Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 19-49
Revised April 2021
November 2019
https://doi.org/10.21799/frbp.wp.2019.49

Fast Locations and Slowing Mobility

Patrick Coate
National Council on Compensation Insurance

Kyle Mangum
Federal Reserve Bank of Philadelphia Research Department

ISSN: 1962-5361

Fast Locations and Slowing Mobility

Patrick Coate* Kyle Mangum†

March 2021

Abstract

Declining internal migration in the United States is driven by increasing home attachment in locations with previously high rates of population turnover. These "fast locations" were the population growth destinations of the 20th century, where home attachments were once low but have increased as regional population growth has converged. Using a novel measure of home 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 mobility decline, and its effect is consistent with the observed spatial pattern. Population aging explains most of the remainder but in a more spatially neutral way. The paper then uses a stylized island economy model featuring endogenous home attachments to show that after a shock, gross migration returns to steady state much more slowly than net population change.

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. We thank Kelly Bishop, Matthias Brachert, Chris Clapp, Tom Cooke, Jeff Lin, Pia Orrenius, and Chris Palmer, as well as numerous seminar and conference participants, for constructive comments at various stages of this research. We are grateful to Solomon Tarlin for tireless research assistance, and David Van Riper for extra help with IPUMS data. We are responsible for remaining errors.

†Federal Reserve Bank of Philadelphia. Contact: kyle.mangum@phil.frb.org. The views of this paper are those of the authors and do not necessarily represent the position of FRB Philadelphia or any other part of the Federal Reserve System.

1 Introduction

Internal migration rates in the U.S. have steadily trended downward in recent decades. The decline is pervasive across demographic strata, housing tenures, and household and family types (Molloy et al. (2011), Kaplan and Schulhofer-Wohl (2017)), suggesting a fundamental shift in the functioning of regions. This observation is alarming for policymakers because migration is considered a primary labor market adjustment mechanism. Americans have typically been regarded as 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)). While the general notion that "people are moving less these days" has entered the Zeitgeist in both the academic and popular social sciences, the causes of the decline remain poorly understood–which is problematic, of course, because knowledge of the mechanisms driving the decline is needed to inform policy and future research.

This paper finds that declining internal migration in the U.S. is primarily due to the increasing prevalence of home attachment. Migration propensity depends strongly on preference for one's place of origin, and the average attachment has increased because regional population growth has converged.[1] 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 higher degrees of home attachment instead of weakly attached transplants. Consequently, migration out of these places has declined. Because fast locations were the source of the majority of migrants, their decline has driven down the national average. Thus, the decline is the result of not just household demographics, but *spatial demographics:* where people are, and from where they came.

We demonstrate the role of home attachment in the national decline of mobility and that it uniquely fits the spatial pattern. We proceed to do this in three parts. First, we collect a set of stylized facts on the mobility decline, its spatial pattern, and the history of population change across regions. Second, we use a structural model of migration behavior to measure the role of evolving home attachment in the migration rates over a three decade period. Third, we use a stylized model to show generically how population convergence would precede a decline in gross migration rates when people exhibit attachment to original locations.

In the first part of the paper, our stylized facts begin with 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, the evidence indicates a *preference* for home, not selection on unobserved moving costs, drives the reluctance

[1] Home (or place) attachment is not a new idea to the social sciences (see the literature review below), but this paper is the first to connect evolving home attachments to the national trend in mobility.

to leave. Those living away from home are significantly more likely to return there, showing an inclination for a home location above alternatives. Moreover, the *intensity* of one's home attachment predicts differential migration rates among natives. More deeply "rooted" natives, measured as those born to locally-born parents, are less likely to leave than unrooted natives.

We then connect the general notion of home attachment to the spatial heterogeneity in migration decline. The distinguishing feature of fast locations is their relatively recent population growth; they are predominantly in the West and South regions of the U.S., the centers of population growth in the 20th century. The high share of transplants and unrooted natives resulted in high rates of population turnover in these places. 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. As population growth rates converged, these places became increasingly populated by regional natives with increasing levels of rootedness.

We close our descriptive analysis showing the relationship between turnover and other local labor market attributes, including city size, demographic composition (age and education) and income opportunities. These features will be important controls in our model but do not drive the spatially heterogeneous trends in migration rates.

With the stylized facts as a foundation, in the second part of the paper we develop and estimate a structural model of migration which allows us to jointly account for multiple factors that affect migration propensity and 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 fixed costs varying by agent type and marginal costs by distance between locations. 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") in our empirical model, we use a state-cohort matching method leveraging geographically harmonized historical census data. Using an individual's state and year of birth, we derive the probability that his or her 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 nonnatives.

A multinomial discrete choice model naturally applies to a location choice problem such as ours (with interest in the effects of both individual and spatial heterogeneity), but there are two complications that discourage the use of a standard conditional logit model. First, there is a marked asymmetry in the elasticities of location attributes in "move out" versus "move to" decisions, especially with regard to home preference.[2] In particular, the marginal effect of the home premium for choosing a location conditional on moving is substantially larger (in

[2]See Monras (2018) for a discussion of the asymmetry of in-moving and out-moving elasticities to local labor market attributes.

percentage terms) than its marginal effect on the decision whether to move at all. To account for this, we use a nested formulation whereby agents first choose whether to move and then, if moving, which location to choose. Second, because relocating comes at a cost, migration is a dynamic decision (Sjaastad (1962)), but agents of different ages and birthplaces, living in different locations, face materially different sets of opportunities and continuation values. This heterogeneity introduces an omitted variable problem if agents' future option values are not accounted for–and further complicated in a nested choice problem such as ours. To address this issue, we leverage the properties of finite dependence in conditional choice probability estimation (Arcidiacono and Miller (2011)) to derive moment conditions for agents facing a dynamic discrete choice problem. We derive a tractable linear estimator, which is to our knowledge the first application of finite dependence in a nested logit model, and the first to use method of moments estimation so that the estimator can be applied to aggregated data.[3]

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 depth of roots help identify the intensity of home attachment. We estimate the model on cross-sectional data from the American Community Survey of 2005-2017, and 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.

A key validation of the model, which derived its estimates from recent 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. This is to test the potential impact of changing decision environments without forcing changes in primitive preferences (whether move costs or the dispersion of idiosyncratic preferences). 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.

Several factors contribute to the aggregate decline, and the model permits a decomposition of the sources to quantify the contribution of each. Demographic factors–principally, an aging populace–matter nontrivially for the national decline but cannot produce the observed magnitude and cannot rationalize the spatial pattern. Rising home attachment, however, can explain the

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

majority of the decline and better fits the spatial pattern. 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 in 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. In other words, 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. To gain further intuition for the mechanism, we use a stylized version of the model to illustrate the role of home attachment in the joint dynamics of migration and population change.

The third, and briefest, step of the paper is a simulation exercise of a hypothetical island economy undergoing spatial population transition. The simulation provides a laboratory environment that allows us to trace the endogenous time profile (loosely, the impulse response) 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 population redistribution and elevated migration rates that have since moderated–with gross turnover rates lagging behind net migration.

This perspective is important for interpretation and evaluation, and we conclude with some normative implications from our analysis. 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 show 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.

Synthesizing the evidence from the descriptives, structural estimation, and illustrative model, we see a nuanced connection of population change to migration flows. Economists typically discuss migration in the spatial equilibrium paradigm–migration as an equilibrating force, with households leaving bad places for better ones. Migration in the real world, however, like many other labor market flows (see, e.g., Davis et al. (2012)), has gross flows far exceeding the resulting net changes (Davis et al. (2013)). Most concerns about declining migration predicate on the logical leap from the counts of people changing places to the health of regions and allocative efficiency of the economy.

Our paper connects gross and net flows in order to understand trends in each. The degree of place attachment largely determines the ratio of gross to net, so that in the long run, migration flows and population changes are jointly causal in an autocorrelated process. Our analysis makes explicit that past population changes affect this degree of place attachment, and the gross flows converge after population changes have resolved. Thus, one message from our paper is that amount of migration per se (how many are changing places) is not necessarily important, but regional population trends (where, on net, movers are locating) is indeed quite important for analyzing the health of regions, just as the canonical spatial equilibrium model would suggest. Thus, papers on rising barriers to population allocation, such as Gyourko et al. (2013), Ganong and Shoag (2017), Herkenhoff et al. (2018), and Hsieh and Moretti (2019), are relevant for understanding mobility generically even if not directly addressing a decline in gross migration.

Much of the concern about mobility decline arose because of its apparent secular nature– a vague friction affecting everyone at once, the specter of America's "lost mojo." Studying its cause is important so that we understand whether the migration decline represents an increase in barriers or a decrease in the economic incentives to relocating. Our findings are relevant on both counts. The cause is not quite "secular" in that it is spatially specific, and the cause is not so much an increase in cost as a rise in the prominence of a *specific form of location preference.* By understanding the mechanisms causing the decline, we will be better able to evaluate its welfare and macroeconomic implications and, if necessary, address it with more targeted policies.

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 work that emerged in the wake of 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]

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

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 (see Blanchard and Katz (1992), Bound and Holzer (2000), Carrington (1996), Zabel (2012), Hornbeck (2012)) and individuals (see Topel (1986) or Kennan and Walker (2011)), a natural concern arose that low mobility will result in spatial mismatch and lower aggregate productivity.[5]

Several studies have offered explanations but do not address the geography of the decline. 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. Some studies suggest that structural changes in the labor market have altered migration incentives. For example, Karahan and Rhee (2014) present a job search model in which the equilibrium with an aging workforce is for all workers to search locally. Bayoumi and Barkema (2019) argue that widening dispersion across metro areas in income and house prices has reduced the ability of workers to move, especially "uphill" to richer locations. Graham and Pinto (2021) connect low mobility by individuals to their low subjective well being. Johnson and Kleiner (2020) find occupational licensing is a barrier to interstate mobility, but its marginal effect is not large enough (nor is its prevalence wide enough) to account for the trend. None of these explanations contends with the spatial heterogeneity driving the migration decline, or the historical population movements underlying it, which are the focus of our paper.

There are a few studies of changes in population responsiveness to local shocks (e.g., Partridge et al. (2012), Hood (2013), Dao et al. (2017)), although this is a subtly distinct question from ours. We focus on long term trends in gross migration, not (necessarily) population changes resulting from net migration. Ultimately, though, our findings on gross migration flows being the result of past population changes mean the two issues are related. One notable hypothesis within this strand of literature is that locations are becoming more similar over time. For example, Kaplan and Schulhofer-Wohl (2017) argue (in a second component to their explanation) that returns to occupations have become more similar across space, causing migration to be less necessary. Accordingly, we allow for trends in labor market heterogeneity in our analysis. And again, these papers do not address the spatial pattern of the decline.

One strand of the migration literature explicitly ties geographic mobility to job mobility, since most long distance moves also involve (and are often motivated by) job changes (Molloy et al. (2014), Molloy et al. (2017), Ihrke (2014)), associating the trend in migration to the larger

migration we study.

[5] 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), Notowidigdo (2011), Autor et al. (2013), Dao et al. (2017)), and Yagan (2019)).

literature on declining labor market dynamism, which has found reductions in job mobility, flows in labor market status, firm growth rates, and entrepreneurship (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)). Our paper, motivated by the spatial heterogeneity of the decline, 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 at least our technique of leveraging local labor market heterogeneity to study national trends, is informative for that literature as well.[6]

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 (see Greenwood (1975) and Greenwood (1985) for reviews) to incorporating nontradable amenities (Graves and Linneman (1979)), idiosyncratic preferences, and move costs (Kennan and Walker (2011), Bayer et al. (2009), Moretti (2011), Coen-Pirani (2010), Lkhagvasuren (2012), or Diamond (2016)). This includes recent work relating migration to home attachments (Dahl and Sorenson (2010), Kennan and Walker (2011), Coate (2014), Zabek (2018), Koşar et al. (2021)) and social capital (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), and Büchel et al. (2020)),[7] which also have substantial literatures in the population sciences outside economics (Dawkins (2006), Michielin et al. (2008), Mulder and Malmberg (2014), Belot and Ermisch (2009), and Clark and Lisowski (2019)). Our study is the first to show how long run population dynamics affect local attachments, with implications for current migration rates.

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.[8] 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

[6]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.

[7]Using data on cellular phone call records, Büchel et al. (2020) find a strong role for social connections (both family and friends) in determining the probability of a move and in directing its destination once it occurs.

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

aggregated choice data,[9] and the first application via a nested logit model.[10] We discuss the model estimation in detail in Section 4 and Appendix C.

2 Migration Decline: Its Geography and History

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.

2.1.1 Data Sources and Uses

Our principal measures of migration come from two sources. The first is the migration flows tables from the U.S. Treasury's Internal Revenue Service (IRS) for 1991 to 2016 (IRS (2018)). 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. Because of its consistent reporting (annual since 1990), geographic detail (counties), and its large sample size (a near universe of taxpayers), the IRS data is the principal source for describing trends in mobility across different geographies.[11] However, it lacks demographic information that will be essential to a microfounded model of migration flows.

The second source of migration flow information is the American Community Survey (ACS, obtained from Ruggles et al. (2019)) for 2005 to 2017. The ACS microdata contains rich demographic information such as age, education, and importantly for our purposes, place of birth, so it will be the primary source for quantifying our model. The ACS reports the respondent's current and one-year ago Public Use Microdata Area (PUMA) of residence, from which we can elicit

[9]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. Theirs relies on time series variation, while ours relies on variation across location-age-education population cells.

[10]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 outside of conditional logit. We show that with some additional algebra, one can still reap the computational benefits of finite dependence.

[11]The IRS data underwent a change in method in the 2011-2012 tax year that resulted in noticeable, but as-yet unexplained, differences in the sample represented and possibly the migration concept, sowing doubt about the reliability of the data for these later years (DeWaard et al. (2020)). We present the data for the period 2012-2016 but rely only on the consistent sample of 1990-2011.

migration probability (move or not) and direction (origin-destination pairs). We also measure local labor market incomes using the ACS.

The 1990 and 2000 censuses also report a retrospective migration question, but at a five-year instead of one-year lookback. This is problematic for constructing time series, because the conversion rate of one year to five year change in residence location is complicated by return and repeat ("onward") moves, which are fairly frequent (Kennan and Walker (2011), DaVanzo (1983)); the five year rate is not simply five times one year. We make an effort to correct for the probability of a one year move event not being observed at a five year window. This effort, detailed in Appendix A, uses location histories in the longitudinal Panel Study of Income Dynamics (PSID, Institute for Social Research (2021)) to derive the return and onward migration rate by age, education, and birthplace status. Attaching this one-to-five year conversion to the ACS microdata, we present time series of implied five year rates. This provides corroborating evidence for our stylized facts, but because it is based on assumptions about the underlying conversion procedure, we do not rely on the five year rates for the model.

Finally, for historical measures of population, we use harmonized decennial census records from Manson et al. (2018), and for its distribution by age, education, and birthplace, we use microdata samples from 1880 to 2000 from Ruggles et al. (2019).

2.1.2 Geographic Concept for Defining Migration

A fundamental issue in a study of migration is the definition of location: At what distance does a move become a "migration?" Many previous studies use interstate migration, presumably because state is available in most datasets. But states are coarse measures of location and vary substantially by size, with larger states containing ample within-state variation.[12] In this paper, we use a metropolitan area in order to more closely correspond to a local labor market.[13] Locational heterogeneity is central to our analysis, and different cities within states vary in rates of mobility, population composition, labor market opportunities, and amenities.[14] However, when defining one's home location, we are confined to using the state definition, as the ACS and censuses ask for respondent's state of birth. We account for multi-state metropolitan areas and examine sensitivity of results to the alternative mappings of state to local labor market in the assignment of home.

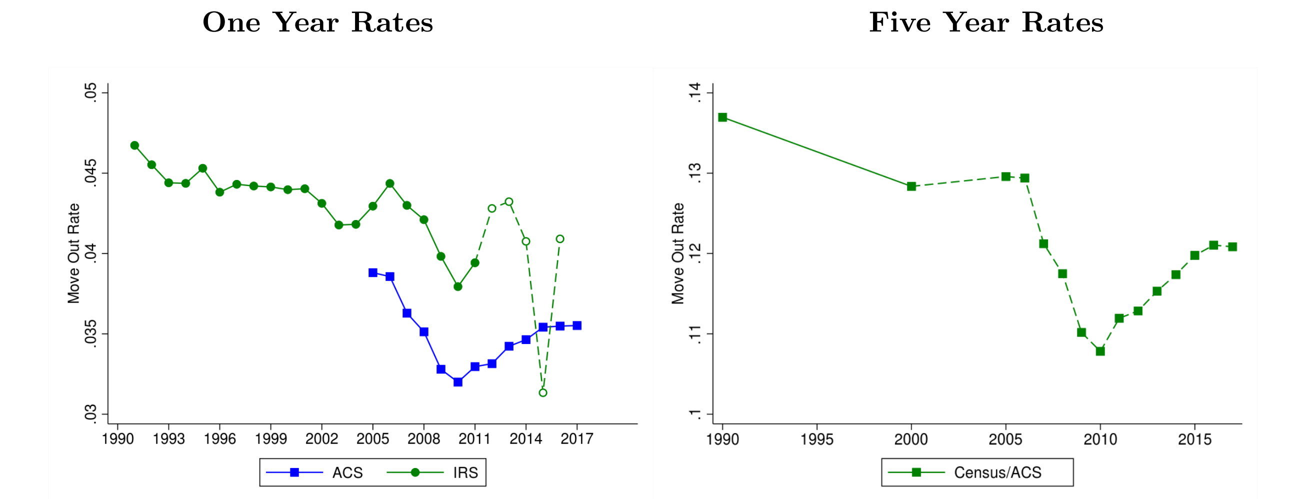
We use a new definition of metropolitan area that we term a local labor market (LLM), which is a close cousin to a commuting zone (CZ) in that both fully partition the U.S. and

[12]This is one major reason we do not rely on the CPS, which provides only state of residence. Another is that the CPS sample is too small to split into subgeographies crucial to our analysis.

[13]In the appendix, we show that the major patterns we document are robust to using state definitions.

[14]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.

Figure 1: Local Labor Market Migration Rates, 1991-2017



NOTES: The lefthand figure plots the one-year outmigration rate over time as defined by LLM and observed in the IRS and ACS microdata samples; Compare Figure B1 for the state-to-state definition. 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.)
The righthand figure plots the five-year outmigration rate over time as defined by LLM and observed in the Census in 1990 and 2000, and the imputed five-year rate using the ACS microdata sample. The imputation uses the five times the one year rate, discounted by the probability the new location is maintained within a five year window. The five year adjustment is made using a hazard model of return and onward migration, estimated on longitudinal PSID data with age, education, and home status as covariates. Compare Figure B1 for the alternative models of the five-year adjustment. (Source: ACS, Census, and PSID data.)

delineate urban areas. We derived LLMs from CZs and modified when necessary to standardize geographies over time or to correspond to current definitions of metropolitan areas (such as a Core-Based Statistical Area, or CBSA). We can map PUMAs and counties into LLMs for each year of data, 1880 to 2017.[15] We focus on the urbanized LLMs. The interested reader is referred to Appendix F for methodological details on data construction and Appendix G for LLM definitions and summary statistics; the uninterested reader can approximately think of CZs or CBSAs without serious threat to interpretation.

With geographic units defined, migration is then defined as exiting one LLM for another, irrespective of change of state. Moves within counties or PUMAs, or among counties or PUMAs of the same LLM, are non-migration events.

Figure 1 displays the national average migration rate over time in the IRS and ACS data samples according to this geographic concept. (For comparison, Appendix Figure B1 displays migration over time for state-level geographic concepts.) There is a clear downward trend (amid some cyclical fluctuations) best seen in the longer IRS time series. The drop from 1991 to the 2005-2011 average was 0.5 percentage point, or a 20 percent drop in the rate at which households change LLMs. The righthand figure plots the five-year outmigration rate using direct summary of the 1990 and 2000 census and implied five year rate using the ACS microdata sample.[16] The five year rate also shows a clear downward trend.

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

[16]See Figure B1 for alternative one-to-five year adjustment methods, all of which make little difference to the findings.

The rest of this section is devoted to unpacking the geography of the decline as a pathway to understanding its causes, starting with a prominent determinant of migration choice.

2.2 The Importance of Home Attachment in the Migration Decision

A critical fact to establish at the outset is that one's home location 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 here 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. Appendix Table B2 standardizes the comparisons using odds ratio regressions for each age and education group using the ACS microdata. A person in their birthplace is only about two-fifths as likely to move as someone away from home–and the probability is remarkably consistent across age and education groups, even when controlling for local labor market factors.

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 move costs (perhaps unobserved) 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. There is already a substantial literature showing the importance of place attachment (see Section 1.1),[17] but again, we want to check whether this pattern holds under our particular definitions.

One testable hypothesis is that if home is deemed especially valuable as a location attribute, then home will be chosen more frequently as a destination when living away from it. In the remainder of Table 1 we examine the rate of returning home. Column 5 reports the conditional choice probability to moving into one's birth state when living away from it (i.e., those who have already left their birthplaces) when migrating somewhere. Roughly one-fifth of moves are returns

[17] Koşar et al. (2021) is especially notable here. That paper finds strong preference for both family proximity and community connections, across income groups and even among those who self-identify as highly mobile.

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

	Move Out Rate (%)				Move Home Rate (%)	
Education/ Age	Total	US-born, In Birthplace	US-born, Not In Birthplace	Foreign-Born	(Conditional Choice Probability) Actual	Synthetic
	1	2	3	4	5	6
Noncollege						
20s	5.79	4.41	12.01	4.12	14.45	0.70
30s	3.52	2.62	6.89	2.66	11.86	0.68
40s	2.35	1.68	4.40	1.77	10.89	0.64
50s	1.91	1.29	3.41	1.53	9.35	0.57
College						
20s	10.02	7.33	15.02	9.89	12.47	1.10
30s	5.07	3.09	7.57	5.68	10.06	1.00
40s	2.56	1.52	3.77	2.70	7.79	0.91
50s	2.20	1.41	3.16	2.03	6.96	0.73

NOTES: The table reports mobility rates by birthplace. A respondent is in birthplace if residing in an LLM within his/her reported state of birth. Move home rates (column 5) are LLM move-in rates, conditional on a migration occurring, weighted by the probability the LLM is the respondent's birth LLM. Alternative definitions are provided in Table B1. The synthetic move home probability is a weighted average of conditional choice probabilities of moving into the home LLM for respondents not born in the LLM; for example, the probability of choosing Los Angeles by people not born in California. (Source: ACS data.)

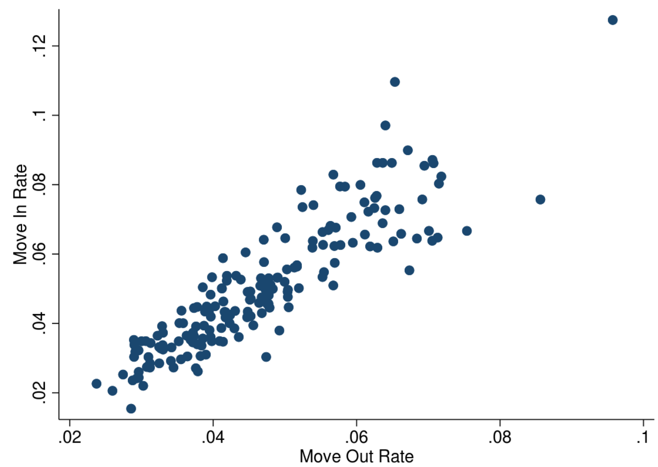
home. For comparison, in column 6 we report the average probability in the general population of migrants of selecting one's home–the expected probability if home were not special.[18] These are all orders of magnitude smaller; a destination is chosen 9 to 21 times more often when it is home than when it is not. Appendix Table B3 verifies this pattern, using the ACS microdata to estimate regressions of the odds ratio on choosing home relative to other possible destinations, conditional on migrating. Like the move-out rates, there is a remarkably robust pattern of home status in destination choice. Applying regression models with controls for other location attributes, a destination being home makes it roughly 6 to 19 times as likely to be chosen than if it were not home.

Table B1 reports on some robustness checks on the move home results. The move home probability is sensitive to the definition of home LLM (when the ACS actually reports birth state), and Table 1 uses a relatively conservative definition that assigns moves home less often than they could be. The appendix table shows that move home rates are as high as one-third under a more accommodative definition of home LLM. The appendix table also addresses a competing hypothesis, that migration is coincidentally towards home because migration networks are relatively local and any mover does not migrate far from her initial location–i.e., home is usually nearby. An odds ratio calculation controlling for origin location shows that moves home are still disproportionately likely when adjusting for move proximity.

These patterns provide initial evidence of the existence of home attachment via a utility premium, and we will provide additional evidence after developing further concepts of home attachment. This explanation is important to bear in mind as we review the spatial pattern of

[18] A random choice probability, one out of 275 LLMs, is about 0.36 percent. In column 6, we are adjusting the probability for the relative sizes of the LLMs. For example, because of New York's market size, there are mechanically more New Yorkers living about the country, and it is a popular destination for migrants from all birthplaces.

Figure 2: Local Labor Market In- and Out-Migration Rates (IRS 1991-1993)



NOTES: The figure plots average LLM in- versus out-mobility rates in the early IRS samples. Compare Figure B3 for additional samples. (Source: IRS data.)

the mobility decline and the evolution of population that preceded it.

2.3 Fast Locations Drive the Migration Decline

The next important fact–and one that is new to this literature–is that the national decline in migration is heterogeneous across space. In this subsection, we show that previously high turnover, "fast" locations are responsible for the national decline by converging towards their "slow" counterparts.

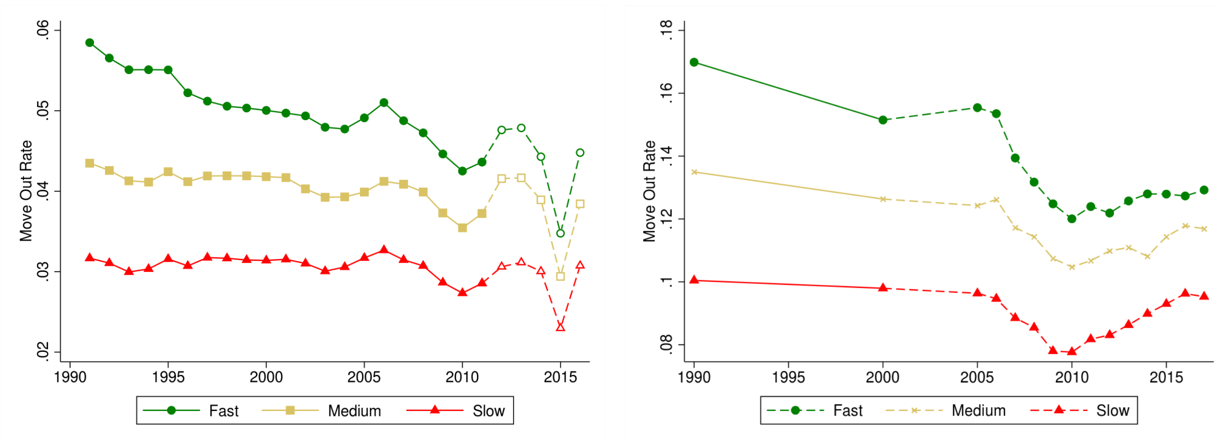
As an initial matter, we first clarify the language of "fast" vis-a-vis "slow" locations. Gross migration rates are correlated–that is, locations with high degrees of inflow also exhibit high degrees of outflow–and there is a large variance across places in the degree of turnover. This fact is not new (see seminal work by Ravenstein (1885), Sjaastad (1962), Miller (1973), to more recently, Coen-Pirani (2010) and Mangum (2016)), but for completeness, Figure 2 displays scatterplots of out-mobility to in-mobility rates in the 1991 IRS data. Appendix Figure B3 displays alternative observations of the IRS at various points in time and the ACS data. The strong positive relationship is evident, with correlation coefficients near 0.9, and 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. Our use of "fast, medium, and slow" is a categorization strictly on the basis of population turnover, which we will fix at 1991 rates.[19]

The fast/slow distinction is important for introducing our primary motivating fact: The decline in mobility is occurring predominantly among fast LLMs. Figure 3 shows the annual out-migration rates for LLMs split into terciles by their mobility rate–fast, medium, and slow.

[19]Despite convergence in rates, the relative rankings by turnover are mostly preserved even in the 2010s, as Appendix Figure B3 indicates.

Figure 3: Changes in Out-Migration Rate Over Time by Average Initial Turnover

One Year Rates **Five Year Rates**



NOTES: The lefthand plot shows the annual out-migration rates in the IRS data 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.)
The righthand plot shows the five-year out-migration rates by initial mobility tercile in the census, spliced with the implied five-year migration rates using ACS migration data adjusted with the repeat migration model from the PSID. The data become dashed series 2005 and following to reflect switch from actual to implied five-year rates. (Source: ACS, Census, and PSID data.)

The lefthand panel focuses on the one-year rate from the IRS. 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.[20] As an alternative view, Appendix Figure B4 shows scatter plots of early (1991) LLM migration rates to the change in rates 20 years later. There is a clear negative correlation across the distribution of initial mobility rates and among LLMs of various size.

The righthand panel shows the five year rates using our preferred one-to-five year conversion procedure for the ACS-era years (2005-2017) (See Figure B4 for alternative conversion procedures). Like the one year rates, five year rates drop substantially more in the initially faster locations. Though this relies on a correction procedure based on external data, the fact that the proportional changes are so similar in census and IRS data provides strong corroborating evidence that the fast/slow distinction is not an artifact of a particular dataset.

Appendix Figure B5 shows the pattern of changes when parsing the migration flows network of origins to destinations. 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. The fast tercile of locations make up one-third of population but by definition a greater share of out-migrants–

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

about 44 percent of them in the early 1990s. We project how many migrants there would be in 2010 if all cities moved at 1990 levels, and we then take the difference from actual as the number of "lost migrants." This accounting indicates that, with a larger share of migrants and a larger decline in rates, the fastest third of LLMs is responsible for 64 percent of the national decline.[21] In contrast, slow locations account for only 13 percent of the decline despite comprising 40 percent of population.[22]

The spatial pattern invites two natural questions: What is different between fast and slow locations? And then, what is changing?

2.4 Fast Locations: Centers of Continental Population Expansion

Which locations are fast? To begin, we show there is a strong regional component to population turnover. Table 2 lists the location of fast, medium, and slow LLMs by region of the U.S., where we have collected contiguous regions of LLMs to balance current population mass; each region comprises (as close as possible) one-fifth of population in 2010.[23] The regions roughly correspond to standard census regions, with one notable difference in the areas of the Rocky Mountains, Central Plains, and Southwest, which are denoted the "Frontier" to distinguish from the South Atlantic and Pacific West. See Appendix G for details.

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 none in the Midwest and only two in the Northeast. (These are Washington DC and Manchester, New Hampshire.)

The regional pattern hints at what is different about fast places: These are LLMs and regions with a recent history of high population growth. Figure 4 makes this plain by plotting decadal rates of population growth by region (in the lefthand plot, all places within region, regardless of metro status) and by LLM category (in the righthand plot, all metro areas, regardless of region). The 20th century was a time of continental expansion,[24] as population growth in the Frontier and especially the West drastically outpaced the national as a whole. These regions transformed from sparsely populated desert to urban growth engines. The Midwest and Northeast lagged 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.[25]

Consequently, fast LLMs 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

[21] 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.

[22] The single largest LLM, New York City straddles the boundary of the bottom third of mobility, with the associated lumpiness making the bottom tercile actually more than one-third of population.

[23] The regions are based on LLMs instead of state boundaries.

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

[25] Texas is contained in the Frontier region and is responsible for a large fraction of recent population growth.

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

	Northeast	Midwest	South	Frontier	West
Panel A: Count of LLMs					
Region Total	38	71	69	61	36
Within Region:					
Fast	2	0	16	16	27
Medium	8	22	14	15	2
Slow	18	29	10	4	0
Rural/Omitted	10	20	29	26	7
Panel B: Population Share					
Region Total	0.24	0.24	0.18	0.16	0.17
Within Region:					
Fast	0.09	0.00	0.24	0.31	0.87
Medium	0.22	0.19	0.32	0.32	0.06
Slow	0.63	0.55	0.14	0.04	0.00
Rural/Omitted	0.07	0.26	0.30	0.33	0.06

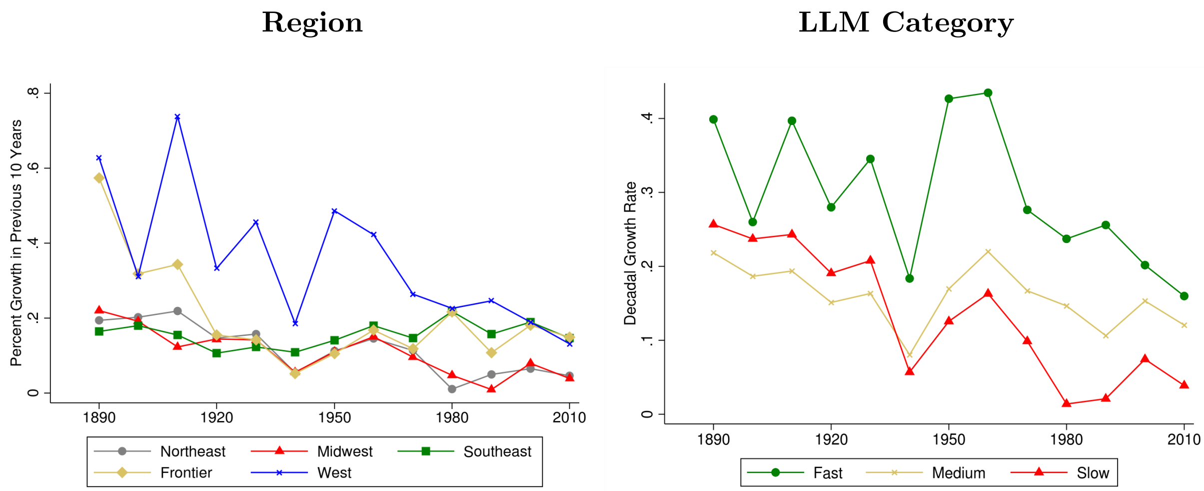
NOTES: See Figure G1 for regional definitions. The omitted category collects select excluded LLMs (military and college towns, as described in Appendix F) and rural areas unclassified as LLMs. (Source: IRS and census data.)

at the start of the 20th century to major urban areas at its end. It is particularly notable that population growth in fast cities peaked in the post-war period–about a generation before the migration decline began–and declined sharply thereafter (though they still are growing somewhat faster than medium and slow cities).

Expansion to these new regions was relatively sudden (by historical standards), as new technologies made developable areas that were once too remote or difficult to inhabit at a large scale. The reasons for the growth of these areas are varied, but one frequent theme is water technology. In the arid West, urban and rural areas developed simultaneously (in contrast to the slower rural-to-urban development of the wetter eastern U.S.), 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)). 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)).

Moreover, advances in transport technology and climate control made accessible newly 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 produc-

Figure 4: 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 G. (Source: Census county population estimates, harmonized over time by Manson et al. (2018), and IRS data for LLM categorization.)

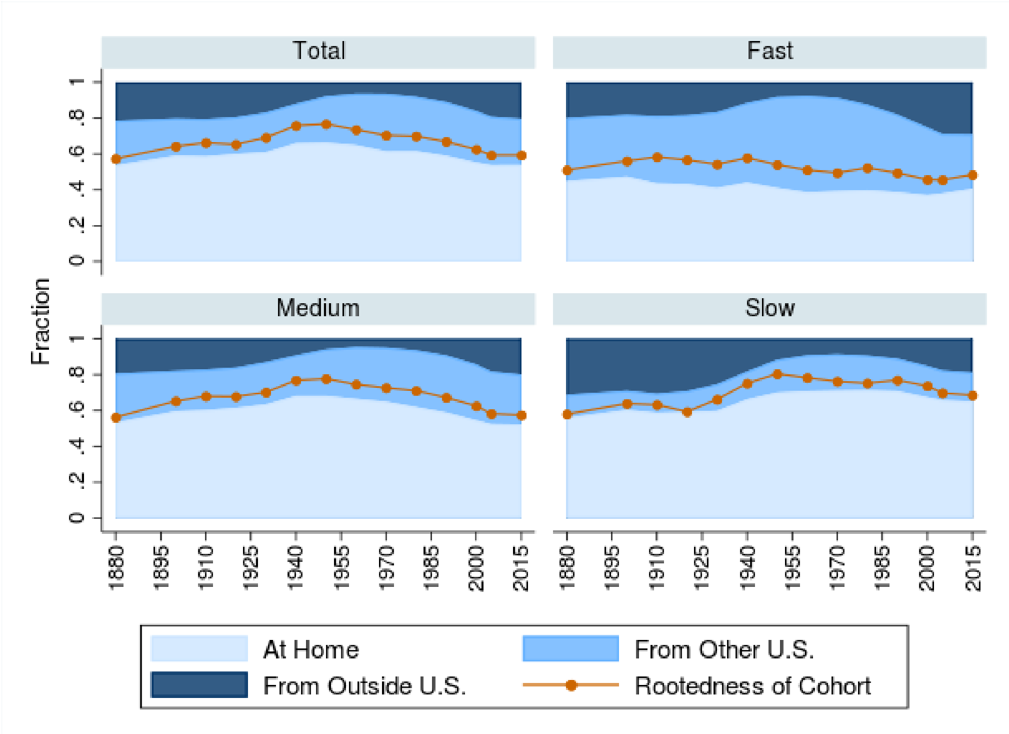
tion. Overall, the technological trends meant the 20th century was a particular phase of history characterized by an opening of the American continent to urbanization at a scale not previously experienced–and one that had converged by century's end.

2.5 The Evolution of At-Home Status Across Space

The history of population growth had implications for the birthplace source composition of these cities–in other words, where people living there had come from originally. Figure 5 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 denominator) to measure the share of U.S. natives who are in the LLM of their birth.

One obvious pattern in the national total in the upper left (though not the focus of this paper) 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 fraction 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 5 shows. That is, the large growth rates in Figure 4 were substantially made up of people moving from other regions and populating the West, the

Figure 5: Population Share From Birthplace Source, by LLM Mobility Category



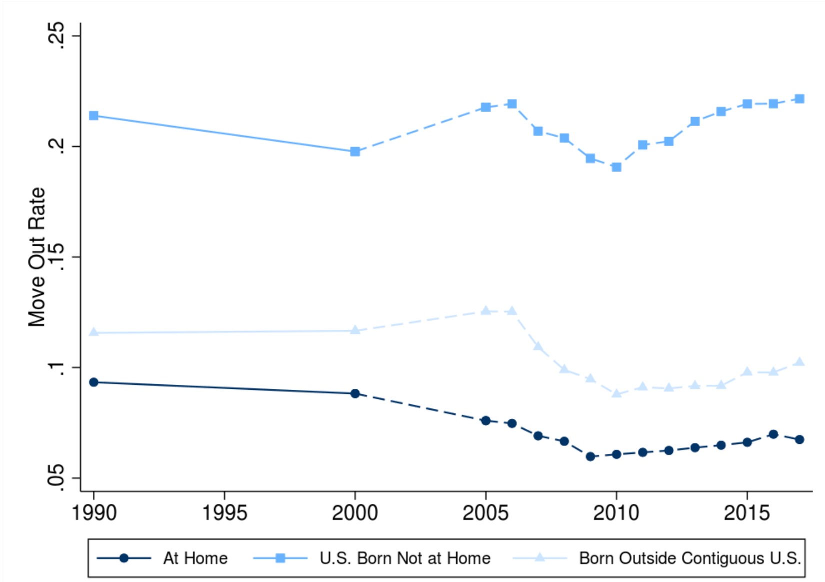
NOTES: At Home refers to living in an LLM in one's state of birth. Other U.S. refers to a birthplace in another U.S. state not covered by one's LLM. Outside U.S. are foreign born from any other country (or non-continental U.S. states and territories). The figure panels are summarized by LLM mobility 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.)

Frontier, and Florida. 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 speed LLMs, in contrast, have at-home shares at higher levels but flat or downward trends.

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 shallowly 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.

One test of this theory is to examine move rates by home status over time, which is unfortunately complicated by the change from a five year retrospective question in the census to a one year retrospective in the ACS. The best-available version of the test is to use the five year conversion rates mentioned earlier and discussed in detail in Section A. Figure 6 shows the migration propensity over time by birthplace status under this conversion. The only discernible

Figure 6: Migration Rate Over Time By Birthplace Status



NOTES: The figure plots implied five year migration rates by home status. The years using the ACS sample (2005 onward) convert to five year rates, as explained in Section A, using the conversion factors in Table A1, hazard model. The samples have been balanced to have a constant age/education composition over time. (Source: Census, ACS, and PSID data).

trend is a slight decrease in the move rates among the at-home group. (Figure B6 shows the pattern by LLM speed, which reveals that the migration rate for the at-home residents of fast locations has trended down, a feature we explain below.) Otherwise, the within-group consistency suggests the aggregate change results from a shift in population composition from the mobile, not-at-home category to the at-home category. Table B4 uses an odds ratio regression to standardize the comparison across the different migration horizons in the census and ACS. The results confirm the impression from Figure 6.

To formally test the hypothesis of home attachment requires a model of migration propensity, a measurement of home attachment, and an accounting of coincident factors. We develop the model in Section 3 after examining the LLM attributes to feed the model.

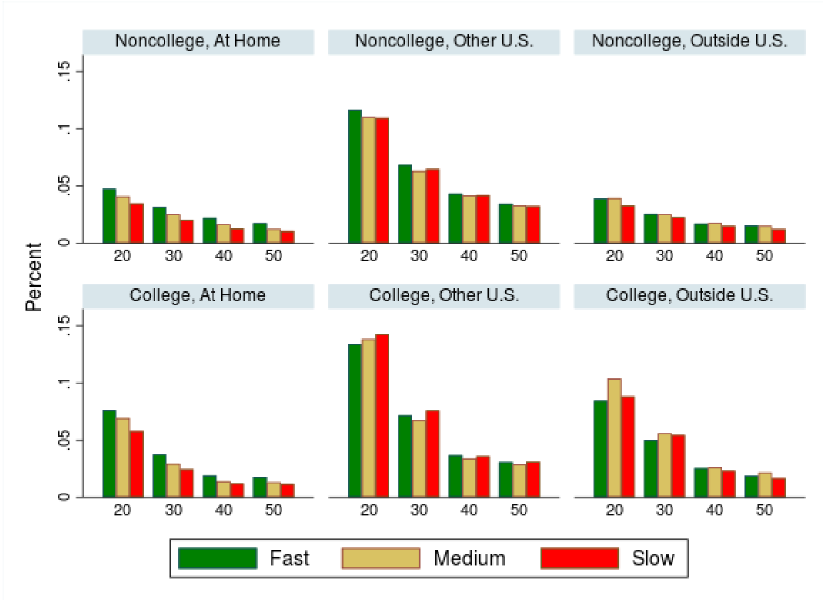
2.6 Local Labor Market Attributes

With population growth history as background, we now review some of the modern attributes of LLMs in our data.

2.6.1 The Degree of Local Home Attachment

Figure 5 indicated that fast locations have smaller shares of at-home 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 shows there is even more to the story. Figure 7 plots the migration propensity by birthplace status for each age and

Figure 7: 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; IRS data for LLM categorization.)

education group, as in Table 1, but split by origin LLM mobility tercile. Within the U.S. born, regional transplant group "Other U.S.," fast, medium, and slow cities send away their transplanted residents at similar rates, and patterns among the foreign born are mixed. However, among the "At Home" group, 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, but also send away their natives at higher rates. This suggests the *intensity* of home attachment varies across space.

Recollection of Figure 5 provides a clue to the source of home attachment intensity. 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 to be born to native Californian parents.

This feature motivates our measure of home attachment, which we will call "rootedness."[26] Appendix F.3 contains a detailed discussion of implementation. In brief, 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

[26]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.

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 20-30 years earlier, even when possibly living in other locations by the 2000s.

Rootedness by our definition is feasible to measure, 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 directly observed for any one individual. Thus, it might be an understandably poor descriptor of home attachment. But in fact, rootedness is highly predictive of migration propensity–and in a pattern consistent with Figure 7. We run odds ratio regressions of move rates using rootedness as an explanatory variable, along with other controls for local labor market attributes. The results are long, so we relegate them to Appendix Table B2. In summary, we find high rootedness to predict low migration rates among natives of an LLM–and only natives–in line with Figure 7. Natives leave more rooted places at much lower rates than less rooted places.[27] If rootedness were a place effect, we would expect the same correlation among nonnatives, but instead nonnatives leave rooted LLMs at similar or higher rates than less rooted places.

We conclude that rootedness will be a suitable proxy for measuring home attachment and its variance by geography and cohort. We discuss additional advantages of the rootedness measure for model estimation in more detail in Section 4.

2.6.2 Demographics and Income

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 first column is initial mobility (1991) and the second is change from 1991 to 2010. A negative correlation of a variable X with changes in migration means a higher value of X is associated with a greater decline in the mobility rate.

The top panel considers obvious population attributes. The first issue to note is that market size has nothing to do with the decline. Population size is on average negatively correlated with turnover rates, but there are cities of all sizes at each place in the distribution of changes (see also Appendix Figure B4). Population growth rate, however, is positively associated with mobility levels and declines, consistent with the historical analysis above.

[27] Return rates are more similar across places; that is, while natives are very likely to return home, natives of more rooted LLMs are no more likely to return than natives of less rooted places. We find this to be a place effect to some degree–in-migration is lower to more rooted places among all migrants, natives and otherwise.

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

		Migration Rate	
Variable	Sample/Statistic	Initial (1991)	Change (1991-2010)
Population	Log Size	-0.25	-0.04
	Growth	0.50	-0.36
Aged Under 40	Share	0.15	-0.35
	Change in Share	0.30	-0.18
College Educated	Share	0.03	-0.18
	Change in Share	-0.36	0.31
Mean Income	Noncollege	-0.22	0.09
	College	-0.07	-0.19
Income Growth	Noncollege	0.16	-0.09
	College	0.12	-0.16
Income Dispersion (SD)	Noncollege	0.28	-0.37
	College	0.20	-0.16

NOTES: The table reports correlation coefficients of local labor market attributes by row to migration levels and changes, by column. Declines in migration are negative changes, so a negative correlation indicates a larger value of the variable is associated with a larger magnitude decline. (Source: IRS migration data; ACS and 1990 census data for attributes except 2010 and 1990 county data for population.)

Next we document population composition characteristics that are predictive of individual migration propensity, age and education. Faster cities are only slightly younger, on average, and have seen a slight increase in the population aged under 40 years. But these younger cities have, if anything, seen larger declines, which suggests aging alone cannot explain the spatial pattern to the decline. Faster locations have no more college graduates, and higher levels of education are weakly associated with larger declines on average. Growth in the college educated share has tilted towards slightly less mobile places, and consequently, places with smaller 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. The correlations, however, suggest that composition by age and education will do little to explain the spatial pattern in the mobility decline.

A model of local labor market migration presupposes the importance of labor opportunities in determining location choice. Accordingly, the lower panel compares local mobility rates with local income distributions by education category. The associations with mean incomes are mixed and relatively weak. LLM mobility rates are negatively correlated with mean income for the noncollege educated, but the correlation with change in mobility is just above zero. Mean incomes for the college educated are weakly correlated with declines. Growth in average income is weakly correlated with mobility rates and decline. Income level of course will be a component of the model, but given these patterns, we do not expect these to be driving the aggregate trends.

On the other hand, there are stronger associations of mobility to income dispersion, especially among the noncollege educated. Places with higher income dispersion exhibit higher turnover and greater decline in migration. We take this point seriously for two reasons. First, 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. We note that migration is more likely among individuals at the higher or lower points of the income distribution compared with the middle, as shown in Appendix Table B5. Perhaps more importantly, the association of LLM income dispersion is strongest among lower income individuals, as shown in Appendix Figure B7.

The second reason is that 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, leading to a trend in the effect of dispersion on migration rates. 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.

3 Model

The descriptive evidence paints a clear picture that the spatial pattern of the migration decline is consistent with converging population growth rates coupled with regular home preferences. However, many things in the economy were changing over this horizon, so we rely on a model to quantitatively test the contribution of home attachment vis-a-vis other margins.[28] 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 alternatives. We write such a model in the tradition of Kennan and Walker (2011), broadly, with adjustments for our focus on spatial heterogeneities and bearing in mind that the model will be applied to cross-sectional data.[29]

3.1 Environment

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

[28]Moreover, data limitations preclude a direct test of home attachment. The best we can do is convert one year migration rates to five year and compare across censuses. The exercises in which we do this (see Figures 3 and 6), are consistent with the explanation of rising home attachments, but the conversion of move rate frequency comes with caveats.

[29]Longitudinal datasets contain personal information and moving histories that would facilitate estimation of a structural model, but are not large or rich enough to include the spatial heterogeneity that is also important to our analysis. Specifically, the measure of the home premium relies on variation in move rates by home status within age, education, and location groups–the latter being especially constraining in smaller datasets such as the PSID.

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 $u_{j=H}(r)$.[30] The rootedness of a location affects the utility offered to its natives but not any other workers born elsewhere living in the location.

Workers face moving costs to relocating from their origin at the start of the period, whether home or not. We make a careful distinction between moving costs and horizontal preferences for particular locations. While the home premium inclines workers to prefer their own birthplace *ceteris paribus*, moving costs introduce frictions. These ideas are often conflated in the literature, either as shorthand or because of data limitations.[31] 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 in counterfactual simulations, where we alter the distribution of home preferences but keep adjustment costs fixed by assumption (we never assume moving costs increase).

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, workers decide whether to stay in their current location. Second, and contingent on deciding to migrate, they choose a location. In either case, a last stage in which income shocks are drawn is also modeled to allow for expected income to have an impact on the location decision.

3.2 Migration: Choice Across Locations

We begin in the middle stage, 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 symmetrically dependent on the pair of locations, representing a generic notion of "distance" (so that $mc_{j,o} = mc_{o,j}$).

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

[30] 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, or aid in job search (see Krolikowski et al. (2020)).

[31] For example, Moretti (2011) introduces a model with a whole distribution of location-specific preferences to study their impact on labor mobility in response to local market shocks. Bayer et al. (2009), Diamond (2016), and Bryan and Morten (2019) put a measure of home 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.

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 workers live past today, 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.[32] With these shocks, the probability of choosing destination j conditional on moving from current location o is given by

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

3.3 Choice of Migrating or Staying

The upper nest is the binary stay (s) or move (m) decision. The value of staying is consuming flow utility in the current location (origin o) and being faced with the same decision next period,

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

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

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

The respective probabilities are

$$Pr(stay) = \sigma_s = \frac{exp[V_{s|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}} \tag{5a}$$

$$Pr(move) = \sigma_m = \frac{exp[V_{m|o}]^{\frac{1}{\delta}}}{exp[V_{s|o}]^{\frac{1}{\delta}} + exp[V_{m|o}]^{\frac{1}{\delta}}}, \tag{5b}$$

[32] Under our data constraints, we abstract from evolving state variables like the future path of incomes, although in principle we could introduce them if data allowed. As we describe in the estimation section, this is not a serious threat to bias in our results because of the reduced-form way in which we capture continuation values.

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. Our emphasis is on the potential effects of local income distributions on gross mobility flows, not the effects of individual income changes on idiosyncratic migration decisions. We do not observe the joint dynamics of income and location in our data, but we include this component in the model to allow for differences in income distributions between locations or over time to affect gross migration rates. We devise a formulation that could generate higher mobility in and out of higher dispersion areas (as Table 3 suggests). In particular, this specification allows workers with especially low or high income draws to migrate at rates higher than those of a mean worker (see Table B5 and Figure B7). Moreover, the formulation incorporates an "information friction," a la Kaplan and Schulhofer-Wohl (2017), which can evolve over time and therefore may affect the trend in migration.

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 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

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

A contact is the better outcome for the worker in the sense that she can reject the poorer offer. Combining these yields the expected utility from beginning the period at income state n

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

a γ-weighted average of income possibilities. 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_n 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 (as in 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 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 $\pi^{nonlocal}$.[33] Using data on income dynamics, we can identify differences in income transitions between those who move to new locations and those who do not.

The income search component serves several purposes in modeling the heterogeneity in spatial dynamics of migration. 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. Moreover, the effects of such income growth 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_n 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. 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 will depend on the parameters and variance across locations.

3.5 Worker Types

Up to this point, for ease of notation we have suppressed worker types, but the relative value of functions (4) and (6) may depend on characteristics of the worker. Birthplace and rootedness

[33] An alternative would be to use different offer arrival parameter γ, but this parameter is already very abstract and difficult to discipline with data, whereas income changes can be calibrated with longitudinal data.

are endowed characteristics. We also treat age and education as immutable characteristics because we observe one location choice event per individual. Location is the endogenous state variable–we observe workers in one state, which they can alter by choosing a new location.

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

$$v_{j,\tau}(o,n) = \underbrace{\mu_{j,\tau} + \omega_{nj}(n,o,\tau)}_{\nu_{j,\tau}} + u_{j=H|m}(r_{j,\tau}) - mc_{o,j,\tau} + \beta EV_{\tau}(j,n'|n). \tag{10}$$

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.[34]

One concern with mixing all these types together in a joint estimation 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 heterogeneous continuation values in a flexible way.

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 (such as Bayer et al. (2009) or Diamond (2016)), which emerged from the demand estimation literature in the theme of Berry et al. (1995), 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) or Bishop (2008), the estimation uses aggregate choice probabilities (market shares) instead of, e.g., maximum likelihood estimation on panel microdata. 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.[35] 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

[34] In other words, the data may be coming from a period of transition–say, net in-migration to a location because of relatively positive amenities. Controlling for this nonparametrically will help identify the parameters of interest, home preference and moving costs, in the usual way–by accounting for omitted variable bias.

[35] 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.

the sources of identification in our model is variation by geography and cohort instead of over time.

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 6.

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 τ. 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. 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 $u_{j=h}(r) = \alpha_\tau I(j = h)$ or a linear function of roots, $u_{j=h}(r) = \alpha_\tau R_j I(j = h)$, the former being a specification check and the latter our preferred specification.[36] 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 there are several advantages. First, it offers a source of 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

[36] 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. Home preference in Las Vegas appears weak but is not inverted.

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. However, reduced form results have showed it to be predictive, and it is readily constructible for a wide swath of geography. Hence, we employ it as a reasonable proxy for the deeper concept of "attachment."

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 kilometers between LLM 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.[37] We further allow distance cost to vary if the destination is one's home location via an interaction of distance terms with the home indicator function. The move cost function is

$$mc = \underbrace{\sum_{\tau} I(\tau) mc_{\tau}}_{types} + \underbrace{\sum_{d}^{D} mc_d d_{o,j}}_{distance} + \underbrace{\sum_{d}^{D} mc_{dh} d_{o,j=h}}_{distance\,X\,home} .$$

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. We briefly describe the method here, and details and derivations are reported in Appendix C.

The model is dynamic discrete choice with multiple types of agents and a large number of locations (and therefore birthplaces), making for a very large state space. Fortunately, there is a very tractable way to estimate the model. As a memory-less discrete choice specification, this

[37]We have experimented with specifications in which LLMs have entry/exit costs, as if there were a political border, to further account for heterogeneity in turnover rates in excess of what is captured by distance, age/education composition, home status. The results are quite similar to what we present here, but are much harder to interpret because the magnitudes depend on the normalization location.

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 and the model's parameters.

Specifically, our model has the property of finite dependence (Arcidiacono and Miller (2011)).[38] 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, $s=1$, 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. From this, we can yield an estimating equation in terms of current and expected future choice probabilities.

Besides providing major computational savings, the CCP method is especially convenient because our cross-sectional data contains a single location choice event per person, making it difficult to estimate a fully dynamic microeconometric model. Yet, ignoring the forward-looking component would risk misspecification bias, with the continuation values essentially being omitted variables. Expressing the future value in the form of expected choice probabilities performs something like a control function in a standard regression model by accounting for this omission.

The derivation of the estimating equations is straightforward but tedious, so we relegate the details to Appendix C. The outcome choice probabilities are twofold: (1) the probability of moving ("move/stay"), and (2) conditional on moving, the probability of choosing a particular location. For the move/stay probability, the mappings between (6) and (5) provide the following expression relating choice probabilities to the utility function:

$$ln\frac{\sigma_{so}}{\sigma_{mo}} - ln\frac{\sigma_{sz}}{\sigma_{mz}} + \frac{\beta}{\delta} ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}} - \frac{(\beta-1)\lambda}{\delta} ln\frac{\sigma_{ko|m}}{\sigma_{kz|m}} = \frac{1}{\delta}\left[\left(u(x_o)-u(x_z)\right)\theta_u + (\beta-1)\left(mc_{j,o}-mc_{j,z}\right)\theta_{mc}\right] \tag{11}$$

where σ_{so}, σ_{mo} are, respectively, the probabilities of staying in or moving away from a location o, and $\sigma_{ko|m}$ is the choice probability of moving to location k from o. Here, we have used the properties of finite dependence to iteratively substitute out the future value functions with conditional choice probabilities. The normalizations of using the odds ratio relative to an arbitrary place z is necessary because the discrete choice model will not identify the scale of utility, only the difference in choosing one option versus another. The major advantage of this expression is that it has data on the left and (functions of) parameters on the right, yielding a fairly standard-looking estimating equation.

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

A similar derivation of the destination choice probability of location j versus k, using (4) and (2), with normalization to location z and iterative substitution of choice probabilities for future value functions, yields the following expression:

$$ln\frac{\sigma_{j|o}}{\sigma_{k|o}}+\beta ln\frac{\sigma'_{m|j}}{\sigma'_{m|k}}+\frac{\beta}{\delta}\frac{1}{\lambda}ln\frac{\sigma_{z|j,m}}{\sigma_{z|k,m}}=\frac{1}{\lambda}\big[\big(u(x_j)-u(x_k)\big)\theta_u+\big(mc_{jo}-mc_{ko}+\frac{\beta}{\delta}(mc_{zj}-mc_{zk})\big)\theta_{mc}\big] \quad (12)$$

which, as before, has choice probabilities on the left and functions of parameters on the right. These equations can be stacked to form the vector equation

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} \frac{1}{\delta} & 0 \\ 0 & \frac{1}{\lambda} \end{bmatrix}}_{\Delta} \underbrace{\begin{bmatrix} u(x_j)-u(x_z) & (\beta-1)(mc_{jo}-mc_{jz}) \\ u(x_{j,o})-u(x_{k,o}) & mc_{jo}-mc_{ko}+\frac{\beta}{\delta}(mc_{zj}-mc_{zk}) \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \theta_u \\ \theta_{mc} \end{bmatrix}}_{\theta}. \quad (13)$$

Vector Y contains the observed choice probabilities (Y_1 the lefthand side of 11 and Y_2 the lefthand side of 12). Matrix Δ is composed of the scaling parameters between nests. 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.). Finally, θ is the vector of parameters to be recovered. In short, Y is data, X is model structure, and Δ and θ are parameters. There are $J-2$ choice probability observations (the origin o and the normalizing location z are trivially excluded) for each location and worker type, and these equations can all be stacked together as a large set of moment conditions. With choice probabilities on the lefthand side, and LLM attributes and parameters on the righthand side, estimation proceeds much like a standard regression: The data matrix X is inverted on the choice probabilities to recover the estimand, θ.[39]

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 origin and type ($ln\frac{\sigma_{so}}{\sigma_{mo}}$), such as workers at home migrating out at lower rates, will imply higher utility ($u(x_o)$); if this varies with rootedness, the differences will load onto the rootedness parameters in $u(x)$. Likewise, higher propensity to choose home as a destination ($\frac{\sigma_{i=H|o}}{\sigma_{j|o}}$) when away from will imply greater utility in home i than alternative destination j. A higher destination choice probability ($\frac{\sigma_{i|o}}{\sigma_{j|o}}$), all else equal, will imply closer proximity and load onto the move cost function. Inclusion of the additional conditional choice probability terms in Y adjusts for differences in the continuation value terms that might otherwise bias estimation of the utility terms. For example, if move out rates from home are lower than when not at home, the $ln\frac{\sigma'_{m|o}}{\sigma'_{m|z}}$ term will be low for $o = home$, and the lefthand side is reduced, accounting for the fact that the odds ratio $ln\frac{\sigma_{so}}{\sigma_{mo}}$ includes expectation

[39] It is convenient but not strictly necessary that utility is linear in parameters. The separability of the preference shocks is required, however, to derive the mapping of value functions to CCPs.

of continued utility premia in the future. Finally, note that an attractive feature of the regression-like specification is that utility function matrix 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 (13) identifies the parameters of interest off of differences in choice probabilities and their relationship to difference between locations. However, scale parameters are not identified here and must be calibrated elsewhere, as we describe next. The set of scaling parameters includes β, λ, δ, and type-specific intercepts of the move cost function.

First, we set β to 0.95 a priori to conform to 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. As noted above, the odds ratios for home/not-home are substantially different between the move/stay decision and the destination choice decision. In order to best match this asymmetry, we use a ratio of parameters on a home indicator in each side of the flow equation. Details are provided in Appendix D.3.2.

Third, a separate estimator for the move cost intercepts can be derived using substitutions similar to the derivation of (13), as detailed in Appendix C.[40] The estimator is:

$$\begin{aligned}
ln\sigma_s - ln\sigma_m =& \frac{1}{\delta}\big(V_o - \lambda ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big) \\
=& \frac{1}{\delta}\big(V_o - V_j - mc_{j|o} - \lambda ln\sigma_{j|o}\big) \\
ln\sigma_s - ln\sigma_m + \lambda ln\sigma_{j|o} =& \frac{1}{\delta}\big(V_o - V_j - \underbrace{mc_\tau}_{cost\,intercept} + \sum_d^D mc_d d_{o,j} + \sum_d^D mc_{dh} d_{o,j=h}\big).
\end{aligned} \tag{14}$$

This equation is identified by average move rates in levels, not relative move rates across different origins. The difference between (13) and (14) 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. However, we can control generically for these unobserved origin and destination factors via fixed effects, treating V_o, V_j as ancillary parameters. Then, entering the distance terms of the move cost function recovered in (13), (14)

[40] This step in estimation average frictions bears the closest resemblance to Artuc et al. (2010). Monras (2018), on the other hand, 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.

will yield an estimate of the move cost intercept for each age/education type, mc_τ in the data from residual average move rates.[41]

4.1.4 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 that we could not discipline with data, 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.

$$\sigma_{m|o,\tau} = f_1(r_{o,\tau}, \mu^W_{j,\tau}, \sigma^W_{j,\tau})$$

$$\sigma_{jo,\tau} = f_2(r_{j,\tau}, \mu^W_{j,\tau}, \sigma^W_{j,\tau})$$

We flexibly estimate the functions f_1, f_2 in order to approximate the choice probabilities of moving or not ($\sigma_{m|j,\tau}$) and choosing a particular destination conditional on moving from j, ($\sigma_{kj,\tau}$) as functions of locations' rootedness and income distributions ($W_{j,\tau} \sim N(\mu^W_\tau, \sigma^W_\tau)$. Note that each variable is indexed by worker type τ. Details are available in Appendix D.4. We emphasize 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 in counterfactual situations.

4.1.5 Auxiliary Model: Income Dynamics

We use an 11-point discrete approximation to the income distribution of each location for each education group. The points are centered at the mean and consist of five one-half standard deviations above and below. Population sizes for each point are assigned by the observed distribution of workers of type τ within each bucket.

Workers transition between steps in the income distribution. We assume these follow a normal distribution and follow Tauchen (1986) in discretizing it. We calibrate the parameters governing income dynamics for incumbent residents compared with migrants, 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.7.

[41] Similar to (13), 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 ($V_o - V_j$).

The information cost parameter, γ, is more difficult to pin down. Intuitively, the parameter represents something simple: higher γ means more opportunity for voluntary income draws and lower γ means less. However, we have no information on the arrival of voluntary income draws over time in which to discipline this parameter. In practice, we will set the baseline to $\gamma = 0.5$ and move it incrementally over time to get a sense of its marginal effect on migration propensity (it turns out to be small).

4.2 Forming the Moment Conditions

Lastly, we need to form the choice probabilities that make up the lefthand side of the estimating equations. Because we need one-year migration probabilities by worker type, home location, and origin, we use the ACS microdata. With age group by education by birthplace (including foreign) and by origin, there are $4 \times 2 \times (J + 1) \times J$ 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.[42] At $J = 70$, there are 39,760 types and 2,738,200 choice probabilities to estimate from the data.

Even with the relatively large cities, however, we still encounter some small cell problems. While the move-stay decision rate is virtually always well-measured, some destination choices (i.e., choice probability conditional on a move) are not observed. We use a smoothing procedure, detailed in Appendix D.6, to impute some of these missing cells. The basic idea is to first estimate the move probability for each cell, but then combine data from similar cells to estimate conditional destination choice probabilities. We then impute the destination choices as the product of the marginal and conditional probabilities: probability of a move (from finer cells) times the probability of a mover choosing particular destination (from coarser cells).[43]

[42] 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.

[43] 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.

5 Results: Estimates and Model Fit

In this section we discuss the coefficient estimates. Our main purpose here is to evaluate the preferred model specification versus alternatives.

5.1 Parameter Estimates and Comparison of Specifications

Table 4 reports the structural parameter estimates and model fit for several specifications of the utility and move cost functions. We report standard errors for the home premium, and suppress the others for brevity. We begin with a discussion of parameter estimates in Panel A and then compare specifications using descriptions of fit in Panel B. The table runs through several specifications of utility and move costs.[44] We report move cost intercepts for the US born types, while those for foreign born are included in the model but suppressed from the table.

Columns 1 and 2 use the most basic model with only cost intercepts by type and a home premium as an indicator variable (column 1) or as a linear function of rootedness (column 2). In either form, the home premium is significant statistically and economically. 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 its income distribution exceeded other locations' by roughly 4 standard deviations–clearly, a massive premium. This comparison is not literally a statement about the elasticity of migration to local income (as realized labor market outcomes are always idiosyncratic); it is a stark signal of the enormity of home preference for the average person.[45]

The average size of the home premium is the same whether we use the simple indicator variable or the roots-dependent model.[46] The difference between the models is evident in their ability to fit the spatial features of the data–specifically, predicting the degree of heterogeneity in move out rates among the at-home population. We evaluate fit in several ways as reported in Panel B, starting with B1 reporting the correlation in predicted to actual move rates across LLMs. Comparing columns 1 and 2 at the level of LLM composite average, either specification can predict reasonably well the heterogeneity between cities, with correlation of predicted and actual at about 0.61. However, among the at-home population, the prediction of the roots varying model is much higher, with a predicted-to-actual correlation of 0.41. Similarly, from B2, the roots-varying model can better generate the variance in move rates across LLMs–and especially

[44] All specifications include for the residual location an "amenity" and mover-entry indicator, by age/education type. This is done to account for differences in both the propensity to stay and the propensity to enter that arise mechanically out of its nebulous geographic definition: Not only is it an order of magnitude larger than even the biggest city, it is geographically proximate to any LLM.

[45] Similarly, in a survey design featuring elicited move probabilities, Koşar et al. (2021) found the average person willing to pay 43 percent of their income to live near family and 36 percent of income to stay in their home community, and that such preferences were strong even among those who self-identified as highly mobile.

[46] The roots coefficient may appear bigger, but in the model it is multiplied by the rootedness share–usually about 0.74 (see Table B4).

among the at-home population. Without a roots-varying home preference, the indicator model cannot meaningfully predict the spatial heterogeneity in move rates among the at-home. In summary, only the roots-varying preference model can generate the qualitative pattern displayed in Figure 7.

Columns 3 and 4 add a distance function in the specification to allow for bilateral move costs. The bilateral costs are also interacted with home status to allow the distance sensitivities into home to differ from a generic move.[47] The move cost parameters indicate that farther migration events, measured in kilometers, are more costly, with additional "discounts" taken for within-region, within-state, and neighboring LLM moves. Note the mean squared error is lower when the distance function is included. The distance cost function shifts up the distribution of move costs (intercepts by type increase, though the ranking is preserved[48]) by allowing the more frequent, close-by move to have substantially lower cost than an arbitrary distance pair; otherwise the mean move cost is driven lower to match the frequency of these more common moves. Moreover, the home premium estimates are slightly larger when accounting for bilateral move frictions, although, interestingly, moves to home show far less sensitivity to distance, with the coefficients reversing sign in the to-home interaction. The same patterns of cross-LLM correlation and move rate variance obtain between the roots-varying and fixed home preference between 3 and 4 as in the comparison between 1 and 2.

Columns 5 and 6 add location by age by education utility fixed effects to account for differences in the average attractiveness of a location to different worker types. This is more flexible both mathematically (an additional 544 terms in the function) and economically (allowing for heterogeneity in city quality). However, it does virtually nothing to affect the home utility or move cost estimates. That is, gross migration flows yield information about home premia and distance sensitivity, even after controlling for the perceived location quality gaps across the demographic groups which would drive net flows to better locations. Panel B2 indicates the models with location-type fixed effects are the best at generating the dispersion in move rates across locations (comparing columns 5 and 6 to 3 and 4 or 1 and 2), especially among the not-at-home, for whom the differences in distribution of income are the only feature left to generate spatially heterogeneous move rates. That is, the model of specifications 1 to 4 generates the correlation across LLMs but not the magnitude of variance. Allowing for unobservable local amenities in 5 and 6 captures the rest, but at some cost in terms of disrupting the fit in overall move out rate (Panel B1).[49] And as before, the roots-varying preference for home fits better.

So far, we have reported on our preferred nested formulation. As a contrast, columns 7 and

[47]The results are quite similar if the home interaction is omitted.

[48]The distance function is mean zero by construction, but this result shows that accounting for geographic distance does not materially change the move rate gaps between groups. If, for instance, older people moved substantially shorter distances than the young, the differences in intercepts might have been affected.

[49]This is largely because the destination choice ("move-to") rates are identifying the fixed effects/location quality terms more so than differences in move out rates.

Table 4: Parameter Estimates and Model Statistics

Fixed Effects Model		LLM=0 Nested		Age X Edu. X (LLM=0) Nested		Age X Edu. X LLM Nested		Age X Edu. X LLM Non-nested	
		1	2	3	4	5	6	7	8
				Panel A: Parameter Estimates					
A1. Home Preference									
Home	Noncollege	3.98		4.14		4.14		3.25	
		(0.007)		(0.006)		(0.004)		(0.006)	
	College	3.14		3.30		3.30		2.38	
		(0.007)		(0.006)		(0.004)		(0.006)	
Roots	Noncollege		5.40		5.44		5.59		4.20
			(0.01)		(0.009)		(0.006)		(0.008)
	College		4.24		4.28		4.43		3.03
			(0.01)		(0.009)		(0.006)		(0.008)
A2. Move Cost Intercepts (US Born)									
Noncollege	20s	-14.24	-14.24	-14.43	-14.43	-14.38	-14.38	-8.52	-8.47
	30s	-16.47	-16.47	-16.66	-16.66	-16.61	-16.61	-9.06	-9.01
	40s	-18.20	-18.20	-18.40	-18.40	-18.35	-18.35	-9.55	-9.50
	50s	-18.97	-18.97	-19.16	-19.16	-19.11	-19.11	-9.76	-9.71
College	20s	-12.10	-12.10	-12.30	-12.30	-12.25	-12.25	-7.68	-7.63
	30s	-14.92	-14.92	-15.11	-15.11	-15.06	-15.06	-8.50	-8.45
	40s	-17.58	-17.58	-17.77	-17.77	-17.72	-17.72	-9.24	-9.19
	50s	-18.22	-18.22	-18.41	-18.41	-18.36	-18.36	-9.49	-9.44
Distance Function									
Main	Neighbors			0.56	0.56	0.37	0.37	0.39	0.39
	Same State			1.57	1.57	1.37	1.37	1.41	1.41
	Same Region			1.11	1.11	0.78	0.77	0.85	0.85
	Log (km)			-0.21	-0.21	-0.43	-0.44	-0.49	-0.49
Distance X To Home	Neighbors			-0.36	0.38	-0.36	0.25	-0.34	-0.34
	Same State			-1.92	-1.63	-1.92	-1.66	-1.85	-1.85
	Same Region			-0.77	-0.05	-0.77	-0.10	-0.68	-0.68
	Log (km)			0.27	0.80	0.27	0.76	0.39	0.39
A3. Specification Details									
Calibrated $\frac{\delta}{\lambda}$	Noncollege	3.80	3.80	3.80	3.80	3.80	3.80	1.00	1.00
	College	3.59	3.59	3.59	3.59	3.59	3.59	1.00	1.00
No. Parameters		34	34	38	38	582	582	582	582
MSE		1.01	1.01	0.77	0.77	0.39	0.39	0.61	0.61
				Panel B: Model Fit:					
B1. Correlation of LLM Move Rates									
All		0.61	0.62	0.61	0.62	0.40	0.48	0.25	0.21
At Home		0.09	0.41	0.04	0.44	-0.04	0.42	-0.13	0.18
Not Home (US)		0.19	0.12	0.24	0.20	0.20	0.18	-0.09	-0.10
Foreign Born		0.07	0.06	0.14	-0.13	0.33	0.34	0.23	0.23
B2. S.D. of LLM Move Rates									
All	Data: 1.16	0.61	0.79	0.63	0.81	0.94	1.01	10.79	10.82
At Home	Data: 1.02	0.08	0.61	0.10	0.62	0.64	0.72	0.73	1.11
Not Home	Data: 2.36	0.25	0.26	0.36	0.38	2.04	2.06	39.89	40.22
Foreign Born	Data: 2.76	0.19	0.19	0.25	0.25	1.48	1.48	40.65	40.72
B3. Choice Probability Projection									
Stayers	(Ideal=1)	0.99	0.99	1.00	0.99	1.00	1.00	1.64	1.55
Movers	(Ideal=1)	0.82	0.78	0.91	0.88	0.75	0.81	0.04	0.04
B4. Odds Ratios, Home/Not Home									
Move-Stay	Data: 0.42	0.35	0.37	0.33	0.37	0.33	0.36	0.01	0.01
Move Home	Data: 10.2	19.83	19.66	18.09	17.75	25.09	23.92	8.74	8.49

NOTES: The table reports coefficient estimates from the structural model estimated according to (13) in Panel A and model fit statistics in Panel B. Standard errors are suppressed except for the home premium estimates, which report GMM errors. The model includes, but the table does not report, move cost intercepts for the foreign born and location fixed effects for the residual LLM location. (Source: Authors' calculations.)

8 use the standard conditional logit specification, non-nested, with only one elasticity parameter. Being based on the same choice probabilities, this model still finds substantial preference for home, and move costs increasing in distance and age and falling in education. (That the magnitude of the coefficients is different is not meaningful, since they are on different scales.) However, forcing the same elasticity between levels of choice has severe consequences in terms of fit. Panel B3 reports the projection value of choice probabilities from the model (an ideal fit, $E(\hat{\sigma}) = \sigma$, is one). The non-nested model overpredicts the value of staying home (B3 "stayers"), and in particular greatly overstates the value of the home premium so that the odds ratio in the move-stay decision is far off the mark (B4).[50] The non-nested model is also oversensitive to differences in migration inflows between places, attributing to them great economic significance, resulting in a poor fit of the destination choice projection (B3 "movers"), and overstates the variance in the not-at-home move rates (B2). Consequently, the non-nested model does a substantially worse job of fitting the spatial heterogeneity in move rates across LLMs, as judged by either the correlation of predicted to actual (B1) or variance across LLMs (B2).

We conclude that the nested model with roots-varying preferences is the preferred choice for predicting move rates by type, origin location, and home location. In what follows, we will use the model without location fixed effects (column 4 instead of 6) because of its better performance in predicting move/stay rates, our main focus in simulations, at some loss of fit regarding specific destination choice, although we have found results with respect to time trends to be very similar either way.

5.2 Model Fit in Baseline Period Simulation

The estimation recovered parameters via the estimating equations and not an explicit targeting of simulated values. We next check that the model is able to replicate the main features of the data. We have already given some sense of the fit of the model in comparing difference specifications in Table 4, but we will now show how well the preferred model produces the patterns described in Section 2.

Table 5 reports the average moving rates by age and education category and at-home status for the data and our baseline simulation. 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 from Section 2, without assigning 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

[50] The non-nested model by construction cannot simultaneously generate the odds ratio gaps between the at-home and not-at-home in the move/stay decision and the to-home and not-to-home in the destination choice decision, as seen in Panel B4. We could reweight the observations to target the move/stay decision explicitly, although this would then bias the move-home rates, which, by undervaluing the option of returning home, negatively affects the prediction of choice values for those living away from home.

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

Type		Move Rate				Conditional
		Total	At Home	Not at Home	Foreign Born	Move Home Rate
Panel A: Data						
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
Panel B: Model						
Noncollege	20s	4.32	2.98	9.52	4.10	32.60
	30s	2.73	1.75	5.55	2.63	28.72
	40s	1.88	1.14	3.64	1.83	27.95
	50s	1.63	0.97	2.95	1.64	26.81
College	20s	8.27	4.80	12.79	10.97	23.42
	30s	4.51	2.33	6.31	6.06	21.49
	40s	2.24	1.16	3.13	2.82	20.79
	50s	1.86	1.00	2.63	2.18	20.40

NOTES: The table reports actual and model-predicted choice probabilities for a closed system of 70 locations (69 largest LLMs plus one residual locale). All figures are in percentages (%). (Source: ACS data and model-generated data.)

model also predicts the inordinate share of moving that is a return to home, actually overstating its magnitude. This is due to the way we define moves home in the data. In estimation, to avoid "stacking the deck" in favor of finding move home preference, we use the more conservative definition of home inflows that likely undercounts the fraction. For comparison, see Table B1, in which the accommodative definition compares favorably with our prediction.

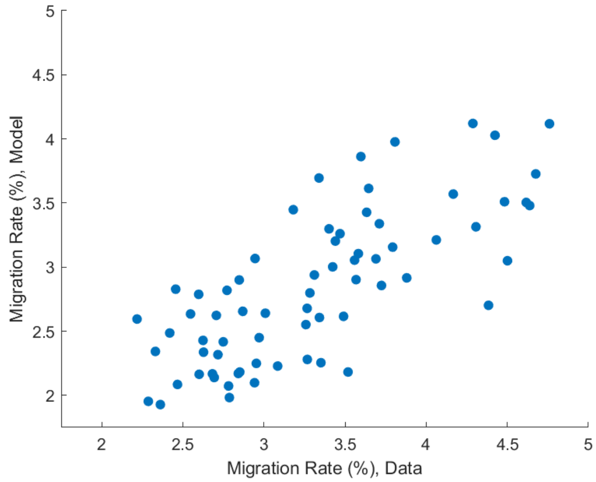
Figure 8 plots the actual and predicted out-migration rates for each metro area in our analysis. The model is able to match (albeit with a slight understatement of the variance across LLMs) the spatial heterogeneity through a combination of differences in demographic composition, income distributions, shares of the at-home and foreign-born, and degree of rootedness.

6 Simulations

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

A given simulation of the model predicts choice probabilities for every state (origin, income step) and agent type (age, education, birthplace) in the economy, amounting to millions of

Figure 8: 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 an LLM in the estimation/simulation sample. (Source: ACS data and model-generated data.)

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).[51] We do not mean to contend that the primitive parameters in utility could not have possibly changed; merely, we want to see how far the model goes in explaining the decline when constrained in this way.

The alternative scenarios could affect the composition of the "mover at-risk" population and/or the migration incentives presented to the populations. Compositional shifts include changes in the distribution of population by age, education, origin location, step in the income distribution, and the fraction at home. Incentive changes include the rootedness of home locations by birth cohort, parameters of the income distributions, and the income search cost. We first combine all factors to show how the model predicts migration changes over time and then decompose the factors one by one to elicit the contribution of each.

6.1 Simulating Mobility Over Time

We begin with the aggregate change to migration (the combined simulation with all changes taking place) in Table 6. Panel A reports the migration rates in the model, and Panel B takes the differences over time compared to the data. Recall that the model was estimated on cross-

[51]In Appendix E.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

Panel A: Simulated Migration Rates						
	Move Rates (%)					Change,
LLM Group	1980	1990	2000	2005-2011	2012-2017	1980 - 2005:17
Combined	3.23	3.17	2.97	2.87	2.91	0.35
Fast	4.21	3.96	3.55	3.35	3.32	0.89
Medium	3.24	3.20	3.02	2.92	2.96	0.30
Slow	2.56	2.49	2.40	2.35	2.42	0.18

Panel B: Log Difference in Migration Rates					
Source	Model 1 Yr.	Model 1 Yr.	IRS 1 Yr.	Model 1 Yr.	Census/ACS 5 Yr.
LLM Group	1980 - 2005:17	1990 - 2005:11	1991:93 - 2005:11	1990 - 2005:17	1990 - 2005:17
Combined	0.113	0.099	0.109	0.095	0.168
Fast	0.238	0.170	0.202	0.178	0.252
Medium	0.097	0.090	0.093	0.082	0.179
Slow	0.074	0.059	0.033	0.048	0.121

NOTES: The table presents the simulated migration rates as indicated. Panel A reports rates in percentage points, and Panel B reports differences in log points. The model was estimated on pooled ACS data, 2005-2017. The census/ACS 5 year estimate uses the one-to-five year adjustment procedure described in Sections 2.1 and A. (Source: Model simulations, IRS, ACS, and census data.)

sectional data and without targeting any dynamics in the migration rate. Nevertheless, by accounting for major factors affecting migration propensity, the model is able to generate the decline. Simulated migration rates fall by 0.35 percentage point from 1980 to the 2005-2017 average. 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 0.18 percentage point, about a seven percent change for slow LLMs from 1980 to 2017.

The simulated change compares favorably with the data. The lowest panel reports the log difference in rates compared with the 2005-2011 average in the model and the IRS data, the source from which we have the most reliable long time series. 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, and we ignore the post-2012 period in which the IRS data became suspect. The model's 9.9 log point drop from 1990 to 2005-2011 compares closely with the IRS drop of 10.8 log points for LLMs in the estimation sample from the early 1990s to 2005-2011. The model also reflects the proportionally larger drop in migration from fast cities comparable to the data–a decline of 18 log points for fast cities, 8 for medium, and 4.7 for slow, compared with 20, 9, and 3 log point declines, respectively, in the IRS data. The model-simulated decline also compares favorably to the drop in implied five year rates, although the change in five year rates is somewhat larger, likely because the one-to-five-year imputation overstates the drop in migration during the Great Recession (see Figure B1).[52]

[52]The conversion of one to five year rates does not account for cyclical changes in the propensity of migration reversal.

6.2 Decomposing the Sources of Decline

The simulated decline is an aggregated result that we can then unpack via *ceteris paribus* breakdowns of the sources of change. Table 7 reports the changes generated by simulations of the model when altering one feature at a time. That is, for each simulation, we hold fixed at their estimation period (2005-2017) values all features except for those denoted by their row. The simulations pose questions such as, what if the age distribution (for example) evolved since 1980, but everything else remained as in the 2010s?

Because we are using a nonlinear model with interactions among many features, this exercise is not literally a "decomposition," so the values need not sum to the total effect (neither within nor between locations). Yet, these breakdown simulations give an indication of the importance of the factor when it alone is present in the model. Some simulations combine one or more changes. To summarize, we show only the difference between 1980 and the estimation period. For reference, the table reports the combined change in the top row.

We begin in the first panel with the most obvious candidates, shifts in population composition. These represent changes to relative group weights–the change in migration rates if the population looked like 1980 but the migration incentives looked like the 2010s. Nationally, the composition changes would predict one-quarter of the total decline, or 0.087 percentage point of the 0.35 total change in migration. The simulation does predict slightly larger declines in fast cities but population changes alone are insufficient to account for the spatial pattern of the decline, comprising only 23 percent of the fast city decline in comparison to 67 percent of slow cities. So while population shifts are relevant as part of the aggregate decline, the compositional changes are too similar across space to explain the much of variance in the decline across cities.

We can further delve into the compositional factors, changing one dimension of the population at a time. Population aging alone actually could explain a greater share of the decline if it had not coincided with countervailing trends. Changes in age group composition alone could produce a 0.43 percentage point decline nationally–actually greater than the predicted total. This is substantially offset by the increasing share of population with college degrees, which would cut the decline directly by about one third (a 0.15 percentage point *increase*) when entering alone and in half when in combination with age changes (the age and education combined subtotal is just 0.20 percentage point). The trend in aging is further offset by continued population growth in faster locations, which would predict a 0.11 percentage point increase in national average rates. Interestingly, the trend towards faster cities is true even within our pre-designated fast/medium/slow categories.[53]

The last compositional change the table reports is the change in the foreign born population.

[53]The model does indeed generate differences in migration probability across LLMs, even within population type cells. The migration probability for an agent of a given type is still a function of the incentives (income opportunities, move costs, and continuation values) offered by his place of residence relative to other locations.

This feature causes opposite effects in different types of cities, which ultimately nets to nearly zero in the national average change. As Figure 5 shows, the share of foreign born living in the U.S. has been rising, and the foreign born tend to have migration probabilities between the at-home and away-from-home domestic population. In slow cities, the foreign born have increased in share relative to the at-home population, driving migration rates up on the margin. In fast cities, the foreign born have been growing relative to the away-from-home, driving migration rates down. The last line of the panel reports the total for all compositional factors including the foreign born. The combination tilts the spatial distribution by slightly more but reduces the aggregate decline.

In summary, we find a moderately important role for population changes in explaining the aggregate decline, but insufficient magnitude and spatial variance to account for the story of Section 2.

The second panel reports the changes induced by shifts in home attachment. These make up the single largest change to migration rates and account for the majority (about 60 percent) of the national decline. Moreover, home attachment affects fast cities to a far greater degree and thus explains much of the spatial heterogeneity in the decline. There are two components to home attachment, the share at home (which operates like a compositional feature) and rootedness of natives (which operates like a change in migration incentives). Share at home is the first-order effect of converging population growth rates, and rootedness is the second-order, lagged effect that occurs when population growth slowed in the previous generation. The share at home dominates, although the increase in rootedness in fast cities is also a nontrivial contributor. As before, the effect on fast cities has outsized influence on the national total because these origins contribute a larger proportion of migrants.

Lastly in the second panel is the combined effect of home attachments and population composition. Here, the interactions between factors show themselves. The combined effects of rising at home share, rising rootedness, and aging generate the magnitude of the national decline and its spatial heterogeneity across cities. The impact of a single factor depressing migration is amplified by concomitant depressors. Aging is relatively more important in places with rising home attachments and vice versa.

Finally, we simulate changes to gross migration resulting from shifts in income opportunities, as Table 3 suggested a relationship between LLM migration rates and income distributions (especially variance). The income simulations pertain to changes in migration incentives, except for one margin, the size of the population in each income bucket.

The distribution of income across cities ("LLM Distributions") has changed, with mean income growth and increases in dispersion weakly associated with declines in outmigration rates (see Table 3). Changing LLM income distributions produce modest declines in migration, primarily out of fast cities.[54] The "Population Distributions" simulation addresses the fact that the distribution of income has also changed within cities, with inequality rising in most places, while migration is most prevalent at the upper and lower ends of the individual income distribution (see Table B5). However, widening inequality within cities does little to change total migration rates, even in combination with the change in LLM Distributions themselves. Lastly, changes in information availability (increases in γ) alter the incentives to migrate as higher income draws become relatively more attainable, and the value of higher draws varies across locations (and is on average higher in fast locations). This effect is directionally consistent with a mobility decline, but in this model has limited quantitative impact on the migration rates.

Taken together, we find a nontrivial but relatively limited role for changes in income affecting migration rates, although the changes indeed bias towards fast locations. Though idiosyncratic income may well be very important for individual households' migration decisions, at the level of LLM average, the trends in income can explain just a fraction of the decline in migration. Instead, it is home attachments in fast cities, in concert with national demographic changes, that are driving migration rates downward.

To summarize the simulations over time, Figure 9 plots the composite simulation with all changes to the population and environment, and then three subtotals–population composition, home attachment, and income changes. The plots show that the effects of aging were most prominent from 1990 to the early 2000s, as the Baby Boom cohorts aged into the low mobility stages of life. The effects of income distribution changes, relatively small in the aggregate, are most present from 1980 to 2000. Home attachment imposes a long and steady downward pressure on the migration rate throughout the simulation period.[55]

6.3 Contributions to the Aggregate Migration Rate

With an understanding for the reasons for the decline, we close this section by showing how trends in at-home status affect the aggregate migration rate. Table 8 reports the move rates over time by home status: at home, not at home, and foreign born. For brevity we sum over all age and education groups but fix the demographic composition at 2010 levels to focus on changes in home status and local labor market incentives, not trends in aging or educational attainment. The not-at-home group is further split into two choice categories, moves home (returns) and moves onward to new locations, and hence their sum will equal the total move rate for the category. We split these to check whether there are any trends in returns home (which would increase,

[54] Changing income opportunities also account for some notable outlier cases, such as San Francisco, Austin, and Raleigh-Durham (see Table E2), where migration rates have not moved in the direction or to the extent that demographics and home attachment would predict.

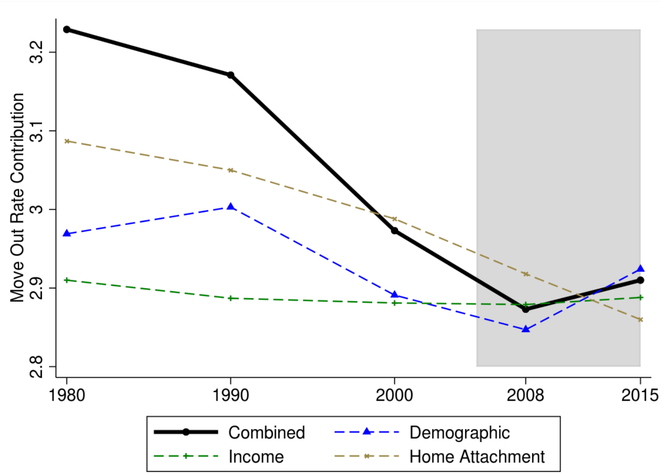
[55] Appendix Figure E1 indicates that the hump-shaped migration trajectory since 1950 was the product of mid century demographic transition and weakening home attachments–effects which are past their peak in the time period of Figure 9.

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

		Decline in Migration Rate (percentage points, $x > 0 \rightarrow$ decline)			
Simulation	LLMs:	Fast	Medium	Slow	All
Total		0.892	0.300	0.182	0.346
Population Composition					
· Combined		0.201	0.186	0.122	0.087
· Age		0.456	0.472	0.365	0.428
· Education		-0.147	-0.145	-0.167	-0.154
· Age + Ed		0.247	0.243	0.123	0.202
· Location		-0.036	-0.041	-0.036	-0.112
· Foreign Born		0.071	0.014	-0.108	-0.008
Population Composition (incl. foreign)		0.263	0.197	0.009	0.053
Home Attachment					
· Combined		0.358	0.071	0.142	0.205
· Share at Home		0.303	0.088	0.127	0.183
· Roots		0.102	-0.009	0.018	0.043
Pop. Composition + Home Attachment		0.837	0.286	0.188	0.329
Income Distributions					
· Combined		0.055	0.015	0.007	0.027
· LLM Distributions		0.053	0.013	0.006	0.026
· Population Distributions		0.0004	0.0003	0.0002	0.0003
· LLM+ Pop Distribution		0.053	0.013	0.006	0.026
· Information		0.002	0.002	0.002	0.002
Pop. Composition + Income		0.253	0.202	0.122	0.107

NOTES: The table reports the change in migration rates from 1980 to the estimation period (2005-2017) when conducting simulations in which only a subset of model factors is changing. The simulated factor(s) is(are) indicated by the row title. All figures are in percentage points (e.g., "0.5" corresponds to a one-half percentage point change). (Source: Authors' calculations.)

Figure 9: 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. (Source: Authors' calculations.)

Table 8: Simulations: Migration Rates by Home Status, No Age Effects

	1980	1990	2000	2005-2011	2012-2017
Move Rate	3.11	3.05	2.99	2.91	2.87
Rate by Home Status					
At Home	1.77	1.77	1.77	1.80	1.79
Not Home	5.07	4.99	4.96	4.91	4.92
Not Home: To Home	1.20	1.18	1.15	1.13	1.14
Not Home: Onward	3.88	3.82	3.81	3.78	3.78
Foreign	3.07	3.06	3.06	3.04	3.04
Contribution by Move Type					
At Home	0.78	0.80	0.83	0.88	0.91
Not Home	1.56	1.49	1.39	1.27	1.20
Not Home: To Home	0.37	0.35	0.32	0.29	0.28
Not Home: Onward	1.19	1.14	1.07	0.98	0.92
Foreign	0.77	0.77	0.77	0.76	0.76

NOTES: The table reports model simulated migration rates for a constant demographic sample. All figures are in percentage points (e.g., "0.5" corresponds to a one-half percentage point change). The destination-based subcategories under "Not Home" add to the total for that home status group. (Source: Authors' calculations.)

ceteris paribus, under stronger home attachments) vis-a-vis more general changes among the "mover class" of away-from-home households. For more details, see Appendix Table E1, which reports the move rates from LLMs of fast, medium, and slow origins and population shares by at home category.

The table shows level differences in move rates by at-home status, but within statuses, migration rates remained remarkably stable since 1980.[56] That is, the trend in the aggregate is mostly the result of a greater share of people represented in the at-home category, not trends in choice probabilities conditional on category. The small trends in rates that do exist are among the not-at-home group, showing their greater susceptibility to migration incentives in absence of home attachments. On net, both returns home and moves elsewhere have declined (proportional declines of 4.5 and 2.7 percent, respectively). However, the quantitative driver of the aggregate trend is the drop in the share of people in the loosely-attached status group: fewer onward opportunities arising and fewer returns home to be made.

To make this point explicit, the second panel of the table shows contribution to total migration by at home status, which is essentially the population share in the category times the move rate. The contribution from the at-home group has risen because of rising population share in that category. This is overwhelmed by the drop in contributions from the not-at-home group.

[56]The totals obscure some spatial differences: Appendix Table E1 shows that in fast LLM origins, migration rates declined among all groups, while in medium and slow LLMs, the trends in migration rates by status vary across groups.

7 An Illustrative Model of Spatial Demographics

The quantitative model shows that home attachment is an important factor in the mobility decline observed in the U.S. One disadvantage of the quantitative model, however, is that it is limited to the change in home attachments observed in the data. We now use a stylized version of the model to illustrate how home attachment evolves over a long horizon and experiment with the impact it has on regional responsiveness.

7.1 Home Preference, Regional Shocks, and Transitional Dynamics

The following thought experiment illustrates how a preference for home mediates a shock to the equilibrium distribution of population in an economy. In the simulation that follows, we generate an economy, qualitatively like the postwar U.S., that experiences shocks to location attractiveness and trace the evolution of population, home attachment, and migration.

We use an overlapping generations (OLG) framework over a system of J distinct locations, or "islands." Each agent lives for A periods, and the economy is simulated for $T >> A$ periods. Agents are given an opportunity to move at a cost. For simplicity, we fix move costs (the college educated 20-somethings point estimate from our quantitative model) and ignore their changing with age or distance (although these can be incorporated in a straightforward way). We start with an initially equal distribution of population ($\frac{1}{J}$) 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 either there is 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").[57]

We simulate each version of the model until it reaches a steady state, where population sizes of the locations are constant, all cohorts have the same strength of preference for home, 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.[58]

[57] 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.

[58] 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

Plot I of Figure 10 shows the relative 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 $1 - (A - B)$. The No Home Preference scenario arrives at the new steady state most quickly, as agents move only on the basis of relative amenities. The Fixed Home Preference scenario arrives at the steady state somewhat more slowly, as home attachments make home-inclined agents reluctant to relocate even for better amenities. The Roots-Varying scenario arrives the most slowly, since the endogenous home preferences are initially very strong. Plot II shows the net reallocation of population that produces the gaps in plot 1. Note that reallocation is not immediate in any scenario, since agents face move cost in all scenarios, but that reallocation is faster in the scenario without home preference, since effectively all agents agree on the quality of each location.

Plot III shows the share of residents at home, which affects the overall migration rate and population adjustment. As more people leave the low amenity locations for higher 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.[59] In other words, 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.

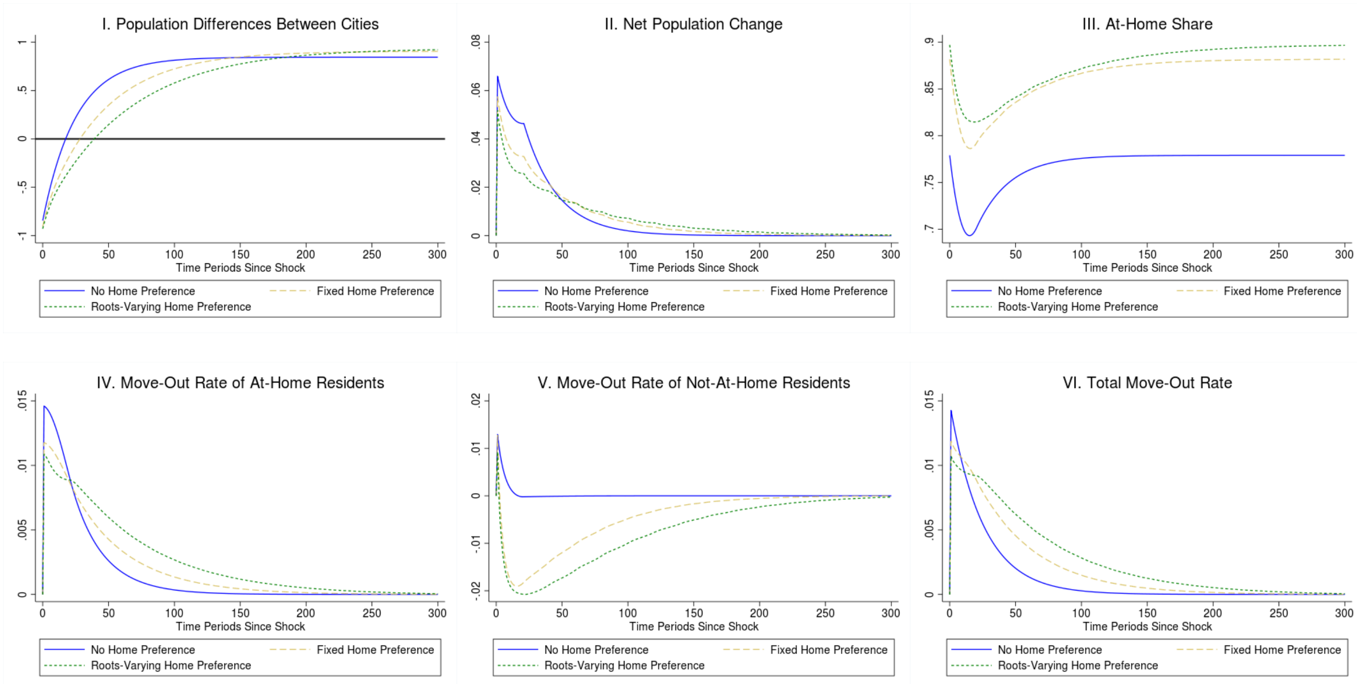
The lower panel plots the migration rates for at-home and not-at-home agents and the total. The No Home Preference scenario has the largest increase and quickest convergence in move rates for the at-home agents in plot IV–and for the not-at-home in plot V, since they have identical incentives in that scenario. Relocation incentives are blunted for the Fixed Home and Roots-Varying scenarios, so the move rate for the at-home increases by less in plot IV. The behavior of the not-at-home in plot V under the home preference scenarios 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 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 VI. 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.

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. This particular implementation is not important for generating the qualitative patterns we describe, as the objective is to see how shocks to the steady state population distribution are mediated over the transition path.

[59]There is also a level difference in at home population between scenarios, because of course, in the No Home Preference scenario, the at home residency is not particularly desirable but only coincidentally obtains because of move costs.

Figure 10: 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 7.)

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, the thought experiment exhibits 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.

7.2 Home Preference and Steady State Migration

The previous simulation exercise suggests–consistent with the concerns of many economists, demographers, and policy makers–that home preference is a threat to desirable population reallocation. We next use this stylized model to show that home preference affects the total migration rate in steady state and its elasticity to shocks–and the subtle ways these differences come about.

Using a model similar to that of Section 7.2, we compare steady state migration rates under different degrees of home attachment.[60] 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 but for the home preference, which varies in strength between scenarios:

[60] Ours is a stylized exercise to illustrate qualitatively 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 9: 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 percentage points.

either low or high. The home preferences are calibrated to be one standard deviation below and above, respectively, the mean rootedness times the average preference parameter from Table 4, and move costs are again the point estimate for college educated 20-somethings.

Table 9 reports the economy's moving rate and the share living at home under higher and lower preferences for home. 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.

The differences have 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, in total more people relocate immediately compared with the high. This total, however, is the net of two effects. First, by construction, the at-home are individually more sensitive to the shock in the low preference scenario. Second, the steady state in the low preference scenario has more agents in the more susceptible not-at-home status. In fact, the not-home group responds by *more* in the high home preference scenario, as the shocks enable their desires to return to their home locations.

8 Evaluation: Has America Lost Its Mojo?

In summary, we find total population adjustment is the product of micro incentives–an individual preference for home–through the macroeconomic history of population movements and their resultant effects on the distribution of population. Understanding the reasons for the

[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.

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. In this section, we offer some concluding thoughts to frame the academic and policy discussion going forward.

The literature studying the migration decline has looked for a structural mechanism in the economy, one pervasive across demographic strata, that has caused a downward trend in mobility. 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 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. Regardless, we hope to refocus the broader discussion on migration to the core issues for welfare and policy.

8.1 The Migration Trend Is Not Occurring in Distressed Areas

One common misconception is a connection of the mobility decline to the relatively poor performance of certain local labor markets. We have 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 11, 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.[63] We continue to use the definition of a move as an exit from an LLM, even if staying within the region we define.

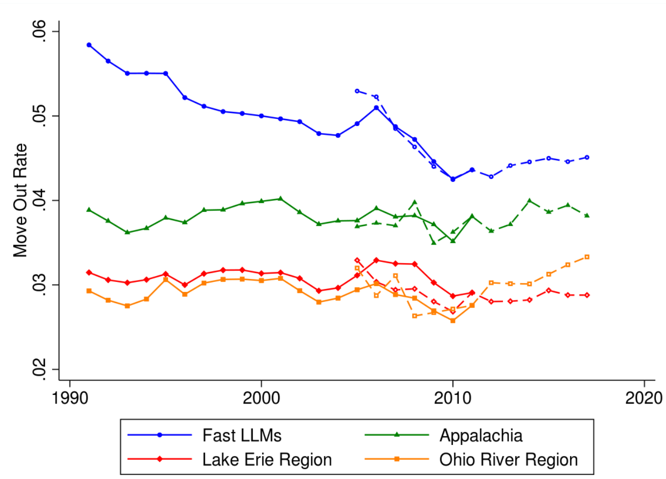
For each of these distressed regions, in contrast to fast LLMs, migration rates are low but not trending. Hence, for whatever problems these local labor markets may have–and perhaps more people "should," in some sense, be exiting these regions–they have been slow for some time and are not particularly afflicted by the mobility decline. A national trend cannot be blamed for further diminishing the poor circumstances of workers in the areas where no trend is exhibited.

8.2 Population Convergence Preceded the Migration Decline

A concern beyond particular lagging regions is whether the population in general is less nimble. Indeed, our illustrative model showed that migration elasticities to local shocks are

[63] We also examined the "Eastern Heartland" at-risk area of Austin et al. (2018) (also a focus in Graham and Pinto (2021)), 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 11: 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.)

lower in the presence of local attachments, so we agree this is in principle a well-founded concern. However, declines in gross mobility do not necessarily mean declines in net population change, as gross flows remain orders of magnitude larger than net flows. Population shift can still go in the "right way," even if the turnover producing it is lower.[64] This is an area worthy of future study. Our message is simply to be precise about gross versus net flows, which are often conflated in discussion of migration trends.

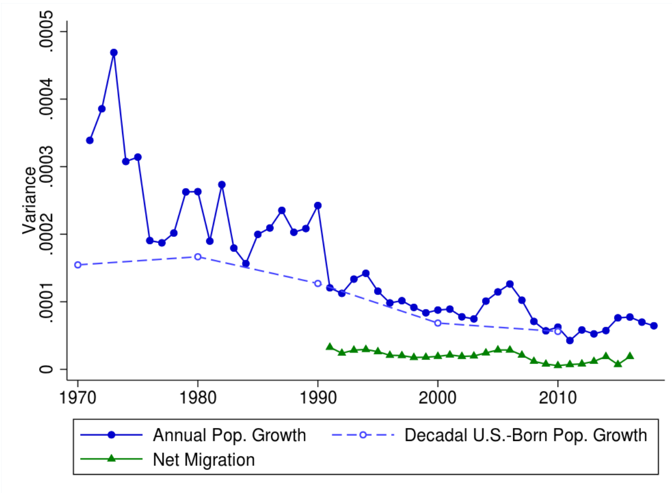
In particular, a direct implication from our analysis is that the concern about falling mobility leading to insufficient reallocation seems to presume the order of events. Part of our contribution to this debate is to understand the autocovariances between gross and net population flows, parsing the lead and the lag. The stabilizing of population growth and, by consequence, home attachment, caused a decline in gross migration rates–not the other way around.

Figure 12 highlights this issue. The figure plots the variance across LLMs in (1) 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.[65] 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, to an extent less than a continuation in the preceding trend in population growth. Cyclical variations in migration variance are still apparent. This does not suggest that a broad migration slowdown

[64] Using state-level data, Kaplan and Schulhofer-Wohl (2017) find that the trend in net migration is flat. Dao et al. (2017) find a lower net migration responsiveness to local labor market shocks from the 1990s on compared with before 1990, but no trend since the break in the early 1990s.

[65] 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.

Figure 12: 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.)

has stopped population growth from happening, but rather the convergence in the long run population trends has preceded (and, in our view, caused) the gross mobility slowdown, with year-to-year fluctuations still occurring.

Hence, the trend in migration is a consequence, not a cause, of abating population shifts. So, our study of migration somewhat ironically ends in an appeal to study changes in population trends (in the spirit of, for instance, Gyourko et al. (2013) and Hsieh and Moretti (2019)) instead of migration per se. At issue for economic growth and equality of opportunity is not whether people are moving around, generally, but whether people can access the best markets - or whether the worst markets need place based assistance. Either way, we are not talking about migration for migration's sake.

8.3 Local Ties and Welfare

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. Proximity to home, friends, and family is an idiosyncratic nontraded good, meaning that heterogeneous households have preferences for some locations over others, and increasing fractions of households are finding it optimal to stay in place. 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.

Yet, rising home attachment is possibly a headwind to the labor market in that it reduces elasticities to local economic shocks. Further research is warranted to understand labor dynamics–gross flows and the net changes they produce–in a more spatially-tied economy.

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. *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, 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.

Bayoumi, T. and J. Barkema (2019). Stranded! How Rising Inequality Suppressed U.S. Migration and Hurt Those Left Behind. Technical report, International Monetary Fund.

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.

BLS (2018). Consumer Price Index, Series CUSR0000SA0L2.

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.

Büchel, K., M. V. Ehrlich, D. Puga, and E. Viladecans-Marsal (2020). Calling From the Outside: The Role of Networks in Residential Mobility. *Journal of Urban Economics 119*, 103277.

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. *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.

Dao, M., D. Furceri, and P. Loungani (2017). Regional Labor Market Adjustment in the United States: Trend and Cycle. *Review of Economics and Statistics* 99(2), 243–257.

DaVanzo, J. (1983). Repeat Migration in the United States: Who Moves Back and Who Moves On? *Review of Economics and Statistics* 65(4), 552–559.

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. D. Fisher, and M. Veracierto (2013). Gross Migration, Housing and Urban Population Dynamics. Federal Reserve Bank of Chicago.

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.

DeWaard, J., M. Hauer, E. Fussell, K. J. Curtis, S. Whitaker, K. McConnell, K. Price, and D. Egan-Robertson (2020). User Beware: Concerning Findings from Recent US Internal Revenue Service Migration Data. Technical report.

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, R. Rodgers, S. Ruggles, and J. R. Warren. (2018). Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset].

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

Ganong, P. and D. Shoag (2017). Why Has Regional Income Convergence in the US Declined? *Journal of Urban Economics 102*, 76–90.

Glaeser, E. L., D. Laibson, and B. Sacerdote (2002). An Economic Approach to Social Capital. *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.

Graham, C. and S. Pinto (2021). The Geography of Desperation in America: Labor Force Participation, Mobility, Place, and Well-Being. *Social Science & Medicine 270*, 113612.

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.

Herkenhoff, K. F., L. E. Ohanian, and E. C. Prescott (2018). Tarnishing the Golden and Empire States: Land-Use Restrictions and the US Economic Slowdown. *Journal of Monetary Economics* 93(1), 89–109.

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.

Hsieh, C.-T. and E. Moretti (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics* 11(2), 1–39.

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.

Ihrke, D. (2014). Reason for moving: 2012 to 2013. Technical report, US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Institute for Social Research (2021). Panel Study of Income Dynamics, Public Use Dataset. Technical report, University of Michigan.

IRS (2018). IRS Migration Data. Technical report, Internal Revenue Service, Statistics of Income.

Johnson, J. E. and M. M. Kleiner (2020). Is Occupational Licensing a Barrier to Interstate Migration? *American Economic Journal: Economic Policy* 12(3), 347–73.

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.

Koşar, G., T. Ransom, and W. Van der Klaauw (2021). Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities. *Journal of Econometrics*.

Krolikowski, P., M. Zabek, and P. Coate (2020). Parental Proximity and Earnings After Job Displacements. *Labour Economics 65*, 101877.

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. *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]. Technical report, 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.

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.

U.S. Census (2011). Centers of population, 2010.

U.S. Census (2018a). County-to-county migration flows.

U.S. Census (2018b). Intercensal population estimates.

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). Employment Hysteresis from 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 Conversion of One Year to Five Year Migration Rate

The census and ACS microdata have the great advantage of containing demographic information including home status. Unfortunately, the census changed from a five year retrospective question in the decadal census to a one year question in the ACS, making it difficult to draw comparisons of migration rates over time. Fairly often moves are reversed (return to origin) or repeated (moved again, to a new location); see DaVanzo (1983) and Kennan and Walker (2011). Hence, in a closed system of discrete locations, it is too simplistic to treat moving rates as independent arrival rate in which five times the one year arrival would deliver the five year rate. Some one year moves would not be observed using a five year retrospective asking where the household lived five years ago.

In an effort to leverage the information on move frequency by demographic and home status, we use auxiliary data to convert one year rates to implied five year rates. The conversion uses the following formula

$$\sigma_{5yr} = 5 \cdot \sigma_{1yr}[1 - p_{reverse} - 0.5 \cdot p_{onward})].$$

The arrival rate is therefore reduced by (i) the probability of a reversal ($p_{reverse}$) occurring within five years, meaning the move would not be observed at all in a survey, and (ii) the probability of another move event to a new location (p_{onward}) occurring within five years, meaning only one of the two moves would be observed.[66] These adjustments deflate the projection to something that would match a five year retrospective question.

What remains is to find estimates of the reversal and onward probabilities. For this, we need longitudinal data detailing the location history of individuals, so we turn to the PSID. This provides a long record of individuals' location histories in addition to demographic information. For each move in the data, we can observe whether the individual moved back to the original or onward to a third location, and if either occurred, the time elapsed since the first move. This allows us to estimate a probability of reversal and onward migration within a t-year window, including estimates by age, education, and home status at the time of move.

However, the PSID is a small sample compared to the census, and since moves are rare, the sample of movers is smaller yet. Data limitations therefore force us to make a few assumptions. First, the reversal and onward rates are assumed constant over time.[67] Second, the reversal and

[66] A person could, of course, move more than twice in five years, meaning the p_{onward} could be scaled by $1/n$, where n is the number of onward moves. In practice, two moves in a five year window is sufficiently close. Moreover, in the biennial years of the PSID, two moves in five years is the most we can observe.

[67] We find reversal rates are higher in the 1980s than in other decades, but we do not detect a general trend in the rates of return or onward moves conditional on an initial move.

onward rates are assumed constant across geographies. Third, the reversal and onward rates at a state level of geography are assumed representative of moves at an LLM level of geography.

Table A1 reports the results using the full 1968-2017 PSID sample, limiting to observations of the working age population, so that initial moves before age 20 or after age 60 are ignored. There are three methods reported: one using simple cell probabilities and two using statistical models. Starting with the cell probabilities, the table shows repeat movers are common: roughly one quarter of moves are reversed within five years, and another quarter are censored by a move to a new location within five years. Furthermore, there are noticeable differences by age and education. Moves initiated later in life, though less frequent, are more likely to stand in five years. The noncollege educated are more likely to return to the place they initially left, and the college educated are more likely to move to a new location.

We find further differences by the home status at time of move–whether the move was leaving home, returning home, or neither. We infer home status as the state in which the individual spent the most time as a minor (under age 18), although the results are similar using other definitions. Because the cells get relatively thin at this level of detail, we use the following statistical models to project the probability. The first is a nested logit specification that estimates the probability of either reversal or onward moves relative to remaining in the first move's location, with the two repeat moves occurring in the same nest. The nesting would account for correlation in "dissatisfaction" in the first move's destination, for example; although we found a simple multinomial logit treating the two repeat move options as independent gave similar results. The model predicts these outcomes as a function of age, education, and home status of the initial move. The second model is a survival time model measuring the probability that the initial move "survives" to the fifth year as a function of age, education, and home status, with the "failures" being either the reversal or onward move events. We use a hazard model with exponential distribution, and model the conditional probability of reversals versus onward moves using the same covariates in a first step logisitic model.

Each of these models, presented in Table A1, reveals differences by home status. Moves away from home are relatively likely to be undone, especially by a reversal back to home. On the other side, moves to home are unlikely to be undone by either reversal or onward move. Moves onward are likely to be repeated by another move onward and also fairly likely to be reversed back to the (not-home) origin.

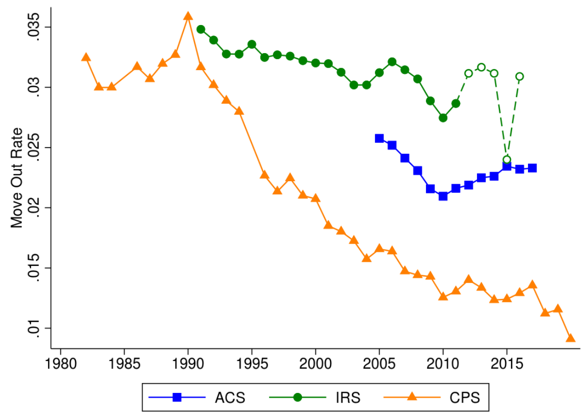
Accounting for these factors allows us to better predict for the ACS the rate at which a move observed one year will persist into the fifth year according to the population characteristics of the LLMs. This allows us to construct time series of five year rates, as in Figures 1, 3, and B6, where we use the hazard model prediction. The adjustment factors are very important, as five times the one year rate would yield an implausibly high five year rate, but as Figures B1 and B4 show, any of the models produces fairly similar results for the time series. While the

Table A1: Repeat Migration

	Count of Moves		Model:	Group Average		Nested Logit		Hazard Model	
	Any	By Status	Repeat Move Type:	Reverse	Onward	Reverse	Onward	Reverse	Onward
Noncollege									
20s	5,010	3,405	At Home	0.352	0.250	0.274	0.237	0.436	0.192
		1,001	Not Home: To Home	0.352	0.250	0.178	0.154	0.184	0.199
		604	Not Home: Onward	0.352	0.250	0.324	0.280	0.131	0.301
30s	2,906	1,580	At Home	0.244	0.226	0.213	0.184	0.310	0.149
		665	Not Home: To Home	0.244	0.226	0.127	0.110	0.135	0.159
		661	Not Home: Onward	0.244	0.226	0.263	0.227	0.091	0.231
40s	1,507	820	At Home	0.221	0.147	0.180	0.155	0.287	0.092
		341	Not Home: To Home	0.221	0.147	0.103	0.089	0.132	0.104
		346	Not Home: Onward	0.221	0.147	0.227	0.196	0.094	0.159
50s	458	261	At Home	0.111	0.079	0.105	0.091	0.127	0.054
		97	Not Home: To Home	0.111	0.079	0.055	0.048	0.052	0.054
		100	Not Home: Onward	0.111	0.079	0.139	0.120	0.038	0.084
College									
20s	1,991	1,407	At Home	0.261	0.347	0.272	0.235	0.320	0.284
		275	Not Home: To Home	0.261	0.347	0.176	0.152	0.110	0.237
		309	Not Home: Onward	0.261	0.347	0.321	0.277	0.074	0.345
30s	1,476	702	At Home	0.190	0.324	0.215	0.186	0.301	0.217
		245	Not Home: To Home	0.190	0.324	0.127	0.110	0.125	0.221
		529	Not Home: Onward	0.190	0.324	0.264	0.228	0.074	0.280
40s	589	246	At Home	0.144	0.261	0.176	0.152	0.249	0.174
		88	Not Home: To Home	0.144	0.261	0.099	0.086	0.107	0.182
		255	Not Home: Onward	0.144	0.261	0.223	0.192	0.060	0.221
50s	164	68	At Home	0.104	0.091	0.095	0.082	0.143	0.052
		20	Not Home: To Home	0.104	0.091	0.050	0.043	0.056	0.050
		76	Not Home: Onward	0.104	0.091	0.128	0.110	0.044	0.084

NOTES: The table reports estimated repeat migration rates within five years of the initial move, by education, age, and home status at time of initial move. The repeat migration events can be "reverse," a return to the origin, or "onward," a move to a new destination. The Group Average model is a simple cell probability, pooling over home statuses. The Nested Logit model is a three-option logistic regression, nesting of the two potential repeat actions in one choice nest. Home status (away from or return to) are entered as controls in the regression, along with intercepts by age-education group. The Hazard Model is a survival time model using an exponential hazard rate and home status controls in the regression, along with intercepts by age-education group. The table reports the probability of "failure," i.e., the repeat move event, within a five year window, with the form of repeat migration (whether reverse or onward) predicted by a logit model conditioned on a failure occurring and containing the same attributes as controls. (Source: PSID data.)

Figure B1: State-to-State Migration Rates



NOTES: The figure reports the time series of migration rates when migration is defined as a change in state of residence. (Source: IRS, ACS, and CPS data.)

data limitations inherent in this conversion process dissuade us from relying too heavily on the results of these figures, it is reassuring that the patterns are similar between one and five year migration rates, even drawing from different data sources.

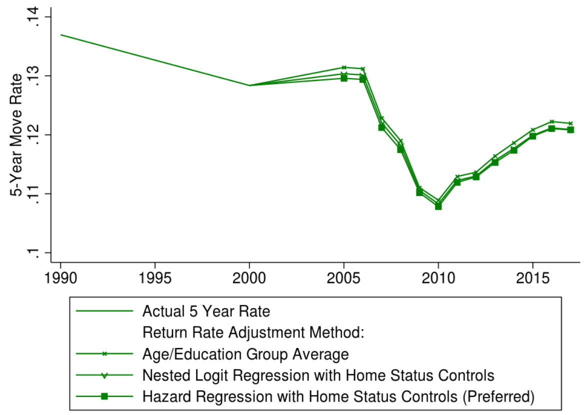
B Additional Facts on the Migration Decline

B.1 Interstate Migration Rates

State to state migration rates are the focus of many of the seminal papers on the migration decline, including Molloy et al. (2011), Kaplan and Schulhofer-Wohl (2017), and Karahan and Rhee (2014). Figure B1 presents the one year state to state migration rate as derived from the IRS, ACS microdata, and March CPS samples. Each shows a downward trend comparable to the LLM rates in Figure 1.

The CPS does not allow for assignment to LLMs, which is one reason it is not our preferred source (though its limited sample size is the main reason). The figure shows the CPS to be somewhat out of step with the other data series, however. Each show a decline, but the drop in the CPS is especially large. While the IRS data (derived from consistent taxpayers) may understandably be a different sample than the CPS, the ACS data should represent a sample similar to the CPS. Yet, the interstate migration rates are half the magnitude in the CPS compared with the ACS, which presents a puzzle we have not been able to resolve. (This series excludes imputed moves (see Kaplan and Schulhofer-Wohl (2012)).) We have checked that it is not due to age composition in the datasets, which are similar. It seems notable that one-third of the drop occurs between 1994 and 1996 (there is no 1995 data point) over the course of a survey redesign.

Figure B2: Robustness of Five Year Migration Rate, 1990-2017



NOTES: The figure displays five year migration rates. For 1990 and 2000, the five year rates are directly from census data. For 2005-2017, the five year rates are the implied rate under conversion by each of the three different models of Table A1. Compare with Figure 1, which uses the hazard regression method. (Source: Census, ACS, and PSID data.)

Table B1: Move Home Rates by Age and Education: Robustness to Geographic Definitions

Education/ Age	Move Home Rate (%) Lower Bound Actual L-5	Lower Bound Synthetic L-6	Upper Bound Actual U-5	Upper Bound Synthetic U-6	Odds Ratio, Move Home/Other In-Movers, Geography-Adjusted Average Ratio Lower Bound L-5/6	Upper Bound U-5/6
Noncollege						
20s	10.23	0.70	39.45	4.71	7.76	116.48
30s	8.84	0.68	34.30	4.48	6.59	88.44
40s	8.42	0.64	31.12	4.18	6.18	76.81
50s	7.37	0.57	28.50	3.86	5.94	69.76
College						
20s	9.78	1.10	31.36	4.84	5.48	63.24
30s	8.16	1.00	27.00	4.45	5.71	59.27
40s	6.58	0.91	22.10	4.08	4.97	46.04
50s	5.90	0.73	21.22	3.73	5.02	44.79

NOTES: The table reports move home rates under alternative definitional scenarios; compare Table 1. The baseline in Table 1 weighted LLMs within the respondent's birthstate by the population-based probability the LLM was his/her birth LLM. The lower boundary assumes other LLMs in the respondent's birthstate are not his/her home LLM. The upper bound version assumes any LLM in the respondent's birthstate is affirmatively a home LLM. The synthetic choice probabilities correspond to their respective definition of home LLM.
Odds ratios are the relative choice probabilities of a destination LLM by LLM natives compared with non natives, conditioning on the origin LLM in order to account for geographic networks of migration flows. For example, the odds ratio compares the choice probability of New York City by native New Yorkers living in Tampa to the choice probability of non-New Yorkers living in Tampa. There are $J(J-1)$ pairs for each age-education group, and the table reports the average odds ratio among non-missing cells. (Source: ACS data.)

B.2 Five Year Migration Rates

As a robustness check on Figure 1, Figure B2 reports the five year rate with alternative models used to make the one-to-five year conversion in the ACS-era data; the pattern is very similar under the various models. See Section A and Table A1 for more information on the one-to-five year conversion models.

B.3 Robustness of Home Status Definitions

We observe in which state a respondent was born, but not his or her LLM of birth. Table 1 used a relatively conservative definition in which we downweight the probability an LLM is home by the distribution of population in the respondent's birth year. For example, if in 1980, one-half of population of the state of Georgia lives in Atlanta and the other half lives in the rest of Georgia, a move of a 30-something Georgia native in the 2010s to Atlanta gets a move home weight of 1/2 instead of 1. Table B1 presents two alternatives. The first (L-5) is an even more conservative definition in which the other LLMs are assigned a weight of zero (e.g., the Georgia native moving to Atlanta is assigned a 0 for his failure to move to the rest of Georgia). The second (U-5) is a more accommodative definition in which a move to any LLM in the home state is assigned a value of 1 (in the Georgia native example, both Atlanta and residual Georgia could qualify as moves home with weight of 1). These two effectively bound the range of move home rate in our data. If moves to a home LLM are relatively more likely conditioning on a move to one's home state, the more liberal assignment of home could be doing a better job of assigning the move home rate. If moves to LLMs other than home within one's home state are more likely (e.g., the native of rural Georgia selects Atlanta more often than someone not from Georgia), then the accommodative definition may technically overstate the rate of move home but still be qualitatively correct regarding place attachment in that proximity to home is valued by migrants.

Columns L-5 and U-5 show the main estimate reported in Table 1 is conservative in that it is nearer the lower bound, although moves home could be as high as one-third of all moves under the more accommodative definition. Under either definition, the actual moves home dominate by far the synthetic versions of move home (L-6 and U-6), indicating that home is chosen by a far greater probability than would be predicted by chance.

The table also addresses a competing hypothesis, that migration is coincidentally towards home because migration networks are relatively local and any mover does not migrate far from her initial location–i.e., home is usually nearby. The table reports the odds ratio, conditioning on the location of origin, of choosing a destination when it is home versus when it is not. This measures, for example, the probability that a person native to Michigan will choose Detroit when living in Cleveland compared to the probability that another Cleveland resident not native to Michigan would choose Detroit. Columns L-5/6 and U-5/6 report, respectively, the median odds ratio under the lower and upper bound assumptions about home LLM. In either case, the statistics show that moves home are disproportionately likely when adjusting for move proximity.

B.4 Measuring the Marginal Effect of Home Status on Migration Choice

Table B2 presents odds ratio regressions of the probability of moving between at-home and not-at-home residents. The regression is $\sigma_{mo,it} = \alpha_0 + \alpha_1 I(o = H) + X_o\beta$, where α_0 is a baseline rate and α_1 is the effect of home status, and o, i, t index origins, individuals, and

Table B2: Out Migration by Home Status: Odds Ratio Regressions

Education / Age	Home Status Method: Controls: Model:	At-Home Indicator None Logit	At-Home Indicator Income Logit	At-Home Indicator LLM FEs LP	At-Home Indicator + Roots None Logit	At-Home Indicator + Roots LLM FEs LP	At-Home Indicator + Roots None Logit
Noncollege							
20s	At Home	0.339	0.343	0.419	0.371	0.413	0.396
		(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)
	At Home X Hi Roots				0.354	0.392	0.379
					(0.001)	(0.003)	(0.004)
	Not Home X Hi Roots						1.090
							(0.016)
30s	At Home	0.365	0.364	0.424	0.378	0.412	0.399
		(0.001)	(0.001)	(0.003)	(0.001)	(0.003)	(0.006)
	At Home X Roots				0.346	0.366	0.365
					(0.001)	(0.004)	(0.005)
	Not Home X Roots						1.072
							(0.020)
40s	At Home	0.371	0.371	0.404	0.375	0.387	0.410
		(0.002)	(0.002)	(0.003)	(0.002)	(0.004)	(0.006)
	At Home X Hi Roots				0.330	0.311	0.361
					(0.002)	(0.005)	(0.006)
	Not Home X Hi Roots						1.128
							(0.023)
50s	At Home	0.368	0.370	0.420	0.368	0.402	0.389
		(0.002)	(0.002)	(0.004)	(0.002)	(0.004)	(0.005)
	At Home X Hi Roots				0.310	0.316	0.329
					(0.002)	(0.005)	(0.005)
	Not Home X Hi Roots						1.081
							(0.020)
College							
20s	At Home	0.448	0.444	0.500	0.493	0.490	0.629
		(0.002)	(0.002)	(0.004)	(0.002)	(0.004)	(0.013)
	At Home X Hi Roots				0.485	0.430	0.619
					(0.003)	(0.006)	(0.013)
	Not Home X Hi Roots						1.353
							(0.033)
30s	At Home	0.390	0.388	0.413	0.408	0.399	0.541
		(0.002)	(0.002)	(0.005)	(0.002)	(0.005)	(0.009)
	At Home X Roots				0.374	0.305	0.497
					(0.002)	(0.007)	(0.009)
	Not Home X Roots						1.425
							(0.030)
40s	At Home	0.394	0.395	0.436	0.403	0.427	0.464
		(0.003)	(0.003)	(0.007)	(0.003)	(0.007)	(0.010)
	At Home X Hi Roots				0.365	0.343	0.420
					(0.004)	(0.010)	(0.010)
	Not Home X Hi Roots						1.203
							(0.033)
50s	At Home	0.439	0.439	0.483	0.446	0.475	0.510
		(0.003)	(0.003)	(0.006)	(0.003)	(0.006)	(0.010)
	At Home X Hi Roots				0.395	0.382	0.452
					(0.004)	(0.009)	(0.010)
	Not Home X Hi Roots						1.196
							(0.030)

NOTES: The table reports odds ratio regressions of move out rates as a function of home status and the controls as indicated in the column headings. The odds ratio is taken with respect to the at-home relative to the not-at-home, US-born groups, within each age and education group. The foreign born are included in the regression sample but excluded from the odds ratio presented. The model in columns 1 to 3 is $\sigma_{m,iot} = \alpha_0 + \alpha_1 I(o = H) + X_o\beta$ and in 4 to 6 is $\sigma_{m,iot} = \alpha_0 + \alpha_1 I(o = H) + \alpha_2 R_o I(o = H) + X_o\beta$, the difference being the inclusion of the location's rootedness for the age cohort. The models are either logit or linear probability (LP), as indicated in the column headings. Year of observation dummies are included in all specifications. (Source: ACS data.)

time, respectively. The odds ratio is then $\frac{\hat{\sigma}_{m,o=H}}{\hat{\sigma}_{m,o\neq H}} = \frac{\alpha_0+\alpha_1}{\alpha_0}$. (The foreign born are included and controlled for in the regression sample but excluded from the odds ratio.) A number below one indicates that people at home are less likely to move (i.e., $\alpha_1 < 0$). The odds ratio has the advantage of standardizing the relative sizes between home statuses across different demographic groups, which are estimated separately in their own regression model, so that results can be compared across groups with different baseline moving rates. Compared with the descriptive results in Tables 1 and B1, the regression allows control for individual or location factors besides home status.

The table presents several specifications of increasing controls. The first column is a simple logit regression. It shows the at-home are only 34 to 45 percent as likely to move as the not-at-home. The next column controls for local income opportunities (mean and variance of income) with very similar results. The third column includes LLM fixed effects to soak up any fixed local features. Here, the odds ratio rises somewhat, indicating that there are generally more at-home people in less mobile locations–which is of course consistent with the population growth patterns presented in Section 2. But the odds ratios are still consistently below one-half across all age and education groups.

The second set of columns expands the model to include a home status indicator and a roots coefficient: $\sigma_{m,iot} = \alpha_0 + \alpha_1 I(o = H) + \alpha_2 R_o I(o = H) + X_o\beta$. This will measure whether there is variation in the move rates among the at-home group that is explained by rootedness, our measure of the intensity of home attachment. To maintain the odds ratio presentation, we report the odds ratio at mean roots for the sample and one standard deviation above ("Hi Roots"). If $\alpha_2 < 0$, then higher rootedness means lower probability of a move among the at-home.

This is indeed what we find: higher rooted places have lower mobility among their LLM natives, which is consistent with Figure 7. The odds ratio at a standard deviation above the mean roots is two to six percentage points lower than at the mean. Notably, this finding is not the result of rooted places being less mobile for other, unobserved reasons. The result remains even when controlling for origin fixed effects, so higher roots still explain variation in move rates among the at home when generically accounting for LLM features. The final column then interacts rootedness with not-at-home (but US-born) status. The odds ratios above one indicate that nonlocal residents actually leave rooted areas at higher than normal rates.

Table B3 conducts a very similar exercise as Table B2, but focusing on the move-to probability, conditional on moving. To do so, we need to construct a dataset of all movers with all possible alternatives, so the regression model is $\sigma_{jo,it} = \alpha_0 + \alpha_1 I(j = H) + X_j\beta$, with $J - 1$ observations (of $j \neq o$) for each individual i. The intuition is much the same: we look at choice probability ratios by home status, controlling for features of the destination or origin-destination pair.

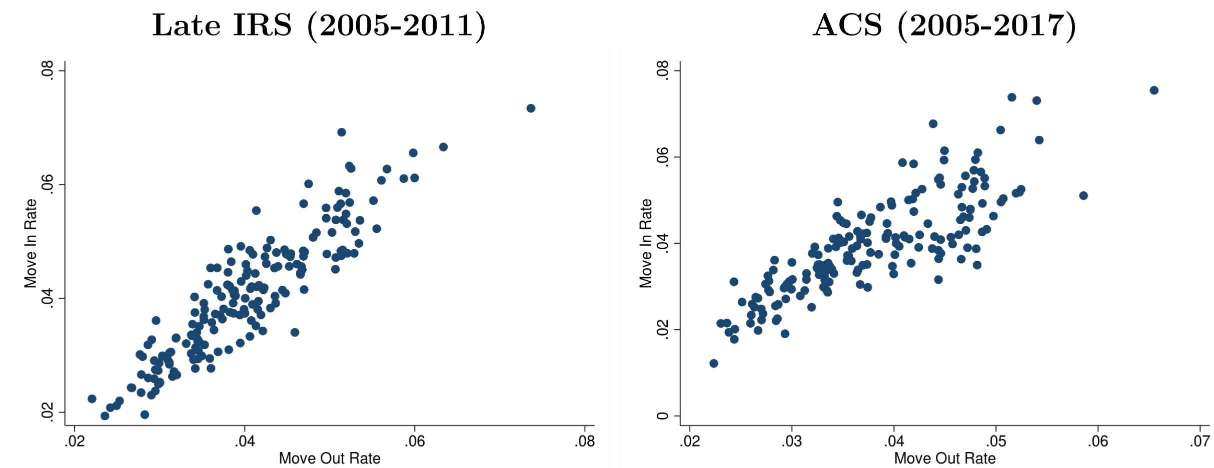
The table shows in the first column that a destination being one's home makes it nine to 19

Table B3: In Migration by Home Status: Odds Ratio Regressions

Home Status Method	At-Home Indicator				At-Home Indicator + Roots		
Education / Controls:	None	Income	Dest. LLM FEs	LLM Pair Fes	None	LLM Pair Fes	None
Age Model:	Logit	Logit	LP	LP	Logit	LP	Logit
Noncollege							
20s To Home	18.921	18.509	11.129	11.109	15.985	11.022	6.735
	(0.294)	(0.289)	(0.145)	(0.150)	(0.215)	(0.149)	(0.121)
To Home X Hi Roots					14.781	10.195	6.228
					(0.284)	(0.205)	(0.141)
Not Home X Hi Roots							0.339
							(0.005)
30s At Home	15.627	15.328	9.267	9.431	13.630	9.406	5.074
	(0.314)	(0.309)	(0.162)	(0.170)	(0.241)	(0.170)	(0.115)
At Home X Roots					13.151	9.160	4.895
					(0.323)	(0.233)	(0.139)
Not Home X Roots							0.284
							(0.005)
40s At Home	14.762	14.586	8.989	8.772	13.080	0.185	4.905
	(0.338)	(0.335)	(0.177)	(0.183)	(0.265)	(8.790)	(0.125)
At Home X Hi Roots					13.267	9.005	4.975
					(0.375)	(0.260)	(0.160)
Not Home X Hi Roots							0.279
							(0.005)
50s At Home	13.240	13.102	8.095	7.936	11.912	7.962	4.216
	(0.317)	(0.314)	(0.167)	(0.176)	(0.256)	(0.177)	(0.111)
At Home X Hi Roots					12.231	8.326	4.329
					(0.371)	(0.253)	(0.147)
Not Home X Hi Roots							0.250
							(0.005)
College							
20s At Home	15.545	15.509	10.316	10.538	13.353	10.537	9.713
	(0.315)	(0.315)	(0.187)	(0.190)	(0.233)	(0.190)	(0.229)
At Home X Hi Roots					13.043	10.191	9.488
					(0.317)	(0.261)	(0.275)
Not Home X Hi Roots							0.674
							(0.013)
30s At Home	12.571	12.603	8.498	8.679	11.136	8.686	5.702
	(0.248)	(0.248)	(0.154)	(0.158)	(0.193)	(0.158)	(0.126)
At Home X Roots					10.806	8.448	5.533
					(0.255)	(0.211)	(0.151)
Not Home X Roots							0.430
							(0.007)
40s At Home	9.646	9.651	6.603	6.605	8.826	6.213	3.361
	(0.294)	(0.294)	(0.187)	(0.193)	(0.244)	(0.182)	(0.112)
At Home X Hi Roots					8.278	7.467	3.152
					(0.306)	(0.226)	(0.131)
Not Home X Hi Roots							0.287
							(0.007)
50s At Home	9.303	9.284	6.138	6.091	8.600	6.092	3.034
	(0.303)	(0.302)	(0.182)	(0.192)	(0.256)	(0.191)	(0.109)
At Home X Hi Roots					8.389	6.074	2.960
					(0.349)	(0.265)	(0.137)
Not Home X Hi Roots							0.254
							(0.007)

NOTES: The table reports odds ratio regressions of destination choice probabilities, conditional on moving, as a function of home status and the controls as indicated in the column headings. The odds ratio is taken with respect to the at-home relative to the not-at-home, US-born groups, within each age and education group. The foreign born are included in the regression sample but excluded from the odds ratio presented. The model in columns 1 to 4 is $\sigma_{jo,it} = \alpha_0 + \alpha_1 I(j = H) + X_j \beta$, with $J - 1$ observations (of $j \neq o$) for each individual i. Columns 5 to 7 include the location's rootedness for the age cohort. The models are either logit or linear probability (LP), as indicated in the column headings. Year of observation dummies are included in all specifications, and standard errors are clustered by i. (Source: ACS data.)

Figure B3: In and Out-Migration by Local Labor Market; Alternative Samples



NOTES: The figures plot in- versus out-mobility rates for U.S. LLMs in the late IRS and ACS microdata samples. Compare Table 2. (Source: IRS and ACS data.)

times more likely to be selected than if it were not home, depending on the demographic group. Controlling for destination income (itself a significant determinant) delivers very similar results. Controlling for destination and destination pair effects reduces the ratio somewhat, but home status still remains an especially strong predictor.

The second set of columns introduces the roots coefficient separately from the at home indicator. Here, the roots coefficient for the at-home is actually negative, which makes the destination a slightly less likely choice for movers when the LLM's rootedness is high. The gap closes slightly when accounting for LLM-pair fixed effects, but the last column is perhaps most telling. A highly rooted destination is substantially less likely to be chosen among non-natives, while only slightly less likely to be chosen among natives. This suggests that rootedness itself is not a deterrent on its own for natives, but rather highly rooted places are less attractive destinations in general–which is of course consistent with their being the places with the lowest population growth over the last few decades.

B.5 In and Out-Migration Rates: Robustness of Figure 2

Figure B3 presents sample robustness tests of Figure 2, displaying the late IRS (2009-2011) period and the LLM averages derived from ACS microdata covering 2005-2017. In all samples, LLM move rates show a high degree of correlation between inflow and outflow rates.

B.6 Additional Geographic Patterns in the Migration Decline

Figure B4 reports on robustness and alternative viewings of the decline in migration rate by initial migration speed. The upper panel shows scatter plots of the migration rate in the IRS data in order to see more of the heterogeneity across LLMs. The lefthand figure shows early

Figure B4: Changes in Out-Migration Rate Over Time; LLM Level Observations and Five Year Robustness

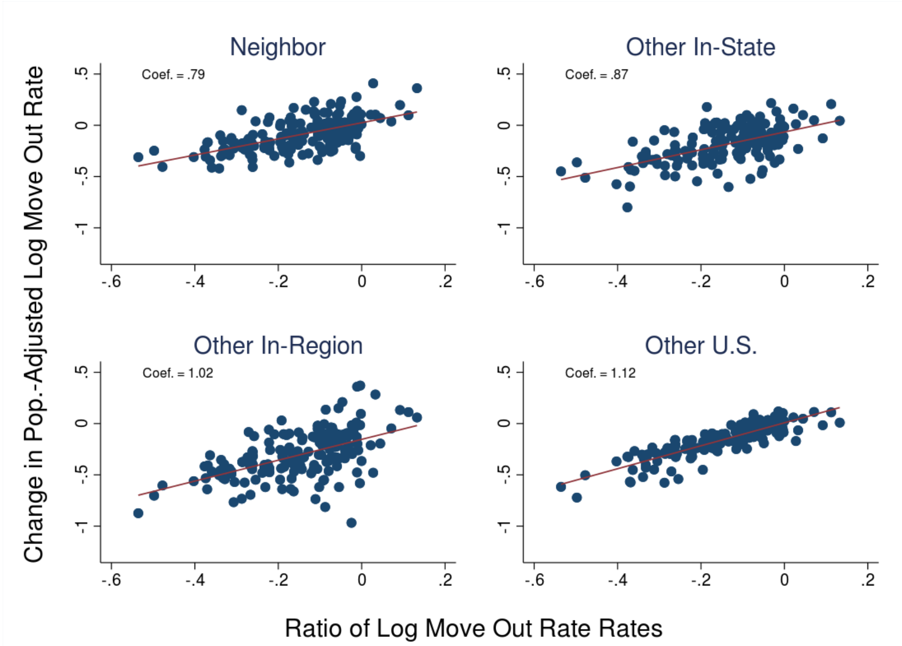
NOTES: The upper lefthand figure plots out-mobility rates for U.S. LLMs in the early and late IRS data series, as noted on the axes. A 45-degree and best-fit line are also plotted. The upper righthand figure plots the change in out-mobility rates for U.S. LLMs to the early period average, with bubble size proportional to population size. (Source: IRS migration data.)
The lower figure plots each LLM speed series under each one-to-five-year conversion model from Table A1. (Source: Census, ACS, and PSID data.)

and late out-migration rates; that the best fit line lies below the diagonal shows that the larger changes occur in initially more mobile places, although the relative rankings by speed are mostly preserved. The righthand figure plots the difference, the change in migration rate against the initial out-migration rate after adjusting for the average scale effect (larger cities being less mobile on average). The bubble sizes are proportional to city size. This plot also shows a clear pattern of more mobile locations exhibiting greater declines across the population-size distribution of cities.

The lower figure plots robustness checks for the five year migration rates, displaying for each series the implied five year rates under each one-to-five-year conversion model from Table A1. The result that fast LLMs drive the migration decline is not affected by the conversion method.

Figure B5 plots the change in migration for a partitioning of potential destinations. The

Figure B5: 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: IRS data).

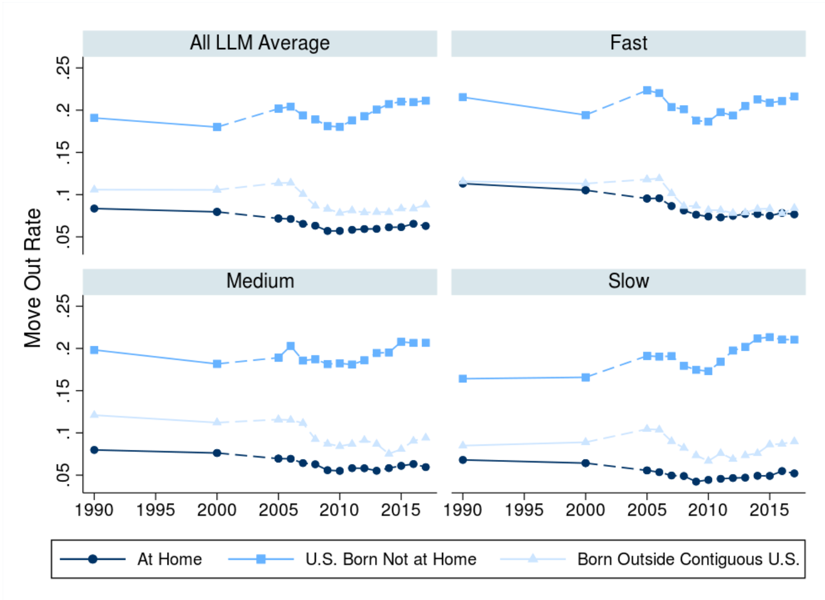
destinations are grouped by "Neighbors" (adjacent LLMs) to the origin LLM, other LLMs in the same state as the origin LLM, others in the same region, and the remainder. The scatter plots show that declines in migration to each kind of destination are associated with declines in out-migration from the 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.

B.7 Move Rates by Home Status Over Time

Figure B6 shows the migration propensity over time by birthplace status. There are no discernible trends in the migration rate within the at-home, not-home, or foreign born groups (with one exception), which suggests that the change is the shift in population composition from the mobile, not-at-home category to the at-home category. The exception is the migration rate for the at-home residents of fast locations, which has trended down as their intensity of attachment has increased, as described in Section 2.6.1. This feature is also consistent with converging population growth driving the migration trend.

Figure 6 showed the migration propensity over time by birthplace status. Table B4 reports odds ratio regressions over time for models of move out rate (see also Table B2 for the ACS

Figure B6: Migration Rate Over Time By Birthplace Status



NOTES: The figure plots implied five year migration rates by home status. The years using the ACS sample (2005 onward) convert to five year rates, as explained in Section A, using the conversion factors in Table A1. The samples have been balanced to have a constant age/education composition over time. (Source: Census and ACS Microdata).

Table B4: Out Migration by Home Status Over Time: Odds Ratio Regressions

Education / Age	Sample Model	ACS 2005-17 Logit	ACS 2005-17 LP FE	Census 2000 Logit	Census 2000 LP FE	Census 1990 Logit	Census 1990 LP FE
Noncollege							
20s	Avg. Roots	0.740		0.764		0.778	
	At Home (at Avg. Roots)	0.371	0.413	0.399	0.415	0.386	0.399
		(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
30s	Avg. Roots	0.753		0.770		0.773	
	At Home (at Avg. Roots)	0.378	0.412	0.389	0.402	0.371	0.379
		(0.001)	(0.003)	(0.001)	(0.001)	(0.000)	(0.001)
40s	Avg. Roots	0.764		0.769		0.798	
	At Home (at Avg. Roots)	0.375	0.387	0.380	0.390	0.372	0.372
		(0.002)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
50s	Avg. Roots	0.768		0.795		0.816	
	At Home (at Avg. Roots)	0.368	0.402	0.399	0.416	0.385	0.403
		(0.002)	(0.004)	(0.001)	(0.002)	(0.001)	(0.002)
College							
20s	Avg. Roots	0.734		0.757		0.759	
	At Home (at Avg. Roots)	0.493	0.490	0.600	0.589	0.609	0.597
		(0.002)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)
30s	Avg. Roots	0.736		0.745		0.748	
	At Home (at Avg. Roots)	0.408	0.399	0.476	0.456	0.485	0.461
		(0.002)	(0.005)	(0.001)	(0.001)	(0.001)	(0.001)
40s	Avg. Roots	0.737		0.741		0.761	
	At Home (at Avg. Roots)	0.403	0.427	0.430	0.404	0.450	0.413
		(0.003)	(0.007)	(0.002)	(0.002)	(0.002)	(0.002)
50s	Avg. Roots	0.738		0.758		0.778	
	At Home (at Avg. Roots)	0.446	0.475	0.469	0.439	0.473	0.438
		(0.003)	(0.006)	(0.002)	(0.003)	(0.004)	(0.004)

NOTES: The table reports odds ratio regressions of move out rates as a function of home status and the controls, for each data sample, as indicated in the column headings. Note that the ACS uses one year retrospective question to measure migration, but the census uses a five year retrospective. The odds ratio is taken with respect to the at-home relative to the not-at-home, US-born groups, within each age and education group. The foreign born are included in the regression sample but excluded from the odds ratio presented. The model in columns 1 to 3 is $\sigma_{m,iot} = \alpha_0 + \alpha_1 I(o = H) + X_o\beta$ and in 4 to 6 is $\sigma_{m,iot} = \alpha_0 + \alpha_1 I(o = H) + \alpha_2 R_o I(o = H) + X_o\beta$, the difference being the inclusion of the location's rootedness for the age cohort. The models are either logit or linear probability with origin fixed effects (LP FE), as indicated in the column headings. (Source: ACS microdata.)

baselines). The odds ratio helps to standardize the comparison across time, despite the ACS data reporting one year rates and census data reporting five year rates. The table shows that odds ratios by home status are quite consistent over time; at home residents move out at rates one-third to two-fifths as often as the not-at-home at each time period of observation. Some of the college educate groups exhibit a slight increasing attachment to home. While rootedness increased in the fast LLMs, the average city over time shows a flat or even declining trend. Thus, it does not appear the sensitivity to home has changed much over time, while the share of people in the at home category has increased, mostly in fast locations.

B.8 Migration Trends by Geography in the CPS

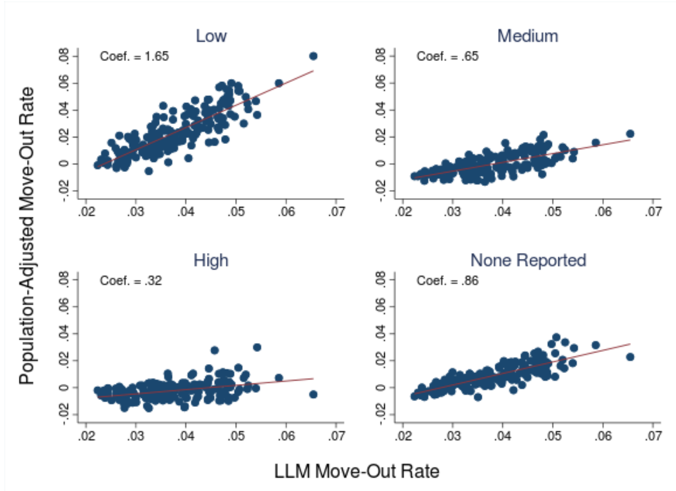
As noted above, the CPS is not our preferred dataset primarily because it is too small to splice by subgeography, and the finest subgeography it offers is state (not LLM, our preferred boundary definition). The CPS also stands out in the magnitude of the migration decline it indicates. However, subject to these limitations, we can check whether any of the spatial heterogeneity we find in the main analysis holds up in the CPS. We do find that interstate rates have declined the most in the Frontier, West, and South regions, and the least in the Northeast, and correlate very highly with changes in the IRS. The CPS disagrees with the IRS slightly on the degree of change in the Midwest region–the CPS shows a moderate decline in interstate migration (comparable to the South), while the IRS shows a much smaller decline (comparable to the Northeast). Among states that are large enough to be feasibly measured, the CPS shows larger declines out of Texas, Florida, and California, and smaller declines in New York, Pennsylvania, and Illinois. These all correlate with findings from the IRS.

B.9 Migration and Income Dispersion

Figure B7 plots the migration rate out of an 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 B5 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 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

Figure B7: LLM Mobility Rate by Point in the Income Distribution



NOTES: The figure shows scatterplots of within-income category out-mobility rate (adjusted for group size) against the average LLM out-mobility rate. The "low" income 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 B5: Individual Migration Rate by Point in the Income Distribution

Short-Run	Moved in Period		Long-Run	Moved in Sample	
Position	Stay	Move	Avg. Posiition	Never	At Least Once
Total	32,182	914	Total	45,057	9,904
rate	97.24	2.76	rate	81.98	18.02
Low	6,924	201	Low	9,701	1,668
rate	97.18	2.82	rate	85.33	14.67
Average	17,035	427	Average	22,318	4,960
rate	97.55	2.45	rate	81.82	18.18
High	8,223	286	High	13,038	3,276
rate	96.64	3.36	rate	79.92	20.08

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.)

"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.

C Modeling Details

The following appendix contains derivations used in the model.

C.1 Model Setup: Choice Probabilities and Value Functions

For convenience, 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

$$
\begin{aligned}
ln\sigma_{j|o} =& \frac{1}{\lambda}\big[v_{j|o} - ln\big(\sum_i exp[v_{i|o}]^{\frac{1}{\lambda}}\big)\big] && \text{(C6a)}\\
ln\sigma_{j|o} - ln\sigma_{i|o} =& \frac{1}{\lambda}\big(v_{j|o} - v_{i|o}\big) && \text{(C6b)}
\end{aligned}
$$

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

[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.

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

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),

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

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

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

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

The following appendix contains details of the model estimation procedure, including derivations used in the construction of the estimating equations.

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. Equation (D5) becomes

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

This equation has expressed the expected value of a move $(ln\big(\sum_k exp[v'_{...}]^{\frac{1}{\lambda}}\big)$ as the choice-specific value of some location, v_k, minus the probability of moving there, $ln\sigma_k$. 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

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

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

$$(ln\sigma_{so} - ln\sigma_{mo}) - (ln\sigma_{sz} - ln\sigma_{mz}) + \frac{\beta}{\delta}(ln\sigma'_{m|o} - ln\sigma'_{m|z}) =$$
$$\frac{1}{\delta}(u_o - u_z) + (\beta - 1)\frac{\lambda}{\delta}\Big(\frac{1}{\lambda}(mc_{ko} - mc_{kz}) - (ln\sigma_{ko|m} - ln\sigma_{kz|m})\Big), \quad \text{(D9)}$$

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

D.2 The Estimation Procedure

Estimation relies on information both in move versus stay decisions and in the propensity to choose one location over another, conditional on 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]

[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.

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

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

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

where differences have been made to ratios for readability, and we have suppressed the $mc_{zi} - mc_{zj}$ from (D3) term in the lower right block of X because the move costs to the normalizing location will be the same by assumption, making the term zero. 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 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).

D.3.1 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. Using (C7c) and substituting (C6a), we have

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

Unlike the expressions in Sections D.1.1 and D.1.2, this does not difference out the intercept of the move cost term in order to recover it in estimation. However, doing so 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

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

where the left-hand side is composed of choice probability data and previously recovered moving costs with respect to distance.[71] 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.3.2 Scale Parameters

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.[72] 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 whether a move is a move home

[71] 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.

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

decision, taking their ratio as $\frac{1}{\delta}$, effectively setting λ to one. This procedure will not identify 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:

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

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

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.4 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:

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

$$\sigma_{jo|m} = g(\mu^W_{j,\tau}, \sigma^W_{j,\tau}, n, I(j = h)). \tag{D16}$$

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.[73] These projections therefore work in concert with the flow utility differences, not in place of them. The CCP-based approach 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. Instead, this 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.

Table D1 reports the coefficients among the US-born population from the models used in the counterfactual simulations.

D.5 Forming the Cell Sizes

We next describe how we split the data into type cells (τs). 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, each location plus foreign born. Thus, we have $4 \times 2 \times (J + 1)$ types. Interacting these with origin (the state variable), we have $4 \times 2 \times J \times (J + 1)$ 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. 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.

As first mentioned in Section 2, 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

[73]If the choice probabilities are observed, we can enter them into a simulation directly.

Table D1: Auxiliary Model Results

Edu.	Noncollege				College			
Age	20s	30s	40s	50s	20s	30s	40s	50s
				Move/Stay Probability Model				
Origin Attributes								
Roots X Home	-0.0853	-0.0472	-0.0293	-0.0238	-0.0877	-0.0498	-0.0239	-0.0183
Home is Residual LLM	-0.0370	-0.0196	-0.0108	-0.0084	-0.0663	-0.0304	-0.0131	-0.0086
In Residual LLM	-0.0384	-0.0253	-0.0189	-0.0168	-0.0020	-0.0099	-0.0074	-0.0102
Inc. Terc 1 X μ_j^W	-0.0121	-0.0084	-0.0158	-0.0102	0.0012	0.0038	-0.0035	-0.0038
Inc. Terc 2 X μ_j^W	-0.0091	-0.0026	-0.0119	-0.0088	-0.0247	0.0005	-0.0054	-0.0025
Inc. Terc 3 X μ_j^W	-0.0254	-0.0056	-0.0123	-0.0092	-0.0541	-0.0058	-0.0084	-0.0077
Inc. Terc 1 X σ_j^W	0.0508	0.1097	0.0683	0.0262	-0.3901	-0.0329	-0.0286	-0.0078
Inc. Terc 2 X σ_j^W	-0.0016	0.0227	0.0084	-0.0018	-0.1099	-0.0160	-0.0141	-0.0340
Inc. Terc 3 X σ_j^W	0.1923	0.0584	0.0131	0.0036	0.2762	0.0678	0.0287	0.0392
No Income	-0.0920	-0.0005	-0.1070	-0.0848	-0.3468	-0.0024	-0.0648	-0.0492
Type Constant	0.1899	0.0602	0.1468	0.1176	0.4763	0.0677	0.0982	0.0792
				Move-To Probability Model				
Destination Attributes								
Roots X Home	0.1864	0.1722	0.1518	0.1339	0.1352	0.0978	0.1163	0.0864
Home is Residual LLM	0.0515	0.0360	0.0278	0.0282	0.0222	0.0268	0.0110	0.0169
To Residual LLM	0.4720	0.2626	0.4743	0.2840	0.4979	0.3263	0.5212	0.4098
Inc. Terc 1 X μ_k^W	0.0087	0.0403	0.0049	0.0347	0.0044	0.0303	0.0041	0.0216
Inc. Terc 2 X μ_k^W	0.0085	0.0393	0.0047	0.0340	0.0036	0.0294	0.0042	0.0210
Inc. Terc 3 X μ_k^W	0.0071	0.0403	0.0033	0.0356	0.0026	0.0287	0.0032	0.0209
Inc. Terc 1 X σ_k^W	0.0238	0.0466	0.0221	0.0350	0.0117	0.0029	0.0149	0.0185
Inc. Terc 2 X σ_k^W	0.0258	0.0597	0.0253	0.0438	0.0220	0.0137	0.0135	0.0263
Inc. Terc 3 X σ_k^W	0.0430	0.0455	0.0408	0.0227	0.0330	0.0223	0.0256	0.0272
No Income	0.1092	0.4755	0.0691	0.4050	0.0550	0.3303	0.0548	0.2490
Type Constant	-0.1000	-0.4645	-0.0600	-0.3943	-0.0463	-0.3200	-0.0463	-0.2396
LLM Pair FEs	y	y	y	y	y	y	y	y

NOTES: The table reports the coefficient estimates (standard errors suppressed) from regression models of the move-out probability (upper panel) and move-to/destination choice probability (lower panel). Each column is a separate specification. These coefficients are inputs to the model simulations using counterfactual CCPs. See Section D.4 for discussion. (Source: ACS data.)

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 preference 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 (see Table B1 and associated discussion in Section B).

D.6 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]

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.7 Income Distributions and Dynamics

To simulate the model, we need measures of income offer distributions and income dynamics. These determine the value 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

[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.

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 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 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 PSID.[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

[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 via indirect inference or method of moments. We did not pursue this because, aside from complicating the analysis, it would essentially assume the workers know the primitive distribution apart from the observed distribution.

[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 CPS and then merge this to the PSID by survey year.

data to a maximum likelihood estimation. If the step size of discretization is s, the probability of the change to income is

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

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 Additional Simulation Results

The following section contains additional results from the model simulations.

E.1 Contributions to the Aggregate Migration Rate By LLM Speed

Table E1 reports on an expanded version of Table 8, splitting the migration by home status series by LLM speed categories. This shows again that migration decline is concentrated among the Fast LLMs. Within each category of LLM, migration rates over time by home status are fairly steady, indicating that most migration decline comes from a switching of status from the not-at-home to the at-home population.

There are some exceptions, however. Notably, Fast LLMs show declines among the at-home (as their rootedness rises over time), and to a lesser extent, among the not-at-home, which is largely a function of the Fast LLMs being slightly more attractive destinations over time. Medium LLMs show some decline in onward migration among the not-at-home, counteracted in

Table E1: Simulations: Migration Rates by Home Status, No Age Effects

	1980	1990	2000	2005-2011	2012-2017
Panel A: All LLMs					
Move Rate	3.11	3.05	2.99	2.91	2.87
Rate by Home Status					
At Home	1.77	1.77	1.77	1.80	1.79
Not Home	5.07	4.99	4.96	4.91	4.92
Not Home: To Home	1.20	1.18	1.15	1.13	1.14
Not Home: Onward	3.88	3.82	3.81	3.78	3.78
Foreign	3.07	3.06	3.06	3.04	3.04
Panel B: Fast LLMs					
Move Rate	3.73	3.64	3.50	3.38	3.29
Move Rate by Home Status					
At Home	2.42	2.40	2.34	2.37	2.31
Not Home	5.14	5.04	4.97	4.86	4.84
Not Home: To Home	1.24	1.22	1.19	1.15	1.16
Not Home: Onward	3.89	3.83	3.78	3.71	3.69
Foreign	3.03	3.02	3.01	2.98	2.98
Panel C: Medium LLMs					
Move Rate	3.03	3.02	3.01	2.95	2.93
Move Rate by Home Status					
At Home	1.65	1.65	1.69	1.73	1.77
Not Home	5.10	5.05	4.99	4.91	4.89
Not Home: To Home	1.27	1.25	1.20	1.18	1.17
Not Home: Onward	3.84	3.80	3.78	3.73	3.71
Foreign	2.98	2.98	2.98	2.98	2.97
Panel D: Slow LLMs					
Move Rate	2.52	2.45	2.43	2.39	2.38
Move Rate by Home Status					
At Home	1.51	1.51	1.50	1.50	1.48
Not Home	4.89	4.81	4.92	5.03	5.12
Not Home: To Home	1.01	0.98	0.99	1.01	1.03
Not Home: Onward	3.89	3.83	3.93	4.02	4.08
Foreign	3.21	3.18	3.20	3.19	3.19

NOTES: All figures are in percentage points (e.g., “0.5” corresponds to a one-half percentage point change). Among the Not Home category, the returns “To Home” and moves elsewhere (“Onward”) sum to the total rate for the category. The simulation has been balanced to have constant age-education population weight over time within each LLM. (Source: Authors’ calculations.)

the aggregate by a rise among the at-home. Slow LLMs show no pattern among the at-home and slight increases in the rate of onward migration.

E.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 9 for the longer time horizon in Appendix Figure E1. 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 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.

E.3 Model Simulation Results by LLM

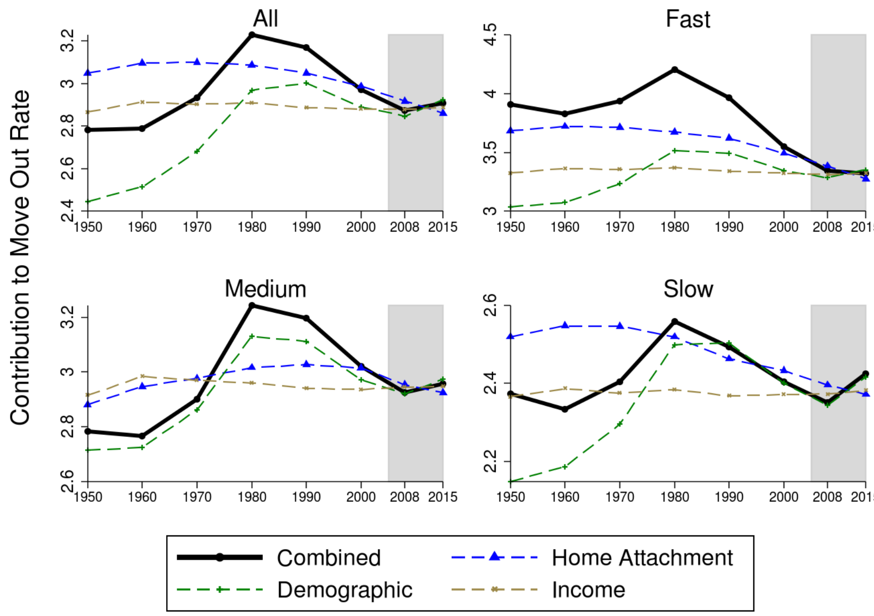
Table E2 reports the migrations changes for each separate LLM as well as the breakdown by categories of decomposition type.

Table E2: Simulated Migration Rates, With Decomposition of Change, by LLM

	Level in 1980 1	Total Change 2	Demographic 3	Home Attachment 4	Income 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: The table reports simulation results by LLM. All figures are in percentages (e.g., “0.5” corresponds to a one-half percentage point change). (Source: Authors’ calculations.)

Figure E1: 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. (Source: Authors' calculation.)

F Details on Data Construction

The following section contains additional details on the processing procedures for datasets used in this paper.

F.1 Sources

Our migration data come from two sources, the American Community Survey (Ruggles et al. (2019)) and the migration flows tables from the U.S. Treasury's Internal Revenue Service (IRS (2018)). The ACS reports the respondents' 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. For the destination choice probability used in the imputations for moment conditions (described in Section D.6), we use aggregated flow tables of ACS flows, provided by U.S. Census (2018a).

We process the 1990 and 2000 census data in the same way. These datasets are similar to the ACS, but the retrospective window for the migration inference question ('Where did you reside t years ago'?) is five years instead of one.

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. DeWaard et al. (2020) caution that the "user beware" in analyzing IRS migration data after 2011. 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, used in the calculations of roots, home status, and income distributions, were also obtained from Ruggles et al. (2019). Intercensal year population estimates were obtained from the U.S. Census Bureau webpage (U.S. Census (2018b)).

Information on location and income dynamics were provided by Institute for Social Research (2021). Estimates of relative income position by state of residence were obtained by merging reported PSID labor income with state level income distributions by year, education, and foreign-born status obtained from the Annual Social and Economic Supplement of the Current Population Survey (March CPS, Flood et al. (2018)). All incomes are deflated by the consumer price index from the Bureau of Labor Statistics (BLS (2018)).

F.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

[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. Van Riper provided particular assistance by constructing the 1960 file.

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 CZ 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 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 G3, and a dataset of the mapping of counties and PUMAs over time is available on our webpage.

Table G3 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 an 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.

Distances between LLMs were calculated using Great Circle distance from central county population centroids (U.S. Census (2011)).

F.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 20-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 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.

[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.

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 an LLM in a state with several large cities, such as Texas, to an 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 include only the U.S. born (the dotted line in Figure 5, 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.

G Data Definitions

The following appendix contains exhibits with data definitions.

Figure G1: Regional Definitions

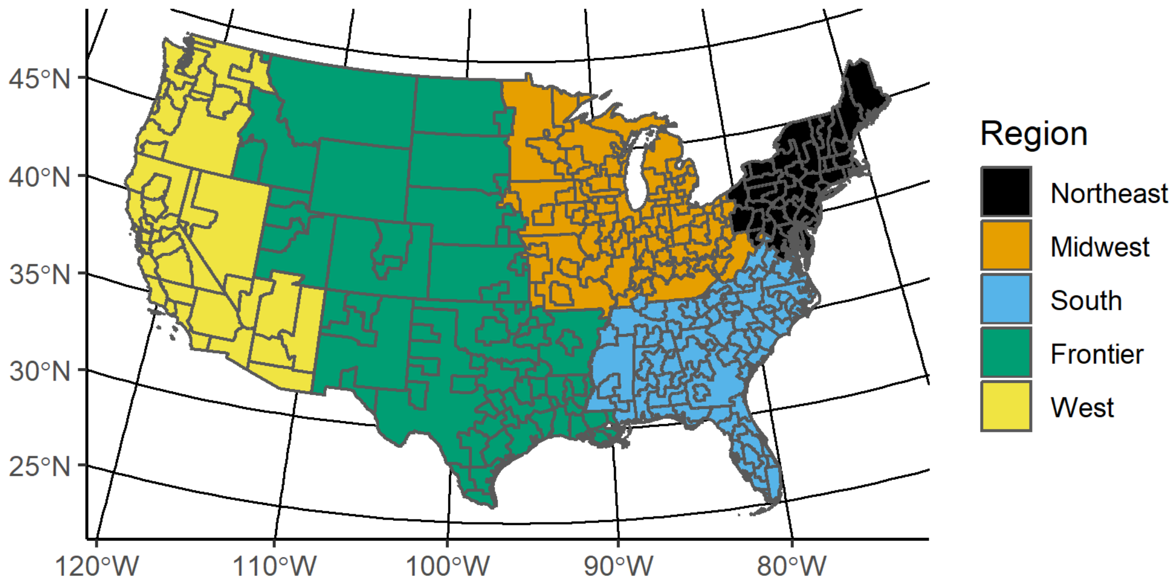


Table G1: Local Labor Markets (LLM) Definitions

			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

Table G1: Local Labor Markets (LLM) Definitions

			Included In:	
LLM	States (# Counties)	Region	Model	Descriptives
Aniston	Alabama (1)	South	No	Yes
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

Table G1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Descriptives
Dothan	Alabama (5)	South	No	No
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

Table G1: Local Labor Markets (LLM) Definitions

			Included In:	
LLM	States (# Counties)	Region	Model	Descriptives
Lansing	Michigan (3)	Midwest	No	No
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

Table G1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Included In: Descriptives
Provo-Orem	Utah (3)	Frontier	No	No
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

Table G1: Local Labor Markets (LLM) Definitions

			Included In:	
LLM	States (# Counties)	Region	Model	Descriptives
Roanoke	Virginia (11), West Virginia (1)	South	No	No
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

Table G1: Local Labor Markets (LLM) Definitions

LLM	States (# Counties)	Region	Included In: Model	Descriptives
Wichita	Kansas (5)	Frontier	No	Yes
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 G2: 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

Table G2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Charlottesville (VA)	0.50	0.78	330,316	194,059	123,881	119,790	120,754
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

Table G2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Houma (LA)	0.86	0.92	286,249	263,213	138,663	105,984	73,971
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

Table G2: LLM Population Shares At Home, Rootedness, and Population Histories

	Share	Roots of	Population in:				
LLM	At Home	At-Home	2010	1980	1950	1920	1880
Montgomery (AL)	0.68	0.87	395,483	307,620	236,046	182,783	133,791
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

Table G2: 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 Nebraska (NE)	0.64	0.81	664,629	706,759	739,460	815,105	274,698
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

Table G2: 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 (IL)	0.79	0.84	275,275	256,037	210,610	179,976	119,182
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 G3: 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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Asheville (NC)	Medium	No Decline	4.32	4.30	4.44
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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Dubuque (IA/IL)	Slow	Small Decline	4.30	3.94	4.02
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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Laredo (TX)	Slow	Big Decline	4.12	2.94	2.59
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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Provo-Orem (UT)	Omitted	Omitted	7.78	6.96	6.30
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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Riverside-San Bernardino (CA)	Fast	Big Decline	7.18	4.96	3.53
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

Table G3: LLM Categorization and Migration Rates

	Category of:		Migration Rate:		
LLM	Speed	Decline	IRS, 1991-1993	IRS, 2008-2011	ACS, 2005-2017
Waterloo (IA)	Medium	Small Decline	4.52	4.16	4.46
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