of locations that are (1) born in a state represented by their current LLM ("At Home"), (2) born in some other U.S. state, or (3) born outside the U.S. Separately, a line reports the ratio of at-home population to other U.S.-born (i.e., dropping the foreign born from the calculation) to measure the share of U.S. natives who are in the LLM of their birth.

The dominant pattern in the national total in the upper left is the fluctuation in the share of foreign-born population, which compressed substantially after immigration restrictions in the early-to-middle 20th century (see Abramitzky et al. (2019)). In recent decades, an increasingly larger proportion of population growth is due to the arrival of the foreign born. However, as Table 1 indicates, understanding the "mover class"– the U.S.-born, away-from-home population– is most relevant for total migration rates.

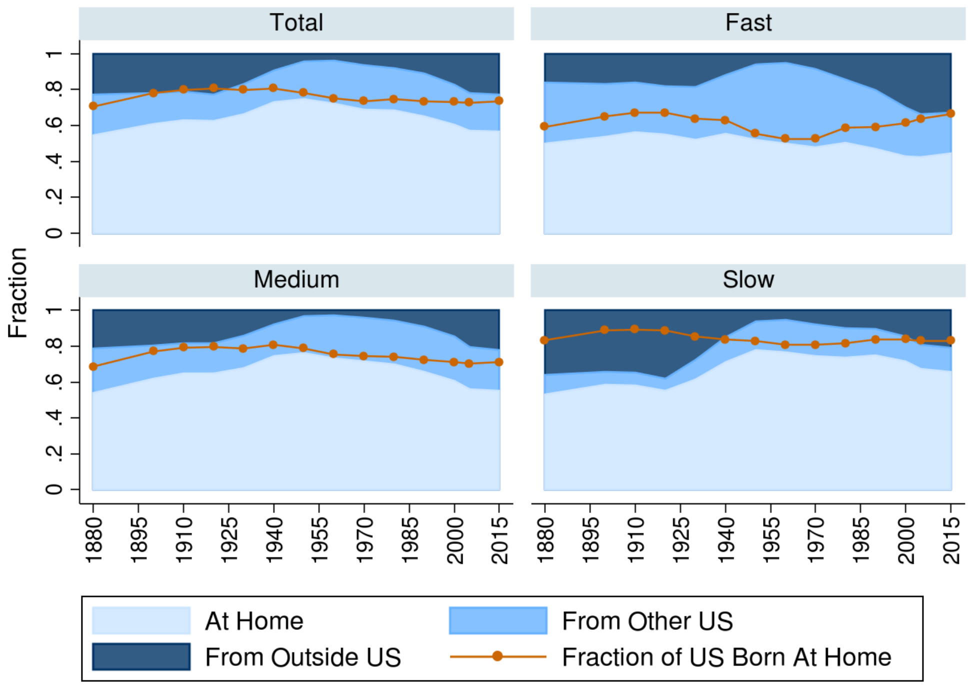
Among this group, the national pattern masks the important heterogeneity by LLM speed category. In the middle century, fast locations had a widening share of U.S.-born population sourced from other states, as the plot of the upper right of Figure 4 shows. That is, the large growth rates in Figure 3 were substantially made up of people moving from other regions and populating the West and Frontier. This growth slowed, and the population share of the at-home gradually increased, from about 60 percent in 1960 to 80 percent by 2010. As one example, in 1960, 20 percent of the U.S.-born Los Angeles residents were from California; by 2010, that share was 71 percent. Medium and slow places have tracked the national average in share of U.S. versus foreign born, and among U.S. born, medium LLMs have nearly flat at-home shares while slow LLMs have shown a small increase.

Knowing that home attachment matters greatly for migration propensity, this pattern is the critical clue in understanding the mobility decline and its spatial pattern. The population growth trends of the 20th century left fast locations with large shares of residents who were not originally from those locations, and hence they were lowly attached to these new locations and more likely to move away again. As the population growth converged, ever-larger proportions of residents were native to these cities, putting greater shares of their populations into a more attached status. To formally test this hypothesis requires a model of migration propensity, a measurement of home attachment, and an accounting of coincident factors.

2.6 Local Labor Market Attributes

With population growth history as background, we now review some of the modern attributes of LLMs in our data. Figure 4 indicated that fast locations have smaller shares of at-home

Figure 4: Population Share From Birthplace Source, by LLM Mobility Category

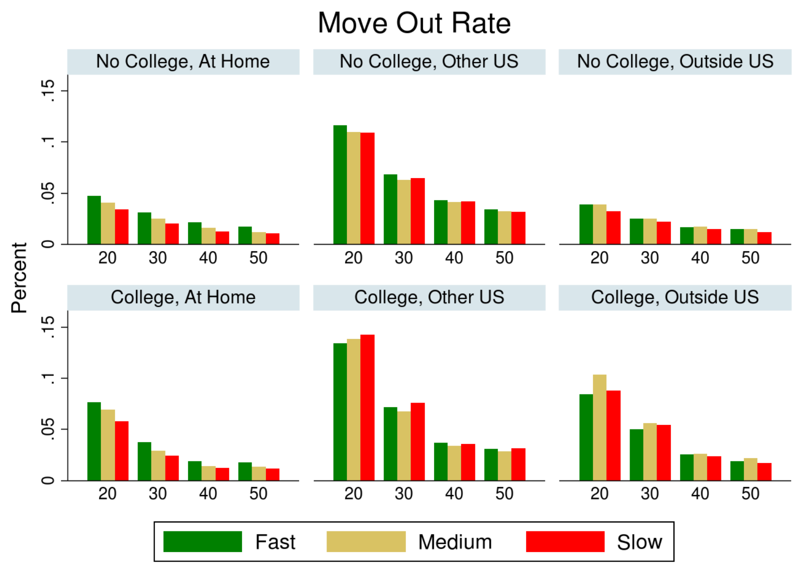


NOTES: At Home refers to living in an LLM in one's state of birth. Other U.S. refers to a birthplace in another U.S. state not covered by the LLM. Outside U.S. are foreign born from any other country (or non-continental U.S. states and territories). The figure panels are summarized by LLM mobility categorization (fast, medium, and slow). Total includes all LLMs and rural areas/excluded small cities. (Source: Authors' calculations using census data; IRS data for LLM categorization.)

residents, which, combined with the migration propensities reported in Table 1, suggests one major reason for their high rates of turnover. But looking within home status across LLMs suggests there is more to the explanation. Figure 5 plots the migration propensity by birthplace status for each age and education group, as in Table 1, but split by origin LLM mobility tercile. Within the from "Other U.S." group, fast, medium, and slow cities send away their transplanted residents at similar rates, and patterns among the foreign born are mixed. Among the "At Home" residents, however, a clear slope emerges–natives of fast locations move away at higher rates than natives of slower locations. Thus, the fast locations not only have more nonnative residents, they send away their natives at higher rates, suggesting the intensity of home attachment varies across space.

Figure 4 provides a clue here as well. The lines plotted in that figure show that fast locations have lower (but increasing) rates of the U.S.-born native to their current LLM. If home is preferable because of social and family networks, as the literature reviewed in section 1.1 indicates, stronger connections could produce stronger preferences, and hence shifting spatial populations might impact migration incentives for several generations. Given post-war U.S. population trends, a Boston native, for example, is far more likely to be born to parents who were also Massachusetts natives than a Los Angeles native is likely to be born to native Califor-

Figure 5: LLM Migration Rates by Age, Eduction, Birthplace Status



NOTES: The figure plots migration propensities by age and educational group and at-home status, split by LLM speed category. The national averages by person type are reported in Table 1. (Source: ACS data.)

nian parents. This motivates our measure of home attachment, which we will call "rootedness."[36] To build the measure, we employ the decennial census microdata with birthplace back to 1880. We define a measure of rootedness to be the probability of being born to parents native to one's own location of birth, applying the measure to current generations by matching birth cohorts for each LLM by place of residence in the child's first census. For example, a 30-something in 2010 was born in the 1970s and was therefore under 10 in the 1980 census. We take children under 10 in 1980 living in, for instance, the Boston LLM, and then, using the family relationship variables in the census, summarize the proportion of their parents who report a home state of Massachusetts. This proportion is a proxy of how highly attached to Boston is someone born there some t years earlier, even when possibly living in other locations by the 2000s. Appendix E.3 contains a detailed discussion of implementation details.

Rootedness by our definition is feasibly measured, but of course is only a proxy. Birthplace may be an imperfect measure of one's sense of "home," and besides, the measure is a cohort-matched propensity and not a directly observed object. But in fact, it is highly predictive. Table 3 shows correlations of LLM attributes with mobility rates and changes thereof, beginning with rootedness. Across LLMs, rootedness is strongly inversely correlated with migration rates, but positively with decline, meaning already highly rooted places in 1990 declined by less. Increasing rootedness over the two decades is correlated with greater declines. Moreover, rootedness is correlated with the mobility of at-home residents, but not that of the not-at-home,[37] rationalizing

[36] We certainly did not invent the terms "roots" or "rootedness,", as these are used in a variety of contexts across the population sciences, but we mean to use it here in a particular way.

[37] Appendix Figure A3 displays scatterplots of LLM move rates by at-home status to rootedness.

Table 3: Correlation of Mobility Rates and Local Labor Market Attributes

		Out-Migration Rate*	
		Level, 1991-1993	Change, 90s-2010s
Roots			
LLM Working Age Avg. Roots, 1990		-0.56	0.56
		ACS, 2005-2017	
Avg. Roots, 2005-2017	*All Population*	-0.31	na
	At-home	-0.36	na
	Not-at-home	0.29	na
Change in LLM Working Age Avg. Roots, 1990-2010		0.15	-0.37
Incomes			
Mean	Noncollege	-0.22	0.09
	College	-0.07	-0.19
Change in Mean	Noncollege	0.16	-0.07
	College	0.15	-0.16
Dispersion	Noncollege	0.28	-0.37
	College	0.20	-0.16
Change in Dispersion	Noncollege	-0.10	0.08
	College	0.09	-0.17
Population Characteristics			
LLM Size $ln(pop)$		-0.25	-0.04
Percent College		0.03	-0.18
Change in Percent College		-0.36	0.31
Percent Under 40		0.15	-0.35
Change in Percent Under 40		0.30	-0.18

*Using IRS migration data, unless otherwise noted.

NOTES: The table reports correlation coefficients of income statistics and changes to migration levels and changes. Declines in migration are negative, so a negative correlation indicates an increase in the variable is associated with a larger magnitude decline. (Source: IRS and ACS data; 1990 census county and microdata.)

the pattern seen Figure 5 and providing evidence that rootedness is a good candidate to measure home attachment. We discuss additional advantages of this measure in estimation in more detail in section 4.

Before proceeding, we check for other local labor market differences between fast and slow cities to incorporate into a formal model. For brevity, we simply summarize some of the key patterns via correlation statistics reported in Table 3.

The top panel considers the income distributions. The associations with mean incomes are mixed. LLM mobility rates are negatively correlated with mean income for the non college educated, but the change in mobility is not. Mean incomes for the college educated are weakly correlated with declines. Growth in average income is weakly correlated with mobility rates and decline. Certainly a model of location choice needs a role for local income opportunities or else risks an omitted variable problem, but given these patterns, we do not expect mean incomes to be driving the aggregate results.

There are stronger associations with income dispersion and mobility rates and declines. Places with higher income dispersion exhibit higher turnover and greater decline in migration (and among the college educated, increases in dispersion are associated with mobility declines). We take these points seriously for two reasons. First, migration is more likely among individuals at the higher or lower points of the income distribution compared with the middle. Furthermore, at all points along the income distribution, migration is more likely out of LLMs with higher dispersion, but the correlation is stronger for lower income individuals. These patterns are depicted in Appendix Table A1 and Figure A4. If fast places have more disperse and uncertain income distributions, workers in these places may face more frequent or more severe shocks to income, which could in part explain the higher tendency to out migrate. Second, changing information availability (as in Kaplan and Schulhofer-Wohl (2017)) may make it easier for workers to avoid or cope with these types of shocks. Therefore, we include in the model a role for heterogeneous income distributions across places and the ability for workers at different points in those distributions to migrate at different rates. We describe an income search process within a location after introducing the basic model.

Finally, we document population characteristics that are intuitive but actually *not* associated with the mobility decline, despite being predictive of an individual's mobility propensity. The lower panel of Table 3 shows correlations of migration levels (and changes) to rates of educational attainment and youth in the local workforce. Faster locations have no more college graduates, and actually, higher levels of education are associated with larger declines on average. Increases in college-educated share are associated with less initial mobility and smaller declines. Faster cities are slightly younger, but cities with younger workforces show larger declines. Increases in the share of young workers are associated with larger declines. Faster cities have seen smaller increases in college-educated share and relative increases in (that is, smaller declines in) the share of young workers, but relative increases in young workers are associated with larger declines.

Nationally, the U.S. has experienced an aging but more educated workforce, and given the large differences in average move rates by worker type (Table 1), we will account for these types in the model, but Table 3 shows that composition by age and education does virtually nothing to explain the spatial pattern in the mobility decline. Finally, the table shows larger cities are less mobile, on average, although there is a fair amount of dispersion among cities of a given size. There is no association in city size and mobility decline.

3 Model

To quantitatively test for the role of home attachment in determining the migration decline, we now write down a model of location choice that incorporates the key aspects described above: home utility and its intensity, local labor market attributes, and various types of workers. A dynamic discrete choice model is well suited to the task of explaining costly migration decisions for heterogeneous workers over a set of heterogeneous alternatives. Such a model is written in the tradition of Kennan and Walker (2011), broadly, with adjustments for our focus on spatial heterogeneities and to be applied to cross-sectional data in which we observe one location choice event.

3.1 Environment

The economy consists of a closed set of J distinct locations. There are a discrete M types of people, "workers," with individual types denoted by m, who each live for $A > 1$ periods. All workers are employed. Each location offers a N-pointed discrete distribution of income. We abstract from labor supply and cost-of-living differences between locales, though in the empirical implementation we adjust for the latter.

Individual workers are endowed with a home location (which may or may not be their current location) that provides them with a utility flow available nowhere else. The size of this flow utility (the "rootedness") is also endowed at time of birth and remains constant throughout the workers' lives. This flow utility is provided at any point in the income distribution and hence is not affected by the income search process detailed below. We denote utility from home, a function of rootedness, simply as $u_{j=H}(r)$.[38] The rootedness of a location affects the utility offered to its natives, and not any other workers living in the location but born elsewhere.

Time is discrete. Workers begin a period in an initial location and face the full set of J alternatives, including their origin. Within a period, two to three substages may occur. First, in the upper nest, workers decide whether to stay in their current location. Then, conditional on deciding to migrate, they choose a location. In either case, a last stage in which incomes shocks are drawn is also modeled to allow for expected income to have an impact on the location decision. Because we do not observe income dynamics jointly with location decisions, we do not estimate this component directly, although we can allow for it in simulation.

Workers face moving costs to relocating from their origin at the start of the period, whether home or not. We carefully make a distinction between moving costs and idiosyncratic preferences for particular locations. While the home premium inclines workers to prefer their birthplace *ceteris paribus*, moving costs introduce frictions. These ideas are often conflated in the literature, either as shorthand or because of data limitations.[39] It is important for our model to separate a relocation cost from an agent's preference for a particular place. Home attachment is a dimension of horizontal location quality, not an adjustment cost to relocation. This is especially important

[38] Our referring to home attachment as "preference" suggests nonpecuniary benefits, but we do not mean to limit the scope. There can be economic benefits from trusting relationships, such as child care provided by grandparents.

[39] For example, Moretti (2011) introduces a model with a distribution of location-specific preferences to study their impact on labor mobility in response to local market shocks. Bayer et al. (2009), Diamond (2016), Bryan and Morten (2019) put home state and/or region in the utility function (as we do), but refer to the effects as "moving costs." Morten and Oliveira (2016) make a more precise distinction, using relocation costs in their model but in empirics checking for robustness to use of both distance costs and specific preferences for birthplace.

in counterfactual simulations, where we alter the distribution of home preferences but keep adjustment costs fixed by assumption.

3.2 Migration: Choice Across Locations

We begin in the lower nest, assuming that a worker has decided to migrate and model the choice of location j conditional on it being different from the origin o. To account for heterogeneity between locations in mobility rates in ways not captured by home preference, income, or other characteristics, we allow moving costs to be dependent on the pair of locations, representing a generic notion of "distance," and hence we label them symmetrically $mc_{j,o} = mc_{o,j}$.

We denote types and describe the income search process below; for now, it suffices to label the common value of a location as ν_j. Workers are presented with the flow value for a location, the home utility premium if applicable, and pair-specific moving costs. Let v denote the value of making a locational choice,

$$v_{j,o} = \nu_j + u_{j=H}(r) - mc_{j,o} + \beta EV(j). \tag{1}$$

Equation (1) shows the sources of utility, the moving cost wedge, and because works live past today ($A > 1$), the continuation value from choosing j, $V(j)$. As is common in discrete choice models, we allow for a temporal idiosyncratic preference shock term distributed Type I extreme value with a variance determined by parameter λ. Preference shocks are important for rationalizing the gross flows at the focus of our analysis. Expectation E is taken over future preference shocks.[40] With these shocks, the probability of choosing destination j conditional on current location o is given by

$$Pr(j|o) = \sigma_{jo} = \frac{exp[v_{j|o}]^{\frac{1}{\lambda}}}{\sum_i exp[v_{i|o}^{\frac{1}{\lambda}}]}. \tag{2}$$

3.3 Choice of Migrating or Staying

The upper nest is the binary stay or move decision. The value of staying is consuming flow utility in the current location, $\nu_j + u_{j=H}(r)$, and being faced with the same decision next period,

$$V_{s|o} = \nu_o + u_{o=H}(r) + \beta EV(o). \tag{3}$$

[40] For simplicity, we abstract from evolving state variables like the future path of incomes, although in principle we could introduce them. As we describe in the estimation section, this is not a serious threat to bias in our results.

The value of moving is the expected value of choosing a destination optimally from $max_j\{v_{j,o}\}$ (the 'Emax'), from which using standard results is

$$V_{m|o} = \lambda ln\big(\sum_k exp[v_{k|o}^{\frac{1}{\lambda}]}\big). \tag{4}$$

The respective probabilities are

$$\begin{aligned} Pr(stay) =& \sigma_s = \frac{exp[V_{s|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}} && \text{(5a)} \\ Pr(move) =& \sigma_m = \frac{exp[V_{m|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}}, && \text{(5b)} \end{aligned}$$

where the elasticity of the upper nest is governed by δ. The expected value of being faced with a move/stay decision in some origin o gives the continuation value of locating there,

$$EV(o) = \delta ln\big[exp[V_o]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}\big]. \tag{6}$$

3.4 Income Search Within Locations

We now specify the income search process. We include this component in the model to allow for differences in income distributions between locations or over time to affect gross migration rates; in particular, allowing workers with especially low or high income draws to migrate at rates higher than those of a mean worker. The estimation described below does not rely on this feature of the model. Our goal in including an income search component is to account for potentially heterogeneous effects on gross migration of income opportunity across locations.

The worker begins the period at some point in the income distribution, y_n. Let $W(n)$ denote the utility afforded by income at some point n (suppressing location notation). Each period in every location, there is a some probability γ that the worker is contacted with a new offer. If the worker fails to get a contact, occurring with probability $1-\gamma$, she is left to a non-optional lottery where her new wage is drawn from the probability distribution $\pi_{n'|n}$. Note the distribution is conditional on her current state. The expected value of the non-optional lottery is

$$w_0(n) = \sum_{n'}^{N} \pi_{n'|n} W(y_{n'}). \tag{7}$$

Were the worker to receive a new contact, she is allowed to choose between her current income y_n and the new income $y_{n'}$. We specify this as a discrete choice subject to an idiosyncratic shock. The new income is not only temporal but a change in her state variable to enter the next period.

Hence, the choice conditional on a new contact is $max\{W(n), W(n')\}$. New offers are drawn from the same probability distribution, so the expected value conditional on making a new contact is

$$w_c(n) = \sum_{n'}^{N} \pi_{n'|n} E(max\{W(y_n), W(y_{n'})\}). \tag{8}$$

Combining these yields the expected utility from beginning the period at income state n

$$\omega_{nj} = \gamma w_c(n) + (1 - \gamma) w_0(n), \tag{9}$$

where ν_{nj} is the expected income term, a component of the commonly available location utility ν_j. The parameter γ represents the "ease of information" in that greater values provide more contacts and more options, though not necessarily higher incomes in all cases. The combined value ω_{nj} is a function of γ, current income, y_n the income available in the location, $y_{n'}$, and the distribution of income changes, $\pi_{n'|n}$.

We allow for the possibility that local and nonlocal searches are not equivalent (see also Karahan and Rhee (2014)). Specifically, we allow the probability of drawing points in the income distribution to depend on whether the income search is conducted by an incumbent resident or someone currently residing in a different location. For example, we think it reasonable that searching in one's current location is in some sense "easier" than searching in a faraway location, as even with increasing availability of information technology, local networks remain important. Notice that this distinction is based on origin, not birth location. To operationalize this idea, we allow for two distributions on new income shocks: π^{local} and $\pi^{nonlocal}$.[41] Using data on income dynamics, we can identify differences in income transitions between those who move to new locations and those who do not.

As we are interested in the heterogeneity in spatial dynamics of migration, the income search component of the model serves several purposes. The first is that income distributions have evolved in different ways across LLMs, which might impact out-mobility. From Table 3, mobility declines are somewhat associated with increasing means of the income distribution (and to a lesser extent, increasing variances). Moreover, the effects of income evolution on out-mobility could be amplified in a model with a local market bias in search opportunities. Second, income distributions are heterogenous across locations, and a common trend in information availability could affect some more than others. Specifically, it can be readily shown (see Appendix C) that ω_{nj} is increasing in the variance of income because of the presence of option value in w_c, and the value of this optionality is convex in γ. Hence, more disperse income distributions (which appear in fast LLMs, shown in Table 3) could be more affected by a change in information frictions.

[41] An alternative would be to use different offer arrival parameter γ, but this parameter is already very abstract and difficult to discipline with data. Using π^{local}, and $\pi^{nonlocal}$ allows us to treat γ as a parameter to study symmetrically across locations.

Thus we are allowing for the possibilities that heterogeneous (or heterogeneously evolving) income distributions could affect the model's predictions about migration rates. Ultimately, the quantification of the income channel–which accompanies our main story of increasing rootedness and nativity–will depend on the parameters and variance across locations. Finally, by incorporating local income opportunities in the model, we allow for potential nonlinear effects between income and home preference.

3.5 Worker Types

Up to this point, we have suppressed worker types for exposition, but the relative value of functions (4) and (6) may depend on characteristics of the worker. Birthplace and rootedness are endowed characteristics, and because we observe one location choice event, we also treat age and education as immutable characteristics. Location is a state variable–we observe workers in one state, which they can alter by choosing a new location. One concern with mixing all these types together is that they may have different continuation values of a particular choice. As we describe below, the use of conditional choice probabilities elides the solution of the model but still accounts for continuation values in a flexible way.

Equation 10 writes out the choice specific values from (1) with all state variables: origin k and the worker's endowed type m, which is her birthplace, age/cohort, and education. k changes endogenously, but m is fixed over time.

$$v_{j,m}(k,n) = \underbrace{\mu_{j,m} + \omega^{nj}(n,k,m)}_{\nu_{j,m}} + u_{j=H|m}(r_m) - mc_{k,j,m} + \beta EV_m(j,n'|n). \tag{10}$$

Note the introduced amenity parameter μ that can vary by type. Differences in amenities or cost of living (conditioning on income), and how different ages and education levels might view these, will drive location net growth (as in Gyourko et al. (2013), Moretti (2013), Diamond (2016)). While our focus is on gross migration, accounting for net migration patterns in estimation will help identify the parameters of interest because our data are conditioned on initial states.

4 Applying the Model to U.S. Data

We now describe how we take the model to data on U.S. LLMs. The estimation strategy proceeds more like a location demand model than a dynamic migration model because our data report one location choice event. While our model was written as a dynamic microeconometric model like Kennan and Walker (2011), Bayer and Juessen (2012), or Bishop (2008), the estimation uses aggregate choice probabilities (market shares) instead of, e.g., maximum likelihood estimation on panel microdata. Hence, estimation more closely resembles a location demand model (such as Bayer et al. (2009) or Diamond (2016)), which emerged from the demand estimation literature in the theme of Berry et al. (1995). We account for forward-looking behavior by exploiting the structure of the logit choice model, which has a closed form solution for continuation values in a dynamic optimization problem.[42] The generic idea of deriving estimating equations relating gross flows to adjustment costs has an antecedent in Artuc et al. (2010), although the nested structure of our model requires us to derive these in a new way, and the sources of variation in our geographic model are quite different.

We estimate the model on a single cross section of U.S. cities, using the 2005-2017 ACS data. This allows us to identify the preference parameters for the economy at that time. Then, taking the primitive parameters as fixed, we simulate the economy in previous periods as location features evolve, which we will describe in section 5.

4.1 Estimation Strategy

The main parameter of interest is the size of the home premium and its dependence on rootedness. The parameters of necessity are the move costs and location amenities, which can vary by person type, origin, and destination.

4.1.1 Utility Parameterization

The utility function seen in (10) contains the parameter $\mu_{j,m}$, which represents mean preferences for a location j held by workers of type m. In our preferred specification, we split types into eight categories, four decadal age groups, 20s to 50s, for each of college and non-college-educated workers. Attributes of the location, such as amenities or cost of living, will be subsumed in this parameter, but in estimating by type, workers of different ages or education levels can have heterogeneous preferences for these.

For home preferences, we use a simple utility function in which utility from home is an indicator variable $u_{j=h}(r) = \alpha_m I(j = h)$ or a linear function of roots, $u_{j=h}(r) = \alpha_m R_j I(j = h)$, the former being a specification check and the latter our preferred specification.[43] In either, we allow α to vary by education level.

The distinction in the specifications highlights the use of rootedness as a measure of the intensity of home attachment, for which we see a couple advantages. First, it offers a source of

[42] Bayer et al. (2016) and Davis et al. (2017) estimate a location demand model accounting for future values in a model of neighborhood choice within a single metro area. Though there are significant differences in context and emphasis, our estimation strategy bears some similarities to these in that we exploit computational savings from logit demand models.

[43] We experimented with several functional forms, and the results are roughly similar, but this single parameter specification is the simplest way to ensure a nonnegative value for home in all markets in all time periods. Adding an intercept, for instance, causes the lowest-rooted city, Las Vegas, to have a negative projected home preference. While home preference in Las Vegas appears weak, a projected inversion in move rates by at-home status is in conflict with the data.

variation that a simple at-home indicator cannot. Variation in birthplace and cohort provide identifying variation in rootedness. Some locations are more rooted than others, and some generations are more rooted than others. For example, young college educated workers may prefer to live and work in San Francisco (captured by ν) on average, but natives of San Francisco also draw a home premium from there that workers born in, say, Boston, do not. The variation in choice probabilities by birthplace identifies the parameter. Similarly, if the rootedness of San Francisco varies between the 20 and 30 year old cohorts, for example, heterogeneity in their propensity to choose San Francisco helps to identify this parameter. Second, it is more plausibly exogenous (in the microeconometric sense that it is predetermined for the agent), as it is an endowed characteristic at birth and is not subject to the person's choices the way an at-home status indicator would be. Thus, rootedness allows us to test for the presence of a home premium even without longitudinal data. We do, however, acknowledge that this will be a noisy measure of social attachment, partly because birth state may not actually measure well what one considers as "home," and partly because, even if it measures "home" well, we do not actually observe the actual rootedness of the 2010-era individual and instead match by cohort. But it is readily constructible for a wide swath of geography, and we suggest it as a reasonable proxy for the deeper concept of "attachment." If it is a meaningless proxy or just very noisily measured, then we will not detect an impact.

The estimation of home preference could be biased if other frictions are ignored, so we turn to the estimation of move costs. The move cost function has an intercept shifter for each education and age category to account for the profile of migration over the life cycle and by worker education level. Then, to account for the spatial component of migration probability, we enter the distance in miles between the MSA centroids. Since migration rates fall off with distance, we expect this term to be negative (i.e., increasing distance means less moving). We also allow a discrete shift in distance for "neighboring" LLMs (those with counties sharing a border), for LLMs in the same state, and for LLMs in the same region. We allow distance cost to vary if the destination is one's home location via an interaction of distance terms with the home indicator function.

Because there is heterogeneity in move rates by LLM even after conditioning on distance and age/education composition, in some specifications we allow for a vector J of move cost shifters for each location. We call these "toll costs" because they are paid whether entering or leaving a city. For example, a move from New York to Boston requires an "exit toll" for leaving New York, and an "entry toll" for arriving in Boston. We use a toll cost because it captures both heterogeneity between cities and the correlation of inflows and outflows.[44] Note that these will be measured conditional on heterogeneity in move rates induced by any differences in income distribution, type composition, and spatial orientation to other markets. (This specification

[44]It is further important that these are symmetric so as not to be conflated with location quality.

provides a useful robustness test of the home premium estimates, although it turns out not be our preferred model.) All together, the move cost function is

$$mc = \underbrace{\sum_{\hat{m}} I(\hat{m}) mc_m}_{types} + \underbrace{\sum_{d}^{D} mc_d d_{o,j}}_{distance} + \underbrace{\sum_{d}^{D} mc_{dh} d_{o,j=h}}_{distance\,X\,home} + \underbrace{\sum_{j}^{J} I(orig = i) mc_i}_{exit\ toll} + \underbrace{\sum_{j}^{J} I(dest = j) mc_j}_{entry\ toll} .$$

Table 4 below will compare specifications to demonstrate the importance of each component of the move cost function.

4.1.2 Estimation Method

We next describe the estimation method. The basic idea is to use the model structure to derive a set of estimating equations. Each equation applies to one choice event, and we stack equations across types and locations to estimate the parameters. Details and derivations are relegated to Appendix C.

The data leverage choice probabilities to recover the utility parameters. We estimate the model on a single cross section of data. All parameters are jointly identified, but we can loosely describe what moments of the data help to target which parameters. Variation in move rates by type, distance, origin, and destination identify the move cost parameters. Variation in net migration (in minus out) by age/education identifies the location preference parameters μ. The home premium is identified by the variation in out-mobility by natives and non-natives, and the in-migration rate for those returning home versus those entering new locations, and how each of these vary with home rootedness.

The model is dynamic discrete choice with multiple types of agents and a large number of locations (and birthplaces). Fortunately, there is a very tractable way to estimate the model. As a memory-less discrete choice specification, this model is well suited for estimation via conditional choice probabilities (CCPs). CCPs arise in the logistic model because of the mapping between the continuation value and choice probabilities (Hotz and Miller (1993)). The advantage is that the model need not be solved to arrive at parameter estimates. Instead, one needs to derive the mapping between choice probabilities conditioned on state variables, which are observed in the data, and the model's parameters. From this, we can yield an estimating equation.

Specifically, this model has a simple derivation exploiting finite dependence (Arcidiacono and Miller (2011)).[45] That is, because the choice problem is memory-less by some point s (i.e., it does not depend irreversibly on the whole sequence of choice), two disparate choices in some period t can be returned to some normalized choice by some future period $t + s$. In our model,

[45]Finite dependence and CCP estimation is also lucidly described in Bishop (2008) and Ma (2019).

$s = 1$, allowing expression of the model parameters in terms of current period and next period (i.e., the person's expected choices after aging one period) choice probabilities.

The derivation of the estimating equations is straightforward but tedious, so we relegate the details to Appendix C. The overview is that we use the mappings between (6) and (5) and between (4) and (2), subject to a normalizing location choice, in order to derive an equation relating utility parameters to choice probabilities, accounting for continuation values via iterative substitution of expected future choice probabilities.

The stacked matrix equation is

$$\underbrace{\begin{bmatrix} Y_1 \begin{cases} ln\frac{\sigma_{so}}{\sigma_{mo}} - ln\frac{\sigma_{sz}}{\sigma_{mz}} + \frac{\beta}{\delta}ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}} - \frac{(\beta-1)\lambda}{\delta}ln\frac{\sigma_{1o|m}}{\sigma_{1z|m}} \\ \cdots \\ ln\frac{\sigma_{so}}{\sigma_{mo}} - ln\frac{\sigma_{sz}}{\sigma_{mz}} + \frac{\beta}{\delta}ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}} - \frac{(\beta-1)\lambda}{\delta}ln\frac{\sigma_{J-2,o|m}}{\sigma_{J-2,z|m}} \end{cases} \\ Y_2 \begin{cases} ln\frac{\sigma_{1|o}}{\sigma_{z|o}} + \beta ln\frac{\sigma'_{m|1}}{\sigma'_{m|z}} \\ \cdots \\ ln\frac{\sigma_{J-2|o}}{\sigma_{z|o}} + \beta ln\frac{\sigma'_{J-2,|1}}{\sigma'_{m|z}} \end{cases} \end{bmatrix}}_{Y} =$$

$$\underbrace{\begin{bmatrix} \Delta_1 \begin{cases} \frac{1}{\delta} & \cdots & \frac{1}{\delta} & 0 & \cdots & 0 \\ \frac{(\beta-1)\lambda}{\delta} & \cdots & \frac{(\beta-1)\lambda}{\delta} & 0 & \cdots & 0 \end{cases} \\ \Delta_2 \begin{cases} 0 & \cdots & 0 & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\ 0 & \cdots & 0 & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \end{cases} \end{bmatrix}}_{\Delta} \underbrace{\begin{bmatrix} X_1 \begin{cases} u(x_1) - u(x_z) & mc(d_{1o}) - mc(d_{1z}) \\ \cdots & \cdots \\ u(x_{J-2}) - u(x_z) & mc(d_{J-2,o}) - mc(d_{J-2,z}) \end{cases} \\ X_2 \begin{cases} u(x_1) - u(x_z) & mc(d_{1o}) - mc(d_{1z}) \\ \cdots & \cdots \\ u(x_{J-2}) - u(x_z) & mc(d_{J-2,o}) - mc(d_{J-2,z}) \end{cases} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \theta_u \\ \theta_{mc} \end{bmatrix}}_{\theta}, \tag{11}$$

where z is a normalizing reference location that in practice we take to be the residual location. Vector Y contains the observed choice probabilities, matrix X is composed of functions of utilities and moving costs (e.g., an indicator for whether a location is home, how far two locations are from each other, etc.), matrix Δ consists scaling parameters, and θ is the vector of parameters to be recovered. In short, Y is data, X is model structure, and Δ and θ are parameters. The "1" blocks come from the move/stay probability equation, and the "2" blocks come from the move-to destination probability, conditional on moving. There are $J-2$ equations in each block, with the origin o and the normalizing location z excluded.

With choice probabilities on the lefthand side, LLM attributes and parameters on the right-hand side, estimation proceeds much like a standard regression: The data matrix is inverted on the choice probabilities to recover the estimand. An attractive feature is that the inversion accounts for the covariances in the data. For example, if shallow-rooted places happen to be younger or less remote, the covariance between move costs and home attachment is accounted for in the X matrix.

4.1.3 Scale Parameters and Move Cost Intercepts

Equation (11) identifies the main parameters of interest off of differences between locations and a normalizing locale z. Hence, scale parameters are not identified here and must be calibrated elsewhere. The set of scaling parameters include β, λ, δ, and type-specific intercepts of the move cost function. That is, equation (11) is specified in relative choice probabilities and thus identifies the move cost parameters that vary with distance between locations, but the intercepts that pin down average move rates by type must be estimated elsewhere.

First, we set β to 0.95 a priori to match an annual discount rate.

Second, we can calibrate the ratio $\frac{\lambda}{\delta}$ by using the relative differences in inflow and outflow rates to a given location for a set of workers of the same type. In order to best match the asymmetry in move-home, move-from-home probabilities, we use a ratio of parameters on a home indicator in each side of the flow equation. Details are provided in Appendix D.4.

Third, a separate estimator for the move cost intercepts[46] can be derived using substitutions similar to the derivation of (11), as detailed in Appendix C.[47] It is

$$\begin{aligned} ln\sigma_s - ln\sigma_m =& \frac{1}{\delta}\big(V_o - \lambda ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big) \\ =& \frac{1}{\delta}\big(V_o - V_j - mc_{j|o} - \lambda ln\sigma_{j|o}\big). \end{aligned} \tag{12}$$

This equation is identified by average move rates, not relative choice probabilities. The difference between (11) and (12) is that the former removed value functions to yield an equation in only choice probabilities and parameters, while the latter is forced to retain value functions we do not wish to solve or estimate. Controlling for a suite of origin and destination fixed effects–treating V_o, V_j as ancillary parameters–and entering the distance terms of the move cost function will yield an estimate of the move cost intercept for each age/education type in the data. Similarly to (11), the regression can account for covariances in the data, if, for instance, young people tend to reside more often in places far apart in space or utility gap $(V_o - V_j)$.

[46] Monras (2018) does not use a move cost term, instead calibrating an average difference in elasticities between nests, i.e. λ versus δ. We could have gone this route, although we prefer using moving costs to compare across types (who have substantially different move rates, as in young versus old) without imposing assumptions about between-destination elasticity. A move cost specification is also consistent with our environment in a model with geography, where some locations are closer in space than others.

[47] This step in particular bears the closest resemblance to Artuc et al. (2010).

4.1.4 Auxiliary Model: Income Dynamics

We calibrate the parameters governing income dynamics, π^{local} and $\pi^{nonlocal}$, using observed income changes for migrating and non-migrating employed workers in the Panel Study of Income Dynamics (PSID). More details are available in Appendix D.8.

4.1.5 Auxiliary Model: Continuation Values

In addition to major computational savings, the use of CCPs to approximate the value function has the advantage of avoiding structure on agent's expectations, which is especially important in our context, since we observe only one choice event but not how choices change as states evolve. Whatever workers might believe about the future is subsumed in the CCP term. However, when we simulate the model at counterfactual environments, we must allow the expectations to change accordingly. Instead of putting structure on the expectations, we flexibly estimate the choice probabilities and use functional projections in counterfactual CCPs. We run an auxiliary model of choice probabilities using the ACS microdata and environmental features of rootedness and income. Details are available in Appendix D.5. Note that these auxiliary models are not used for identification of utility parameters but as an input into the simulations by allowing for a substitution of projected CCPs for the expected value function.

4.1.6 Forming the Moment Conditions

With age group by education by birthplace by origin, there are $4 \times 2 \times J \times J$ cells, with J choices for each. Because spatial heterogeneity is important for our analysis, we would like to include as many locations as possible, but a larger J leads to two practical problems. First, the sample sizes for small cities become too small to reliably estimate choice probabilities. Second, because there are $J \times J$ choice probabilities and J birthplaces, the memory requirements of our stacked estimator increase cubically in J. We choose a cutoff of $J = 70$, which is the number of LLMs with at least one million residents in 2010. There are 69 named cities, and a residual location aggregating the remaining smaller places.[48]

At $J = 70$, there are 39,760 types and 2,738,200 choice probabilities to estimate from the data. In principle, one could form all the moment conditions necessary from ACS, but in practice, the many cells become small, even among these relatively large places. Hence, a smoothing procedure is in order, which we detail in Appendix D.7.[49] In brief, we take di-

[48]The two smallest included cities are Fort Myers, Florida and Manchester, New Hampshire. The two largest excluded are Poughkeepsie, New York, and Baton Rouge, Louisiana.

[49]We have verified that the results hold up for several different weighting schemes that account for measurement error in the moments and differences in market shares (i.e., larger cities having more observations and hence getting larger weights). The differences across specifications were slight, so for simplicity we proceed with the unweighted version.

rectly from the data the conditional probabilities after combining small cells and then impute joint probabilities as the product of (smoothed) conditional and marginal probabilities. The starting point is to estimate move probabilities by origin, age, education, and at-home status as a coarse categorical variable, combining away-from-home locations,[50] to arrive at $Pr(move_{m,k})$. We then combine origins and estimate return-home and move-elsewhere probabilities, $Pr(choose\, j_m | move = 1)$.[51] The full matrix is then derived from the conditional probabilities of these two estimates: $Pr(move_{m,k}) \cdot Pr(choose\, j_m | move = 1, home = h)$. The probabilities are then differenced to form the estimating equation (11).

4.2 Estimates

Table 4 reports the structural parameter estimates for several specifications of the utility and move cost functions. We report standard errors for the home premium, and suppress the others for brevity. The lower panel reports details on the specification and model simulation fit (to be discussed more below).

The table runs through several specification checks. Columns 1 and 2 use the most basic model with only a home premium, move cost intercepts by type, and an entry toll for the residual location.[52] The home premium is large and significant, whether measured as an indicator variable (column 1) or a linear function of rootedness (column 2). (We compare these below.) For some context, an estimate of 3-4 utils is worth about 3.5 to 5 standard deviations of LLM income; that is, the average worker prefers home as if it offered incomes roughly 4 standard deviations above other markets. Move costs by type reflect the differences in average mobility rates across age and education; this is not altogether surprising, although the estimation demonstrates the pattern holds even when accounting for differing rates between age and education categories of living away from one's birthplace.

Columns 3 and 4 repeat 1 and 2, adding a distance function for move costs and a dummy variable for non-residual locations for each age/education type. The home premium estimates are very similar, though slightly larger. As Table 1 suggested, preference for home is not a coincidence of geography. The move cost parameters indicate farther migration events, measured in kilometers, are more costly, with discounts taken for within region, state, and neighboring LLM

[50] The main cell to smooth over is birthplace, in the case it is not the current or destination location. We pool the not-at-home types instead of interacting the full 68 possible types (J except for home, residual, or foreign-born). The cells get particularly sparse for small LLM birthplaces.

[51] We have tried several different specifications and source datasets for the destination probabilities, including the ACS microdata, ACS aggregate flows data from Census, and the IRS data. Our preferred specification uses the ACS aggregate flows tables, which allow us to cut the data by birth state. Results are qualitatively similar using other datasets.

[52] This is an entry toll, but there is no exit toll. Some form of entry toll is appropriate for the residual location, because it is geographically nothing like an actual metro area. Not only is it an order of magnitude larger than even the biggest city, it is geographically proximate to any LLM.

moves. Interestingly, moves to home show far less sensitivity to distance, with the coefficients reversing sign in the to-home interaction. Move cost intercepts are scarcely affected.

Columns 5 and 6 add toll costs, or LLM-by-education specific intercepts for the move cost function. This changes the interpretation of the move cost intercepts, but otherwise changes very little. Thus, estimates of home attachment are robust to a flexible local move cost function, even though lower mobility places have higher home attachments on average. Similarly, adding location by age by education utility fixed effects (560 additional parameters in all) does virtually nothing to affect the home utility or move cost estimates. That is, gross migration flows yield similar information about home premia and distance sensitivity, even after controlling for the average attractiveness of a location to different worker types.

Looking across the specifications of Table 4, it is apparent that the home premium coefficients are robust to many formulations of the problem, indicating the proclivities to remain at home, or chose home when away from it, are strong and nearly independent patterns in the data. Adding more parameters allows the model to better fit the data (mechanically, but also in terms of adjusted R^2) but does not alter these key parameters.

4.3 Model Fit in Baseline Simulation

We next check that the model is able to replicate the main features of the data. Note that the parameters are obtained via the estimating equations in (11), not an explicit targeting of the actual values. That is, the simulation will fit only as well as the model structure represents the data generating process.

Table 5 reports the average moving rates by age and education category and at-home status for the data and our baseline simulation of the model using specification 4.[53] The model is able to match the age profile in moving rates as well as the differences between college and non-college educated workers. It also matches quite well the difference between workers at home and not at home, the main qualitative pattern in the data important for our tests of demographic shifts, without assigning them different move cost terms. This is accomplished through presence of a home preference in the current flow utility and its effect on continuation values.

The spatial heterogeneity in migration rates is obviously important for our analysis. Figure 6 plots the actual and predicted out-migration rates for each metro area in our analysis. The model is able to match the spatial heterogeneity through a combination of differences in demographic composition, income distributions, shares of transients, and degree of rootedness.

Returning to Table 4, the lower panel reports some statistics on model fit when using the various specifications as the data generating process. The last four rows show the correlation in predicted out-migration rates to actual for LLMs in total and for the birthplace subgroups

[53] The simulation results between specifications 4 and 8 are similar, but 4 uses many fewer parameters.

Table 4: Parameter Estimates

Fixed Effects		None		Residual		Moving Toll		Destination X Age X Edu.	
		1	2	3	4	5	6	7	8
Home Preference									
Home	Noncollege	3.98		4.14		4.14		4.14	
		(0.007)		(0.006)		(0.004)		(0.004)	
	College	3.14		3.29		3.29		3.29	
		(0.007)		(0.006)		(0.004)		(0.004)	
Roots	Noncollege		5.40		5.51		5.64		5.65
			(0.01)		(0.009)		(0.006)		(0.006)
	College		4.24		4.34		4.48		4.48
			(0.01)		(0.009)		(0.006)		(0.006)
Move Cost									
US born									
Noncollege	20s	14.24	14.24	14.39	14.39	8.28	8.28	14.36	14.36
	30s	16.47	16.47	16.62	16.62	10.51	10.51	16.59	16.59
	40s	18.20	18.20	18.36	18.36	12.25	12.25	18.32	18.32
	50s	18.97	18.97	19.12	19.12	13.01	13.01	19.08	19.08
College	20s	12.10	12.10	12.26	12.26	4.96	4.96	12.22	12.22
	30s	14.92	14.92	15.08	15.08	7.78	7.78	15.04	15.04
	40s	17.58	17.58	17.73	17.73	10.44	10.44	17.70	17.70
	50s	18.22	18.22	18.37	18.37	11.08	11.08	18.34	18.34
Foreign born: US Born +									
Noncollege	20s	3.28	3.28	3.29	3.28	3.29	3.28	3.29	3.28
	30s	2.82	2.82	2.82	2.82	2.82	2.82	2.82	2.82
	40s	2.47	2.47	2.47	2.47	2.47	2.47	2.47	2.47
	50s	2.12	2.12	2.12	2.12	2.12	2.12	2.12	2.12
College	20s	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75
	30s	0.23	0.23	0.24	0.23	0.24	0.23	0.24	0.23
	40s	0.44	0.44	0.45	0.45	0.45	0.45	0.45	0.45
	50s	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72
Distance Function									
Main	Neighbors			0.65	0.65	0.44	0.44	0.44	0.43
	Same State			1.40	1.40	1.25	1.25	1.26	1.26
	Same Region			1.06	1.06	0.72	0.72	0.71	0.70
	Log (km)			-0.26	-0.26	-0.47	-0.47	-0.47	-0.48
To Residual	Noncollege	3.21	3.21	3.21	3.21			3.21	3.21
	College	3.83	3.83	3.84	3.84			3.84	3.84
Distance	Neighbors			-0.61	0.07	-0.61	-0.02	-0.61	-0.03
X To Home	Same State			-1.69	-1.51	-1.69	-1.53	-1.69	-1.53
	Same Region			-0.74	-0.06	-0.74	-0.10	-0.74	-0.10
	Log (km)			0.27	0.71	0.27	0.69	0.27	0.69
Specification Details and Fit:									
Calibrated $\frac{\delta}{\lambda}$			Noncollege, 3.80		College, 3.59				
No. Parameters	No. Parameters	34	34	38	38	174	174	582	582
Fit:	MSE	1.01	1.01	0.77	0.77	0.38	0.38	0.39	0.39
Adj. R^2	Adj. R^2	0.83	0.83	0.87	0.87	0.94	0.94	0.93	0.93
Correlation in LLM Move Rate, Model and Data									
	All	0.36	0.38	0.40	0.41	0.07	0.11	0.42	0.48
	At Home	-0.02	0.16	0.01	0.21	-0.10	0.00	0.14	0.33
	Not Home (US)	0.28	0.23	0.26	0.23	-0.14	-0.15	0.33	0.31
	Foreign Born	0.03	0.03	0.10	0.10	-0.48	-0.48	0.42	0.42

NOTES: The table reports structural parameter estimates under five versions of the model; standard errors in parentheses, but suppressed for most coefficients. Point estimates for metro-area specific move costs (columns 5 and 6) and average location quality (columns 7 and 8) are available upon request. "Neighboring MSA" refers to metro area pairs in the same U.S. state and/or less than 100 miles apart. Regions are defined as in Figure F1.

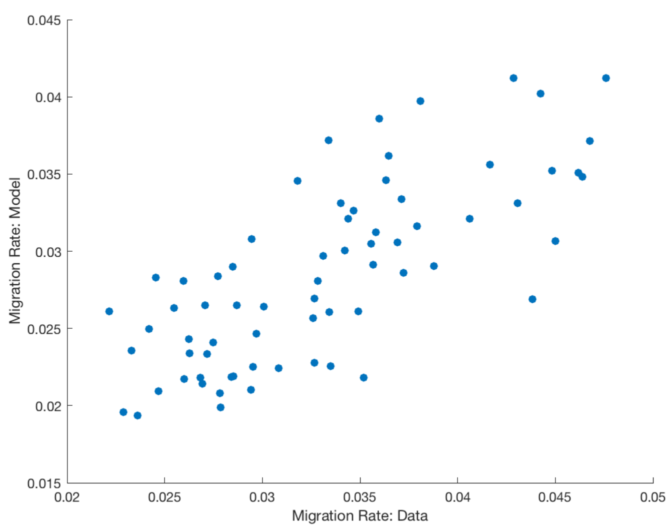
Table 5: Actual and Predicted Move Rates by Age, Education, and At-Home Status

Data		Total	At-Home	Not-at-Home (US)	Foreign-Born	Move To Home
Noncollege	20s	4.83	3.68	10.94	3.41	18.72
	30s	3.07	2.36	6.32	2.28	15.98
	40s	2.07	1.50	4.09	1.54	14.44
	50s	1.76	1.21	3.27	1.31	13.01
College	20s	8.52	5.97	13.05	8.40	17.26
	30s	4.62	2.78	6.86	5.04	14.20
	40s	2.30	1.34	3.41	2.35	10.62
	50s	2.00	1.30	2.89	1.76	9.69

Model		Total	At-Home	Not-at-Home (US)	Foreign-Born	Move To Home
Noncollege	20s	4.32	2.97	9.59	4.12	33.05
	30s	2.73	1.74	5.59	2.64	29.23
	40s	1.89	1.14	3.67	1.84	28.57
	50s	1.64	0.96	2.98	1.65	27.42
College	20s	8.31	4.78	12.90	11.04	23.63
	30s	4.53	2.32	6.37	6.10	21.70
	40s	2.25	1.16	3.16	2.84	21.07
	50s	1.87	1.00	2.66	2.20	20.70

NOTES: All figures are in percentages (%).

Figure 6: Model Fit: Predicted and Actual Out-Migration Rates by LLM



NOTES: The figure plots the model's baseline predicted out-mobility rates by metropolitan area to actual rates. Each marker represents a LLM in the estimation/simulation sample. (Source: ACS data and model-generated data.)

(home, not home, and foreign born). One key pattern is worth highlighting. Most of the specifications can generate a fairly good match of predicted to actual LLM out-migration in the aggregate by accounting for birthplace shares and moving cost, but only the specifications with home premium as a function of roots can fit a correlation *within* birthplace groups. The pattern of slower places having lower out-migration of natives (in Figure 5) requires a form of between-LLM heterogeneity that applies to the at-home differently from the not-at-home, and rootedness serves this purpose.

5 Simulations

The purpose of the empirical model is to conduct simulations at different scenarios in the same geographic setting. By simulating alternative population distributions, home attachments, or income offerings, the model allows us to see how each channel affects mobility trends.

A simulation of the model predicts choice probabilities for every state (origin, income draw) and agent type (age, education, birthplace) in the economy, amounting to millions of predicted values. To summarize the main findings, we report in the tables below the average mobility rates for the complete set of LLMs and also breakdowns by the initial mobility rates (fast, medium, and slow) among the 70 LLMs in the estimation. We simulate the model at the same primitive parameter values but use the environments and type weightings for five time periods: 1980, 1990, 2000 censuses, the early ACS (2005-2011), and the late ACS (2012-2017).[54] We do not necessarily contend that the primitive parameters in utility could not have changed; merely, we want to see how far the model goes in explaining the decline when constrained in this way.

The environmental factors are the local income distributions, search cost, and the rootedness of the birth cohorts. These factors can actually change the model's predicted mobility rates for the agents. In contrast, shifts in composition (e.g., age/education, population share in each location, at-home share) maintain the baseline probabilities but change the weights that make up the aggregate. We first combine all factors to show how the model predicts migration changes changing over time and then decompose the factors one by one to elicit the contribution of each factor. Recall that a maintained assumption is that utility primitives do not change. One slight exception is the search cost parameter γ, which by assumption is larger now than in previous decades. We have little sense as to how to calibrate this over time, so we guess that it is most recently at 0.5 and declines by 0.1 per year. As we will show, this parameter ends up being of little consequence at our parameter values.

[54] In Appendix B.2, we simulate the model back to 1950 instead of 1980, maintaining the assumption of fixed primitives, and altering age/education shares, income, and home attachments. The model is able to replicate the hump-shaped pattern in postwar migration rates described in Fischer (2002) and Molloy et al. (2017).

Table 6: Simulated Trends in Migration Rates

		Level					Difference		
		1	2	3	4	5	6	7	8
Year(s)		1980	1990	2000	2005-11	2012-17	1980-2017	1990-2011	1990-2011
		(%)	(%)	(%)	(%)	(%)	(Δ %)	(Δln(%))	(Δln(%))
LLM Type		Model	Model	Model	Model-Est	Model-Est	Model	Model	IRS Data
All	Named	3.23	3.17	2.97	2.87	2.91	-0.32	-0.10	-0.10
By Speed:	Fast	4.21	3.96	3.55	3.35	3.32	-0.89	-0.17	-0.19
	Medium	3.24	3.20	3.02	2.92	2.96	-0.29	-0.09	-0.06
	Slow	2.56	2.49	2.40	2.35	2.42	-0.13	-0.06	-0.03

NOTES: Units are indicated in the third row; most are percentage points. The difference of natural log of the migration rates is taken to compare the the changes in the IRS data to those in the model (which was estimated on ACS baseline data). The row for "All Named" LLMs excludes the residual location.

5.1 Simulating Mobility Over Time

We begin with the aggregate change to migration–the combined simulation with all changes taking place–in Table 6. The census years and ACS analogs are listed in columns 1 to 5. A long difference from 1980 to 2017 is taken in 6, and in 7, the difference of 1990 to 2011 so as to compare with available IRS data for that period in column 8. The model generates a 0.32 percentage point decline in mobility from 1980 to 2017. For the IRS data comparison period, this matches well at a 10 log point drop in predicted migration.[55] Importantly, the model successfully matches the heterogeneity in the decline, with more mobile cities declining more: 0.89 percentage point, about a 27 percent change, for fast LLMs, compared with 0.13 percentage point, about a 6 percent change for slow LLMs from 1980 to 2017.

5.2 Decomposing the Sources of Decline

The simulated decline is a net result that we can then unpack via *ceteris paribus* breakdowns of the sources of change. Table 7 reports in columns 2 to 9 the changes generated in the model from the first to last simulation period when altering one feature at a time, and the combined change is reported for reference in column 1. For each simulation, we hold all factors except one fixed at their estimation period (2005-2017) values.[56] Table 7 lists two panels, one for all population (including foreign born) and one for U.S.-born only, summarizing across LLMs as in Table 6, while Table B1 in the appendix reports a subset of results for each LLM separately.

We begin with the demographic changes. First, in column 2, we simulate as if the population aged as we observe in the data. The model estimates that had there been no other changes to the economy since 1980, this feature alone was large enough to have generated the full decline

[55] We take logs to compare proportions, as the IRS rates are at a persistently higher level than the ACS data on which the model is estimated.

[56] Because we are using a nonlinear model with interactions among many features, this exercise is not literally a decomposition, but it meaningfully demonstrates the magnitude of the sources.

nationally, although the change would have been too similar across types of cities to correspond with the data. Furthermore, the effect of aging was substantially mitigated by an increase in college attainment, which shown in column 3 has generated a 0.18 percentage point increase in mobility rates, canceling out almost half of the effect of aging. The effect of college attainment is also present at comparable magnitudes across all types of cities.

Cities continue to grow in size at different rates, so in column 4, we also change the share of the population residing in each LLM, holding all else fixed, including subgroup size. This would generate an 0.11 percentage point increase in migration rates, indicating that population growth has still trended toward more mobile locations, a factor that has also mitigated the decline. Interestingly, this is even present within LLM categories. The effects are slightly larger among the U.S. born.[57]

Columns 5, 6, and 7 of Table 7 show the change in predicted mobility when altering just variables related to home attachment. Column 5 shows that an increasing share of at-home population accounts for 0.21 a percentage point in the national decline but is by definition larger among the U.S. born. The effect is concentrated in fast cities but is also present in medium and slow as well. Column 6 shows that an increase in the foreign-born population share has resulted in small migration declines out of fast cities as they replace (by proportion) the not-at-home, U.S.-born population, but an increase out of slow cities as they replace the at-home. The net effect nationally is nearly zero. Column 7 shows the effect of changing rootedness of the at-home (holding fixed the share at home), which is a change in the strength of their home attachment. This produced a 0.10 percentage point drop in migration in fast cities (especially in California). Rootedness has changed little in slow and medium speed cities.

Finally, the last two columns of Table 7 evaluate the change in migration as if income distributions changed as they did in the data but population composition and home attachment held steady. This would generate a small amount of decline in mobility, centered in fast cities. A decrease in the information parameter γ would add to the decline, in a rank ordering in agreement with the data, but its marginal contribution is quite small. That is, while qualitatively the information feature of the model performed as expected, the quantitative power of this channel is limited, given the distribution of income opportunities across cities and income draw parameters π. However, as Kennan and Walker (2011) have shown, while the effect of expected income on migration is significant, it is highly idiosyncratic in that it depends on individual match components more than local labor market averages. It is possible that there is more impact of information availability than is picked up by LLM aggregate data on income distributions. However, it does not seem to be driving the dominant spatial patterns we document.

Figure 7 summarizes the results by plotting simulated migration rates when subtotaling the

[57]Population growth among the foreign born centers on two types of destinations: large gateway cities (e.g. New York and Chicago), which tend to be slow or medium, and strongly growing locations (e.g., Las Vegas and Phoenix), which tend to be fast.

Table 7: Simulations: Counterfactual Changes in Migration Rates, 1980-2017, by Source of Change

LLM Type		Total 1	Age 2	Education 3	Location 4	At Home 5	Foreign Born 6	Roots 7	Income 8	Information 9
All Population										
All	Named	-0.319	-0.409	0.180	0.113	-0.206	0.017	-0.043	-0.019	-0.0021
By Speed:	Fast	-0.889	-0.439	0.178	0.051	-0.343	-0.059	-0.102	-0.044	-0.0024
	Medium	-0.287	-0.457	0.168	0.057	-0.108	-0.005	0.009	-0.011	-0.0022
	Slow	-0.135	-0.341	0.192	0.014	-0.133	0.115	-0.017	0.000	-0.0018
U.S.-Born Population										
All	Named	-0.365	-0.413	0.141	0.144	-0.274		-0.057	-0.017	-0.0023
By Speed:	Fast	-0.868	-0.458	0.129	0.074	-0.496		-0.148	-0.043	-0.0028
	Medium	-0.299	-0.465	0.142	0.064	-0.139		0.011	-0.013	-0.0023
	Slow	-0.280	-0.330	0.149	0.040	-0.167		-0.022	0.003	-0.0018

NOTES: All figures are in percentage points (e.g., "0.5" corresponds to a one-half percentage point change).

changes into categories of changes. The "demographic" simulation combines the effects of aging, education, and the spatial distribution of population (columns 2, 3, and 4 from Table 7), "home attachment" combines 5, 6, and 7, and "income" combines 8 and 9. The figure contains a plot for all LLMs and one each for fast, medium, and slow LLMs. Note that each subfigure is on its own vertical scale.

For each set of LLMs, demographic explanations generate a declining trend, with most of the downward movement occurring between 1990 and 2010, as Baby Boomers aged into the slow mobility groups of 40 and 50. The share of decline varies across the fast, medium, and slow subsets. The demographic changes generate 14 percent of the total decline (1980-2017) but 55 to 61 percent of the relatively smaller declines in medium and slow cities, respectively.

The collective effect of changing home attachment accounts for about 0.40 percentage point of the decline in fast LLMs, and consequently a large fraction of the national decline–0.23 of the simulated 0.32 percentage point, about 70 percent.[58] Figure 7 shows how the decline caused by rising home attachment is a steady, gradual trend, and was the largest component of the trend in fast cities and national average but a more modest contribution to slower places.

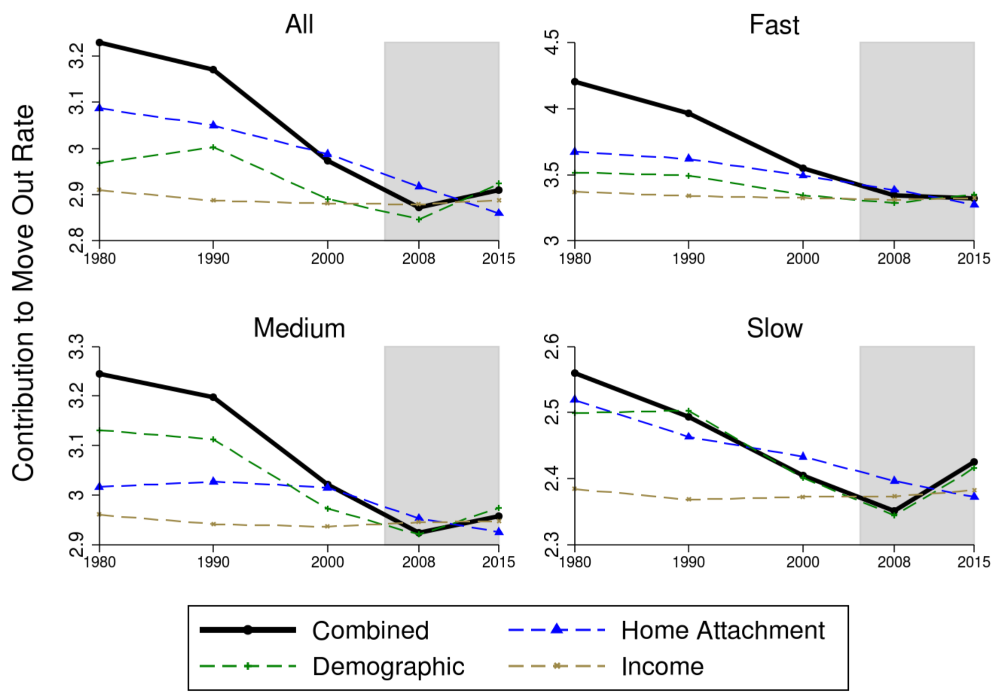
Changing income opportunities produce a gradual downward trend in migration rates, more so in fast LLMs, but from Figure 7, it is apparent the other effects are quantitatively larger. However, we note that changes in income distributions do help explain migration declines out of some individual cities that are otherwise not as well accounted for, such as Seattle, Tampa, Phoenix, and San Francisco (see Table B1).[59]

In summary, we find that an aging population accounts for some of the aggregate migration decline, but it is rising home attachment that accounts for the larger share and produces the

[58] It is here that the nonlinearities become obvious. For fast cities in particular, the subtotal effects fall short of the actual total. Coincidentally, in the national series, the components sum nearly to the total, so using the fraction is not misleading in that case.

[59] Likewise, income growth mitigates the predicted increase in migration out of growing LLMs such as Austin, Charlotte, and Raleigh-Durham.

Figure 7: Simulated Trends in Mobility, Decomposed



NOTES: The figures plot the time path of mobility generated by the model in total and counterfactual subtotal simulations for each category of LLMs. Note that each panel has its own vertical scale. Subtotals may not add to the combined value because of nonlinearities in the model. The shaded region denotes the estimation period.

spatial heterogeneity described in section 2.

6 An Illustrative Model

Our quantitative model was limited to the change in home attachments observed in the data. Having shown that the change in home attachment is an important factor explaining the mobility decline, we now use a stylized version of the model to illustrate how home attachment can evolve over a long horizon, and the impact it has on regional responsiveness.

6.1 Home Preference and Steady State Migration

We will first consider migration rates under different scenarios of home attachment.[60] This begins with a very simple version of the illustrative model, an economy of J identical locations and one agent type. At time $t = 0$ a measure 1 of agents is born across the locations, distributed equally at $\frac{1}{J}$. Agents are endowed with a preference for their birthplace, constant across the identical locations, and are given an opportunity to move at a cost. All moves are idiosyncratic

[60] Ours is a highly stylized exercise to illustrate the effects of evolving home attachments. Zabek (2018) considers a quantitative model of local labor market elasticity in the presence of locally-tied workers.

Table 8: Illustrative Model: Simulated Move Rates Under Various Parameterization Scenarios

Home	Pop. Share	Migration Probability at Steady State			Change in Migration Probability in Response to Shock		
Preference	At Home	At Home	Not Home	Total	At Home	Not Home	Total
Low	24.78	5.37	10.69	9.37	0.72	1.35	1.19
High	62.81	3.55	12.06	6.71	0.48	1.49	0.86

NOTES: The table reports shares of agents residing at home and aggregate gross migration rates for an economy with a single cohort of agents with many successive choice periods under different strengths of preference for home. All values are percents (%).

but for the home preference. We ignore aging, cohort, and distance effects to emphasize the evolution of the state variables over successive choice opportunities. After sufficient opportunities, the share at home or away from home reaches a steady state (in practice, we simulate the model for 200 periods). Table 8 reports the economy's moving rate and the share living at home under higher and lower preferences for home, which we calibrate to be one standard deviation below and above, respectively, the mean rootedness times the average preference parameter from Table 4. Move costs are the point estimate for college educated 20-somethings.

The table illustrates the direct and indirect ways home preference can affect mobility in the economy. Migration rates are always higher for the not-at-home than the at-home, but the size of the gap depends on the intensity of home preference.[61] Less obvious, perhaps, is that weaker home preferences mean fewer agents living at home in the steady state. The high preference scenario has the lowest mobility because of lower mobility among the at-home *and* a higher share of agents in the lower move propensity, at-home state. In contrast, in the low home preference scenario, the total weighted average migration rate is closer to that of the not-at-home group.

This has implications for adjustment to regional shocks. To see this, we conduct some comparative statics at the steady state. In the last three columns of the table, we report the change in migration after making half the locations more desirable for all agents and half less desirable.[62] In the low preference scenario, more people relocate immediately compared with the high, both because they are individually more sensitive and because more agents are in the more susceptible not-at-home status.

6.2 Home Preference and Regional Shocks

The previous exercise clarifies the mechanism of the model but does not directly illustrate its dynamics. A more illuminating exercise is to show how the preference for home mediates a shock to the equilibrium distribution of population in an economy. The next exercise lets us simulate an economy (qualitatively like the postwar U.S.) that experiences shocks to location attractiveness, tracing the evolution of home status and preference.

[61] The not-at-home move more in the high preference scenario because of a higher rate of returns to home.

[62] The size of the shock is arbitrary for making this point. The important feature is that all scenarios receive the same shock.

In the simulation that follows, we use an overlapping generations (OLG) framework. Each agent lives for A periods, and the economy is simulated for $T >> A$ periods. We start with an initial distribution of population for the very first cohort, $t = 0, a = 0$, and simulate their behavior over A periods; they spread out across locations, and when they "die" at A, the distribution of their population forms the strength of the home preference for the next generation to be born. Thus, the home preference moves endogenously with population history under roots-based preferences and an OLG structure. We then compare this model, "Roots-Varying Preferences," with others in which there is either no preference for home ("No Home Preference"), or the preference for home is constant and not endogenously generated by prior spatial distributions of population ("Fixed Home Preference").[63]

We simulate each version of the model until it reaches a steady state, where all cohorts have the same strength of preference for home, population sizes of the locations are constant, and all migration is idiosyncratic. Then, we introduce an unanticipated permanent shock to location attributes in order to cause population reallocation across space. We split the locations into sets A and B, "good" and "bad" amenities, and the shock reverses the good to bad and bad to good.[64] This particular implementation is not important for generating the qualitative patterns we describe. The objective is to see how shocks to the steady state population distribution are mediated over the transition path.

Plot 1 of Figure 8 shows the population sizes of the two types of locations. As good and bad locations interchange, their populations transition from a negative difference $A - B$ to a positive $1 - (A - B)$. Note that the No Home Preference scenario arrives at the new steady state more quickly, as agents move only on the basis on amenities. The Fixed Home Preference scenario arrives at the steady state somewhat more slowly, and the Roots-Varying arrives the most slowly, since the endogenous home preferences are initially very strong. Plot 2 shows the net reallocation of population that produces the gaps in plot 1. As in the exercise from Table 8, these show that home preference can result in slower population adjustment.

Plot 3 shows the share of residents at home, which, as Table 8 showed, affects the overall migration rate and population adjustment. As more people leave the once-high amenity locations for the new high-amenity places, the share living in their birthplace drops. This resolves to its steady state value again more quickly in the No Home Preference scenario because the populations of A and B stabilized more quickly without ties of home preference (and because it

[63]To generate similar migration rates in initial steady state, the "No Home Preference" simulation has higher migration costs than the others. The constant "Home Preference" simulation is set to have a constant flow utility of being at home that equates migration rates to the steady state value in the "Roots Preference" simulation. These adjustments are made to start from the same migration rate value for ease of illustration, but they are not essential to make our point, as it is the dynamics in response to shocks that are the emphasis of this exercise.

[64]An economy with heterogeneous locations has a different steady state value of migration than one with homogenous locations, since the out-migration probabilities differ and not all locations are the same size. The shock we introduce changes which locations have which attributes, but maintains equally heterogeneous locations. That is, the cross sectional variance between locations are the same before and after the shock.

does not have to rise as high to reach steady state). That is, when agents prefer their birthplaces, they are more likely to leave the good locations to return to their homes, a reinforcement that further lengthens the path to population adjustment.

The lower panel plots the migration rates for at-home and not-at-home agents and the total. The No Home Preference scenario has the largest increase and quickest convergence in move rates for the at-home agents (and not-at-home, since they have identical incentives here), but incentives are blunted for the Fixed Home and Roots-Varying scenarios. Under the home preference scenarios, the behavior of the not-at-home shows oscillation because it depends on whether the not-at-home population is away for idiosyncratic or common reasons. Move rates initially go up in response to the shock and then fall as more of them make their way to the newer, better places and have less incentive to leave (even for home). Then, move rates slowly rise again back to the steady state, as more of the migration of the not-at-home becomes idiosyncratic instead of directed by a shock to the common valuations of each place.

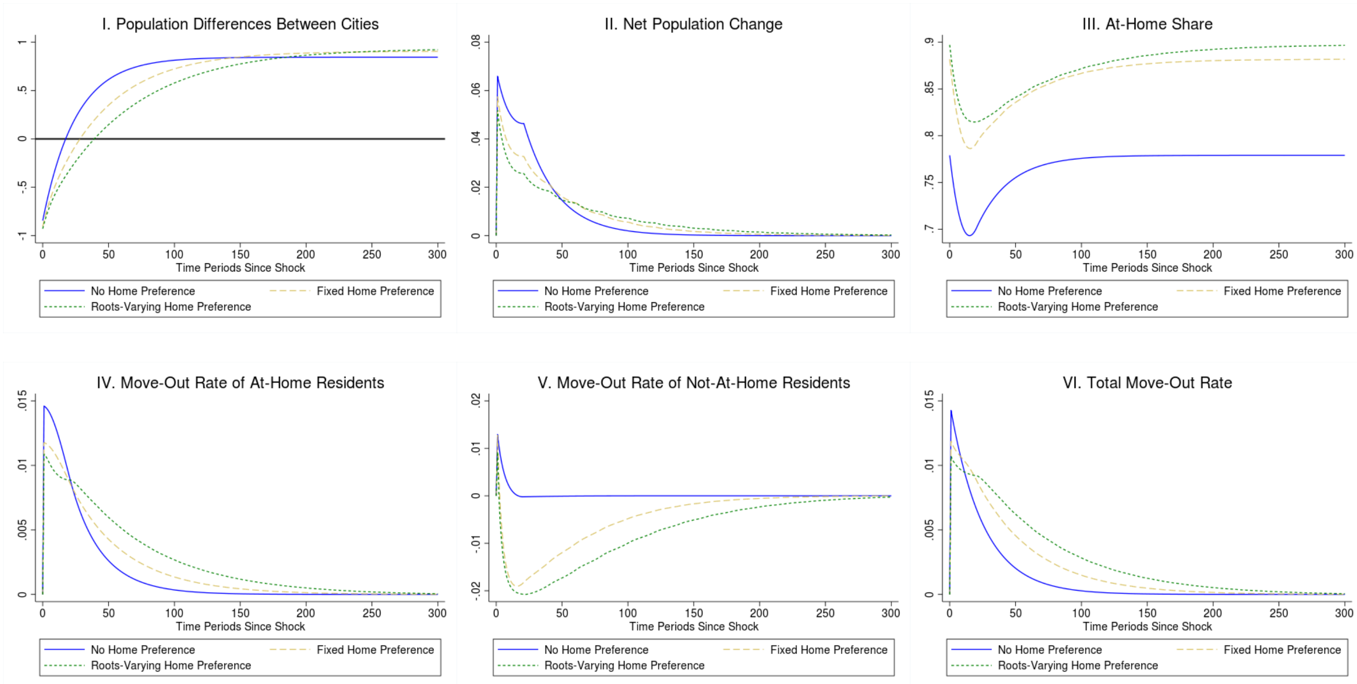
The sum of the at-home and not-at-home rates produces the impulse in total migration rates seen in plot 6. Hit with the same shock, the economies with home preferences experience elevated migration rates for a longer period of time, with longer, slower returns to steady state. When the strength of home preference varies with the history of population, the shock produces transition-era cohorts with weak home preferences, and hence population churn is high and convergence to steady state is especially long and slow. Thus, our model formulation generates the qualitative pattern in gross mobility consistent with the spatial evolution of population in the U.S. and the subsequent migration trend, as shown in Section 2.

7 Evaluation: Has America Lost Its Mojo?

Understanding the reasons for the mobility decline is critical for evaluating the risk it poses to the efficiency of labor markets and whether policy intervention is warranted–and if so, which policy. The literature studying the migration decline has looked for a structural mechanism in the economy that has caused a downward trend in mobility, pervasive across demographic strata. This paper has argued that an important underlying structural change is the long-run settling of the spatial distribution of population, which has generated an increasing degree of home attachment–a slowly trending pull factor common to many types of households, but changing unequally across space. We close by offering an evaluation of the mobility decline in light of these findings. We see several reasons not to be worried, but with some important qualifications.

We begin with the major qualification: There is, hypothetically, good cause for concern. The reasoning comes from the illustrative model in section 6. An individual at home is less likely to respond to other incentives to relocate. In the aggregate, an economy with home preference–especially the evolving kind that we develop in this paper–will adjust to shocks more slowly and

Figure 8: Illustrative Model: Migration and the Transition Between Steady States



NOTES: The figure shows the time paths of population reallocation and migration rates after a shock to the steady state distribution of population under three home preference regimes. "No Home Preference" means no home preference at all, "Fixed Home Preference" means a constant value of home preference in all time periods, and "Roots-Varying Home Preference" means a preference for home that varies endogenously over time by the share of the dying cohort of agents living at home. (Source: Model-generated data described in section 6.)

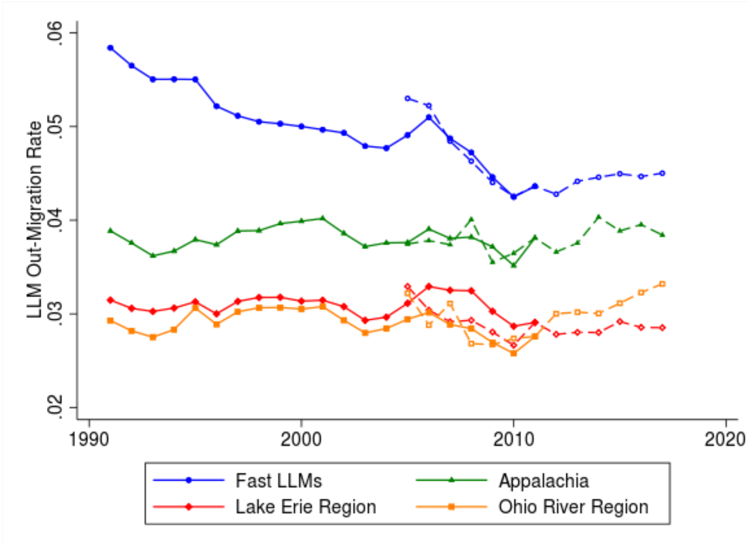
return to steady state at a longer lag.

Conceptually, these are risks, resulting in a broad concern that falling mobility means the economy will be unable to make population adjustments when they are warranted. But before crafting a policy around this hypothetical, we need to ask: Is population adjustment in the U.S. impeded today? Our views on the immediate risks are more sanguine.

One possible misconception is a connection of the mobility decline to the relatively poor performance of certain local labor markets. We showed, in contrast, that actually the growing cities are more often the sources of mobility decline. We can also look directly at migration rates out of regions and LLMs generally considered to be underperforming. In Figure 9, we plot the out-migration from lagging regions we organize by topographic features for convenience: (i) the Lake Erie region, containing the Rust Belt cities of Detroit, MI, Toledo, Cleveland, and Youngstown, OH, Erie, PA, and Buffalo, NY, (ii) the Ohio River Valley, including Pittsburgh, Cincinnati, and Louisville, and (iii) the cities and rural areas of the Appalachian mountains from West Virginia to North Georgia.[65] (Here, we use the same definition of a move being an exit from a LLM, even if it means staying within the region we define.)

[65] We also examined the "Eastern Heartland" at-risk area of Austin et al. (2018), which comprises the noncoastal states east of the Mississippi River, plus Missouri, Arkansas, and Louisiana. We found the regional trends dominating, in that northern and midwestern parts of the Eastern Heartland showed flat migration rates, while some southern areas trended down moderately.

Figure 9: Migration Rates Out of Distressed Areas



NOTES: The figure plots out-migration rates for fast LLMs and the three regions with relatively underperforming labor markets: (i) the Lake Erie region, containing the Rust Belt cities of Detroit, MI, Toledo, Cleveland, and Youngstown, OH, Erie, PA, and Buffalo, NY, (ii) the Ohio River Valley, including Pittsburgh, Cincinnati and Louisville, and (iii) the cities and rural areas of the Appalachians mountains from West Virginia to North Georgia. The IRS and ACS series (solid and dashed lines, respectively) are overlaid and benchmarked to the same value in 2011. (Source: IRS and ACS data.)

While fast LLMs show declining out-migration, in each of the three regions, the rates of mobility are quite flat. Hence, for whatever problems these local labor markets may have–and perhaps more people "should," in some sense, be exiting these regions–they have been slow for some time and are not particularly afflicted by the mobility decline. Hence, the national trend cannot be blamed for further diminishing the poor circumstances of workers in these areas.

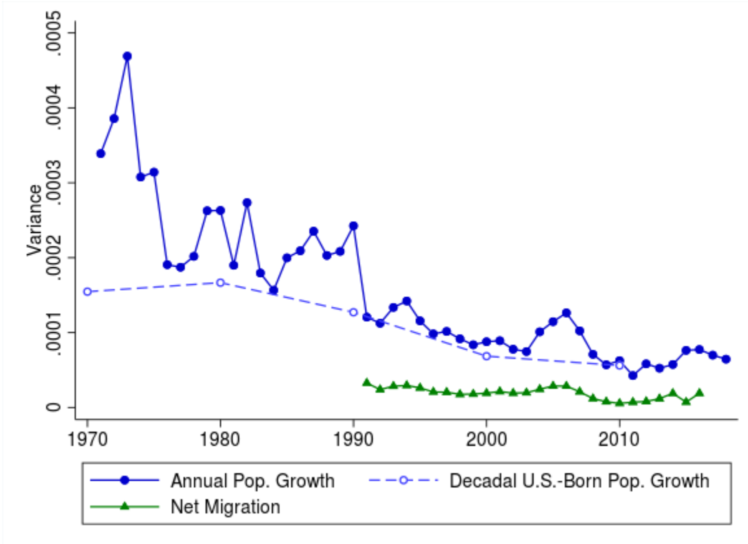
Beyond particular lagging regions, is the population less nimble? Declines in gross mobility do not necessarily mean declines in net population change, as gross flows remain orders of magnitude larger than net flows. More directly, the concern about falling mobility leading to insufficient reallocation seems to presume the order of events. Part of our contribution to this debate is to understand the autocovariances between gross and net population flows, parsing the lead and the lag. Our analysis indicates that the stabilizing of population growth and, by consequence, home attachment, caused a decline in gross migration rates. Concern over gross mobility seems directed at the wrong object.

Figure 10 highlights this issue. The figure plots the variance across LLMs in (1) in annual population growth from census intercensal estimates, (2) annualized U.S.-born population growth from decennial census, and (3) net migration rates from the IRS data. These would directly indicate how much population reallocation is occurring.[66] The variance in population growth has been trending downward for several decades, and the trend has if anything moderated more recently. The variance in net migration is trending downward only slightly,[67] to an

[66]The variance of population growth may depart from the variance in net migration to the extent there are differences in birth and death rates and the arrival of foreign migrants.

[67]Using state-level data, Kaplan and Schulhofer-Wohl (2017) find that net migration is flat.

Figure 10: Cross Sectional Variance in Population Growth and Net Migration



NOTES: The figure presents the unweighted cross sectional variance in population growth and net migration. The spatial unit is urban LLMs. The annual population growth series uses intercensal estimates of population growth. The decadal series uses an annualization of the long difference, $\frac{1}{10} ln(P_{t+10} - P_t)$. The net migration series is the variance across LLMS in inflow minus outflow as a proportion of initial size. Two outliers have been excluded: New Orleans, Louisiana and Biloxi, Mississippi in 2006, both of which experienced large population loss in the aftermath of Hurricane Katrina. (Source: Census and IRS data.)

extent less than a continuation in the preceding trend in population growth. Cyclical variations in migration variance are still apparent. This does not suggest some great migration slowdown has stopped population growth from happening, but rather the convergence in the long run population trends have preceded (and, in our view, caused) the gross mobility slowdown, with year-to-year fluctuations still occurring.

Finally, we step back and consider welfare implications of our findings. One of the key conclusions of attributing declining mobility to rising home attachment is the implication that agents are making optimal, unconstrained choices to stay in place. Alarmist views of the decline in dynamism fear that some friction is preventing people from making moves they would otherwise like to make. If instead people simply prefer to be near family and friends and in a familiar place, the perspective changes considerably. In a sense, we agree with the argument of Kaplan and Schulhofer-Wohl (2017), though for different reasons, that the mobility decline is not particularly concerning and could actually be evidence of a well-functioning market with declining incentives to move. Moreover, if locations are offering more similar occupational opportunities, as Kaplan and Schulhofer-Wohl (2017) argue, then aspects like home attachment may have become more important at the margin.

References

Abramitzky, R., P. Ager, L. Boustan, E. Cohen, and C. Worm-Hansen (2019). The Effects of Immigration on the Economy: Lessons From the 1920s Border Closure. Working paper, Stanford University.

Alesina, A., Y. Algan, P. Cahuc, and P. Giuliano (2015). Family Values and the Regulation of Labor. *Journal of the European Economic Association* 13(4), 599–630.

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica* 79(6), 1823–1867.

Arsenault, R. (1984). The End of the Long Hot Summer: The Air Conditioner and Southern Culture. *The Journal of Southern History* 50(4), 597–628.

Artuc, E., S. Chaudhuri, and J. McLaren (2010). Trade Shocks and Labor Adjustment: A Structural Empirical Approach. *American Economic Review* 100(3).

Austin, B., E. Glaeser, and L. H. Summers (2018). Saving the Heartland: Place-Based Policies in 21st Century America. In *Brookings Papers on Economic Activity Conference Drafts*.

Autor, D. 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.

Autor, D., D. Dorn, and G. Hanson (2019). When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Young Men. *American Economic Review: Insights forthcoming*.

Autor, David, H., D. Dorn, and G. H. Hanson (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review* 103(6), 2121–68.

Barnett, C. (2008). *Mirage: Florida and the Vanishing Water of the Eastern US*. University of Michigan Press.

Bayer, C. and F. Juessen (2012). On the Dynamics of Interstate Migration: Migration Costs and Self-Selection. *Review of Economic Dynamics* 15(3), 377–401.

Bayer, P., N. Keohane, and C. Timmins (2009). Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management* 58(1), 1 – 14.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A Dynamic Model of Demand for Houses and Neighborhoods. *Econometrica* 84(3), 893–942.

Belot, M. and J. Ermisch (2009). Friendship Ties and Geographical Mobility: Evidence From Great Britain. *Journal of the Royal Statistical Society: Series a (Statistics in Society)* 172(2), 427–442.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Bishop, K. C. (2008). A Dynamic Model of Location Choice and Hedonic Valuation. Olin School of Business, Washington University in Saint Louis.

Blanchard, O. J. and L. F. Katz (1992). Regional Evolutions. *Brookings Papers on Economic Activity 1*, 1–37.

Bound, J. and H. J. Holzer (2000). Demand Shifts, Population Adjustments, and Labor Market Outcomes During the 1980s. *Journal of Labor Economics* 18(1), 20–54.

Bryan, G. and M. Morten (2019). The Aggregate Productivity Effects of Internal Migration: Evidence From Indonesia. *Journal of Political Economy* 127(5), 000–000.

Carrington, W. J. (1996). The Alaskan Labor Market During the Pipeline Era. *Journal of Political Economy* 104(1), 186–218.

Carrington, W. J., E. Detragiache, and T. Vishwanath (1996). Migration With Endogenous Moving Costs. *The American Economic Review*, 909–930.

Chinitz, B. (1986). The Regional Transformation of the American Economy. *American Economic Review* 76(2), 300–304.

Clark, W. A. V. and W. Lisowski (2019). Extending the Human Capital Model of Migration: The Role of Risk, Place, and Social Capital in the Migration Decision. *Population, Space and Place* 25(4), e2225.

Coate, P. (2014). Parental Influence on Labor Market Outcomes of Young Workers. unpublished.

Coen-Pirani, D. (2010). Understanding Gross Workers Flows Across U.S. States. *Journal of Monetary Economics 57*, 769–784.

Cooke, T. J. (2011). It Is Not Just the Economy: Declining Migration and the Rise of Secular Rootedness. *Population, Space and Place* 17(3), 193–203.

Cooke, T. J. (2013). Internal Migration in Decline. *The Professional Geographer* 65(4), 664–675.

Dahl, M. S. and O. Sorenson (2010). The Migration of Technical Workers. *Journal of Urban Economics* 67(1), 33–45.

David, Q., A. Janiak, and E. Wasmer (2010). Local Social Capital and Geographical Mobility. *Journal of Urban Economics* 68(2), 191–204.

Davis, M. A., J. Gregory, D. A. Hartley, and K. T. Tan (2017). Neighborhood Choices, Neighborhood Effects, and Housing Vouchers.

Davis, S. J., R. J. Faberman, and J. Haltiwanger (2012). Labor Market Flows in the Cross Section and Over Time. *Journal of Monetary Economics* 59(1), 1–18.

Davis, S. J. and J. Haltiwanger (2014). Labor Market Fluidity and Economic Performance. Working paper 20479, National Bureau of Economic Research.

Dawkins, C. J. (2006). Are Social Networks the Ties That Bind Families to Neighborhoods? *Housing Studies* 21(6), 867–881.

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives* 28(3), 3–24.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016). Declining Business Dynamism: What We Know and the Way Forward. *American Economic Review* 106(5), 203–07.

Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review* 106(3), 479–524.

Dorn, D. (2009). *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. Ph. D. thesis, University of St. Gallen.

Falck, O., S. Heblich, A. Lameli, and J. Südekum (2012). Dialects, Cultural Identity, and Economic Exchange. *Journal of Urban Economics* 72(2-3), 225–239.

Fischer, C. S. (2002). Ever-More Rooted Americans. *City & Community* 1(2), 177–198.

Flood, S., M. King, S. Ruggles, and J. R. Warren. (2015). Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-Readable Database]. Minneapolis: University of Minnesota. 2015.

Frey, W. (2009). The Great American Migration Slowdown. Technical report, Brookings Institution, Washington, DC.

Glaeser, E. L., D. Laibson, and B. Sacerdote (2002). An Economic Approach to Social Capital. *The Economic Journal* 112(483), F437–F458.

Glaeser, E. L. and K. Tobio (2008). The Rise of the Sunbelt. *Southern Economic Journal* 74(3), 610–643.

Graves, P. E. and P. D. Linneman (1979). Household Migration: Theoretical and Empirical Results. *Journal of Urban Economics* 6(3), 383–404.

Greenwood, M. J. (1975). Research on Internal Migration in the United States: A Survey. *Journal of Economic Literature*, 397–433.

Greenwood, M. J. (1985). Human Migration: Theory, Models, and Empirical Studies. *Journal of Regional Science* 25(4), 521–544.

Gyourko, J., C. Mayer, and T. Sinai (2013). Superstar Cities. *American Economic Journal: Applied Economics* 5(4), 167–199.

Hood, K. (2013). What Can Regional Labor Markets Tell Us About Declining US Migration? unpublished.

Hornbeck, R. (2012). The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe. *American Economic Review* 102(4), 1477–1507.

Hotchkiss, J. L. and A. Rupasingha (2018). Individual Social Capital and Migration. FRB Atlanta WP 2018-3.

Hotz, V. J. and R. A. Miller (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *Review of Economic Studies* 60(3), 497–529.

Hyatt, H. R. and J. R. Spletzer (2013). The Recent Decline in Employment Dynamics. *IZA Journal of Labor Economics* 2(1), 5.

Hyatt, H. R. and J. R. Spletzer (2017). The Recent Decline of Single Quarter Jobs. *Labour Economics 46*, 166–176.

Kan, K. (2007). Residential Mobility and Social Capital. *Journal of Urban Economics* 61(3), 436–457.

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 (2017). Understanding the Long-Run Decline in Interstate Migration. *International Economic Review* 58(1), 57–94.

Karahan, F. and S. Rhee (2014). Population Aging, Migration Spillovers, and the Decline in Interstate Migration. FRB of New York Staff Report, no. 699.

Kennan, J. and J. R. Walker (2011). The Effect of Expected Income on Individual Migration Decisions. *Econometrica 79*, 211–251.

Lkhagvasuren, D. (2012). Big Locational Unemployment Differences Despite High Labor Mobility. *Journal of Monetary Economics* 59(8), 798–814.

Luckingham, B. (1984). The American Southwest: An Urban View. *The Western Historical Quarterly* 15(3), 261–280.

Ma, L. (2019). Learning in a Hedonic Framework: Valuing Brownfield Remediation. *International Economic Review* 60(3), 1355–1387.

Mangum, K. (2016). Coming and Going: Metropolitan Area Heterogeneity in Migration Rates. Technical Report 2016-1-1, Usery Workplace Research Group.

Manson, S., J. Schroeder, D. V. Riper, and S. Ruggles. (2018). IPUMS National Historical Geographic Information System: Version 13.0 [Database]. Minneapolis: University of Minnesota.

Michielin, F., C. H. Mulder, and A. Zorlu (2008). Distance to Parents and Geographical Mobility. *Population, Space, and Place* 14(4), 327–345.

Miller, E. (1973). Is Out-Migration Affected by Economic Conditions? *Southern Economic Journal* 39(3), 396–405.

Molloy, R., C. L. Smith, and A. Wozniak (2017). Job Changing and the Decline in Long-Distance Migration in the United States. *Demography* 54(2), 631–653.

Molloy, R., C. L. Smith, and A. K. Wozniak (2014). Declining Migration Within the US: The Role of the Labor Market. Technical report, National Bureau of Economic Research.

Molloy, R., R. Trezzi, C. L. Smith, and A. Wozniak (2016). Understanding Declining Fluidity in the US Labor Market. *Brookings Papers on Economic Activity* 2016(1), 183–259.

Molloy, R. S., C. L. Smith, and A. Wozniak (2011). Internal Migration in the United States. *Journal of Economic Perspectives* 25(3), 173–196.

Monras, J. (2018). Economic Shocks and Internal Migration.

Moretti, E. (2011). Local Labor Markets. *Handbook of Labor Economics 4*, 1237–1313.

Moretti, E. (2012). What Workers Lose by Staying Put. *Wall Street Journal May 26*.

Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics* 5(1), 65–103.

Morten, M. and J. Oliveira (2016). Paving the Way to Development: Costly Migration and Labor Market Integration. Working Paper 22158, National Bureau of Economic Research.

Mulder, C. H. and G. Malmberg (2014). Local Ties and Family Migration. *Environment and Planning A: Economy and Place* 46(9), 2195–2211.

Notowidigdo, M. J. (2011). The Incidence of Local Labor Demand Shocks. Technical report, National Bureau of Economic Research.

Partridge, M. D., D. S. Rickman, M. R. Olfert, and K. Ali (2012). Dwindling U.S. Internal Migration: Evidence of Spatial Equilibrium or Structural Shifts in Local Labor Markets? *Regional Science and Urban Economics* 42(1-2), 375–388.

PSID (2014). Panel Study of Income Dynamics, Public Use Dataset. Technical report, Institute for Social Research, University of Michigan Survey Research Center.

Ravenstein, E. G. (1885). The Laws of Migration. *Journal of the Statistical Society of London* 48(2), 167–235.

Reisner, M. (1993). *Cadillac Desert: The American West and Its Disappearing Water*. Penguin.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2019). Integrated Public Use Microdata Series: Version 9.0 [Machine-Readable Database]. Minneapolis: University of Minnesota.

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy* 70(5), 80–93.

Tauchen, G. (1986). Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. *Economics Letters* 20(2), 177–181.

Thompson, D. (2016). How America Lost Its Mojo. *The Atlantic, May 27*.

Topel, R. (1986). Local Labor Markets. *Journal of Political Economy* 94(3), 111–143.

Trippett, F. (1979). The Great American Cooling Machine. *Time Magazine 13*.

Ullman, E. L. (1954). Amenities as a Factor in Regional Growth. *Geographical Review* 44(1), 119–132.

Wiggins, L. (1995). The Birth of the City of Miami. *Tequesta-The Journal of the Historical Association of Southern Florida. Number LV*.

Yagan, D. (2019). Moving to Opportunity? Migratory Insurance Over the Great Recession. *Journal of Political Economy* 127(5), 2505–2558.

Zabek, M. (2018). Local Ties in Spatial Equilibrium. mimeo., University of Michigan.

Zabel, J. E. (2012). Migration, Housing Market, and Labor Market Responses to Employment Shocks. *Journal of Urban Economics* 72(2), 267–284.

Appendices

A Supporting Exhibits

Figure A1 plots the change in migration for a partitioning of potential destinations. The destinations are grouped by "Neighbors" (adjacent LLMs) to the origin LLM, other LLMs in the same state as the origin LLM, others in the same region, and the remainder. The scatter plots show that declines in migration to each kind of destination are associated with declines in out-migration from the origin, and there is not a substitution of one type of destination for another. The pattern also holds when splitting by large and small LLM destinations or by common and uncommon destinations. Thus, the decline appears to be a general drying up of the origin out-mobility rate.

Figure A2 shows the interstate migration rates as measured in the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC, or commonly, the March CPS). The CPS contains the respondents' state of origin, which we can then classify into regions. The CPS is not our preferred dataset because (1) its sample is too small to study small areas over time, which is our focus here, (2) its geography is limited, so we cannot map it into a fully partitioned set of local labor markets, and (3) it is compromised by changes in sample and notably imputation procedures that affect estimates of migration rate (see Kaplan and Schulhofer-Wohl (2012)). However, it is a commonly used source in the literature.

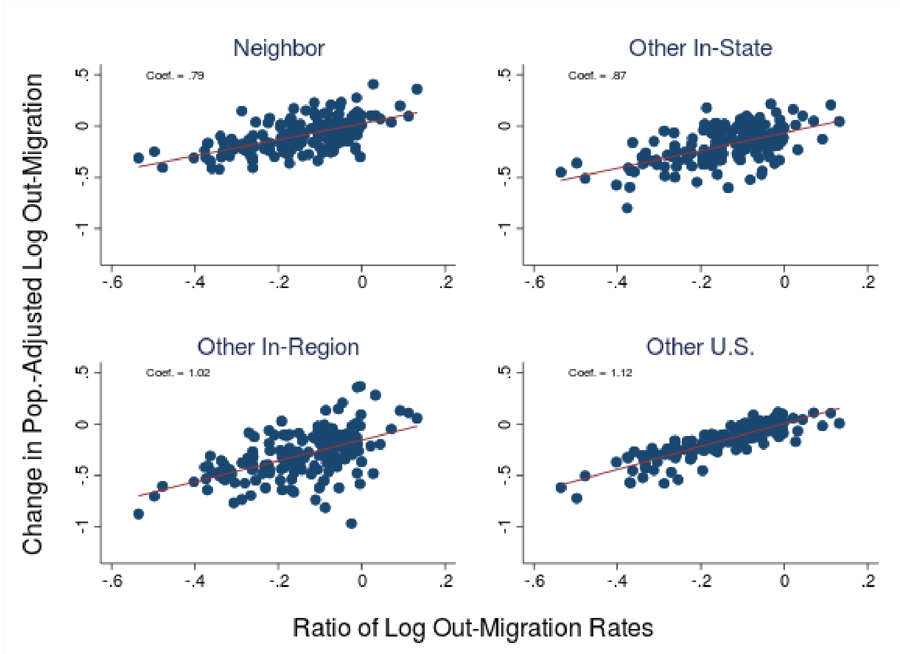
According to the figure, the West and Frontier regions show the greatest extent of the mobility decline. While the populations or definitions of location and migration event do not exactly align between the IRS and CPS, we see this as validation that the migration decline is a spatially concentrated phenomenon in certain regions of the country.

Figure A3 plots out mobility rates against average rootedness for the natives and U.S. born transplants in each of the 183 LLMs in our data. Higher rootedness is associated with lower out-mobility for natives, but shows no association with mobility for non-natives.

Figure A4 plots the migration rate out of a LLM for each tercile of the income distribution (plus those without reported income) against the LLM's average out-mobility rate. Among each income subgroup, higher out-mobility rates are associated with higher average mobility for the location, but the slope is strongest among the low income type. This is one reason we are concerned with controlling for the income distribution of LLMs, and the distribution of income by age/education/at-home status, within our model.

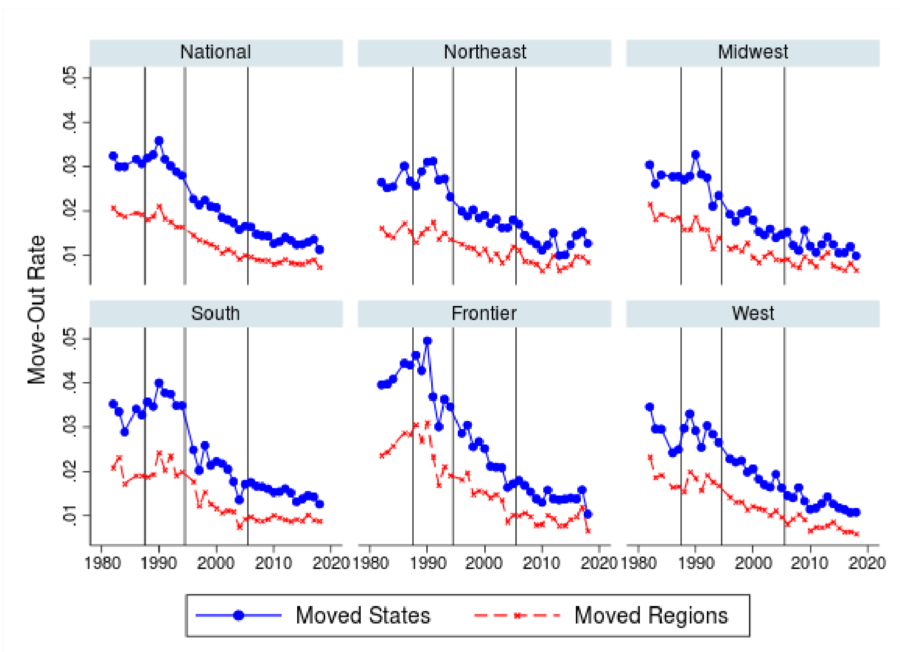
Table A1 shows migration rates by tercile of the income distribution, after adjusting for predicted wage based on age, education, and state of residence. The microdata come from the PSID and the income prediction comes from the CPS. Migration is defined by a change in state of residence. The "short-run position" uses the classification of residual income in the last two

Figure A1: Change in Out-Migration Rates by Destination



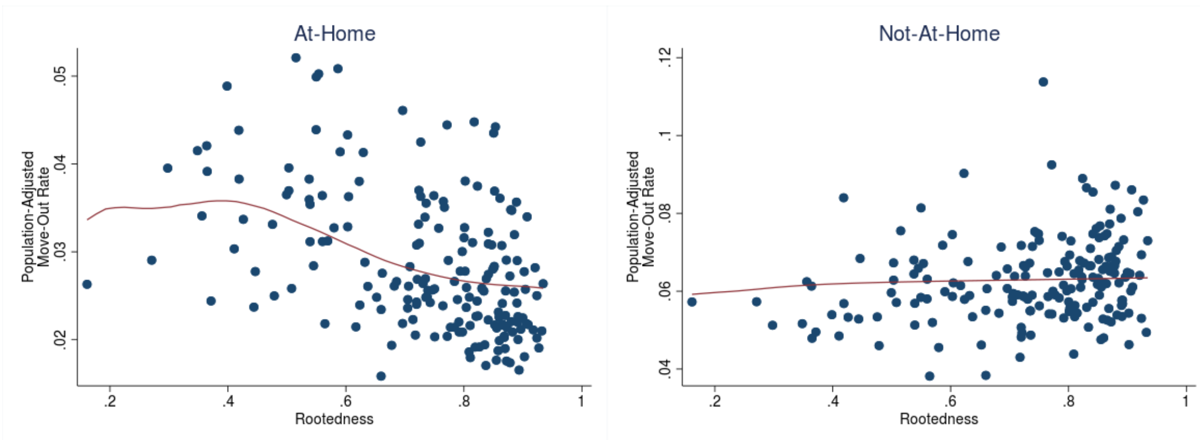
NOTES: The figure plots the log change in migration to a set of destinations to the log total change in out-migration for the origin. The destinations are grouped by "Neighbors" (adjacent LLMs) to the origin LLM, other LLMs in the same state as the origin LLM, others in the same region, and the remainder. Note the horizontal axis is the same data on all four plots. The population adjustment on the vertical axis accounts for population growth by differencing the destination flow and the origin population size, $y = (ln(flow_{late}) - ln(flow_{early})) - (ln(pop_{late}) - ln(pop_{early}))$. (Source: Author's calculations using IRS data).

Figure A2: Changes in CPS State-to-State Migration Rate Over Time, by Region



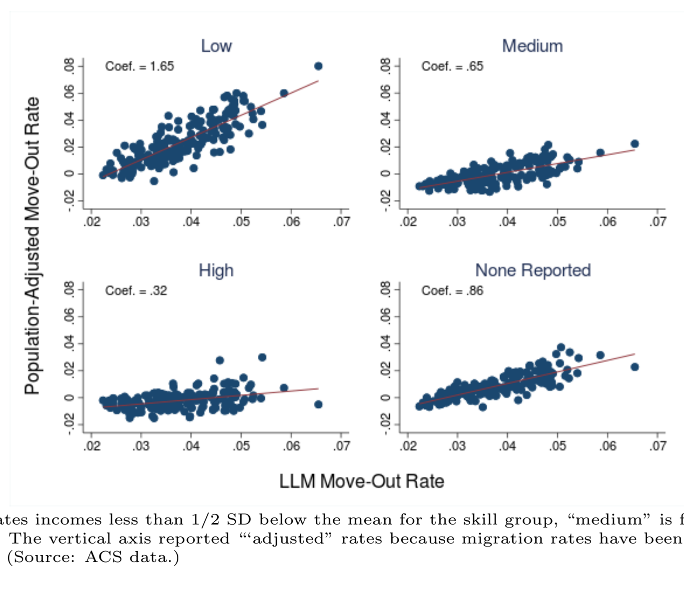
NOTES: The figure shows state-level migration rates nationally and by the five regions shown in Figure F1. An individual changing states within the same region is counted as a migration event. Vertical lines denote changes in imputation methodology in the CPS; see Kaplan and Schulhofer-Wohl (2012) for further detail. (Source: Author's calculations using CPS-ASEC data).

Figure A3: Migration Rates and Local Rootedness, by At-Home Status



NOTES: The figures plot average out-migration rates by LLM for residents flagged as being at-home or not-at-home; that is, in a LLM of their state of birth. (Source: ACS 2005-2017 data.)

Figure A4: LLM Mobility Rate by Point in the Income Distribution



NOTES: The “low” category indicates incomes less than 1/2 SD below the mean for the skill group, “medium” is from -1/2 to 1/2 SD above, and “high” is more than 1/2 SD above. The vertical axis reported “‘adjusted” rates because migration rates have been normalized to reflect a consistent composition by age and education. (Source: ACS data.)

Table A1: Individual Migration Rate by Point in the Income Distribution

Short-Run Position	Moved in Period Stay	Move	Long-Run Average Position	Moved in Sample Never	At Least Once
Low	10,641	267	Low	18,315	6,963
	0.976	0.024		0.725	0.275
Average	25,845	518	Average	32,533	12,493
	0.980	0.020		0.723	0.277
High	13,931	360	High	21,327	8,589
	0.975	0.025		0.713	0.287

NOTES: The table classifies workers in the PSID relative to their predicted income (given education, age, and state of residence) in a given year, categorizing the residual into his/her income position. The first set of columns uses the classification of the last two years and whether the person migrates in the following year. The second set averages over the person's individual average residual and relates their migration history in the sample. Migration events are defined as moves across states. (Source: PSID data.)

years and whether the person migrates in the following year. The "long-run position" averages over the person's individual average residual, creating an estimate of the individual's income "fixed effect," and relates their migration history in the sample. The table shows that higher income types are more likely to have ever moved, but high and low transient shocks to income are associated with higher mobility.

B Additional Simulation Results

B.1 Model Simulation Results by LLM

Table B1 reports the migrations changes for each separate LLM as well as the breakdown by categories of decomposition type.

B.2 Expanding the Simulation to the Entire Postwar Period

High quality data chronicling migration rates are a relatively new development, but using best available data prior to 1990, Fischer (2002), Molloy et al. (2017), and Kaplan and Schulhofer-Wohl (2017) document that migration rates in the U.S. peaked in the 1980s (or possibly late 1970s). Our main simulations start with the 1980 census and proceed to show the decline, but we are curious whether the mechanisms in our model would generate an inverted U-shaped pattern if we walked the model farther back into history. As a supplement to our main analysis, we simulate the model for 1950-1970 as well. We replicate Figure 7 for the longer time horizon in Appendix Figure B1. This comes with the caveat that we maintain the assumption of fixed primitives–an assumption that becomes less credible as we drift father back from our estimation period of 2005-2017. We nonetheless think it interesting to see the model's dynamics when constrained to contain only the mechanisms we focus on.

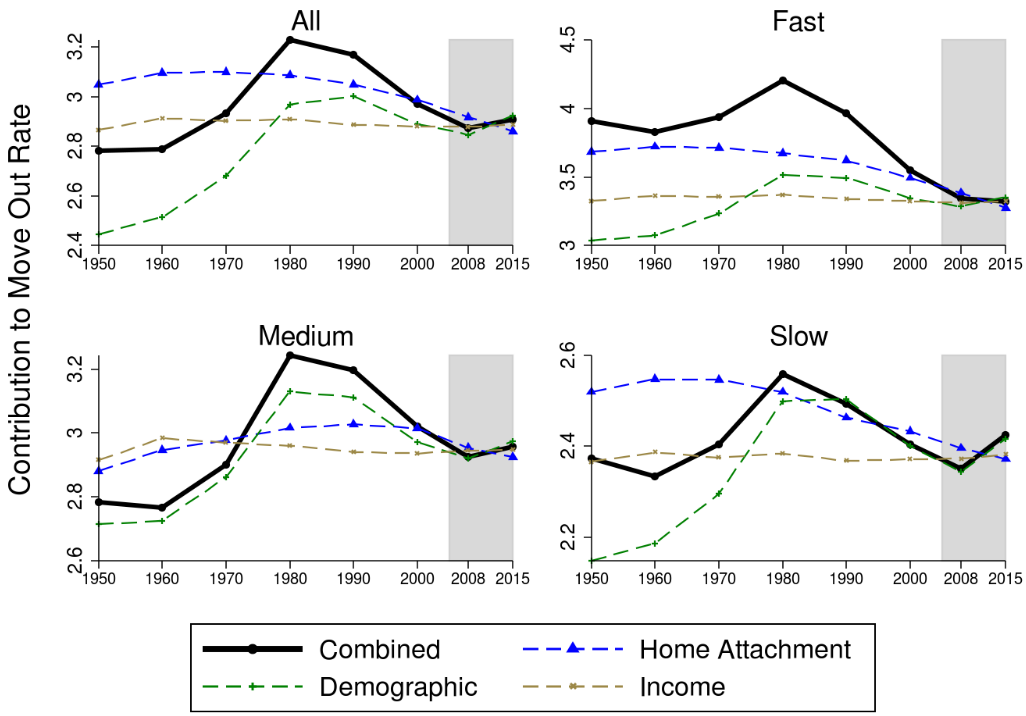
The model simulations indeed show a hump-shaped pattern of migration peaking in 1980. The decomposition indicates that the sharpness of the hump is the result of an educated and

Table B1: Simulated Migration Rates, With Decomposition of Change, by LLM

	Level in 1980	Total Change	Demographic	Home Attachment	Income
	1	2	3	4	5
Akron (OH)	2.943	-0.704	-0.205	-0.601	0.059
Albany (NY)	2.277	-0.193	-0.187	-0.056	-0.037
Atlanta (GA)	3.157	0.160	-0.459	0.365	-0.007
Austin (TX)	3.228	-0.004	-0.326	0.316	-0.167
Baltimore (MD)	3.309	-0.315	-0.016	-0.238	-0.051
Birmingham (AL)	2.166	0.145	-0.194	0.217	0.011
Boston (MA)	2.678	0.039	0.030	0.022	-0.110
Buffalo (NY)	2.279	-0.296	-0.127	-0.282	0.021
Charlotte (NC/SC)	2.418	0.826	-0.203	0.773	-0.055
Chicago (IL/IN/WI)	2.873	-0.372	-0.114	-0.319	0.018
Cincinnati (OH/KY/IN)	2.321	-0.115	-0.129	-0.083	0.008
Cleveland (OH)	3.038	-0.676	-0.180	-0.592	0.077
Columbus (OH)	3.023	-0.319	-0.118	-0.245	-0.008
Dallas-Fort Worth (TX)	3.188	-0.296	-0.304	0.008	-0.046
Dayton (OH)	3.138	-0.569	-0.221	-0.477	0.083
Denver (CO)	4.552	-0.886	-0.478	-0.249	-0.052
Detroit (MI)	2.856	-0.740	-0.162	-0.716	0.115
El Paso (TX/NM)	3.262	-0.532	-0.161	-0.436	0.055
Fort Myers (FL)	5.127	-1.147	-0.618	-0.233	-0.046
Fresno (CA)	3.985	-1.402	-0.175	-1.012	0.025
Grand Rapids (MI)	2.333	-0.193	-0.175	-0.178	0.065
Greensboro (NC/VA)	2.253	0.435	-0.238	0.468	0.032
Greenville (SC/NC)	2.255	0.454	-0.226	0.488	0.026
Harrisburg (PA)	2.276	0.253	-0.272	0.361	-0.004
Houston (TX)	3.339	-0.497	-0.338	-0.168	0.041
Indianapolis (IN)	3.160	-0.311	-0.150	-0.245	0.031
Jacksonville (FL/GA)	3.595	-0.261	-0.301	0.024	-0.017
Kansas City (MO/KS)	2.787	-0.157	-0.225	-0.028	0.017
Lancaster (PA)	1.965	0.265	-0.160	0.288	-0.016
Las Vegas (NV)	5.223	-1.027	-0.614	-0.110	0.024
Los Angeles (CA)	4.396	-1.330	-0.150	-0.837	0.023
Louisville (KY/IN)	2.192	0.103	-0.163	0.134	0.008
Manchester (NH)	4.313	-0.595	-0.514	0.136	-0.160
McAllen (TX)	2.476	-0.135	-0.182	-0.106	-0.026
Memphis (TN/MS/AR)	2.354	-0.045	-0.217	0.109	-0.012
Miami (FL)	4.733	-1.286	-0.101	-0.496	-0.014
Milwaukee (WI)	2.627	-0.157	-0.231	-0.071	0.042
Minneapolis (MN/WI)	2.490	-0.117	-0.226	-0.002	-0.026
Monmouth (NJ)	3.529	-0.689	-0.231	-0.338	-0.027
Nashville (TN)	2.737	0.503	-0.220	0.543	-0.036
New Orleans (LA)	2.551	-0.318	-0.240	-0.174	0.022
New York (NY/NJ/CT)	2.633	0.041	0.077	-0.116	-0.066
Norfolk (VA/NC)	3.633	-0.096	-0.378	0.268	-0.040
Oklahoma City (OK)	3.359	-0.300	-0.319	-0.080	0.009
Orlando (FL)	5.088	-0.965	-0.467	-0.343	0.026
Philadelphia (PA/NJ/DE)	2.579	-0.159	-0.014	-0.212	-0.016
Phoenix (AZ)	5.015	-1.146	-0.420	-0.377	-0.052
Pittsburgh (PA)	1.872	0.182	-0.100	0.090	0.060
Portland (OR/WA)	3.985	-0.591	-0.290	-0.211	-0.010
Providence (RI/MA)	2.256	-0.046	-0.164	0.071	-0.078
Raleigh-Durham (NC)	2.913	0.664	-0.455	0.776	-0.138
Residual	3.386	-0.557	-0.196	-0.271	-0.046
Residual Connecticut (CT)	3.727	-0.504	-0.365	-0.034	0.020
Richmond (VA)	3.025	0.028	-0.230	0.137	-0.015
Riverside-San Bernardino (CA)	4.602	-1.780	-0.214	-1.201	0.032
Rochester (NY)	2.501	-0.289	-0.256	-0.188	0.071
Sacramento (CA)	4.090	-1.277	-0.180	-0.798	-0.029
Salem (OR)	4.614	-1.017	-0.500	-0.503	0.108
Salt Lake City (UT)	3.249	-0.335	-0.293	-0.054	0.000
San Antonio (TX)	2.777	-0.149	-0.149	-0.036	-0.031
San Diego (CA)	5.107	-1.611	-0.106	-0.838	-0.104
San Francisco (CA)	4.172	-1.029	-0.002	-0.557	-0.152
San Jose (CA)	4.143	-1.048	0.041	-0.475	-0.173
Seattle (WA)	3.956	-0.655	-0.226	-0.223	-0.095
St. Louis (MO/IL)	2.410	-0.232	-0.136	-0.196	0.025
Syracuse (NY)	2.296	-0.294	-0.224	-0.156	0.000
Tampa (FL)	4.945	-0.998	-0.332	-0.322	-0.107
Tucson (AZ)	4.762	-1.029	-0.251	-0.503	0.058
Tulsa (OK)	3.498	-0.444	-0.345	-0.218	0.042
Washington (DC/VA/MD)	4.147	-0.647	-0.179	-0.248	-0.075

NOTES: All figures are in percentages (e.g., "0.5" corresponds to a one-half percentage point change).

Figure B1: Simulated Trends in Mobility, Decomposed



NOTES: The figures plot the time path of mobility generated by the model in total and counterfactual subtotal simulations for each category of LLMs. Note that each panel has its own vertical scale. Subtotals may not add to the combined value because of nonlinearities in the model. The shaded region denotes the estimation period.

youthful workforce (the Baby Boom generation) emerging from 1970 to 1990. Increasing home attachment also follows an inverted U pattern, peaking at different times for LLMs of different speeds, but as shown in the main analysis, is more dominant in the trend of later decades of the simulation.

C Modeling Details

C.1 Model Setup: Choice Probabilities and Value Functions

For convenience, here we rewrite the choice probability and value functions from section 3.

C.1.1 Lower Nest: Where to, Conditional on Moving

The probability of choosing destination j conditional on current location o is given by

$$Pr(j|o) = \sigma_{jo} = \frac{exp[v_{j|o}]^{\frac{1}{\lambda}}}{\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}}. \tag{C1}$$

Taking logs yields

$$ln\sigma_{jo} = \frac{1}{\lambda}v_{j|o} - ln\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}. \tag{C2}$$

The expected value of choosing a destination optimally (the 'Emax') is the value of moving out of origin o, $V_{m|o}$:

$$V_{m|o} = \lambda ln\big(\sum_k exp[v_{k|o}]^{\frac{1}{\lambda}}\big). \tag{C3}$$

C.1.2 Upper Nest: Move/Stay Decision

The upper nest is the binary stay or move decision. Letting $V_{s|o}$ denote the value of remaining in the origin this period, the respective probabilities are

$$Pr(stay) =\sigma_s = \frac{exp[V_{s|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}} \tag{C4a}$$

$$Pr(move) =\sigma_m = \frac{exp[V_{m|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}}. \tag{C4b}$$

Similar to (C1), the expected value of being faced with a move/stay decision in some origin o is

$$V_o = \delta ln\big[exp[V_o]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}\big]. \tag{C5}$$

This is the value of being in the location indexed by o.

The closed form logit choice probabilities have the usual convenient features for expressing log choice probabilities and odds ratios. Taking logs in equation (C1) and differencing two destinations yields

$$ln\sigma_{j|o} =\frac{1}{\lambda}\big[v_{j|o} - ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big] \tag{C6a}$$

$$ln\sigma_{j|o} - ln\sigma_{i|o} =\frac{1}{\lambda}\big(v_{j|o} - v_{i|o}\big) \tag{C6b}$$

and from equations (C4b) and (C4a), using C3 in the third line:

$$
\begin{aligned}
ln\sigma_s =& \frac{1}{\delta}V_o - ln\big(exp[V_o]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}\big) && \text{(C7a)}\\
ln\sigma_m =& \frac{1}{\delta}V_{m|o} - ln\big(exp[V_o]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}\big) && \text{(C7b)}\\
ln\sigma_s - ln\sigma_m =& \frac{1}{\delta}\big(V_o - \lambda ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big). && \text{(C7c)}
\end{aligned}
$$

C.2 Locational Heterogeneity in Income Search

Locations are heterogeneous in the incomes they offer, so that utility afforded by the income point n may vary, i.e., $y_n^j > y_n^k$ for some locations j, k. A higher mean shifts the entire distribution, so that both $w_0(n)$ and $w_c(n)$ move proportionately at each n.[68]

The effect of dispersion is perhaps less obvious. Higher dispersion can increase the value function of the worker, ceteris paribus, because of the presence of the optionality in the successful search. The availability of high draws is a good when there is opportunity to reject low ones, as this model allows. The w_0 term is not affected by a mean-preserving spread in income, but the w_c term is convex in the variance of the income distribution. Mathematically, given a Type I EV assumption on the shock in the successful search event, $w_c(n) = \sum_{n'} \pi_{n'|n} ln(exp(w_{n'}) + exp(w_n)) - \Delta$, where Δ is Euler's constant (to adjust for the mean of the T1EV distribution), which is increasing in the spread of the distribution of n'. To see the effect of dispersion on this value, consider a mean-preserving spread of σ separating two income levels, $w_1 = w + \sigma$, $w_2 = w - \sigma$. The expected value of receiving these two options is

$$w_c = ln(exp(w + \sigma) + exp(w - \sigma)).$$

The log transformation is monotonic, so to show the sign, we focus on the sum of the exponential terms. Its derivative is

$$\frac{\partial exp(w_c)}{\partial \sigma} = \sigma(exp(w + \sigma) - exp(w - \sigma)) > 0.$$

Since σ is positive and the exponent is an increasing function, the value is increasing in the spread.

Next, the income search value function is clearly affected by the probability of a successful search in ways that interact with the income distribution. Writing out the expected value of income search, (9),

[68] The w_c term is slightly less sensitive to the mean than w_0 because a successful search induces some reversion by providing a mixture of two draws from the distribution.

$$\omega_n = \sum_{n'} \pi_{n'|n} \Big[\gamma ln(exp(w_{n'}) + exp(w_n)) - \Delta + (1-\gamma) w_{n'} \Big],$$

we see that the derivative of this expected value with respect to the contact probability is

$$\frac{\partial \omega_n}{\partial \gamma} = \sum_{n'} \pi_{n'|n} \Big[(ln(exp(w_{n'}) + exp(w_n)) - \Delta - w_{n'} \Big] > 0,$$

which is the probability-weighted gap between the expected income resulting from the successful and unsuccessful search. This gap is increasing in the income distribution mean (in general) and variance (in the support of actual data for U.S. cities) because of the nonlinearity in the successful search term.

Combining these last two results would show that $\frac{d^2 w_n}{d\sigma\gamma} > 0$. Thus, we allow for the possibility that higher mean and/or higher dispersion locations have been more affected by increasing information availability. As information increases, local search will dominate nonlocal search to a greater extent when the local market has higher mean and/or variance.

D Estimation Details

D.1 Deriving Estimating Equations

D.1.1 Destination Conditional on Moving

The value of a location depends on (i) its current offer of flow utility, (ii) the size of switching costs between the origin and the new location, and (iii) the continuation value offered by placing oneself in a new state. Writing out the components of the difference in values shows that the odds ratio in (C6b) is

$$\begin{aligned} ln\sigma_{j|o} - ln\sigma_{i|o} =& \frac{1}{\lambda}\Big(u_j + mc_{jo} + \beta ln\big(exp[V_{s|j}]^{\frac{1}{\delta}} + exp[V_{m|j}]^{\frac{1}{\delta}}\big)\Big) \\ &-\frac{1}{\lambda}\Big(u_i + mc_{io} + \beta ln\big(exp[V_{s|i}]^{\frac{1}{\delta}} + exp[V_{m|i}]^{\frac{1}{\delta}}\big)\Big). \end{aligned} \tag{D1}$$

Intuitively, this expression says the relative probability of choosing two locations is a matter of (i) the difference in their utilities, (ii) the difference in the move costs in reaching them from the origin o, and (iii) the difference in the continuation value induced by changing one's station to j vis-a-vis i. The latter could matter in a model with geography (including birthplace geography), as j and i may be more or less remote from other locations, or may differ in the home premium they offer the individual. The future value components can be substituted using (C5), which then appears in the denominator of (C4b), allowing for substitution of (C3), yielding

$$
\begin{aligned}
ln\sigma_{j|o} - ln\sigma_{i|o} =& \frac{1}{\lambda}(u_j + mc_{jo} - u_i - mc_{io}) \\
&+ \frac{\beta}{\delta} ln\big(\sum_k exp[v'_{k|j}]^{\frac{1}{\lambda}}\big) - \frac{\beta}{\lambda} ln\sigma'_{m|j} \\
&- \frac{\beta}{\delta} ln\big(\sum_i exp[v'_{k|i}]^{\frac{1}{\lambda}}\big) - \frac{\beta}{\lambda} ln\sigma'_{m|i},
\end{aligned} \tag{D2}
$$

where the prime symbol is used to indicate the value of the variable in the next period. This equation has substituted out the future value terms for the (relative difference in) future moving probabilities and the expected value function in the where-to moving decision. Note that we have elected to take the future value with respect to the moving probability and value rather than that of staying. This allows us to substitute out the value of moving out from the two candidate locations, and normalize relative to an arbitrary third location z. The normalization obtains because the value of ending up in z will be constant for the individual, although the cost of reaching z may differ between j and i. Therefore the probability of reaching z will depend on the next-selected destination, but otherwise the history of choices is "forgotten" once z is reached. In this way, we are leveraging the logic of finite dependence to iteratively substitute out future value terms, returning to a renewal state. Using (C6b), we derive

$$
\begin{aligned}
ln\sigma_{j|o} - ln\sigma_{i|o} =& \frac{1}{\lambda}(u_j + mc_{jo} - u_i - mc_{io} - \beta(ln\sigma'_{m|j} - ln\sigma'_{m|i})) \\
&+ \frac{\beta}{\delta}\big(\frac{1}{\lambda}(mc_{zj} - mc_{zi} - (ln\sigma'_{zj} - ln\sigma'_{zi})\big).
\end{aligned} \tag{D3}
$$

Equation (D3) is now a reduction of (C6b) to parameters–in the utility and move cost functions, and the scale parameters–and moments from the data, the choice probabilities.

D.1.2 The Move/Stay Decision

The log odds ratio of staying (not migrating) is given in (C7c). This can also be converted to a linear estimating equation when differencing relative to a normalizing origin location z. The basic idea is to iteratively apply the forward substitution used to account for the continuation value terms as in (D3). Several more forward substitutions are needed to return the estimating equation, but the logic is the same as that used in the last subsection.

$$
\begin{aligned}
(ln\sigma_{so} - ln\sigma_{mo}) - (ln\sigma_{sz} - ln\sigma_{mz}) =& \frac{1}{\delta}\big(V_o - \lambda ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big) \\
&-\frac{1}{\delta}\big(V_z - \lambda ln\big(\sum_i exp[v_{i|z}]^{\frac{1}{\lambda}}\big)\big) \\
=& \underbrace{\frac{1}{\delta}\big(V_o - V_z\big)}_{staying} \\
&\underbrace{-\frac{\lambda}{\delta}\big(ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big) - ln\big(\sum_i exp[v_{i|z}]^{\frac{1}{\lambda}}\big)\big)}_{moving}.
\end{aligned}
\tag{D4}
$$

We will treat the "staying" and "moving" blocks separately for convenience. The difference in the value of staying in o relative to z is written out as

$$
\begin{aligned}
V_o - V_z =& u_o + \beta ln\big[exp(V'_o)^{\frac{1}{\delta}} + \lambda ln\big(\sum_i exp[v'_{i|o}]^{\frac{1}{\lambda}}\big)^{\frac{1}{\delta}}\big] \\
&-\Big(u_z + \beta ln\big[exp(V'_z)^{\frac{1}{\delta}} + \lambda ln\big(\sum_i exp[v'_{i|z}]^{\frac{1}{\lambda}}\big)^{\frac{1}{\delta}}\big]\Big) \\
=& u_o + \beta\frac{\lambda}{\delta} ln\big(\sum_k exp[v'_{k|o}]^{\frac{1}{\lambda}}\big) - ln\sigma'_{m|o} \\
&-\Big(u_z + \beta\frac{\lambda}{\delta} ln\big(\sum_k exp[v'_{k|z}]^{\frac{1}{\lambda}}\big) - ln\sigma'_{m|z}\Big).
\end{aligned}
\tag{D5}
$$

We first expanded the expression into flow utilities and continuation values and then substituted the continuation value using (C7b) and (C3). The relative continuation values of staying are thus expressed as the expected value of an optimal move less the probability of moving anywhere, conditioning on origin o versus z. (For comparison, this equation looks like (D2) without the moving cost terms.)

We are now in position to employ a substitution for the expected value of a move using (C6a) for the $ln\big(\sum_k exp[v'_{...}]^{\frac{1}{\lambda}}\big)$ terms.

$$
\begin{aligned}
V_o - V_z =& u_o - u_z - (ln\sigma'_{m|o} - ln\sigma'_{m|z}) \\
&+\beta\frac{\lambda}{\delta}\big(\frac{1}{\lambda}(v_{k|o} - v_{k|z}) - (ln\sigma_{ko|m} - ln\sigma_{kz|m})\big).
\end{aligned}
\tag{D6}
$$

We have now expressed the expected value of the move as the choice-specific value of some location k minus the probability of moving there. The relative values between starting this choice from o vis-a-vis z is simply the difference in the cost of reaching the location, as once the agent is in k, there is no impact of the memory of how she got there. That is, we again leverage the property of finite dependence–in one more step, agents can be returned to equivalent places in the state space. Thus, in the same substitution that arrived at (D3), here we have

$$\begin{aligned} V_o - V_z =& u_o - u_z - (ln\sigma'_{m|o} - ln\sigma'_{m|z}) \\ &+\beta\frac{\lambda}{\delta}\big(\frac{1}{\lambda}(mc_{ko} - mc_{kz}) - (ln\sigma'_{ko|m} - ln\sigma'_{kz|m})\big). \end{aligned} \tag{D7}$$

The moving block uses the same technique, substituting out the expected value of a move employing finite dependence. (The staying block merely needed one more step to arrive here.) We thus have

$$\begin{aligned} &\frac{\lambda}{\delta}\big(ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big) - ln\big(\sum_i exp[v_{i|z}]^{\frac{1}{\lambda}}\big)\big) = \\ &\qquad \frac{\lambda}{\delta}\big(\frac{1}{\lambda}(mc_{ko} - mc_{kz}) - (ln\sigma_{ko|m} - ln\sigma_{kz|m})\big). \end{aligned} \tag{D8}$$

One difference of note between (D7) and (D8) is that the former uses one period ahead choice probabilities (and is discounted by β) while the latter uses current period choice probabilities. Our estimation focuses on the geography and ignores aggregate shocks or trends, so that $ln\sigma_{ko|m} = ln\sigma'_{ko|m}$. In other applications, one may need to forecast the differences between these as states evolve. Subject to this caveat, our estimating equation combines (D7) and (D8) to yield

$$\begin{aligned} &(ln\sigma_{so} - ln\sigma_{mo}) - (ln\sigma_{sz} - ln\sigma_{mz}) + \frac{\beta}{\delta}(ln\sigma'_{m|o} - ln\sigma'_{m|z}) = \\ &\qquad \frac{1}{\delta}(u_o - u_z) + (\beta - 1)\frac{\lambda}{\delta}\big(\frac{1}{\lambda}(mc_{ko} - mc_{kz}) - (ln\sigma_{ko|m} - ln\sigma_{kz|m})\big). \end{aligned} \tag{D9}$$

which is a function of only choice probabilities and utility parameters.

D.2 The Estimation Procedure

Estimation relies on information in both move versus stay decisions and in the propensity to choose one location over another, conditional one one's own attributes (such as birthplace). That is, we can stack (D3) and (D9) into one simultaneous equation problem evaluated by standard matrix operations.[69]

To do so, we need to make some practical decisions over how many moments to target. For each origin, there are $J-1$ potential destinations, so with a normalizing destination, there are $J-2$ choice probabilities on the left-hand side of (D3). However, each of them has another $J-1$ choice probabilities for each destination, meaning there would be $J \times (J-2) \times (J-1)$ equations for each type of agent in the data. This becomes a computational problem when J is large and there are many types (we have $A \times E \times (J+1) = 568$ types). In practice, we will ignore the last term comprising the second line of (D3), essentially treating it as specification error. Our reason for doing so is that the computational savings are large (dropping the $J-1$ factor in the number of equations) while these extra moments yield little additional information. They represent the differential move cost and choice probability of reaching the outside option z, which are fairly similar across places. In other words, the value of a destination is chiefly determined by its utility, the move cost of reaching it (geography), and the probability of moving out of it again, but not by how easy or difficult it becomes to reach a rural area from there. Besides, the term is multiplied by $\frac{\beta}{\delta}$, a small decimal in our calibration, that substantially reduces the contribution of this term to the variance.

There is, naturally, one move/stay decision for each origin-type, but there could be as many as $J-1$ equations for (D9), depending on how many k potential destinations we want to include. In contrast to the dropping of the last term in (D3), there is geographical heterogeneity represented in the mc_{ko}, $ln\sigma_{ko|m}$ terms of (D9), so we elected to use all the potential destinations, as they might help in identifying the move cost terms or correcting for differences in option value.[70] Ex post we found the contribution of these terms to be small, and qualitatively, the results are similar either way. The stacked system of equations is

[69] If the utility function were nonlinear in parameters, the objective function could still be evaluated using standard methods, although not simple matrix inversion, obviously. The main point of our procedure is that finite dependence has yielded a simple set of targeted moments.

[70] This also balances the number of equations between the move/stay and moving-to-where contributions, although such a balance could also be accomplished through appropriate weighting.

$$
\underbrace{\begin{bmatrix}
ln\frac{\sigma_{so}}{\sigma_{mo}} - ln\frac{\sigma_{sz}}{\sigma_{mz}} + \frac{\beta}{\delta}ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}} - \frac{(\beta-1)\lambda}{\delta}ln\frac{\sigma_{1o|m}}{\sigma_{1z|m}} \\
\cdots \\
ln\frac{\sigma_{so}}{\sigma_{mo}} - ln\frac{\sigma_{sz}}{\sigma_{mz}} + \frac{\beta}{\delta}ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}} - \frac{(\beta-1)\lambda}{\delta}ln\frac{\sigma_{Ko|m}}{\sigma_{Kz|m}} \\
ln\frac{\sigma_{1|o}}{\sigma_{z|o}} + \beta ln\frac{\sigma'_{m|1}}{\sigma'_{m|z}} \\
\cdots \\
ln\frac{\sigma_{K|o}}{\sigma_{z|o}} + \beta ln\frac{\sigma'_{K|1}}{\sigma'_{m|z}}
\end{bmatrix}}_{Y} =
$$

$$
\underbrace{\begin{bmatrix}
\frac{1}{\delta} & \cdots & \frac{1}{\delta} & 0 & \cdots & 0 \\
\frac{(\beta-1)\lambda}{\delta} & \cdots & \frac{(\beta-1)\lambda}{\delta} & 0 & \cdots & 0 \\
0 & \cdots & 0 & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda} \\
0 & \cdots & 0 & \frac{1}{\lambda} & \cdots & \frac{1}{\lambda}
\end{bmatrix}}_{\Delta}
\underbrace{\begin{bmatrix}
u(x_1) - u(x_z) & mc(d_{1o}) - mc(d_{1z}) \\
\cdots & \cdots \\
u(x_K) - u(x_z) & mc(d_{Ko}) - mc(d_{Kz}) \\
u(x_1) - u(x_z) & mc(d_{1o}) - mc(d_{1z}) \\
\cdots & \cdots \\
u(x_K) - u(x_z) & mc(d_{Ko}) - mc(d_{Kz})
\end{bmatrix}}_{X}
\underbrace{\begin{bmatrix} \theta_u \\ \theta_{mc} \end{bmatrix}}_{\theta}, \quad \text{(D10)}
$$

where differences have been made to ratios for readability. Y is the choice probabilities from the data, X is the function of utilities and moving costs (e.g., whether a location is home, how far two locations are from each other, etc.), Δ are the scaling parameters (determined elsewhere, as will be explained), and finally θ is the vector of parameters to be recovered. A standard regression of Y on $[\Delta X]$, with choice of a weight matrix, as appropriate, will yield θ.

D.3 Additional Estimating Equations

Equation (D10) identifies the main parameters of interest off of differences between locations and a normalizing locale z. Hence, scale parameters are not identified here and must be calibrated elsewhere. The set of scaling parameters includes λ, δ, β and intercepts of a move cost function (whereas D10 identifies the distance parameters that vary between locations).

Move Cost Intercepts

An estimator for the move cost intercepts can be derived using substitutions similar to those employed in sections D.1.1 and D.1.2.[71] Using (C7c) and substituting (C6a), we have

[71] Monras (2018) does not use a move cost term, instead calibrating an average difference in elasticities between nests, i.e., λ relative δ. We could have gone this route, although we prefer using moving costs to compare across types (who have substantially different move rates, as in young versus old) without imposing assumptions about between-destination elasticity. A move cost specification is also consistent with our environment in a model with geography, where some locations are closer in space than others.

$$
\begin{aligned}
ln\sigma_s - ln\sigma_m =& \frac{1}{\delta}\big(V_o - \lambda ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big) \\
=& \frac{1}{\delta}\big(V_o - v_{j|o} - \lambda ln\sigma_{j|o}\big) \\
=& \frac{1}{\delta}\big(V_o - V_j - mc_{j|o} - \lambda ln\sigma_{j|o}\big).
\end{aligned}
\tag{D11}
$$

Unlike the expressions in D.1.1 and D.1.2, this does not difference out the intercept of the move cost term, but it also retains the value functions of residing in o and j, which are unknown. To isolate the move cost intercept, we saturate the equation with type-by-location fixed effects (for each pair of origin and destination) to absorb the value functions. (That is, we effectively can estimate V_o as a composite object, but not its components). We then recover the move cost intercept (by type) from the average amount of migration observed in the data, correcting for differences in relative attractiveness of a location (V_o versus V_j) and its average remoteness ($mc_{j|o}$). The estimating equation is

$$
ln\sigma_s - ln\sigma_m - \frac{\lambda}{\delta} ln\sigma_{j|o} + \frac{1}{\delta} mc_{j|o} = \frac{1}{\delta}\big(V_o - V_j\big) - \frac{1}{\delta} mc_0, \tag{D12}
$$

where the left-hand side is composed of choice probability data and previously recovered moving costs with respect to distance.[72] From this equation we recover $A \times E \times 2 = 16$ move cost terms–one for each decade/education group (8), for the US and foreign born (x2).

D.4 Scale Parameters

The discount rate is set to $\beta = 0.95$.

Finally, we are ready to derive a calibration for the scale parameters, λ, δ. These are the most difficult to identify in our environment, because we are using cross-sectional variation.[73] Our primary objective here is to obtain values that preserve the main feature in the data– large move-home flows and small move-out-of-home flows. To that end, we do this in the most straightforward way: We compare move/stay decisions for home versus not-home locations to move-to decisions for home versus not-home locations, and set a scale parameter once at the outset. That is, we look at the ratios of C6b to C7c to elicit $\frac{\lambda}{\delta}$. In practice, we estimate dummy variables for whether the individual is at home in move/stay, and the frequency of move-home decisions, and take their ratio as $\frac{1}{\delta}$, effectively setting λ to one. This procedure will not identify

[72] The estimates are largely similar when ignoring the relative move cost term and treating it as specification error. That is, while distance impacts greatly the set of destinations one reaches, remoteness does not seem to be driving average mobility rates. Note also that relative move cost is mean zero by construction, although in principle it could be correlated with the value of the location, so it is technically correct to include it.

[73] In contrast, Artuc et al. (2010) and Monras (2018) use time variation for identification.

utility primitives, as it only absorbs value functions without definition of their parameters, but is sufficient to find the average ratio of outflow to inflow rates. Chronologically, this is the first step, and we feed the calibrated δ into the estimating equations (D10) and (D12). We set δ for each education group (although they are coincidentally of similar size), and then use cross-location, cohort, and age variation in choice probabilities to estimate the remaining parameters.

Writing out C6b and C7c shows the idea:

$$
\begin{aligned}
&ln\sigma_{j,o} - ln\sigma_{k,o} = \\
&\quad \frac{1}{\lambda}\Big(u_j - u_k - \beta(ln\sigma_{m,j} - ln\sigma_{m,k}) + mc_{j,o} - mc_{k,o} + \frac{\beta}{\delta}[(mc_{i,j} - mc_{i,k}) - (ln\sigma_{i,j} - ln\sigma_{i,k})]\Big)
\end{aligned} \tag{D13a}
$$

$$
\begin{aligned}
&\frac{ln\sigma_{s,o}}{ln\sigma_{m,o}} - \frac{ln\sigma_{s,k}}{ln\sigma_{m,k}} = \\
&\qquad \frac{1}{\delta}\Big(u_j - u_k - \beta(ln\sigma_{m,o} - ln\sigma_{m,k}) + (\beta - 1)[(mc_{i,o} - mc_{i,o}) - (ln\sigma_{i,o} - ln\sigma_{i,o})]\Big),
\end{aligned} \tag{D13b}
$$

where the first-order terms on the right-hand side are the same. Using indicators/controls for the right-hand side will capture the composite effect of flow utility and continuation value (and therefore does not identify parameters), but can approximately get the scale differences between the move out and move-to decisions. Additional controls help with omitted variables. We use moving costs (in the upper equation) and ignore the trailing terms, treating them as specification errors. The trailing terms merely represent the change in option value created by choosing one location versus another *via its change in the accessibility of other locations.* The change in option value from the move out probability is already captured.

We use the ratio on the home indicator control to best match the elasticity of choosing home from afar versus the reduced likelihood of moving away from home.

D.5 Auxiliary Model: Counterfactual Conditional Choice Probabilities

The continuation value of a location j can be expressed using either from (C7a), $FV_j = \frac{1}{\delta}V_j - ln\sigma_{s|j}$, or from (C7b), $FV_j = \frac{1}{\delta}\lambda ln\big(\sum_i exp[v_{i|j}]^{\frac{1}{\lambda}}\big) - ln\sigma_{m|j}$, which is in turn $\frac{\lambda}{\delta}\big(\frac{1}{\lambda}V_k + \frac{1}{\lambda}mc_{kj} - ln\sigma_{k|j,m}\big) - \ln\sigma_{m|j}$.

Consider defining FV_j by (C7b) and FV_k by (C7a). Then the difference is

$$FV_j - FV_k = \frac{1}{\delta}(V_k + mc_{kj}) - \frac{\lambda}{\delta}ln\sigma_{k|j,m} - ln\sigma_{m|j} - (\frac{1}{\delta}V_k - ln\sigma_{s|k}).$$

The V_k terms cancel to arrive at the relative future values of

$$FV_j - FV_k = ln\sigma_{s|k} - ln\sigma_{m|j} + \frac{1}{\delta}mc_{kj} - \frac{\lambda}{\delta}ln\sigma_{k,j|m}. \tag{D14}$$

This expresses the relative future value of two choices as a function of the move rates, stay rates, and the probability of moving to one location from the other, correcting for the move cost. When choosing a normalizing location to have continuation value of zero (i.e., when $j = k$), we have relative future values for all other places using the choice probabilities and parameters.

What remains is to estimate the choice probabilities from the data. We want to do this in a way that projects counterfactual choice probabilities as we alter states within the model environment, such as income distributions or home attachment. In practice, we flexibly estimate the move out and move in probabilities as interactions of linear functions of income mean, dispersion, and rootedness. When we alter these in simulations, we project new choice probabilities given the new set of income distributions or rootedness.

We need projections of the probability of moving from any origin, $\sigma_{s|k}$, and the probability of choosing an arbitrary location k conditional on living in j, $\sigma_{k,j|m}$. The latter requires a modeling decision on what the arbitrary location k will be. In principle, any place will do, but in practice, it is easier to estimate an auxiliary model on a frequently chosen place. We use the home location for U.S.-born residents, and the residual location for the foreign born.

Then, using the ACS microdata, we estimate two choice probability functions, f, g that take as arguments the income and home preference features of the locations:

$$\sigma_{m|j} = f(y^j_\mu, y^j_\sigma, n, I(j = h)) \tag{D15}$$

$$\sigma_{k,j|m} = g(y^j_\mu, y^j_\sigma, n, I(j = h)). \tag{D16}$$

In practice we use the mean and standard deviation of income, interacted with dummies for whether the individual is a high-, medium-, and low-income type, the rootedness of the location interacted with home status, and indicators for whether the location is the residual. For the destination conditional probability, we use a set of LLM pair dummy variables to capture distance in a flexible way. The equations are estimated separately for each age/education group, and within group, separately for the U.S. and foreign born.

The parameters of these equations allow us to project choice probabilities for alternative values of income distributions or rootedness. This exercise is not used for the purposes of identifying anything in the model primitives. Rather, it serves as a projection of choice probabilities outside the data for use in the CCP substitution of the value function in counterfactual simulations. These projections therefore work in concert with the flow utility differences, not in place of them. (If the choice probabilities are observed, we can enter them into a simulation directly.)

This allows the model to simulate future values without assuming a path of choices, something we are hesitant to do for a long sequence, since doing so would involve imposing expectations about aggregate states on the agents in the data. This rather takes an agnostic approach to the expectations of the agents in the data: Whatever they believe about the future is captured by choice probabilities, and we are simply deriving a flexible function of that object.

D.6 Forming the Cell Sizes

First, we split the data into type cells. To maintain sufficient cell sizes, we use decade age grouping (20s, 30s, 40s, 50s). Education is split into the college educated and the non-college educated. The other dimension of type is birthplace. Thus, we have $4 \times 2 \times J$ types. Interacting these with origin (the state variable), we have $4 \times 2 \times J^2$ cells.

We first calculate the population in each cell so that we can weight appropriately to calculate aggregate statistics in the estimation sample as well as simulations and previous years of data. Note that one reason migration can change over time in the model is by shifting the weight assigned to each type. This is more obvious in some dimensions–for instance, in thinking about the aging of the population–but in our model with heterogeneous locations and preferences for home, the changing composition by origin, cohort, and birthplace will also matter for aggregation.

One complication with assigning weights by birthplace group is that the model is designed around LLMs, but birthplace in the census is reported by state. Some states contain multiple LLMs and some LLMs straddle state political boundaries. Note the entire US is partitioned by our geographic areas, so that by definition every state has at least one LLM (the outside option, location z) and most have two or more. For example, Atlanta is entirely in Georgia, but some Georgia-born residents came from rural areas and smaller unspecified LLMs. Hence, we need to map between state of birth and LLM of birth when assigning weight to an observation.

It was simplest in practice to assign someone living in an LLM within her birth state to be fully at home. An alternative would have been to assign weights based on lifetime migration probabilities, but this required a lot of assumptions about lifetime mobility that we were not actually modeling, including the propensity of repeat migration. At the other extreme, we could split the observation by population in his year of birth, and assign weight by the population shares. For example, say we observe someone living in Houston whose state of birth is Texas. We could also use population (or cohort population) in his year of birth to assign him as (for example) a one-quarter Dallas native, one-quarter Houston native, one-eighth Austin and San Antonio native, and one-quarter other. But we found this drastically understates the at-home share because it effectively assumes full mobility within state. Less than half of the respondents are in the city of their birth, despite a large majority being in their state of birth.

This issue is potentially more serious on the destination side. If we see a Texas-born, out-of-Texas resident move to Houston, we again do not know if she was born in Houston or elsewhere in the state. How we characterize the move has implications for measuring home preference, as a comparison of columns 5 and 6 of Table 1 suggests. Because these gaps were so large, and home preferences is a major piece of our analysis, we opted for the more conservative route and down-weighted the probability that such a move was a move home by the proportion of population in the person's year of birth (i.e., used column 6 instead of 5 from Table 1).

D.7 Forming the Moment Conditions

We then proceed to calculate choice probabilities. Even with a relatively large dataset and large LLMs, migration is infrequent enough that a fully interacted cell definition resulted in many empty cells. Once the data are cut to, for example, 40-something college-educated workers living in Houston but born in Cleveland, there are few individuals populating the cell. We may fail to observe any of this type moving to, say, Kansas City, but do not believe that the probability of that event is literally zero. Our smoothing procedure is designed to make aggregated cells that preserve the kinds of detail in the stylized facts presented above.

We create three tables of move probabilities from the ACS. These are:

1. the probability of migrating (to anywhere), by age, education, origin, and whether home is the origin, the residual, elsewhere in the U.S., or abroad. This will capture the main differences in moving costs by type, accounting for different incentives imposed by one's current labor market, combining all other birthplaces into one "away" category for precision. Call this $p_1(age, edu, birthplace)$.

2. the probability of returning home (i.e., moving from an away-location back to one's birthplace) by age, education and birthplace. This will capture differences in preference for home by location and cohort of birth, and in addition to stay rates for natives and nonnatives in 1, is an important moment for identifying the preference for home as a function of rootedness. Origins are combined for precision. Call this $p_2(age, edu, birthplace)$.

3. the probability of choosing a location as a destination that is *not* one's birthplace, by age and education and origin, combining all other birthplaces for precision, except that U.S. and foreign born are separate categories. This captures the geographic network of migration as well as differences in preferences for destinations for workers with different skills, and helps to identify the income component of utility. Call this $p_3(age, edu, origin, foreign)$.[74]

[74] We experimented with many versions of this estimator, conditioning on different aggregations of origin, and deriving from ACS microdata, aggregate data, and even the IRS data. We elected not to use the IRS data because we could not separate by moves to or not to home. The ACS aggregate data provided the fewest empty cells, since it was not subject to censorship requirements of the public use microdata.

Note that each p_n is computed by age and education, although we drop subscripts for exposition. The full matrix of cell-specific choice probabilities is then formed from the product of these for the corresponding cases, which are as follows.

1. For the reflexive entry (i.e., "stayers," the diagonal in the matrix), the entry is simply the probability p_1.
2. For people living in their birthplace but moving away, the probability is $(1 - p_1)p_3$.
3. For people living away from their birthplace, moves home are a conditional probability, $(1 - p_1)p_2$.
4. For people living away from their birthplace, moves *not* to home are a conditional probability, $(1 - p_1)(1 - p_2)p_3$.

Altogether, this forms a full $J \times J$ matrix (origin to destination) of flow probabilities for each type of worker, which is placed in the left-hand side of (D10) above. The interesting variation comes from the conditioning of the cell probability estimates, and the smoothing results from removing one dimension of conditioning, which is reintroduced when making the moments from the products of the conditional probabilities.

D.8 Income Distributions and Dynamics

To simulate the model, we need measures of income offer distributions and income dynamics. These determine the value of of utility from income as represented in (9) via (7) and (8). There are three sets of parameters to calibrate.

First is the available income distribution of each location. To focus on spatial differences in income opportunities, we construct a measure of the local income distribution having adjusted for differences in the local labor force composition. Specifically, after limiting the data to regularly employed workers, we run a regression of log earnings on controls for sex, race, English proficiency, and household composition in order to strip out compositional differences at national average labor prices.[75] We do this separately for non-college and college-educated workers because each face different labor market opportunities. The resultant income distributions from the ACS (and decennial census for prior decades) form the distribution of income opportunities for each local labor market in our sample.[76] The residualized income for each location has mean $\mu_{j,\tau}$ and variance $\sigma^2_{j,\tau}$. We use an N-pointed discretized distribution where the

[75]The results are largely similar when we also control for industry and occupational categories. Our preferred specification is to leave these out of the regression (so their variance contribution remains in the residual), since these can differ materially across labor markets–the kind of spatial variation we want to retain as a city characteristic.

[76]Note that this measures the *observed* income distribution, though we feed it into the model as if it were the *primitive* distribution. In principle, one might be able to estimate the primitive distribution by choosing

steps between points are one-half standard deviations from the mean, $w_n = \mu^{\omega}_{j,\tau} + 0.5n\sigma^{\omega}_{j,\tau}$, with integer $n \in \{-5, \cdots, 5\}$. Notice that the step in the income distribution, n, is a state variable in the model, and search occurs relative to that point, not a particular dollar value. For instance, a mean income worker in city A will search around the mean in her location and others, even if the nominal income of the mean in city B is, say, higher than the mean in city A. This accounts for average productivity differences between cities that shift the income distribution.

The second set of parameters is the probability of transition between these points, the $\pi_{n'|n}$ parameters from (7) and (8). We assume these follow a normal distribution and follow Tauchen (1986) to discretize it. For example, for someone at the bottom of the distribution, 2.5 standard deviations below, to move to the top, 2.5 deviations above, would require a shock drawn with probability of five standard deviations above the mean of a normal. This introduces persistence to the income process, and indirectly accounts for unobserved types of workers in that, for example, the high productivity workers in one city are more likely drawing from the higher side of the distribution in other cities as well, a necessary simplification given that our migration data allows us to observe only one income draw, not one in each location.

The distinction between local and nonlocal search, equations (7) and (8), is to allow the possibility that workers may face a different distribution of offers from their current location than distant locations. In particular, it may be "easier" in some sense to search locally. We could approach this two ways: through the probability of getting a contact, γ, or by the transition distributions, $\pi_{n'|n}$. The latter is more easily disciplined by data since we never actually observe the "successful" and "unsuccessful" searches. But with data on the joint dynamics of location and income, we can measure whether movers and non movers experience significantly different income dynamics. For this we turn to the Panel Study of Income Dynamics (PSID, PSID (2014)).[77] Using the post-1997 PSID for workers employed for two consecutive surveys, we measure the change in their income in standard deviations, which becomes the input data to a maximum likelihood estimation. If the step size of discretization is s, the probability of the change to income is

$$Pr(\Delta y) = \pi^{stay}_{n'|n} = \Phi(n + 0.5s) - \Phi(n - 0.5s),$$

a statistical distribution (e.g., normal), guessing parameters, simulating the model, and matching the observed distribution by some metric. We abstract from this because: (i) it would greatly complicate our estimation routine, which is otherwise computationally easy, (ii) with available data, we observe income only in the current location and not in past locations, but the model would try to predict the income distribution based on changes over migration events, and (iii) conceptually, we believe that, in a partial equilibrium model such as ours, it is reasonable to assume that individuals use observed income distributions to make migration decisions–otherwise they would have to have some superior knowledge of the true primitive distribution from which incomes are drawn.

[77] The geography in the PSID is state and we need local income distributions at annual frequencies, so we first measure the standard deviation of incomes by education and state using the March CPS (Flood et al. (2015)) and then merge this to the PSID by survey year.

where Φ is the standard normal distribution. For movers, we relax the symmetry assumption, shifting the step by some value ω,

$$Pr(\Delta y) = \pi^{move}_{n'|n} = \Phi(n + 0.5s + \omega) - \Phi(n - 0.5s - \omega),$$

so that income changes can be better or worse on average for movers. The estimation step is to place each worker at the closest point in the discretized distribution of income and then calculate the probability of observing the change in income for a given guess of ω. The value of ω that maximizes the likelihood of the data is used to calculate the income dynamics probabilities π^{mover} vis-a-vis the baseline $\pi^{nonmover}$. Using our discretization, we obtain the estimate of $\omega = -.13$. We did not constrain it to be negative, but because income changes for movers are on average worse than for stayers, which is consistent with our conjecture.

The final parameter is the probability of successful search, λ, which gives the worker the preferable option value search (8), instead of the vulnerable random search (7). As noted, we have little sense how to discipline this with the data, so we will treat it parametrically as a proof of concept exercise. This parameter will change over time to reflect increasing availability of information in the labor market.

E Details on Data Construction

E.1 Sources

Our migration data come from two sources, the American Community Survey (ACS) and the migration flows tables from the U.S. Treasury's Internal Revenue Service (IRS). The ACS reports the respondent's current and one year ago Public Use Microdata Area (PUMA) of residence, from which we can elicit migration probability (move or not) and direction (origin-destination pairs). We use the ACS from 2005 to 2017. Migration is elicited using the *puma* and *migpuma* variables.

The IRS infers migration events from changes in the address on individual tax returns in two successive years, publishing the total county-to-county flows in each year, as well as the total stayers in, inflows to, and outflows from individual counties. One limitation is that the data are censored at flows less than 10 households, meaning many origin-destination pairs are unobserved. On average, about 70 percent of flows are on observed origin-destination routes, and the rest are censored. We measure migration using the internal subtotals of total domestic inflows and outflows, subtracting flows between counties within the same local labor market (which are almost never censored).

The IRS data underwent a change in method in the 2011-2012 tax year that resulted in noticeable differences in the sample represented. Our understanding is that the data were computed and published by the Census Bureau from 1990 to 2011, and the IRS took charge in 2012 and following. The IRS had different methods for tracking addresses across multiple returns (in cases of, for instance, household formation and dissolution), and late filers, which tended to be households with complicated returns. Thus, the set of individuals represented changed, and because of the recursive nature of the data, this introduced year-over-year fluctuations that may take several more years before they can be safely compared across time. We present the data for the period 2012-2016, but only rely on the consistent sample of 1990-2011.

We also leverage aggregated population data at the county level, which we will use to show population growth trends. We obtained the county population estimates from Manson et al. (2018) and relied heavily on that project's harmonization of geographies across census years. Census microdata samples, 1880-2000, were obtained from Ruggles et al. (2019).

E.2 Geography: The Local Labor Market

In this paper, we will work with a unit of analysis we term a local labor market (LLM), which fully partitions the geography of the continental US. The LLM is derived from a Commuting Zone

(CZ) but modified to meet some specific objectives.[78] One objective is geographic consistency over time and across datasets. We were able to define constant boundary LLMs for both counties and PUMAs for use in, respectively, Census aggregate population and microdata, dating from 1880 to current releases. A second objective is to fit more intuitive notions of an integrated labor market area, more like a core-based statistical area (CBSA) or metropolitan statistical area (MSA). In many cases, these line up well with the commuting zone, but in some, the LLM covers a large and heterogeneous area. For example, most of southern California is in one CZ, despite substantial heterogeneity in populations and labor market opportunities between, say, the inland counties, which we split into a Riverside/San Bernardino LLM and the coastal counties, which we further split into Los Angeles and Ventura LLMs, making three local labor market units instead of one. A full list of LLMs is presented in Table F3, and a dataset of the mapping of counties and PUMAs over time is available on our webpage.

Table F3 includes definitions as well as indicators for when the LLM is included in the analysis. In our descriptive analyses in section 2, we report measures of migration flow and population growth for the 183 LLMs that are characterized as urban areas. We aggregate the remainder of the continental 48 states into an omitted category,[79] comprising rural areas and some unusual LLMs–smaller cities dominated by universities ("college towns" such as Athens, Georgia or Bloomington, Indiana) or military bases (such as Jacksonville, North Carolina), which have nonstandard migration behavior. Our empirical model focused on the 70 largest LLMs. Throughout, we define a migration event as an exit from a LLM for a different LLM, so that a move within a county or PUMA, or across counties or PUMAs within the same LLM, is considered staying in place.

E.3 Discussion of the Calculation of Roots

The decennial census data contain detailed geographic information on current residence and birth state for individuals in a household. We use household structure variables and cohort matching to estimate rootedness of a particular cohort for each home LLM. We identify a birth cohort by looking at all individuals who are less than 10 years old in a particular census wave. For example, a twenty-something in 2010 was aged less than 10 in 1990. For the cohort living in each LLM, we calculate the percentage of their parents who were born in a state in which the LLM has a county. For example, children in Dallas are rooted if their parents were born in Texas, and children in Kansas City are considered rooted if their parents were born in Kansas

[78] Our starting definition of LLM derived from CZs as in Dorn (2009), Autor and Dorn (2013), and Autor et al. (2019). The geographic mapping over the census years relied heavily on the documentation provided by the Ruggles et al. (2019) and Manson et al. (2018) projects. We are additionally grateful to Dave Van Riper and Jeff Bloom for assistance.

[79] Our categorization defines "residual states" as LLMs; for example, the population of Oklahoma not in Oklahoma City or Tulsa is in "residual Oklahoma." We combine these here for expositional convenience.

or Missouri. The child must be living in his/her home LLM to be counted in this sample. We ascribe the cohort-LLM combination to have the rootedness measured by this fraction.

There are a few possible concerns given that we use metro area (LLM) for location but state for place of birth. For example, we do not actually know in which *city* a child's parents were born. It is possible that a child was born in Dallas but the parents were born in Houston or Austin, and certainly the percentage of Dallas children's parents born in Dallas is smaller than the percentage born in all of Texas. When comparing across cohorts, if the measurement error is similar, the *change* in rootedness is still accurate. But if we are comparing a LLM in a state with several large cities, such as Texas, to a LLM in a state with only one major city, such as Minnesota, we will likely measure the rootedness of Dallas as too high compared to the Twin Cities. This may be grounds for some within-state migration adjustment. In practice we found that adjustments made trivial impact on our rootedness measures, because out of state birthplaces drove the first order differences between cities. Unless within-state migration is strongly negatively correlated with between-state (which other datasets indicate is not the case), our measure of rootedness will if anything shrink the dispersion of rootedness across LLMs.

We use the location of residence for children under 10 as "home" for the purpose of cohort matching. There is some mobility of young children that introduces uncertainty into our estimates. One possible adjustment is to probabilistically assign children to potential birth cities, but in practice the change to measurements is small.

Note that the rootedness proportions only include the U.S. born (the dotted line in Figure 4, not the shaded areas). The foreign born are treated as a separate birthplace group–not at home, and not from other U.S.–since, among other concerns, they have no domestic "home" location to prefer above others, and they cannot be rooted. However, another concern is first-generation immigrants. In such cases, the child was born in the United States but one or both parents were born outside the US, leaving ambiguity in defining the child's roots. By our strict definition, of course we can say with certainty this child was not born in the same commuting zone as his parents. However, many immigrants move to cities that have an established population of immigrants from their native country already. (Perhaps a Cuban immigrant in Miami, for instance, should be considered "rooted" in a sense.) For our purposes, it was simplest to calculate rootedness only for children of native-born parents, since only they can be "at home."

A final note on roots is that our current method calculates roots for a cohort of children, and later we will match that average rootedness measured at the city level to people born in that city. However, at that point we will have divided our adult sample into college graduates and non-college graduates. Since college graduates are more mobile on average, and there is positive intergenerational transmission of education, we expect that the college graduate subset of any cohort will be less rooted than the cohort as a whole. Our methodology implicitly assigns the relationship between rootedness and education to be the same in every city.

F Data Definitions

Figure F1: Regional Definitions

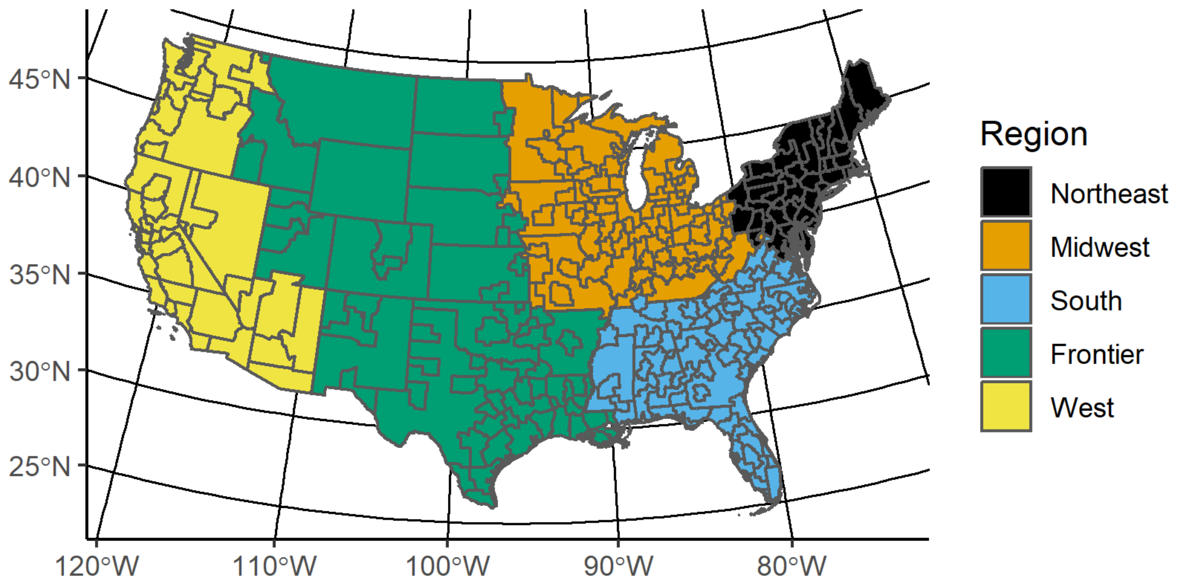


Table F1: Local Labor Markets (LLM) Definitions

			Included In:	
LLM	States (# Counties)	Region	Model	Descriptives
Abilene	Texas (6)	Frontier	No	No
Akron	Ohio (4)	Midwest	Yes	Yes
Albany	Georgia (6)	South	No	Yes
Albany	New York (9)	Northeast	Yes	Yes
Albuquerque	New Mexico (5)	Frontier	No	Yes
Alexandria	Louisiana (3)	Frontier	No	No
Allentown	Pennsylvania (3)	Northeast	No	Yes
Amarillo	Texas (5)	Frontier	No	Yes
Aniston	Alabama (1)	South	No	Yes

Table F1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Descriptives
Ann Arbor	Michigan (1)	Midwest	No	No
Appleton	Wisconsin (7)	Midwest	No	Yes
Asheville	North Carolina (5)	South	No	Yes
Athens	Georgia (5)	South	No	No
Atlanta	Georgia (28)	South	Yes	Yes
Atlantic City-Vineland	New Jersey (3)	Northeast	No	Yes
Auburn	Alabama (3)	South	No	No
Augusta	South Carolina (2), Georgia (7)	South	No	No
Austin	Texas (8)	Frontier	Yes	Yes
Bakersfield	California (1)	West	No	Yes
Baltimore	Maryland (6)	Northeast	Yes	Yes
Baton Rouge	Louisiana (10)	Frontier	No	Yes
Beaumont	Texas (7)	Frontier	No	Yes
Bellingham	Washington (1)	West	No	No
Biloxi	Mississippi (7)	South	No	Yes
Binghamton	Pennsylvania (1), New York (2)	Northeast	No	Yes
Birmingham	Alabama (7)	South	Yes	Yes
Bloomington	Illinois (3)	Midwest	No	No
Bloomington	Indiana (7)	Midwest	No	No
Boise City	Idaho (6)	Frontier	No	Yes
Boston	Massachusetts (6)	Northeast	Yes	Yes
Bradenton	Florida (4)	South	No	Yes
Buffalo	New York (2)	Northeast	Yes	Yes
Burlington	Vermont (6)	Northeast	No	No
Cedar Rapids	Iowa (4)	Midwest	No	Yes
Champaign-Urbana	Illinois (7)	Midwest	No	No
Charleston	South Carolina (4)	South	No	Yes
Charleston	West Virginia (6)	Midwest	No	Yes
Charlotte	South Carolina (3), North Carolina (7)	South	Yes	Yes
Charlottesville	Virginia (9)	South	No	No
Chattanooga	Tennessee (4), Georgia (3)	South	No	Yes
Chicago	Indiana (4), Wisconsin (1), Illinois (8)	Midwest	Yes	Yes
Chico	California (5)	West	No	Yes
Cincinnati	Indiana (3), Ohio (6), Kentucky (6)	Midwest	Yes	Yes
Clarksville	Tennessee (3), Kentucky (3)	South	No	No
Cleveland	Ohio (5)	Midwest	Yes	Yes
College Station	Texas (4)	Frontier	No	No
Colorado Springs	Colorado (4)	Frontier	No	No
Columbia	Missouri (7)	Midwest	No	No
Columbia	South Carolina (7)	South	No	Yes
Columbus	Alabama (1), Georgia (5)	South	No	No
Columbus	Ohio (9)	Midwest	Yes	Yes
Corpus Christi	Texas (9)	Frontier	No	Yes
Cumberland	West Virginia (3), Maryland (2)	Northeast	No	Yes
Dallas-Fort Worth	Texas (16)	Frontier	Yes	Yes
Davenport	Illinois (4), Iowa (1)	Midwest	No	Yes
Dayton	Ohio (9)	Midwest	Yes	Yes
Daytona Beach	Florida (3)	South	No	Yes
Denver	Colorado (14)	Frontier	Yes	Yes
Des Moines	Iowa (8)	Midwest	No	Yes
Detroit	Michigan (6)	Midwest	Yes	Yes
Dothan	Alabama (5)	South	No	No

Table F1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Descriptives
Dover	Delaware (2), Maryland (3)	Northeast	No	Yes
Dubuque	Iowa (4), Illinois (1)	Midwest	No	Yes
Duluth	Minnesota (3), Wisconsin (1)	Midwest	No	Yes
Eau Claire	Wisconsin (7), Minnesota (1)	Midwest	No	Yes
El Paso	New Mexico (1), Texas (2)	Frontier	Yes	Yes
Erie	New York (1), Pennsylvania (6)	Northeast	No	Yes
Evansville	Illinois (1), Kentucky (2), Indiana (8)	Midwest	No	Yes
Fargo	Minnesota (3), North Dakota (2)	Frontier	No	Yes
Fayetteville	Missouri (1), Arkansas (3), Oklahoma (1)	Frontier	No	Yes
Fayetteville	North Carolina (7)	South	No	No
Flagstaff	Arizona (2), Utah (1)	West	No	Yes
Flint	Michigan (1)	Midwest	No	Yes
Florence	Tennessee (2), Alabama (3)	South	No	Yes
Fort Myers	Florida (2)	South	Yes	Yes
Fort Smith	Arkansas (3), Oklahoma (5)	Frontier	No	Yes
Fort Wayne	Indiana (8)	Midwest	No	Yes
Fresno	California (4)	West	Yes	Yes
Gainesville	Florida (6)	South	No	No
Goldsboro	North Carolina (3)	South	No	No
Grand Rapids	Michigan (8)	Midwest	Yes	Yes
Green Bay	Wisconsin (4)	Midwest	No	Yes
Greensboro	North Carolina (11), Virginia (2)	South	Yes	Yes
Greenville	North Carolina (3)	South	No	No
Greenville	South Carolina (11), North Carolina (1)	South	Yes	Yes
Hagerstown	West Virginia (3), Maryland (1), Pennsylvania (2)	Northeast	No	Yes
Harrisburg	Pennsylvania (7)	Northeast	Yes	Yes
Hattiesburg	Mississippi (5)	South	No	No
Hickory	North Carolina (5)	South	No	Yes
Houma	Louisiana (4)	Frontier	No	Yes
Houston	Texas (13)	Frontier	Yes	Yes
Huntington	Ohio (1), Kentucky (5), West Virginia (2)	Midwest	No	Yes
Huntsville	Alabama (4), Tennessee (1)	South	No	Yes
Indianapolis	Indiana (10)	Midwest	Yes	Yes
Iowa City	Iowa (5)	Midwest	No	No
Jackson	Michigan (3)	Midwest	No	Yes
Jackson	Mississippi (7)	South	No	Yes
Jacksonville	Florida (5), Georgia (2)	South	Yes	Yes
Jacksonville	North Carolina (1)	South	No	No
Johnson City	Virginia (4), Tennessee (6)	South	No	Yes
Johnstown	Pennsylvania (4)	Northeast	No	Yes
Joplin	Missouri (2), Oklahoma (1), Kansas (3)	Midwest	No	Yes
Kalamazoo	Michigan (4)	Midwest	No	Yes
Kansas City	Missouri (8), Kansas (6)	Midwest	Yes	Yes
Killeen	Texas (3)	Frontier	No	No
Knoxville	Tennessee (8)	South	No	Yes
LaCrosse	Minnesota (1), Wisconsin (4)	Midwest	No	Yes
Lafayette	Indiana (7), Illinois (1)	Midwest	No	No
Lafayette	Louisiana (7)	Frontier	No	Yes
Lake Charles	Louisiana (6)	Frontier	No	No
Lakeland	Florida (3)	South	No	Yes
Lancaster	Pennsylvania (4)	Northeast	Yes	Yes
Lansing	Michigan (3)	Midwest	No	No

Table F1: Local Labor Markets (LLM) Definitions

			Included In:	
LLM	States (# Counties)	Region	Model	Descriptives
Laredo	Texas (3)	Frontier	No	Yes
Las Vegas	Nevada (1)	West	Yes	Yes
Lawrence	Kansas (1)	Frontier	No	No
Lexington	Kentucky (13)	Midwest	No	No
Lincoln	Nebraska (5)	Frontier	No	No
Little Rock	Arkansas (7)	Frontier	No	Yes
Longview	Texas (6)	Frontier	No	Yes
Los Angeles	California (2)	West	Yes	Yes
Louisville	Indiana (5), Kentucky (7)	Midwest	Yes	Yes
Lubbock	Texas (6)	Frontier	No	No
Lynchburg	Virginia (5)	South	No	No
Macon	Georgia (10)	South	No	Yes
Madison	Wisconsin (6)	Midwest	No	No
Manchester	New Hampshire (4)	Northeast	Yes	Yes
Mansfield	Ohio (5)	Midwest	No	Yes
McAllen	Texas (4)	Frontier	Yes	Yes
Medford	Oregon (2)	West	No	Yes
Melbourne	Florida (2)	South	No	Yes
Memphis	Mississippi (4), Tennessee (3), Arkansas (1)	South	Yes	Yes
Miami	Florida (3)	South	Yes	Yes
Midland	Texas (6)	Frontier	No	Yes
Milwaukee	Wisconsin (7)	Midwest	Yes	Yes
Minneapolis	Minnesota (14), Wisconsin (2)	Midwest	Yes	Yes
Mobile	Alabama (5)	South	No	Yes
Modesto	California (4)	West	No	Yes
Monmouth	New Jersey (2)	Northeast	Yes	Yes
Monroe	Louisiana (6)	Frontier	No	Yes
Montgomery	Alabama (5)	South	No	Yes
Muncie	Indiana (6)	Midwest	No	Yes
Myrtle Beach	South Carolina (7)	South	No	Yes
Nashville	Tennessee (12)	South	Yes	Yes
New Orleans	Louisiana (9)	Frontier	Yes	Yes
New York	New Jersey (12), Connecticut (1), New York (10)	Northeast	Yes	Yes
Norfolk	Virginia (18), North Carolina (1)	South	Yes	Yes
Ocala	Florida (2)	South	No	Yes
Oklahoma City	Oklahoma (11)	Frontier	Yes	Yes
Olympia	Washington (1)	West	No	No
Omaha	Iowa (3), Nebraska (6)	Frontier	No	Yes
Orlando	Florida (5)	South	Yes	Yes
Owensboro	Kentucky (6)	Midwest	No	Yes
Panama City	Florida (3)	South	No	No
Parkersburg	Ohio (1), West Virginia (5)	Midwest	No	Yes
Pensacola	Florida (4)	South	No	No
Peoria	Illinois (8)	Midwest	No	Yes
Philadelphia	Pennsylvania (5), Delaware (1), New Jersey (4)	Northeast	Yes	Yes
Phoenix	Arizona (2)	West	Yes	Yes
Pittsburgh	Pennsylvania (9)	Northeast	Yes	Yes
Portland	Maine (9)	Northeast	No	Yes
Portland	Oregon (4), Washington (2)	West	Yes	Yes
Poughkeepsie	New York (4)	Northeast	No	Yes
Providence	Massachusetts (1), Rhode Island (5)	Northeast	Yes	Yes
Provo-Orem	Utah (3)	Frontier	No	No

Table F1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Included In: Descriptives
Raleigh-Durham	North Carolina (7)	South	Yes	Yes
Redding	California (2)	West	No	Yes
Reno	Nevada (5)	West	No	Yes
Residual Alabama	Alabama (30)	South	No	No
Residual Arizona	Arizona (6)	West	No	No
Residual Arkansas	Arkansas (59)	Frontier	No	No
Residual California	California (14)	West	No	No
Residual Colorado	Colorado (46)	Frontier	No	No
Residual Connecticut	Connecticut (7)	Northeast	Yes	Yes
Residual Florida	Florida (10)	South	No	No
Residual Georgia	Georgia (90)	South	No	No
Residual Idaho	Idaho (34)	Frontier	No	No
Residual Illinois	Illinois (55)	Midwest	No	No
Residual Indiana	Indiana (23)	Midwest	No	No
Residual Iowa	Iowa (65)	Midwest	No	No
Residual Kansas	Kansas (84)	Frontier	No	No
Residual Kentucky	Kentucky (78)	Midwest	No	No
Residual Louisiana	Louisiana (11)	Frontier	No	No
Residual Maine	Maine (7)	Northeast	No	No
Residual Maryland	Maryland (6)	Northeast	No	No
Residual Massachusetts	Massachusetts (4)	Northeast	No	No
Residual Michigan	Michigan (47)	Midwest	No	No
Residual Minnesota	Minnesota (60)	Midwest	No	No
Residual Mississippi	Mississippi (59)	South	No	No
Residual Missouri	Missouri (73)	Midwest	No	No
Residual Montana	Montana (57)	Frontier	No	No
Residual Nebraska	Nebraska (79)	Frontier	No	No
Residual Nevada	Nevada (11)	West	No	No
Residual New Hampshire	New Hampshire (6)	Northeast	No	No
Residual New Mexico	New Mexico (23)	Frontier	No	No
Residual New York	New York (18)	Northeast	No	No
Residual North Carolina	North Carolina (41)	South	No	No
Residual North Dakota	North Dakota (51)	Frontier	No	No
Residual Ohio	Ohio (37)	Midwest	No	No
Residual Oklahoma	Oklahoma (48)	Frontier	No	No
Residual Oregon	Oregon (22)	West	No	No
Residual Pennsylvania	Pennsylvania (9)	Northeast	No	No
Residual South Carolina	South Carolina (6)	South	No	No
Residual South Dakota	South Dakota (60)	Frontier	No	No
Residual Tennessee	Tennessee (56)	South	No	No
Residual Texas	Texas (127)	Frontier	No	No
Residual Utah	Utah (18)	Frontier	No	No
Residual Vermont	Vermont (8)	Northeast	No	No
Residual Virginia	Virginia (50)	South	No	No
Residual Washington	Washington (23)	West	No	No
Residual West Virginia	West Virginia (29)	Midwest	No	No
Residual Wisconsin	Wisconsin (31)	Midwest	No	No
Residual Wyoming	Wyoming (23)	Frontier	No	No
Richland	Washington (4), Oregon (2)	West	No	Yes
Richmond	Virginia (17)	South	Yes	Yes
Riverside-San Bernardino	California (2)	West	Yes	Yes
Roanoke	Virginia (11), West Virginia (1)	South	No	No

Table F1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Included In: Descriptives
Rochester	Minnesota (5)	Midwest	No	Yes
Rochester	New York (9)	Northeast	Yes	Yes
Rockford	Illinois (2)	Midwest	No	Yes
Rocky Mount	North Carolina (3)	South	No	Yes
Sacramento	California (4)	West	Yes	Yes
Saginaw	Michigan (7)	Midwest	No	Yes
Salem	Oregon (6)	West	Yes	Yes
Salinas	California (3)	West	No	Yes
Salt Lake City	Utah (7)	Frontier	Yes	Yes
San Antonio	Texas (10)	Frontier	Yes	Yes
San Diego	California (1)	West	Yes	Yes
San Francisco	California (7)	West	Yes	Yes
San Jose	California (1)	West	Yes	Yes
Santa Barbara	California (2)	West	No	Yes
Santa Fe	New Mexico (4)	Frontier	No	Yes
Santa Rosa	California (3)	West	No	Yes
Savannah	Georgia (3), South Carolina (2)	South	No	No
Scranton	Pennsylvania (7)	Northeast	No	Yes
Seattle	Washington (3)	West	Yes	Yes
Sheboygan	Wisconsin (2)	Midwest	No	Yes
Shreveport	Louisiana (8)	Frontier	No	Yes
Sioux City	Iowa (3), South Dakota (1), Nebraska (3)	Midwest	No	Yes
Sioux Falls	South Dakota (5)	Frontier	No	Yes
South Bend	Michigan (2), Indiana (4)	Midwest	No	Yes
Spokane	Washington (3), Idaho (4)	West	No	Yes
Springfield	Illinois (4)	Midwest	No	Yes
Springfield	Massachusetts (3)	Northeast	No	No
Springfield	Missouri (9)	Midwest	No	Yes
St. Joseph	Kansas (1), Missouri (6)	Midwest	No	Yes
St. Louis	Missouri (9), Illinois (8)	Midwest	Yes	Yes
St. Lucie	Florida (2)	South	No	Yes
State College	Pennsylvania (4)	Northeast	No	No
Steubenville	Ohio (1), West Virginia (2)	Midwest	No	Yes
Stockton	California (1)	West	No	Yes
Sumter	South Carolina (4)	South	No	No
Syracuse	New York (7)	Northeast	Yes	Yes
Tallahassee	Florida (9)	South	No	No
Tampa	Florida (4)	South	Yes	Yes
Terre Haute	Indiana (7)	Midwest	No	Yes
Texarkana	Texas (6), Arkansas (2)	Frontier	No	Yes
Toledo	Ohio (5), Michigan (1)	Midwest	No	Yes
Topeka	Kansas (5)	Frontier	No	Yes
Tucson	Arizona (3)	West	Yes	Yes
Tulsa	Oklahoma (11)	Frontier	Yes	Yes
Tuscaloosa	Alabama (3)	South	No	No
Tyler	Texas (6)	Frontier	No	Yes
Ventura	California (1)	West	No	Yes
Waco	Texas (4)	Frontier	No	No
Washington	Maryland (6), District Of Columbia (1), Virginia (17)	Northeast	Yes	Yes
Waterloo	Iowa (6)	Midwest	No	Yes
Wheeling	Ohio (2), West Virginia (4)	Midwest	No	Yes
Wichita	Kansas (5)	Frontier	No	Yes

Table F1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Descriptives
Wichita Falls	Texas (3)	Frontier	No	No
Williamsport	Pennsylvania (4)	Northeast	No	Yes
Wilmington	North Carolina (5)	South	No	Yes
Yakima	Washington (2)	West	No	Yes
Youngstown	Ohio (3), Pennsylvania (2)	Midwest	No	Yes
Yuma	California (1), Arizona (2)	West	No	Yes

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Abilene (TX)	0.66	0.82	196,843	172,513	134,144	137,116	17,352
Akron (OH)	0.78	0.73	1,107,622	1,064,749	776,219	515,494	151,735
Albany (GA)	0.71	0.84	180,806	167,405	112,442	108,317	54,668
Albany (NY)	0.75	0.86	1,111,956	981,287	770,338	631,217	535,495
Albuquerque (NM)	0.46	0.62	914,290	523,105	188,604	58,247	28,877
Alexandria (LA)	0.75	0.85	195,995	193,378	142,942	109,147	46,498
Allentown (PA)	0.57	0.86	712,481	551,052	441,008	364,172	168,204
Amarillo (TX)	0.59	0.70	265,123	199,141	100,138	27,657	387
Aniston (AL)	0.68	0.85	118,572	119,761	79,539	47,822	19,591
Ann Arbor (MI)	0.58	0.67	344,791	264,748	134,606	49,520	41,848
Appleton (WI)	0.78	0.89	590,250	460,060	323,219	258,144	183,072
Asheville (NC)	0.47	0.80	457,948	306,253	228,671	135,278	60,611
Athens (GA)	0.57	0.86	261,908	134,955	78,870	89,456	48,776
Atlanta (GA)	0.36	0.72	5,189,409	2,330,869	1,104,602	759,095	392,270
Atlantic City-Vineland (NJ)	0.53	0.65	528,712	409,251	258,127	164,722	66,156
Auburn (AL)	0.52	0.85	172,613	113,708	91,688	81,715	73,699
Augusta (GA/SC)	0.61	0.81	582,723	410,163	254,243	225,879	148,926
Austin (TX)	0.46	0.80	1,768,155	623,416	294,154	224,460	109,889
Bakersfield (CA)	0.54	0.56	839,631	403,089	228,309	54,843	5,601
Baltimore (MD)	0.55	0.68	2,662,691	2,174,023	1,457,181	931,413	519,349
Baton Rouge (LA)	0.77	0.83	923,581	672,081	352,539	210,563	130,198
Beaumont (TX)	0.68	0.72	506,079	460,162	313,552	159,446	31,449
Bellingham (WA)	0.45	0.61	201,140	106,701	66,733	50,600	3,137
Biloxi (MS)	0.51	0.72	481,300	368,852	172,497	100,433	25,925
Binghamton (NY/PA)	0.80	0.90	295,081	301,336	246,834	172,585	122,510
Birmingham (AL)	0.69	0.87	1,128,047	930,281	753,630	482,579	100,098
Bloomington (IL)	0.72	0.80	225,083	178,638	131,280	128,429	115,560
Bloomington (IN)	0.71	0.82	300,670	245,028	178,394	157,972	122,705
Boise City (ID)	0.37	0.48	643,599	301,600	147,746	80,175	7,921
Boston (MA)	0.54	0.82	4,932,588	4,309,184	3,611,745	2,927,214	1,347,714
Bradenton (FL)	0.22	0.36	897,121	428,192	77,059	27,492	2,080
Buffalo (NY)	0.82	0.84	1,135,509	1,242,826	1,089,230	753,393	274,057
Burlington (VT)	0.50	0.72	334,199	259,455	172,605	154,255	145,827
Cedar Rapids (IA)	0.71	0.80	274,295	229,254	162,166	135,291	102,398
Champaign-Urbana (IL)	0.71	0.78	399,848	389,856	289,135	218,358	162,032
Charleston (SC)	0.46	0.74	703,499	462,238	245,950	180,364	139,186
Charleston (WV)	0.78	0.87	341,027	385,661	408,641	224,370	64,647
Charlotte (NC/SC)	0.47	0.86	2,066,843	1,086,694	694,290	414,074	196,764
Charlottesville (VA)	0.50	0.78	330,316	194,059	123,881	119,790	120,754

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Chattanooga (TN/GA)	0.65	0.78	541,846	440,327	307,726	180,706	67,393
Chicago (IL/IN/WI)	0.59	0.74	9,355,945	7,978,308	5,720,703	3,679,252	864,564
Chico (CA)	0.62	0.60	436,433	279,971	142,688	71,663	48,282
Cincinnati (OH/KY/IN)	0.77	0.86	2,150,213	1,770,391	1,256,813	890,308	658,336
Clarksville (TN/KY)	0.45	0.78	294,835	187,014	123,611	118,926	107,631
Cleveland (OH)	0.75	0.70	2,077,240	2,173,734	1,680,736	1,103,877	284,499
College Station (TX)	0.61	0.87	255,264	134,134	86,433	89,864	63,805
Colorado Springs (CO)	0.23	0.39	696,692	347,662	97,216	70,778	20,764
Columbia (MO)	0.65	0.81	340,194	238,024	157,818	137,755	130,801
Columbia (SC)	0.58	0.81	835,910	572,662	339,798	281,888	158,272
Columbus (GA/AL)	0.64	0.83	314,980	282,425	210,548	130,010	97,353
Columbus (OH)	0.66	0.73	1,892,010	1,314,441	783,609	526,113	307,388
Corpus Christi (TX)	0.71	0.88	527,888	441,121	291,130	67,886	18,269
Cumberland (MD/WV)	0.69	0.84	159,358	154,520	151,936	128,059	71,153
Dallas-Fort Worth (TX)	0.43	0.72	6,512,481	3,103,335	1,318,069	771,883	255,930
Davenport (IA/IL)	0.76	0.80	385,684	410,633	306,843	239,904	146,874
Dayton (OH)	0.73	0.71	1,168,172	1,151,295	785,793	517,156	331,255
Daytona Beach (FL)	0.31	0.42	664,653	320,224	101,211	40,384	10,269
Denver (CO)	0.32	0.44	3,418,663	1,935,528	755,253	445,732	84,405
Des Moines (IA)	0.66	0.79	693,163	500,160	380,482	300,148	161,748
Detroit (MI)	0.73	0.71	4,296,250	4,353,413	3,170,315	1,407,111	338,194
Dothan (AL)	0.61	0.83	245,838	200,541	142,643	140,977	43,899
Dover (DE/MD)	0.55	0.72	536,112	310,840	182,740	149,840	128,115
Dubuque (IA/IL)	0.82	0.87	166,554	172,404	149,403	140,679	137,094
Duluth (MN/WI)	0.79	0.86	290,637	309,629	285,142	283,804	6,495
Eau Claire (WI/MN)	0.79	0.89	326,790	276,277	226,192	198,224	126,892
El Paso (TX/NM)	0.48	0.76	1,013,356	578,967	238,823	119,387	7,152
Erie (PA/NY)	0.84	0.90	648,739	675,901	584,839	472,406	325,001
Evansville (IN/KY/IL)	0.82	0.91	465,647	424,363	365,729	296,786	226,154
Fargo (ND/MN)	0.75	0.89	238,526	174,614	132,581	109,211	23,363
Fayetteville (AR/OK/MO)	0.45	0.74	504,691	228,845	128,667	115,197	63,443
Fayetteville (NC)	0.53	0.79	699,392	519,881	315,853	178,622	88,762
Flagstaff (AZ/UT)	0.35	0.59	352,579	147,177	51,200	36,052	8,098
Flint (MI)	0.83	0.73	425,790	450,449	270,963	125,668	39,220
Florence (AL/TN)	0.73	0.89	237,731	211,471	162,127	130,034	68,027
Fort Myers (FL)	0.17	0.35	940,274	291,237	29,892	7,981	641
Fort Smith (AR/OK)	0.59	0.80	361,460	274,570	216,069	251,094	43,474
Fort Wayne (IN)	0.68	0.74	585,429	492,705	342,174	264,924	181,488
Fresno (CA)	0.54	0.62	1,676,476	897,213	509,511	222,044	20,759
Gainesville (FL)	0.49	0.62	365,553	214,925	95,453	57,830	29,795
Goldsboro (NC)	0.60	0.89	244,559	187,693	155,121	109,865	66,618
Grand Rapids (MI)	0.76	0.83	1,352,296	997,113	637,924	427,110	264,187
Green Bay (WI)	0.72	0.88	334,026	248,795	162,788	124,157	69,466
Greensboro (NC/VA)	0.62	0.87	1,719,480	1,226,539	812,974	476,264	243,766
Greenville (NC)	0.64	0.89	249,005	166,082	127,766	91,336	47,175
Greenville (SC/NC)	0.62	0.87	1,392,816	976,115	686,861	506,266	259,816
Hagerstown (MD/WV/PA)	0.69	0.89	487,101	327,345	221,019	180,226	136,727
Harrisburg (PA)	0.67	0.87	1,157,172	893,927	622,891	459,576	307,747
Hattiesburg (MS)	0.71	0.85	174,897	129,476	98,924	70,718	16,760
Hickory (NC)	0.59	0.88	524,934	352,995	221,521	127,288	69,076
Houma (LA)	0.86	0.92	286,249	263,213	138,663	105,984	73,971

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Houston (TX)	0.43	0.72	6,058,500	3,249,059	1,168,450	426,029	154,293
Huntington (WV/KY/OH)	0.86	0.93	346,607	362,367	313,034	221,469	116,419
Huntsville (AL/TN)	0.59	0.83	604,783	389,855	214,345	172,898	124,005
Indianapolis (IN)	0.63	0.70	1,735,670	1,198,556	753,130	526,350	286,805
Iowa City (IA)	0.63	0.81	225,217	172,984	125,472	105,664	101,051
Jackson (MI)	0.78	0.77	306,828	283,514	204,470	148,467	123,097
Jackson (MS)	0.75	0.89	591,126	455,328	300,508	186,848	146,127
Jacksonville (FL/GA)	0.45	0.66	1,408,280	758,255	374,617	159,862	42,734
Jacksonville (NC)	0.32	0.55	177,772	112,784	42,047	14,703	9,829
Johnson City (TN/VA)	0.68	0.86	609,299	521,426	396,211	243,667	147,377
Johnstown (PA)	0.85	0.93	398,272	447,911	471,643	446,562	167,590
Joplin (MO/KS/OK)	0.67	0.82	283,276	236,572	224,092	260,542	109,313
Kalamazoo (MI)	0.74	0.73	521,908	466,530	312,887	196,241	128,918
Kansas City (MO/KS)	0.61	0.74	2,048,694	1,515,021	1,021,717	759,148	365,266
Killeen (TX)	0.41	0.70	405,300	226,661	100,037	75,813	35,973
Knoxville (TN)	0.57	0.79	814,914	605,022	450,718	254,088	115,735
LaCrosse (WI/MN)	0.76	0.88	234,775	191,193	160,236	135,495	103,829
Lafayette (IN/IL)	0.76	0.83	367,029	330,285	263,171	231,106	175,999
Lafayette (LA)	0.84	0.93	584,118	476,339	318,239	216,170	91,306
Lake Charles (LA)	0.73	0.85	344,953	313,284	177,752	115,400	20,060
Lakeland (FL)	0.38	0.50	728,612	388,557	147,706	53,252	4,463
Lancaster (PA)	0.73	0.91	1,212,744	944,067	772,717	655,557	430,494
Lansing (MI)	0.75	0.78	464,036	419,750	244,159	134,041	93,001
Laredo (TX)	0.52	0.92	269,622	111,054	65,935	33,995	11,852
Las Vegas (NV)	0.10	0.16	1,951,269	463,087	48,289	4,859	1,286
Lawrence (KS)	0.51	0.67	110,826	67,640	34,086	23,998	21,700
Lexington (KY)	0.62	0.84	594,522	416,639	251,840	208,491	178,010
Lincoln (NE)	0.63	0.75	346,215	256,077	188,618	164,144	74,988
Little Rock (AR)	0.60	0.78	721,030	514,263	306,207	230,519	89,313
Longview (TX)	0.65	0.82	314,342	252,813	201,595	147,134	86,161
Los Angeles (CA)	0.37	0.50	12,828,837	9,410,212	4,367,911	997,830	33,381
Louisville (KY/IN)	0.71	0.86	1,256,868	1,034,761	697,918	465,795	322,989
Lubbock (TX)	0.71	0.83	323,328	262,506	154,276	26,388	152
Lynchburg (VA)	0.66	0.85	252,634	194,178	135,327	116,481	96,244
Macon (GA)	0.65	0.85	433,867	322,858	206,336	174,331	113,893
Madison (WI)	0.61	0.80	663,994	459,186	289,194	204,218	160,737
Manchester (NH)	0.33	0.56	965,532	650,663	341,635	278,326	206,556
Mansfield (OH)	0.83	0.77	322,726	321,912	230,210	173,433	130,409
McAllen (TX)	0.50	0.90	1,264,091	537,717	320,484	86,550	25,296
Medford (OR)	0.35	0.43	285,919	191,311	85,052	28,060	10,639
Melbourne (FL)	0.28	0.30	681,404	332,855	35,525	11,312	938
Memphis (TN/MS/AR)	0.69	0.87	1,316,100	997,844	676,274	404,768	220,185
Miami (FL)	0.21	0.36	5,564,635	3,220,844	693,705	66,542	257
Midland (TX)	0.62	0.77	301,317	225,236	87,702	4,995	505
Milwaukee (WI)	0.68	0.77	1,923,761	1,711,491	1,224,476	787,834	315,406
Minneapolis (MN/WI)	0.64	0.81	3,495,023	2,349,968	1,397,973	957,340	282,794
Mobile (AL)	0.64	0.81	650,341	502,814	337,738	192,615	111,428
Modesto (CA)	0.55	0.60	843,862	445,496	214,740	78,679	26,594
Monmouth (NJ)	0.55	0.66	1,206,947	849,211	281,949	127,080	69,993
Monroe (LA)	0.80	0.85	256,044	252,300	192,233	123,725	63,119
Montgomery (AL)	0.68	0.87	395,483	307,620	236,046	182,783	133,791

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Muncie (IN)	0.77	0.75	407,767	447,760	342,991	252,131	132,542
Myrtle Beach (SC)	0.56	0.88	634,562	420,248	329,155	230,863	127,235
Nashville (TN)	0.46	0.78	1,565,244	897,511	543,999	381,527	281,354
New Orleans (LA)	0.71	0.82	1,237,034	1,348,007	808,561	510,463	287,874
New York (NY/NJ/CT)	0.49	0.83	19,024,827	16,695,493	13,797,957	9,085,160	3,079,306
Norfolk (VA/NC)	0.47	0.69	1,709,794	1,240,802	729,861	443,975	177,752
Ocala (FL)	0.34	0.55	472,534	177,191	44,298	29,188	14,161
Oklahoma City (OK)	0.53	0.70	1,392,545	999,625	618,226	391,923	0
Olympia (WA)	0.41	0.55	252,264	124,264	44,884	22,366	3,270
Omaha (NE/IA)	0.63	0.76	902,041	689,736	471,079	389,349	165,149
Orlando (FL)	0.24	0.36	2,227,831	829,197	200,909	58,666	11,164
Owensboro (KY)	0.74	0.85	182,783	168,848	137,921	135,203	92,534
Panama City (FL)	0.39	0.59	196,264	116,059	55,963	19,389	4,283
Parkersburg (WV/OH)	0.85	0.92	201,716	208,308	150,269	135,434	111,396
Pensacola (FL)	0.34	0.45	684,856	421,002	173,518	84,535	23,002
Peoria (IL)	0.78	0.79	542,766	557,067	443,265	344,136	247,067
Philadelphia (PA/NJ/DE)	0.70	0.84	5,864,235	5,179,609	3,939,435	2,899,082	1,398,427
Phoenix (AZ)	0.25	0.37	4,192,887	1,599,970	374,961	105,706	7,710
Pittsburgh (PA)	0.80	0.89	2,483,851	2,781,748	2,703,797	2,212,645	756,747
Portland (ME)	0.58	0.81	919,237	740,581	573,820	458,405	373,949
Portland (OR/WA)	0.42	0.57	2,126,816	1,286,159	732,584	389,094	49,886
Poughkeepsie (NY)	0.68	0.82	930,341	727,971	422,388	319,733	285,733
Providence (RI/MA)	0.66	0.87	1,600,852	1,421,795	1,173,465	963,402	415,571
Provo-Orem (UT)	0.53	0.67	539,313	232,606	97,280	60,322	25,174
Raleigh-Durham (NC)	0.37	0.80	1,634,847	694,400	419,524	253,721	145,785
Redding (CA)	0.69	0.59	240,686	154,603	55,689	26,243	18,793
Reno (NV)	0.18	0.27	557,548	254,659	64,888	31,276	30,365
Residual Alabama (AL)	0.70	0.89	1,046,171	980,854	940,523	855,457	538,333
Residual Arizona (AZ)	0.41	0.68	478,407	246,950	115,671	85,720	9,715
Residual Arkansas (AR)	0.60	0.79	1,448,307	1,339,224	1,315,041	1,269,705	588,373
Residual California (CA)	0.67	0.58	506,103	348,243	210,175	115,724	104,514
Residual Colorado (CO)	0.41	0.59	913,841	606,774	472,620	423,119	89,158
Residual Connecticut (CT)	0.55	0.66	2,657,268	2,300,433	1,502,938	1,059,695	510,658
Residual Florida (FL)	0.45	0.58	339,655	209,951	113,897	99,516	52,367
Residual Georgia (GA)	0.62	0.86	2,404,365	1,627,718	1,348,199	1,344,691	672,140
Residual Idaho (ID)	0.46	0.63	722,562	530,884	372,112	299,609	23,752
Residual Illinois (IL)	0.74	0.82	1,524,064	1,551,894	1,415,324	1,404,512	1,177,024
Residual Indiana (IN)	0.68	0.75	1,009,889	895,308	652,190	556,035	468,958
Residual Iowa (IA)	0.72	0.84	1,060,510	1,205,667	1,276,864	1,281,160	899,603
Residual Kansas (KS)	0.60	0.75	985,020	1,007,988	1,009,330	1,056,732	623,904
Residual Kentucky (KY)	0.70	0.87	1,845,366	1,664,729	1,547,836	1,337,647	883,164
Residual Louisiana (LA)	0.80	0.86	185,382	201,083	181,374	144,678	101,415
Residual Maine (ME)	0.66	0.85	409,124	384,079	339,954	309,609	274,987
Residual Maryland (MD)	0.54	0.71	272,569	182,003	127,089	119,492	119,911
Residual Massachusetts (MA)	0.63	0.77	373,814	307,064	188,888	146,872	108,956
Residual Michigan (MI)	0.80	0.85	1,288,607	1,121,526	808,749	788,627	365,031
Residual Minnesota (MN)	0.73	0.83	1,291,813	1,216,614	1,151,008	1,034,394	365,636
Residual Mississippi (MS)	0.75	0.89	1,481,914	1,453,985	1,517,605	1,342,133	863,350
Residual Missouri (MO)	0.64	0.77	1,442,368	1,248,716	1,202,139	1,272,572	1,000,483
Residual Montana (MT)	0.49	0.63	989,415	786,690	591,024	548,889	39,159
Residual Nebraska (NE)	0.64	0.81	664,629	706,759	739,460	815,105	274,698

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Residual Nevada (NV)	0.21	0.33	191,734	82,747	46,906	41,272	30,615
Residual New Hampshire (NH)	0.43	0.64	350,938	269,947	191,607	164,757	140,435
Residual New Mexico (NM)	0.51	0.61	700,353	541,752	362,254	234,203	54,537
Residual New York (NY)	0.77	0.87	1,136,649	1,097,247	962,773	847,200	780,104
Residual North Carolina (NC)	0.59	0.87	1,469,843	1,077,126	921,444	696,419	445,460
Residual North Dakota (ND)	0.66	0.83	506,492	545,263	540,894	584,508	24,302
Residual Ohio (OH)	0.81	0.82	1,799,179	1,705,606	1,354,287	1,226,591	1,024,186
Residual Oklahoma (OK)	0.56	0.73	1,070,417	1,016,124	922,894	1,008,840	0
Residual Oregon (OR)	0.42	0.48	723,809	551,242	374,449	213,477	54,861
Residual Pennsylvania (PA)	0.81	0.91	337,363	337,856	329,595	317,557	223,428
Residual South Carolina (SC)	0.70	0.89	145,784	136,640	132,601	144,840	95,302
Residual South Dakota (SD)	0.62	0.77	557,656	527,065	526,350	534,550	65,716
Residual Tennessee (TN)	0.61	0.85	1,778,759	1,347,811	1,098,394	997,498	780,149
Residual Texas (TX)	0.67	0.85	2,077,585	1,757,024	1,603,566	1,387,555	489,413
Residual Utah (UT)	0.56	0.74	522,536	264,514	170,728	149,831	53,916
Residual Vermont (VT)	0.46	0.71	291,542	252,001	205,142	198,173	186,459
Residual Virginia (VA)	0.64	0.82	1,155,178	1,008,230	907,223	753,522	522,942
Residual Washington (WA)	0.46	0.55	1,299,761	854,568	548,973	352,751	27,603
Residual West Virginia (WV)	0.72	0.86	832,259	911,224	1,033,709	784,634	294,385
Residual Wisconsin (WI)	0.69	0.83	1,150,980	997,581	785,650	711,280	277,008
Residual Wyoming (WY)	0.36	0.43	563,626	469,557	290,529	194,402	20,789
Richland (WA/OR)	0.48	0.54	403,261	262,341	156,414	81,975	24,840
Richmond (VA)	0.55	0.75	1,186,501	795,892	497,645	357,779	238,128
Riverside-San Bernardino (CA)	0.50	0.54	4,224,851	1,558,182	451,688	123,698	9,290
Roanoke (VA/WV)	0.66	0.86	500,446	414,297	287,179	198,667	115,790
Rochester (MN)	0.58	0.77	253,060	186,793	134,313	114,614	108,906
Rochester (NY)	0.76	0.82	1,217,156	1,125,717	802,490	628,628	429,912
Rockford (IL)	0.64	0.63	349,431	279,514	169,455	106,251	42,013
Rocky Mount (NC)	0.71	0.92	233,626	186,273	166,059	115,869	59,976
Sacramento (CA)	0.54	0.56	2,149,127	1,099,814	375,636	133,144	71,077
Saginaw (MI)	0.85	0.85	522,007	549,601	368,140	271,464	151,023
Salem (OR)	0.43	0.48	1,043,897	738,159	372,865	156,357	55,385
Salinas (CA)	0.44	0.55	732,708	503,590	211,402	63,244	29,688
Salt Lake City (UT)	0.53	0.71	1,694,911	959,893	418,555	237,189	61,788
San Antonio (TX)	0.58	0.84	2,154,746	1,165,043	614,016	294,212	70,755
San Diego (CA)	0.39	0.42	3,095,313	1,861,846	556,808	112,248	5,829
San Francisco (CA)	0.40	0.54	4,885,219	3,585,032	2,287,370	1,030,145	361,163
San Jose (CA)	0.32	0.52	1,781,642	1,295,071	290,547	100,676	35,039
Santa Barbara (CA)	0.51	0.55	693,532	454,129	149,637	62,990	18,655
Santa Fe (NM)	0.47	0.75	235,303	141,697	90,772	51,352	32,810
Santa Rosa (CA)	0.57	0.60	636,384	402,785	155,740	81,608	45,322
Savannah (GA/SC)	0.47	0.82	534,621	310,596	204,567	148,497	86,346
Scranton (PA)	0.63	0.90	910,959	782,304	790,539	798,300	333,692
Seattle (WA)	0.37	0.51	3,439,809	2,093,112	1,120,448	601,090	11,616
Sheboygan (WI)	0.76	0.89	196,949	183,853	147,790	111,557	71,711
Shreveport (LA)	0.70	0.79	520,016	486,215	371,213	262,379	125,505
Sioux City (IA/NE/SD)	0.73	0.91	187,401	181,825	183,923	173,213	47,633
Sioux Falls (SD)	0.54	0.75	242,125	152,765	115,598	90,898	25,751
South Bend (IN/MI)	0.70	0.73	657,918	632,176	470,503	270,817	151,476
Spokane (WA/ID)	0.49	0.58	696,213	471,470	308,723	214,875	7,311
Springfield (IL)	0.79	0.84	275,275	256,037	210,610	179,976	119,182

Table F2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Springfield (MA)	0.60	0.76	692,942	646,148	508,312	419,265	187,375
Springfield (MO)	0.55	0.72	514,409	317,508	211,630	190,002	111,977
St. Joseph (MO/KS)	0.72	0.85	165,929	152,839	159,934	177,842	142,565
St. Louis (MO/IL)	0.75	0.80	2,819,961	2,516,116	1,926,706	1,392,529	730,141
St. Lucie (FL)	0.26	0.40	424,107	151,196	27,987	5,079	399
State College (PA)	0.78	0.91	326,745	286,894	241,898	249,492	143,219
Steubenville (OH/WV)	0.84	0.88	124,454	163,099	157,787	114,082	43,913
Stockton (CA)	0.52	0.63	685,306	347,342	200,750	79,905	24,349
Sumter (SC)	0.65	0.87	223,344	173,651	145,309	134,143	78,419
Syracuse (NY)	0.80	0.86	1,091,336	1,091,865	856,670	694,686	488,966
Tallahassee (FL)	0.51	0.67	484,972	293,750	175,343	128,248	71,700
Tampa (FL)	0.27	0.41	2,783,243	1,613,603	436,365	129,872	8,947
Terre Haute (IN)	0.74	0.84	265,851	257,619	240,214	256,159	173,150
Texarkana (TX/AR)	0.71	0.86	247,936	208,688	174,734	158,659	66,214
Toledo (OH/MI)	0.82	0.80	831,665	819,982	608,294	426,728	196,423
Topeka (KS)	0.63	0.72	233,870	203,953	147,623	129,449	83,772
Tucson (AZ)	0.33	0.45	1,159,029	637,588	182,048	93,834	14,786
Tulsa (OK)	0.52	0.68	1,052,519	819,904	526,196	417,160	0
Tuscaloosa (AL)	0.73	0.89	219,461	164,166	131,406	96,102	73,441
Tyler (TX)	0.66	0.85	492,092	303,603	212,576	205,538	89,547
Ventura (CA)	0.47	0.50	823,318	529,174	114,647	28,724	5,073
Waco (TX)	0.66	0.85	306,073	227,126	200,036	180,502	70,945
Washington (DC/VA/MD)	0.34	0.55	5,679,291	3,452,103	1,732,083	787,285	477,083
Waterloo (IA)	0.75	0.85	224,524	243,203	200,669	154,704	105,730
Wheeling (WV/OH)	0.84	0.90	188,383	236,142	242,356	247,681	157,400
Wichita (KS)	0.58	0.64	602,269	452,979	297,524	191,064	65,997
Wichita Falls (TX)	0.59	0.69	151,306	137,930	115,205	95,029	6,074
Williamsport (PA)	0.80	0.89	224,399	226,655	196,248	179,341	150,378
Wilmington (NC)	0.51	0.81	455,603	242,991	181,257	120,169	73,830
Yakima (WA)	0.47	0.54	284,146	197,385	157,958	81,447	2,272
Youngstown (OH/PA)	0.87	0.87	764,722	880,371	732,538	532,694	225,826
Yuma (AZ/CA)	0.42	0.62	390,768	182,664	90,981	58,357	4,500

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Abilene (TX)	Omitted	Omitted	9.23	6.69	6.56
Akron (OH)	Slow	No Decline	3.52	3.40	3.13
Albany (GA)	Fast	Small Decline	5.54	5.07	4.75
Albany (NY)	Slow	No Decline	3.30	3.11	2.78
Albuquerque (NM)	Fast	Big Decline	5.61	4.03	3.94
Alexandria (LA)	Omitted	Omitted	6.29	3.79	4.73
Allentown (PA)	Slow	No Decline	3.30	3.61	3.20
Amarillo (TX)	Fast	Big Decline	6.62	5.07	5.20
Aniston (AL)	Fast	Big Decline	6.84	5.13	5.86
Ann Arbor (MI)	Omitted	Omitted	9.86	8.28	8.63
Appleton (WI)	Slow	No Decline	3.53	3.45	3.14
Asheville (NC)	Medium	No Decline	4.32	4.30	4.44

Table F3: LLM Categorization and Migration Rates

LLM	Category of: Speed	Decline	Migration Rate: IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Athens (GA)	Omitted	Omitted	7.51	7.55	6.52
Atlanta (GA)	Medium	Small Decline	4.45	4.12	3.36
Atlantic City-Vineland (NJ)	Medium	Big Decline	4.23	3.50	3.31
Auburn (AL)	Omitted	Omitted	8.47	7.77	7.08
Augusta (GA/SC)	Omitted	Omitted	5.33	4.62	4.16
Austin (TX)	Fast	Big Decline	6.49	5.14	4.08
Bakersfield (CA)	Fast	Big Decline	6.12	4.23	3.47
Baltimore (MD)	Medium	No Decline	3.72	3.71	3.32
Baton Rouge (LA)	Slow	No Decline	3.94	3.75	2.98
Beaumont (TX)	Medium	No Decline	4.66	4.66	4.46
Bellingham (WA)	Omitted	Omitted	6.29	5.88	4.42
Biloxi (MS)	Fast	Big Decline	6.40	5.51	5.23
Binghamton (NY/PA)	Medium	Small Decline	4.36	3.95	3.56
Birmingham (AL)	Slow	No Decline	3.30	3.37	3.37
Bloomington (IL)	Omitted	Omitted	5.44	5.23	5.07
Bloomington (IN)	Omitted	Omitted	5.44	5.23	5.45
Boise City (ID)	Medium	Big Decline	5.23	4.54	3.98
Boston (MA)	Medium	Big Decline	3.79	3.03	2.70
Bradenton (FL)	Fast	Big Decline	5.77	4.96	4.38
Buffalo (NY)	Slow	No Decline	2.60	2.42	2.38
Burlington (VT)	Omitted	Omitted	4.55	4.15	3.80
Cedar Rapids (IA)	Medium	Small Decline	4.51	4.17	4.25
Champaign-Urbana (IL)	Omitted	Omitted	6.17	5.18	5.48
Charleston (SC)	Fast	Big Decline	7.15	5.23	4.89
Charleston (WV)	Slow	Small Decline	3.76	3.20	3.34
Charlotte (NC/SC)	Medium	No Decline	4.12	4.06	3.45
Charlottesville (VA)	Omitted	Omitted	6.24	5.95	5.66
Chattanooga (TN/GA)	Slow	Small Decline	3.97	3.68	3.47
Chicago (IL/IN/WI)	Slow	Small Decline	2.88	2.49	2.44
Chico (CA)	Fast	Big Decline	6.60	5.29	4.77
Cincinnati (OH/KY/IN)	Slow	No Decline	2.96	2.87	2.65
Clarksville (TN/KY)	Omitted	Omitted	11.14	9.28	8.06
Cleveland (OH)	Slow	No Decline	2.94	2.90	2.72
College Station (TX)	Omitted	Omitted	9.78	8.66	7.96
Colorado Springs (CO)	Omitted	Omitted	10.24	7.30	6.80
Columbia (MO)	Omitted	Omitted	5.59	5.24	5.85
Columbia (SC)	Medium	Small Decline	4.77	4.47	4.02
Columbus (GA/AL)	Omitted	Omitted	8.21	7.51	6.77
Columbus (OH)	Medium	Small Decline	3.73	3.52	3.26
Corpus Christi (TX)	Fast	Big Decline	6.19	5.34	4.67
Cumberland (MD/WV)	Slow	No Decline	3.27	3.10	4.00
Dallas-Fort Worth (TX)	Fast	Big Decline	4.80	3.41	2.82
Davenport (IA/IL)	Medium	No Decline	4.11	4.02	3.34
Dayton (OH)	Medium	Small Decline	3.84	3.45	3.31
Daytona Beach (FL)	Fast	Small Decline	6.36	6.00	5.04
Denver (CO)	Fast	Big Decline	4.71	3.60	3.64
Des Moines (IA)	Medium	Big Decline	4.74	3.80	3.46
Detroit (MI)	Slow	No Decline	3.02	2.83	2.44
Dothan (AL)	Omitted	Omitted	6.78	5.91	4.77
Dover (DE/MD)	Slow	No Decline	3.98	3.80	3.22
Dubuque (IA/IL)	Slow	Small Decline	4.30	3.94	4.02

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Duluth (MN/WI)	Slow	Small Decline	3.95	3.64	4.01
Eau Claire (WI/MN)	Medium	Small Decline	4.78	4.36	4.10
El Paso (TX/NM)	Fast	Big Decline	6.29	4.37	4.44
Erie (PA/NY)	Slow	Small Decline	3.25	3.00	2.84
Evansville (IN/KY/IL)	Slow	No Decline	3.10	2.95	3.00
Fargo (ND/MN)	Fast	Big Decline	5.95	5.09	5.07
Fayetteville (AR/OK/MO)	Fast	Small Decline	5.68	5.11	4.19
Fayetteville (NC)	Omitted	Omitted	8.22	7.14	6.30
Flagstaff (AZ/UT)	Fast	Big Decline	9.57	7.37	6.55
Flint (MI)	Medium	Small Decline	4.92	4.59	4.82
Florence (AL/TN)	Slow	No Decline	3.29	3.44	3.41
Fort Myers (FL)	Fast	Big Decline	6.28	5.61	4.50
Fort Smith (AR/OK)	Medium	Big Decline	4.90	4.14	4.19
Fort Wayne (IN)	Slow	Small Decline	3.96	3.37	3.37
Fresno (CA)	Medium	Big Decline	4.52	3.11	2.51
Gainesville (FL)	Omitted	Omitted	8.49	8.20	8.17
Goldsboro (NC)	Omitted	Omitted	5.75	5.11	4.28
Grand Rapids (MI)	Slow	No Decline	3.25	3.20	2.96
Green Bay (WI)	Slow	No Decline	3.89	3.74	4.17
Greensboro (NC/VA)	Slow	No Decline	3.22	3.19	3.00
Greenville (NC)	Omitted	Omitted	5.41	5.62	6.64
Greenville (SC/NC)	Slow	No Decline	3.00	2.91	2.83
Hagerstown (MD/WV/PA)	Slow	No Decline	3.56	3.52	3.92
Harrisburg (PA)	Slow	No Decline	3.13	3.12	3.00
Hattiesburg (MS)	Omitted	Omitted	5.98	6.12	5.60
Hickory (NC)	Slow	No Decline	3.58	4.01	3.44
Houma (LA)	Slow	Small Decline	3.90	3.69	3.08
Houston (TX)	Medium	Big Decline	4.14	2.96	2.43
Huntington (WV/KY/OH)	Slow	Small Decline	3.87	3.47	3.54
Huntsville (AL/TN)	Medium	Big Decline	4.38	3.68	3.59
Indianapolis (IN)	Medium	Small Decline	3.78	3.42	3.27
Iowa City (IA)	Omitted	Omitted	7.19	6.56	6.46
Jackson (MI)	Medium	Small Decline	4.68	4.21	4.44
Jackson (MS)	Medium	Big Decline	4.31	3.65	3.85
Jacksonville (FL/GA)	Fast	Big Decline	6.26	5.16	4.28
Jacksonville (NC)	Omitted	Omitted	22.45	17.13	13.34
Johnson City (TN/VA)	Slow	No Decline	2.89	2.86	2.76
Johnstown (PA)	Slow	Small Decline	2.74	2.53	2.67
Joplin (MO/KS/OK)	Medium	Small Decline	5.17	4.67	4.86
Kalamazoo (MI)	Medium	Small Decline	5.04	4.70	4.44
Kansas City (MO/KS)	Medium	Big Decline	4.17	3.52	3.36
Killeen (TX)	Omitted	Omitted	15.31	12.10	9.08
Knoxville (TN)	Slow	No Decline	3.85	3.82	3.51
LaCrosse (WI/MN)	Medium	Small Decline	4.68	4.22	4.56
Lafayette (IN/IL)	Omitted	Omitted	5.36	5.19	4.96
Lafayette (LA)	Slow	Small Decline	3.63	3.03	2.92
Lake Charles (LA)	Omitted	Omitted	9.13	5.21	4.45
Lakeland (FL)	Fast	No Decline	6.16	6.34	4.82
Lancaster (PA)	Slow	No Decline	2.90	2.99	2.74
Lansing (MI)	Omitted	Omitted	6.07	5.41	5.17
Laredo (TX)	Slow	Big Decline	4.12	2.94	2.59

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Las Vegas (NV)	Fast	Big Decline	6.53	5.30	4.21
Lawrence (KS)	Omitted	Omitted	10.68	9.41	8.25
Lexington (KY)	Omitted	Omitted	4.77	4.46	4.60
Lincoln (NE)	Omitted	Omitted	5.05	4.89	4.61
Little Rock (AR)	Medium	Big Decline	5.03	4.26	4.15
Longview (TX)	Fast	Big Decline	5.70	4.80	5.05
Los Angeles (CA)	Fast	Big Decline	4.74	3.17	2.93
Louisville (KY/IN)	Slow	No Decline	2.91	2.80	2.95
Lubbock (TX)	Omitted	Omitted	7.70	5.89	6.15
Lynchburg (VA)	Omitted	Omitted	4.14	3.95	3.16
Macon (GA)	Medium	No Decline	4.73	4.54	4.33
Madison (WI)	Omitted	Omitted	4.96	4.63	4.41
Manchester (NH)	Fast	Big Decline	5.67	4.15	3.26
Mansfield (OH)	Slow	Small Decline	4.09	3.81	3.67
McAllen (TX)	Medium	Big Decline	4.73	2.77	2.67
Medford (OR)	Fast	Big Decline	5.84	4.49	4.85
Melbourne (FL)	Fast	Big Decline	6.05	5.19	4.78
Memphis (TN/MS/AR)	Medium	Big Decline	4.17	3.47	3.27
Miami (FL)	Medium	Big Decline	4.03	3.39	3.19
Midland (TX)	Fast	Big Decline	7.05	5.19	5.42
Milwaukee (WI)	Slow	No Decline	3.12	2.94	2.63
Minneapolis (MN/WI)	Slow	Small Decline	3.06	2.79	2.61
Mobile (AL)	Slow	Small Decline	3.96	3.73	3.26
Modesto (CA)	Fast	Big Decline	5.78	4.48	3.74
Monmouth (NJ)	Medium	Big Decline	4.52	3.45	2.78
Monroe (LA)	Medium	Small Decline	4.48	4.00	3.74
Montgomery (AL)	Medium	No Decline	5.17	5.07	4.65
Muncie (IN)	Medium	Small Decline	4.56	4.19	4.72
Myrtle Beach (SC)	Medium	Small Decline	4.67	4.24	3.87
Nashville (TN)	Medium	No Decline	4.19	4.10	3.76
New Orleans (LA)	Medium	No Decline	3.86	3.81	3.58
New York (NY/NJ/CT)	Slow	Small Decline	2.86	2.36	2.23
Norfolk (VA/NC)	Fast	Big Decline	6.92	5.55	4.63
Ocala (FL)	Fast	Big Decline	6.40	5.67	5.40
Oklahoma City (OK)	Medium	Big Decline	4.75	3.57	3.56
Olympia (WA)	Omitted	Omitted	7.70	7.32	5.82
Omaha (NE/IA)	Medium	Big Decline	4.64	3.53	3.69
Orlando (FL)	Fast	Big Decline	6.94	5.87	4.75
Owensboro (KY)	Slow	Small Decline	3.42	3.13	3.67
Panama City (FL)	Omitted	Omitted	8.04	7.09	7.12
Parkersburg (WV/OH)	Slow	Big Decline	3.73	3.07	3.32
Pensacola (FL)	Omitted	Omitted	7.97	6.96	6.50
Peoria (IL)	Slow	Small Decline	3.78	3.53	3.74
Philadelphia (PA/NJ/DE)	Slow	Small Decline	2.96	2.67	2.36
Phoenix (AZ)	Fast	Big Decline	5.52	4.25	3.50
Pittsburgh (PA)	Slow	No Decline	2.37	2.20	2.30
Portland (ME)	Medium	Small Decline	3.98	3.42	2.93
Portland (OR/WA)	Medium	Small Decline	4.14	3.84	3.68
Poughkeepsie (NY)	Medium	Big Decline	4.79	4.08	3.25
Providence (RI/MA)	Slow	Small Decline	3.45	3.15	2.86
Provo-Orem (UT)	Omitted	Omitted	7.78	6.96	6.30

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Raleigh-Durham (NC)	Fast	Big Decline	5.40	4.75	4.46
Redding (CA)	Fast	Big Decline	6.11	4.70	4.67
Reno (NV)	Fast	Big Decline	7.07	5.15	4.79
Residual Alabama (AL)	Omitted	Omitted	4.19	4.26	3.84
Residual Arizona (AZ)	Omitted	Omitted	9.76	7.04	6.12
Residual Arkansas (AR)	Omitted	Omitted	4.90	4.35	3.38
Residual California (CA)	Omitted	Omitted	7.48	6.06	6.05
Residual Colorado (CO)	Omitted	Omitted	7.26	6.43	5.41
Residual Connecticut (CT)	Medium	Big Decline	3.76	3.00	2.60
Residual Florida (FL)	Omitted	Omitted	9.00	7.34	4.71
Residual Georgia (GA)	Omitted	Omitted	4.90	5.02	3.89
Residual Idaho (ID)	Omitted	Omitted	6.11	5.37	4.63
Residual Illinois (IL)	Omitted	Omitted	4.23	4.08	3.76
Residual Indiana (IN)	Omitted	Omitted	4.69	4.14	3.84
Residual Iowa (IA)	Omitted	Omitted	4.55	4.37	3.52
Residual Kansas (KS)	Omitted	Omitted	6.62	6.03	4.74
Residual Kentucky (KY)	Omitted	Omitted	4.11	3.83	3.05
Residual Louisiana (LA)	Omitted	Omitted	6.05	5.76	4.22
Residual Maine (ME)	Omitted	Omitted	4.44	3.63	2.97
Residual Maryland (MD)	Omitted	Omitted	4.50	4.64	3.71
Residual Massachusetts (MA)	Omitted	Omitted	5.35	4.18	4.27
Residual Michigan (MI)	Omitted	Omitted	5.40	4.48	3.74
Residual Minnesota (MN)	Omitted	Omitted	4.69	4.31	3.37
Residual Mississippi (MS)	Omitted	Omitted	4.14	3.95	3.54
Residual Missouri (MO)	Omitted	Omitted	5.40	5.04	4.16
Residual Montana (MT)	Omitted	Omitted	5.06	4.22	3.15
Residual Nebraska (NE)	Omitted	Omitted	4.97	4.86	3.26
Residual Nevada (NV)	Omitted	Omitted	10.88	7.75	6.18
Residual New Hampshire (NH)	Omitted	Omitted	6.33	5.66	4.59
Residual New Mexico (NM)	Omitted	Omitted	8.17	5.90	5.03
Residual New York (NY)	Omitted	Omitted	5.17	4.77	4.20
Residual North Carolina (NC)	Omitted	Omitted	5.24	5.34	4.54
Residual North Dakota (ND)	Omitted	Omitted	5.58	4.65	4.18
Residual Ohio (OH)	Omitted	Omitted	3.87	3.64	3.40
Residual Oklahoma (OK)	Omitted	Omitted	7.32	6.25	4.96
Residual Oregon (OR)	Omitted	Omitted	6.96	5.71	5.13
Residual Pennsylvania (PA)	Omitted	Omitted	3.59	3.38	3.10
Residual South Carolina (SC)	Omitted	Omitted	4.67	4.61	4.21
Residual South Dakota (SD)	Omitted	Omitted	5.77	5.06	3.33
Residual Tennessee (TN)	Omitted	Omitted	3.77	4.07	3.34
Residual Texas (TX)	Omitted	Omitted	7.01	6.12	5.04
Residual Utah (UT)	Omitted	Omitted	6.95	6.41	5.86
Residual Vermont (VT)	Omitted	Omitted	5.41	4.67	4.59
Residual Virginia (VA)	Omitted	Omitted	3.90	3.88	3.40
Residual Washington (WA)	Omitted	Omitted	6.95	5.76	5.19
Residual West Virginia (WV)	Omitted	Omitted	3.94	3.35	3.16
Residual Wisconsin (WI)	Omitted	Omitted	4.42	4.03	3.28
Residual Wyoming (WY)	Omitted	Omitted	6.60	5.89	5.57
Richland (WA/OR)	Fast	Big Decline	5.93	4.14	4.45
Richmond (VA)	Medium	No Decline	3.91	3.86	3.43
Riverside-San Bernardino (CA)	Fast	Big Decline	7.18	4.96	3.53

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Roanoke (VA/WV)	Omitted	Omitted	4.46	4.14	4.15
Rochester (MN)	Medium	Small Decline	5.00	4.44	4.74
Rochester (NY)	Slow	Small Decline	3.08	2.78	2.85
Rockford (IL)	Slow	No Decline	4.16	4.06	3.77
Rocky Mount (NC)	Slow	Small Decline	4.24	3.99	4.05
Sacramento (CA)	Fast	Big Decline	5.69	4.08	3.67
Saginaw (MI)	Medium	Big Decline	4.14	3.42	3.41
Salem (OR)	Fast	Big Decline	5.71	4.52	4.18
Salinas (CA)	Fast	Big Decline	8.56	5.21	4.81
Salt Lake City (UT)	Medium	Small Decline	4.19	3.89	3.92
San Antonio (TX)	Fast	Big Decline	5.13	3.95	3.47
San Diego (CA)	Fast	Big Decline	7.01	5.03	4.66
San Francisco (CA)	Fast	Big Decline	5.04	3.88	3.64
San Jose (CA)	Fast	Big Decline	6.74	4.97	4.57
Santa Barbara (CA)	Fast	Big Decline	7.13	5.14	4.91
Santa Fe (NM)	Fast	Big Decline	6.28	5.20	4.89
Santa Rosa (CA)	Fast	Big Decline	5.38	4.07	3.69
Savannah (GA/SC)	Omitted	Omitted	7.30	6.79	6.70
Scranton (PA)	Slow	No Decline	2.90	3.31	2.77
Seattle (WA)	Fast	Big Decline	4.71	3.96	3.48
Sheboygan (WI)	Slow	No Decline	3.72	3.59	3.14
Shreveport (LA)	Medium	Big Decline	5.05	3.88	3.99
Sioux City (IA/NE/SD)	Medium	Big Decline	5.04	4.36	4.70
Sioux Falls (SD)	Fast	Small Decline	5.64	5.13	4.97
South Bend (IN/MI)	Medium	Small Decline	4.47	4.13	4.24
Spokane (WA/ID)	Fast	Big Decline	5.25	4.35	3.97
Springfield (IL)	Slow	Small Decline	4.21	3.82	4.38
Springfield (MA)	Omitted	Omitted	4.06	3.32	2.81
Springfield (MO)	Medium	Small Decline	5.00	4.69	4.70
St. Joseph (MO/KS)	Medium	No Decline	4.77	4.66	5.25
St. Louis (MO/IL)	Slow	Small Decline	3.12	2.67	2.70
St. Lucie (FL)	Fast	Big Decline	6.71	5.98	5.16
State College (PA)	Omitted	Omitted	4.67	4.42	4.00
Steubenville (OH/WV)	Slow	No Decline	3.64	3.46	3.39
Stockton (CA)	Fast	Big Decline	6.51	5.36	4.14
Sumter (SC)	Omitted	Omitted	6.17	5.42	4.72
Syracuse (NY)	Slow	Small Decline	3.55	2.98	2.87
Tallahassee (FL)	Omitted	Omitted	6.17	6.45	5.90
Tampa (FL)	Fast	Big Decline	5.39	4.69	3.78
Terre Haute (IN)	Medium	No Decline	4.48	4.30	4.40
Texarkana (TX/AR)	Fast	Big Decline	5.52	4.61	4.87
Toledo (OH/MI)	Slow	Small Decline	3.82	3.60	3.35
Topeka (KS)	Medium	Big Decline	4.75	4.09	4.63
Tucson (AZ)	Fast	Big Decline	6.25	4.69	4.68
Tulsa (OK)	Medium	Big Decline	4.89	3.86	3.73
Tuscaloosa (AL)	Omitted	Omitted	5.30	5.01	4.46
Tyler (TX)	Fast	Big Decline	6.36	5.23	4.80
Ventura (CA)	Fast	Big Decline	7.54	4.68	4.11
Waco (TX)	Omitted	Omitted	6.24	5.43	5.11
Washington (DC/VA/MD)	Fast	Big Decline	5.19	4.01	3.65
Waterloo (IA)	Medium	Small Decline	4.52	4.16	4.46

Table F3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Wheeling (WV/OH)	Slow	Small Decline	3.41	2.98	3.30
Wichita (KS)	Medium	Big Decline	4.83	3.87	3.64
Wichita Falls (TX)	Omitted	Omitted	7.74	7.44	7.28
Williamsport (PA)	Slow	Small Decline	3.68	3.38	3.34
Wilmington (NC)	Medium	No Decline	4.89	5.25	4.49
Yakima (WA)	Fast	Big Decline	5.53	4.39	3.54
Youngstown (OH/PA)	Slow	No Decline	2.96	2.95	2.94
Yuma (AZ/CA)	Fast	Big Decline	7.06	4.84	3.93