Fast Locations and Slowing Labor Mobility

Patrick Coate[*]   Kyle Mangum[†]

April 2018

Abstract

This paper explains declining internal migration in the United States using a new empirical fact: the decline is driven by locations with typically high rates of population turnover. These "fast" locations were the population growth destinations of the twentieth century, where home attachments were low but increased as regional population growth converged. Using a novel measure of home attachment, the 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 most of the decline not attributable to aging and is consistent with the spatial pattern.

1 Introduction

Internal migration rates in the U.S. have steadily trended downwards since the late 1980s. As labor mobility is thought to be one of the primary mechanisms by which regions adjust to local shocks (Blanchard and Katz (1992)), an obvious concern is that low labor mobility will result in spatial mismatch and lower aggregate productivity. The issue has garnered interest in the literature, yet explanations have been elusive. Molloy et al. (2011) and Kaplan and Schulhofer-Wohl (2017) document that compositional explanations–e.g. an aging population, a rise in two-earner households, disruptions in the housing market or other cyclical fluctuations–do not suffice because the decline is persistent and pervasive across types of households.

This paper offers a new explanation of the decline motivated by a novel empirical fact: the decline in migration is almost wholly due to decreasing out-mobility rates from places with the highest typical rates of population turnover. That is, declining mobility is the result of "fast locations" slowing down. This suggests something is differentially affecting the fast locations, and we use the spatial variation as an opportunity to study the larger trend.

We show that a distinguishing feature of fast-but-slowing locations is that these were the destinations for population growth in the twentieth century: the Sunbelt and West Coast (Chinitz

(1986), Blanchard and Katz (1992), Glaeser and Tobio (2008)). Places growing by the spatial reallocation of population have, by definition, larger contingents of transient residents. This means that they have more non-natives residents, i.e. "transplants," but also that subsequent generations–the natives born to the transplants–have thinner family and social networks than others of the same birth cohort in cities with more established populations.[1] Population growth in some of these fast locations has tapered (e.g. California),[2] although it continues in others (e.g. Texas). More generally, rates of population growth across metropolitan areas have converged since the middle decades of the twentieth century, and regional population reallocation has slowed.[3]

As a result of the settling population growth, an important trend among the fast locations is an increasing home attachment of their populations. To measure this, we introduce an index of home attachment that we term "rootedness." To proxy for the depth of social and family attachments to one's place of birth, we define a person as more rooted to her birthplace if her parents were also native to that location. We derive the index of rootedness using a cohort-based analysis of decennial census data stretching back to the post World War II era. The index varies substantially by location of birth, and also within location by birth cohort. We show that the propensity of natives to out-migrate is strongly predicted by how deep are the roots of their location of birth, but rootedness shows no association for non-natives in the location. Moreover, the rate of return migration–moving from an away-from-home location back to one's place of birth–depends on birthplace rootedness. Together, these patterns indicate that there is a utility premium associated with home that we seek to measure.

In fast locations, there are two senses by which home attachment has increased in recent decades. First, greater shares of current residents are native to fast places than were previous cohorts. Natives are drastically less likely to out-migrate, so increasing rates of nativity have led to lower average rates of out mobility. Second, in cities where population growth slowed in the late twentieth century, recent cohorts of natives are more rooted (by our measure), and therefore less likely to move away than were previous cohorts of natives. Therefore, an increasing level of home attachment could explain the decline in mobility in a manner consistent with its spatial pattern.

These two components demonstrate how the measure of rootedness serves a duel role in our analysis. First, it provides as a plausibly exogenous measure of the preference for home, which might otherwise be subject to bias from selection on unobserved idiosyncratic moving costs.

---

[1]We focus on the U.S. born population. To be clear, when we use the term "native" we mean to distinguish those born in the region from those born in another region of the U.S.

[2]See Herkenhoff et al. (2018).

[3]We must be careful to note that our study is focused on declining rates of gross mobility since 1990. Rates of net mobility have not declined during this period (Kaplan and Schulhofer-Wohl (2017)). Our explanation for gross mobility is that decades-prior changes to net reallocations have impact on more recent changes in gross mobility.

Rootedness varies by location and cohort, so we can leverage variation in the endowment of preference for home to estimate a home preference. Second, there are direct effects of trends in rootedness in explaining the out-migration trends. Places that have become more rooted show some of the largest declines in out-mobility rates.

To quantify the impact of trends in home attachment on aggregate migration rates, we develop and estimate a structural model of migration with heterogeneous locations and agents. The agents vary by location of residence, location of birth, and strength of home preference, as well as standard features like age, education, and income. Locations differ by the incomes and amenities they offer and their mobility cost. We first use a generic calibration of the model to illustrate how weaker home preferences can increase gross migration in steady state by placing more households in the more precarious away-from-home state. Moreover, endogenous home preferences will amplify gross migration rates along a transitional path to new a steady state and lengthen the transition itself. We next estimate the model using Census microdata and aggregated migration flow data from the Internal Revenue Service. The main objective of estimation is to recover parameters governing home preference, leveraging variation by location and cohort in the rootedness index. The richness of the state space of the model and the construction of the rootedness index give confidence that we estimate a utility premium for home, not unobserved move costs.

The model also accommodates complementary explanations. Trends in local labor markets, such as rising real incomes or increasing dispersion of income, could account for some locations' falling rates of out mobility. Moreover, a notable paper by Kaplan and Schulhofer-Wohl (2017) posits that increasing information availability has caused the decline in mobility. We consider a variant of this explanation that could be consistent with the spatial pattern in the mobility decline: that increasing information availability has improved the relative value of searching within a location versus searching across locations, an effect larger in locations with higher income dispersion, another feature of mobile markets. The model can quantify these effects directly and in interaction with other trends.

The estimated model allows us to conduct counterfactual simulations to do an accounting of the sources of mobility decline. Quantitatively, the effect of a greater share of population being at-home is large, accounting for about 40 percent of the observed mobility decline, and most of the decline not accounted for by the aging of the population.[4] The home attachment effect is especially large in the mobilest and least rooted cities, accounting for nearly 60 percent of the decline in shallow-rooted locations. The direct effects of changes in the rootedness index– i.e., the strength of home preference increasing *within* some of the fast locations–is important for explaining changes in those places (California, South Florida), but smaller in the aggregate

---

[4]Statements about "shares" of the decline roughly indicate the magnitude of the relative effects, but are somewhat loose because factors in the model can interact in nonlinear ways.

because the within-trends are present in only a subset of the fast locations. The effects of changes in income distributions and information availability are small in aggregate, although these materially contribute to changes within a small subset of the fast locations (parts of Texas). More generally, we see these effects as the type of primitive change that can be amplified through preference for home, so we hesitate to discount their role. But by including them in the model, we are able to conclude that their direct effects are not first-order.

We conclude that declining gross mobility is a natural consequence of the spatial population trends of the twentieth century U.S. Changing regional populations gave rise to weak home attachments that propagated over several generations. After population reallocation converged, rising home attachments followed and resulted in slowing rates of gross mobility.

The policy implications of this result are mixed. On balance, we do not see the slowing mobility rate as particularly alarming, because like Kaplan and Schulhofer-Wohl (2017) (but for different reasons), we find it to be the outcome of workers optimally choosing not to move. However, the role of home preference does raise the possibility of a disconnect between labor market outcomes and family and social utility outcomes. One implication is that place-based policy could be warranted when home preferences are stronger or more prevalent, as such policies might be necessary to impact households far from the margin of relocation.

Several researchers have offered explanations for the decline in mobility. Partridge et al. (2012) suggest that mobility was artificially high in the postwar U.S. as population reallocated from the Northeast and Midwest to the West and South, but now the U.S. is reaching a new long run spatial equilibrium. To be clear, that study focused on shorter term net mobility (population changes), not the long run trend gross mobility decline documented by Molloy et al. (2011). Similarly, recent work by Ganong and Shoag (2013) has documented a decline in regional income convergence, but does not study gross mobility rates. We nonetheless find these highly relevant, in that related trends might affect gross mobility. Kaplan and Schulhofer-Wohl (2017) posit that converging occupational returns and increasing information availability (through informational technology like the internet and ease of travel) has led to reduced incentives to migrate, quantifying their mechanisms in a two-location, two-sector model of migration and occupational choice over the life cycle.[5] Karahan and Rhee (2014) suggest that an aging workforce has led firms to recruit more workers locally (including the young), resulting in a substitution of local search for national search, demonstrating the concept in an island economy job search model. Hood (2013) finds evidence of decreasing variance in regional shocks.

We cannot reject any of these theories; to the contrary, we make substantial efforts to incorporate versions of these mechanisms into our model. However, none of the previous literature has documented or attempted to explain the geographic heterogeneity in the mobility decline.

---

[5]Kaplan and Schulhofer-Wohl (2017) also document a decline in repeat and return mobility, which they ascribe to fewer "experimental" moves requiring reversal. Changes to home preference could also cause a decline in repeat mobility, especially given the large share of moves that are returns to home.

This is a notable omission when attempting to explain *"the"* decline, when actually, in the data, some locations have slowed and others have not. Moreover, the mechanism we propose can amplify secular changes to gross mobility or overhang trends in net mobility, and therefore complementary explanations nest into our general framework.

There is a larger literature on declining labor market "dynamism" besides locational mobility, including job creation/destruction and occupational switching (see Davis et al. (2012), Molloy et al. (2014), or for a counterpoint, Kambourov and Manovskii (2008)). Our emphasis in this paper is clearly on geographic mobility, motivated by a striking spatial heterogeneity, and we do not directly study these other forms of labor market flows. However, we hope the mechanism and modeling techniques we propose could subsequently inform that literature as well.[6]

In our empirics and methodology, we contribute to the literature on household migration and frictional location choice. Past literature on household migration has found strong proclivities to live in the region of one's birth or childhood (Bayer et al. (2009), Kennan and Walker (2011), Coate (2014)).[7] We introduce a new measurement of home attachment that helps to distinguish between household moving frictions and preferences for past locations, which, due to data limitations, are often conflated (Bayer et al. (2009), Diamond (2016), Morten and Oliveira (2016)). By leveraging spatial and cohort variation in home attachment, the model can be estimated without longitudinal data, frequently a constraint on studies of local labor market migration. Finally, we contribute to work on the estimation of dynamic choice models using conditional choice probability (CCP) estimation (Hotz and Miller (1993), Arcidiacono and Miller (2011), Bishop (2008), Ma (2017), Davis et al. (2016)). We derive for our model a linear method of moments estimator that is highly tractable despite a large state space. This is, to our knowledge, the first implementation of CCP estimation on aggregated choice data.

The rest of the paper proceeds as follows. In Section 2, we describe our data and present a set of stylized facts about internal migration in US cities to motivate the model. In Section 3, we write down a model of migration based on labor market opportunities and home preference. Section 4 conducts generic simulations of the model to illustrate the role of home preference on aggregate migration rates. In Section 5 we set out our estimation strategy and report its results. In Section 6, we conduct counterfactual simulations to quantify how the proposed mechanisms explain recent patterns of migration decline. Section 7 concludes with interpretations for future research and policy. Appendices provide additional detail on data construction (A), derivation of the estimator (B) and its targets (C), and counterfactual simulation results (D).

---

[6]Initially, our goal in embedding an income search model within a geographic location decision was to address the interaction of geographic and job mobility. However, we found no datasets sufficient to speak simultaneously to both issues.

[7]Recently, Zabek (2018) has examined the role of "local ties" (i.e. residing in the location of birth) in determining welfare effects of local labor market shocks and place-based policies, but does not address the decline in mobile or the intensity of the social ties for locals.

2 Empirical Facts

2.1 Data Description

We begin by establishing the motivating empirical facts by employing four types of datasets: two county-level aggregates and two sets of microdata. (When applying the structural model to data, we supplement with some additional data sources to be described below.) The county level data include population estimates from the decennial U.S. Census from 1900 to 2010 and internal migration flows data for 1990 to 2014 from the U.S. Department of Treasury, Internal Revenue Service (IRS). The population data allow us to track the time path of metropolitan area sizes. The IRS data provide migration flows, publishing county-to-county flow aggregates of tax returns filed and exemptions claimed (close approximations of households and persons, respectively), as inferred by the change in tax address. By using county-level data, we are able to form constant boundary definitions of local labor markets, such as Commuting Zones (CZ). We also use census microdata: the 2007-2011 samples of the American Community Survey (ACS), and public use decennial census samples dating from 1940 to 2000 (Ruggles et al. (2015)). The ACS is our main source of information for current local labor market attributes. The older decennial census data is used to build a measure of historical population social connectedness that we describe in more detail below.

Both the ACS and IRS data contain information on migration flows. The IRS dataset provides gross flows at a wide geographic coverage at a long time horizon. For each year, we observe the flows out of and into every U.S. county,[8] so we can produce a time series for every local labor market we define, forming a panel data set of metro level migration rates which permits a study of geographic heterogeneity in migration and its trends.[9] The disadvantage of the dataset is its aggregation, so we observe no demographics or labor market events. We cannot build any profile of movers from this data by any typology other than space.

The ACS data also report migration flow by recording the respondent's current public use microdata area (PUMA) as well as their last PUMA of residence one year ago.[10] From this, we can construct gross inflows and outflows for origins and destinations, and a more limited version of place-to-place flows, because the geographic coverage is limited considerably by sample size,

---

[8] The IRS data censor any origin-destination cells with fewer than 10 migrating households; thus, we do not observe place-to-place flows for many location pairs. However, the data contain total outflows and inflows for each county, meaning the censoring affects only county-pairs and not calculation of total gross flows. Also, the censoring is less severe for the large and mobile metro areas on which we focus our analysis.

[9] Kaplan and Schulhofer-Wohl (2017) mention that state to state flows in IRS are slightly smaller than in the CPS, and with a flatter trend, perhaps because the IRS data would capture fewer low income (and hence non-filer) households who move at higher rates.

[10] We focus on the ACS rather than previous Census public use data in part because the move frequency (one year) matches that of the IRS. Previous Census samples record the five-year migration rate. It is problematic to convert one-year rates to five-year rates since a large proportion of move events are repeated (onward to a new location) or reversed (back to a previous location) (Kennan and Walker (2011).)

whereas the IRS data are effectively a census of taxpayers. While the time variation is very limited and too short to examine the migration trend, the ACS does provide demographic detail. Important variables include age, education, and birthplace. We also use the ACS to derive estimates of income and population cell sizes for input into the structural model.

Our initial stylized facts will use the full set of 271 metro areas available in the IRS and decennial Census data. Our more detailed analysis of social connectedness and the estimated structural model will focus on a subset of 54 large metro areas (plus one residual location).[11]

## 2.2 Mobile Cities and Migration Decline

The decline in migration has been documented by several studies using different datasets and definitions of migration, and we do not attempt a full rehashing here. Our aim is to document a fact new to the literature–that the migration decline has centered on the highly mobile cities. Afterwards, we describe some key features of these mobile locations and what such features suggest regarding mechanisms behind the decline in mobile cities.

As an initial matter, we show that it is fair to characterize cities as "mobile" or "immobile" because locations with high degrees of in-migration also exhibit high out-migration. This fact is not new–it was recognized in studies from Ravenstein (1885), Miller (1973), and Sjaastad (1962) to more recently, Coen-Pirani (2010) and Mangum (2016)–but for completeness we demonstrate it holds in our datasets. Figure 1 displays two scatterplots of out-mobility to in-mobility rates. The lefthand plot reports the total in and out rates for 271 metro areas in the U.S. using the IRS data for 2007-2011. The strong positive relationship is evident, with a correlation coefficient of 0.93. Moreover, these rates are highly persistent over time: the rank correlations of mobility rates for any pair of years from 1991 to 2014 are upwards of 0.88. The righthand plot reports the in- and out-rates for demographic subgroups in decadal age bins from 20s to 50s, for college educated and non-college educated, using 2007-2011 ACS data in our selection of 54 large metro areas. The strong positive association is present within population groups as well. Hence, the positive correlation between inflows and outflows is not due to the substitution of one population subgroup for another.[12] A location's designation of more or less mobile is reference to population *turnover*; hence, "fast" or "slow." Although much of our empirical model focuses on out-migration rates, the emphasis is on gross mobility and not, for instance, a city losing population. There could be many reasons for the spatial pattern in population turnover (as studied in more detail by Coen-Pirani (2010) Mangum (2016)), and we will highlight features of fast locations below.

---

[11]The 54 locations were chosen starting with the 50 largest MSAs by population in 2010, although we added and subtracted a few at the margin of the 50 ranking in order to have a sufficiently diverse set of locations (by mobility rates, population growth, and income dispersion).

[12]Coen-Pirani (2010) reports a high degree of correlation even when using much finer subgrouping cells, including occupational categories, at the geography of U.S. state.

Figure 1: In and Out-Migration by Metro Area
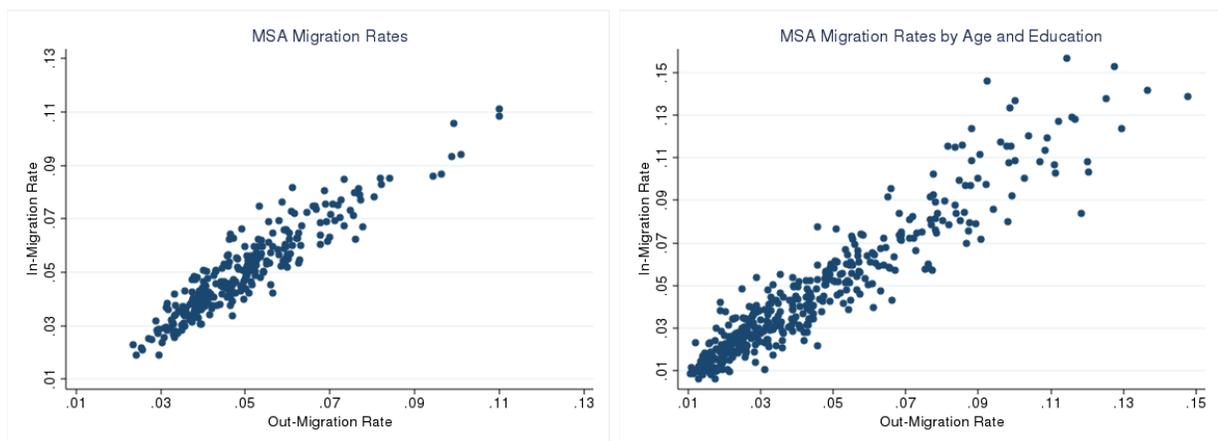
NOTES: The figures plot in- versus out-mobility rates for all US metro areas (IRS data, left) or for metro/decadal age/education bins for large metros (ACS data, right).

The fast/slow distinction is important for introducing our primary empirical fact: the decline in mobility is occurring almost exclusively in the most mobile markets. The plot on the left of Figure 2 compares the change in mobility rates from the first half decade in the data (1991-1995) to the last (2010-2014) to a mobility index for each metro area created from the full panel of IRS flows data.[13] These exhibit a strong negative correlation. The righthand plot in the figure shows the times series of annual out-migration rates for metro areas split into terciles by their mobility rate. The most mobile third of cities show a strong downward trend, dropping from about 5.7 percent to 4.6 percent from 1990 to 2014 (a 20 percent decline). The change for the middle third was much smaller, declining from about 4.28 to 4 percent (a seven percent decline). The least mobile third showed essentially no decline.[14]

We have also used the geographic detail in the IRS data to parse the migration flows network of origins to destinations, allowing us to check whether the slowing out-mobility is associated with to lower inflows to destinations particular to high mobility origins. While the networks of higher and lower mobility cities are not identical, the trends from each type of origin by destination are remarkably similar. (Full results available on request.) Thus, the decline in migration seems to be the result of more staying, generally, in high mobility places, and not apparently due to reduced inflows to any particular type of destination.

---

[13]The index is derived from a gravity regression on location-to-location flows in the IRS data. The gravity regression controls for original and destination size and distance between markets, and take origin and destination fixed effects. The origin fixed effect–the residual amount of out mobility after controlling for the location's place in the network–is used here as the mobility index. The main impact of the index is to make mobility rates "size-neutral." The results are similar when simply using raw mobility rates.

[14]The differences in trends are statistically significant in regressions of out mobility on time and location fixed effects.

Figure 2: Changes in Migration Rate Over Time
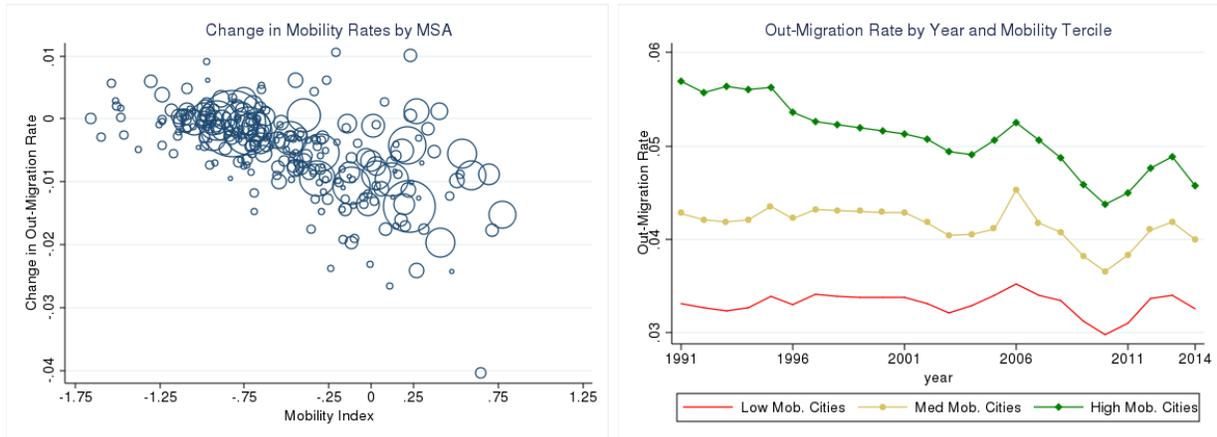
NOTES: The left figure shows the change in metro level out-mobility compared to initial mobility rates; bubble size is proportional to metro population. The right figure shows the annual out-migration rates by initial mobility tercile. (Source: IRS data).

2.3 The Features of Fast Locations

These empirical patterns invite two questions: what is different between fast and slow locations? And then, what is changing? We proceed by describing some features of these markets that are potentially important for mobility patterns.

2.3.1 Population Shifts

Which locations are mobile? Table 1 lists the 54 major cities in our main analysis, ranked by a mobility index.[15] The fast locations tend to be in the Sunbelt and West Coast–with California, Florida, and Texas cities, plus Las Vegas, Seattle, and Phoenix, among the top– while the slow locations are mostly in the Northeast and Midwest. These are not exclusively declining postindustrial Rust Belt cities, however. New York, Boston, and Chicago, vibrant cities in many regards, also exhibit low rates of mobility.

The first distinguishing attribute is that fast locations were the cities growing most rapidly in the 20th century. Using the decennial censuses from 1900 to 2010, we constructed population sizes of constant-geogrpahy metro areas by aggregating the county level estimates.[16] Figure 3 shows the share of 2010 population living in the metro area in prior decades, grouping metro areas by mobility tercile, with the thirds corresponding to one third of U.S. urban population as of 2010, not necessarily counts of cities (although the results are similar either way). Each category of city has obviously grown in size as the U.S. population has grown and urbanized. However, the more mobile locations have grown faster and more recently. In 1950, the least

---

[15]The index derives from origin and destination fixed effects in a gravity regression model. This means the index is distance-adjusted.

[16]This analysis account for counties that split or merge by aggregating to the highest common denominator in terms of spatial area.

Table 1: Move Rates by Metro Area

Mobility Index Rank	FIPS	Name	ACS	IRS	Index (from IRS)
1	7320	San Diego	5.52	5.25	0.77
2	4120	Las Vegas	5.60	5.32	0.70
3	6200	Phoenix	4.14	4.25	0.59
4	7600	Seattle	3.64	3.72	0.54
5	8520	Tucson	5.94	4.83	0.51
6	8280	Tampa	3.99	4.80	0.27
7	4480	Los Angeles-Riverside	2.95	2.60	0.23
8	5000	Miami	5.24	6.95	0.21
9	7400	San Jose	5.48	4.82	0.19
10	5960	Orlando	5.49	6.11	0.18
11	7360	San Francisco	4.25	4.01	0.12
12	200	Albuquerque	4.44	4.21	0.06
13	640	Austin	4.64	5.27	0.03
14	2080	Denver	4.06	4.46	0.03
15	6440	Portland	3.21	3.98	0.01
16	7240	San Antonio	4.23	4.02	0.00
17	6920	Sacramento	4.63	4.17	-0.02
18	3590	Jacksonville FL	5.06	5.22	-0.03
19	1920	Dallas-FW	4.68	4.76	-0.12
20	5720	Norfolk-VA Bch	5.68	6.24	-0.13
21	7160	Salt Lake City	4.19	4.00	-0.14
22	8840	Washington DC	4.60	4.46	-0.30
23	3360	Houston	3.52	3.43	-0.32
24	520	Atlanta	4.69	4.57	-0.40
25	5880	Oklahoma City	4.29	3.90	-0.45
26	6640	Raleigh-Durham	4.45	4.67	-0.46
27	1120	Boston	2.81	3.12	-0.48
28	4920	Memphis	3.80	3.62	-0.56
29	1520	Charlotte	4.03	4.44	-0.62
30	3280	Hartford	2.27	2.94	-0.64
31	3760	Kansas City	4.16	3.56	-0.66
32	5120	Minneapolis	2.51	3.09	-0.68
33	5360	Nashville	4.45	4.23	-0.70
34	1280	Buffalo	2.54	2.47	-0.71
35	1600	Chicago	2.84	2.69	-0.73
36	1840	Columbus	3.46	3.64	-0.74
37	720	Baltimore	3.39	3.92	-0.76
38	2160	Detroit	1.32	2.85	-0.78
39	5600	New York	2.98	2.90	-0.82
40	160	Albany	3.16	3.22	-0.82
41	2000	Dayton-Springfield	3.69	3.56	-0.83
42	3480	Indianapolis	4.54	3.51	-0.88
43	5080	Milwaukee	3.20	3.18	-0.88
44	1000	Birmingham	4.36	3.52	-0.89
45	6160	Philadelphia	2.27	2.69	-0.91
46	6480	Providence	3.08	3.31	-0.93
47	7040	St Louis	2.55	2.95	-0.94
48	6760	Richmond	3.68	3.89	-0.97
49	3120	Winston-Salem	1.61	4.30	-0.98
50	1680	Cleveland	2.57	2.95	-0.99
51	6280	Pittsburgh	2.39	2.30	-1.04
52	4520	Louisville	3.20	2.89	-1.09
53	1640	Cincinnati	2.97	3.00	-1.10
54	3160	Greenville SC	3.83	3.49	-1.14

NOTES: The table lists mobility rates for metro areas sorted by their implied residual mobility index, as measured by origin and destination fixed effects in a gravity regression with distance and size controls. Source: IRS and ACS.

Figure 3: Metropolitan Area Population Accumulation Over Time, by Mobility Rate
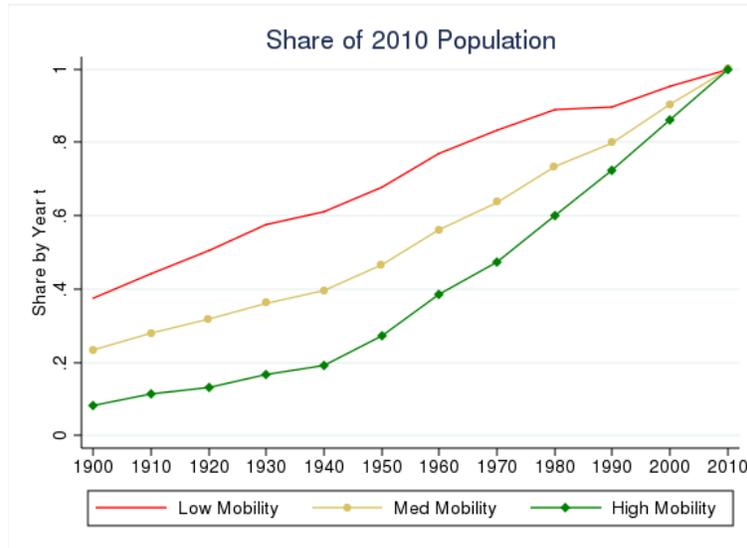
mobile third had already acquired over three-fifths of their 2010 populations, while the highest mobility tercile had acquired just one-fifth. The rates of postwar growth are far steeper in the fast locations.

Population growth rates have equalized in recent decades, however. Figure 4 plots the cross-sectional standard deviation (i.e. across metro areas) of log population growth rates at a decadal frequency. The figure shows a clear downward trend, even restricting to the post World War II era. The variance in growth rates in the 1990-2010 period was about half that of the 1950s and 1960s. Population reallocation across space has therefore "settled down" since the major shifts of the mid-late 20th century.

What implications might the trends in the spatial distribution of population growth have for current gross mobility rates? As the previous literature has shown, repeat mobility is common. That is, people living in their "home" locations are far less likely to migrate than those away from home.[17] Locations with more recent arrivals will have higher mobility rates simply by having a larger share of more precariously situated "transplanted" households.

To test the connection, we measure for each metro area the share of individuals who arrived from elsewhere and the comparison of move rates by past migration status. To do this, we need *three* locations–current and last, to measure migration, and one location prior to this to indicate whether the household had arrived to the origin from elsewhere. The ACS contains only one possible move event (the current and last location), but it does report the respondent's birth state. Thus, we use birth state as a proxy for the respondent's "home." (In this and all analysis

---

[17]This point is a focus of the model in Kennan and Walker (2011) and the primary mechanism generating inflow/outflow correlation in Coen-Pirani (2010).

Figure 4: Variance in Metropolitan Area Population Growth Rates Over Time
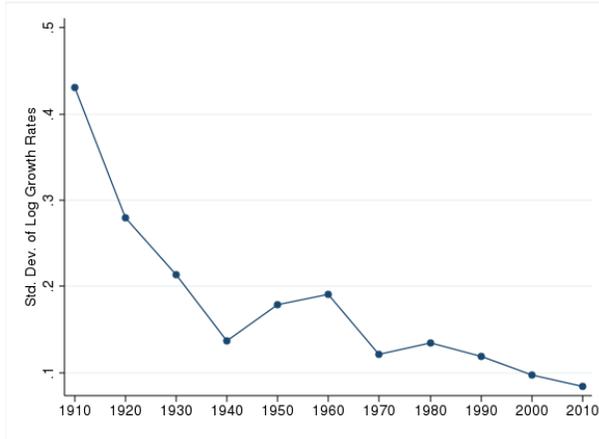
NOTES: The figure reports the unweighted cross sectional standard deviation in log population growth rates by metro area. Source: Authors'
calculations using decennial census county population estimates.

that follows, we consider only U.S.-born population so that "home" is not foreign.) An obvious
but unavoidable limitation is that some persons will have little connection to their birth state,
especially if they moved early or often as a child. Neither do we observe intervening moves
in their adulthood. In either case, however, such would bias the analysis toward finding no
importance of home location.[18]

Table 2 reports shares residing at home and migration rates by home status, split by age
and education (less than a college degree, "non-college," and college degree or more), for fast
and slow locations. A large fraction of people reside in their home state (see also Bayer et al.
(2009)), and there are well known differences in migration propensity over the life cycle and by
educational attainment that apply to both home/not-home categories. But there are much larger
differences in the propensity to migrate between home statuses throughout age and education
categories: those not at home are three to six times more likely to change their metro area of
residence. The table shows that high mobility locations are different from low in two ways.
First (and somewhat mechanically), they have a greater share of residents who are not natives,
meaning some of the difference in out-mobility rates can be attributed to having more in the
high mobility risk group (see Coen-Pirani (2010)). Second, their natives are more likely to move
away (comparing columns 2 and 6), a feature to which we return later.

---

[18]One challenge in these data is that we use metropolitan area as the local area market, but the birthplace
is by state. For metro areas that cross state lines, birth in any state of that metro area is recorded as being
"at-home." For example, a New York City resident born in New Jersey is home, even if they currently live in New
York state. For states with multiple metro areas, we count anyone born in the state as being at home, though
they may have been born in a different metro area in that state. For example, a Los Angeles resident born in
California is home, though some of these persons may have been born in San Francisco or San Diego, etc. For
estimation of the structural model, we assign weights to observations to adjust for the potential misclassification
of residents in large states as being home when they are actually from another metro area. Note, however, that
the two states where at-home status is most likely to be biased upward, California and Texas, still show much
lower at-home rates.

Table 2: Move Rates by Age, Education, and At-Home Status

City Type:	1	2	3	4	5	6	7	8
		Low M	obility			High M	obility	
	At Home Pop. Share	At Home Moved	Not Home Moved	Move Home |Move	At Home Pop. Share	At Home Moved	Not Home Moved	Move Home |Move
Non-college, by age								
20s	77.37	2.38	15.46	59.16	70.75	3.03	14.14	49.82
30s	73.26	1.73	8.89	55.59	65.74	2.27	8.90	46.42
40s	75.52	0.88	6.52	50.90	61.21	1.42	5.78	45.80
50s	75.27	0.68	4.49	47.94	56.21	1.07	4.14	45.85
College, by age								
20s	65.35	4.71	16.20	48.78	54.94	4.65	13.89	46.82
30s	60.55	1.83	7.30	37.40	49.60	2.41	7.43	35.30
40s	60.05	0.81	3.66	30.15	46.66	1.12	3.72	34.72
50s	59.41	0.82	2.97	30.45	43.78	0.95	3.04	36.04

NOTES: The table reports out-mobility rates (in 2010) by for residents flagged as being at-home or not-at-home; that is, in a metropolitan area of their state of birth. Source: ACS data. The low and high mobility metropolitan areas are classified by the index resulting from a gravity regression on IRS data.

Together, all of these facts suggest a role for the 20th century population reallocation in the recent downward trend in mobility in the U.S. Locations that grew precipitously in decades past have had more residents living away from their native home, and hence had greater rates of outmigration by having a more fragile, less-connected population. That the major decliners were West Coast and Sunbelt locations, where population grew most in the latter half of the 20th century seems to support this notion. As population growth stabilized, more generations of people were "from California," for instance, and thus had lower propensities to out-migrate.[19]

To be clear, our point of interest is the downward trend in gross migration since the late 1980s, which concerns not population changes per se but flows in all directions. Our emphasis in this paper is not that population reallocation has slowed (though the trend in Figure 4 seems worthy of future research), but that historical population shifts could generate lower rates of idiosyncratic gross mobility by placing more households in the less precarious at-home status. That is, the "new long run spatial equilibrium" a la Partridge et al. (2012) in net mobility would eventually generate lower rates of gross mobility.

However, this itself could be subject to reverse causality: that at home status is the product of failure to migrate, so any secular factors affecting migration likelihood would also generate at-home status. We have two ways of addressing this concern. First, in many locations the stabilization of population growth occurred long enough ago that the intergenerational deepening of roots provides a way to measure home preference and trends therein. Our analysis of home preference in the next subsection will show this. Second, we will show that at-home status can amplify other factors that reduce out mobility. This will be more evident in the context of a

---

[19]The timing of a city's population growth and its mobility decline is also informative. In a regression of mobility decline on past growth rates, the population share added in decades 1950-1980 is most predictive of a drop in mobility 1990 to 2010. (Results available on request.) That is, the mobility decline is most evident in places whose populations were established one to two generations prior to the one exhibiting a decline in migration. This fact motivates our analysis of rootedness.

location choice model.

2.3.2 Home Preference and "Rootedness"

It is important to know whether the difference in migration propensity by at-home status is because home affords some positive utility flow or whether it merely proxies for some unobserved move cost. That is, the difference in migration propensities between columns 2 and 3 (and 6 and 7) of Table 2 could be selection of people who would rarely move for any reason and those who are unobservably more apt to relocate.[20] If home offers a utility premium, however, then being in an at-home state places the agent at a higher threshold for a migration decision.

The literature has consistently found some preference for home locations (Kennan and Walker (2011), Bishop (2008)), including, importantly, in dynamic models of migration that distinguish between transitory "move costs" and inclinations for an initial location.[21] The interpretation is often that family and social networks are richer at home, and one's initial location could be formative in setting preferences and habits. While there may be a distribution of preference for home, it seems plausible that the average person would prefer their home to a ceteris paribus location.

We see evidence of a preference for home in two features of our data. First, patterns in return migration indicate a preference for home. While the difference in move rates seen in Table 2 between persons at home and not-at-home could be explained by the latter being a selected sample of those with lower mobility costs, columns 4 and 8 of the table show that a disproportionate share of moves made by these persons–from one quarter to almost half, depending on age and education–are returns back to their birthplace.[22] Given a large set of alternatives, many of them return home, indicating their preference for that place relative to all others.[23]

Second, there is more direct evidence of utility for home in the heterogeneity in locations. We noted earlier from Table 2 that native migration is more common out of high mobility places, but repeat mobility from non-natives is not. Why should this be? The long run population trends offer clues. Differences in mobility rates among natives of fast and slow locations could be due to differences in the "rootedness" of their native populations. We think of individuals having "deep roots" in their home location if they have a dense network of family relations in the area (and

---

[20]The data are already split by age and education, but there could be unobserved household factors, including psychic costs.

[21]In static/myopic models of migration (see Bayer et al. (2009), Diamond (2016)), the distinction between preference and transitory utility cost is muddied, but still preferences for home are detectable.

[22]The life-cycle pattern is also suggestive. Younger persons away from home are actually more likely to return, which would be consistent with a higher flow utility (which they have more periods to draw) being obtain in their homes.

[23]Another way to see this is to condition on the destination and compare the entrants' characteristics. Conditional on moving into some location X, the entering household is far more likely to be from X than from somewhere else.

more generally, a more developed social network). If home is preferable because of social and family networks, stronger connections produce stronger preferences, and hence shifting spatial populations might impact migration incentives for several generations. For example, given post-war U.S. population trends, a Boston native is far more likely to be born to parents who were also Massachusetts natives than a Los Angeles native is likely to be born to native Californian parents, and hence may have a deeper family network because she has more successive generations in her home state. The at-home Los Angelan may be more likely to migrate than the Bostonian when faced with ceteris paribus migration incentives.

To empirically test this idea, we construct a measure of such rootedness that can reach across many locations and points in time. To do so, we employ the decennial Census microdata back to 1950. We define "rootedness," as in the example, to be the probability of being born to parents native to one's own location of birth. State of birth has been a question in the decennial census along with PUMA of residence since 1950. A crosswalk of PUMA codes to Commuting Zone allows us to match constant-definition metro areas back to 1950.[24] We apply a measure of rootedness for current generations by matching birth cohorts for each metro area 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 living in the Boston metro area in 1980, 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 our measure of rootedness is matched by cohort and birthplace to current data, yielding a measure of how highly attached to Boston is someone born there some t years earlier, even when possibly living in other locations by 2010.

There are major advantages of using rootedness as a measure of home attachment. First, it offers a source of variance that a simple at-home indicator cannot. Some locations are more rooted than others, and some generations are more rooted than others. Second, it is more plausibly exogenous (predetermined, at least, 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, in large part because we do not actually observe the rootedness of the 2010-era household and instead match by cohort. But it is readily constructible for a wide swath of geography, and we suggest it as a reasonable proxy for the deeper concept of "attachment." If it is excessively noisily measured or meaningless, then we will not detect an impact.

Actually, we find this measure has strong predictive power. Figure 5 shows two scatter plots of mobility rates to our measure of rootedness in 2010. The lefthand plot shows out-migration rates of natives (i.e. those living at-home) to the rootedness of the location, a weighted average

---

[24]The crosswalk was made available to IPUMS by David Dorn.

15

Figure 5: Migration Rates and Local Rootedness by At-Home Status
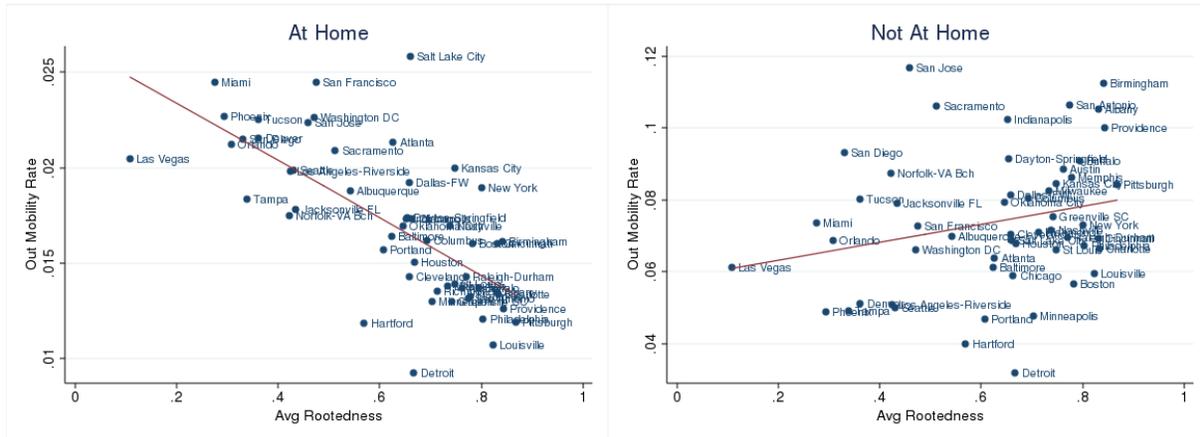
NOTES: The figures plot average out-mobility rates (in 2010) by metro area for residents flagged as being at-home or not-at-home; that is, in a metropolitan area of their state of birth. Source: ACS data.

across cohorts. There is a clear negative correlation–more rooted places have less out-migration from natives. The righthand plot does the same for metro area residents not at home, showing a weakly positive correlation. This servs as a placebo check on our rootedness measure, indicating rootedness measures something about the natives of the location and not the location itself. The premium is statistically significant in a regression and robust to inclusion of individual and location controls. Moreover, the probability of a return-home outcome is also statistically associated with the strength one's home rootedness. (Results available on request.)

This evidence indicates that home status affords some meaningful utility. The implication is that trends in home status can affect gross mobility rates. The estimation of the model, below, provides a more comprehensive measurement of the rootedness utility premium by accounting for other choice-relevant factors such as local labor market opportunities and amenities, move costs, expectations, and the variance in the choice set.

If rootedness matters for mobility, then past trends in population growth could affect current trends in migration rates. The generational lag in population growth and migration decline occurs as the population "settles in" to home and becomes more rooted. Figure 6 displays a plot (on the left) of the change in mobility to the change in the share at-home. Virtually all declining places show an increase in the share of residents having a home attachment. With a measure of the utility premium of home, one can quantify the impact this would have on the cities' out-migration rates. The plot on the right shows metro area's change in mobility to the change in the city's measured rootedness.[25] The rootedness measure has changed little for most cities, but has deepened in several Sunbelt locations, especially California. Much of the migration decline accounts from these cities, so the figure exhibits a negative correlation.[26] A

---

[25]The pattern is very similar if reweighting the cohorts to match a constant composition by age.

[26]We have would also liked to test whether the migration of natives within these locations has decline since circa 1990, but doing so would require converting a five-year rate in older censuses to a one year rate in the ACS,

Figure 6: Change in At-Home Share, Roots, and Mobility, 1990-2010
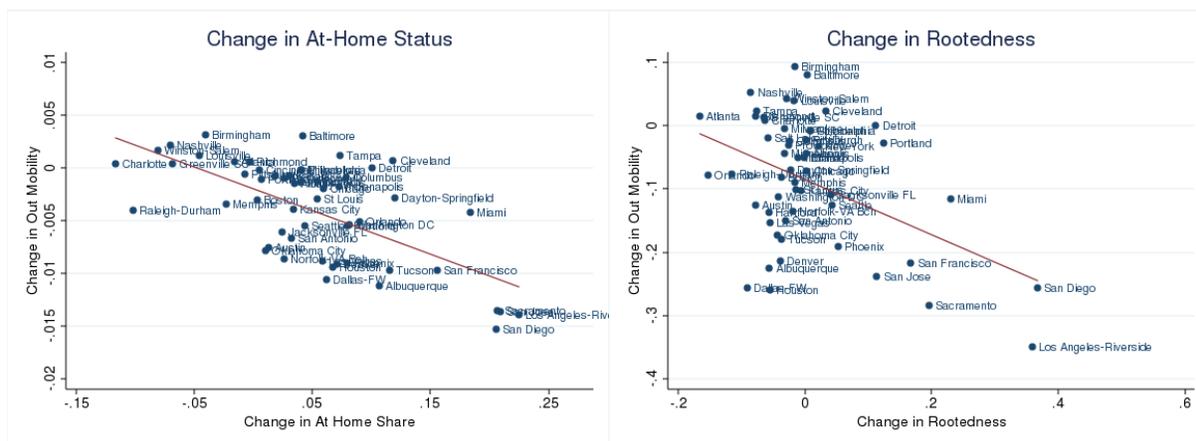
NOTES: The figures plot change in mobility rates (IRS data) to changes in share residing at home (left, Census and ACS data) and changes in rootedness (right, authors' calculations on historical Census data).

few exceptions include cities in Texas, Denver, and Las Vegas, where population growth was somewhat high in the mid-late 20th century but also continues today, so that by our measure, rootedness has not deepened.

In summary, the spatial heterogeneity in mobility trends suggests the decline in gross mobility is a natural consequence of the slowing population reallocation of the 20th century: a population shift "hangover." Deepening roots would cause less out mobility, which causes more at-home status, which creates less out-mobility and further deepens roots, and so on. The decline in gross mobility is then a consequence of the spatial distribution of population moving along the transitional path to a new long run steady state.

2.3.3 Local Labor Market Attributes

Before proceeding, we must recognize some of the other differences in fast and slow cities. In particular, the analysis thus far has had little to say regarding the metro areas as local labor markets. The literature has long recognized the importance of income opportunities for determining migration incentives, so an analysis of migration would not be complete without a consideration of differences in the local labor market conditions. For brevity, we simply summarize some of the key patterns.

Higher mobility cities tend have more disperse income distributions, even after conditioning on worker attributes (such as household structure, education, and industry and occupation). The correlation is strongest among the non-college educated. There is a weak correlation between mobility rates and metro average incomes (also when conditioning on observables). (Mangum (2016) discusses these facts in more detail.) Therefore, mobility declines are somewhat associated with higher mean incomes, but especially with higher income dispersion, although there are

_____

which is problematic since many moves are reversed (repeated or returned).

several notable exceptions. For example, New York City has high income dispersion but no mobility decline.

Nevertheless, it is possible that these features of the income distribution are relevant for the decline of out-migration from these high mobility locations. This speculation is a variation of the proposed explanation in Kaplan and Schulhofer-Wohl (2017), increasing information availability leading to less idiosyncratic mobility, but specifically tailored to the labor market and in light of the spatial variation in the mobility decline. Information about labor market opportunities would be especially valuable in cities with higher earnings. And, to the extent that the earning dispersions is indicative of income uncertainty (temporally) and/or location specific match distributions (statically), increasing information availability may disproportionately improve high income dispersion locations. As information becomes more readily available, workers in these locations are better able to find good matches and avoid negative income shocks without migrating away.[27] Moreover, increasing information availability would change the entire forward-looking decision problem of an individual income maximizer. The option value of remaining in a high income dispersion location could increase relatively more, and hence bias towards high dispersion/high mobility locations. We account for these concerns by incorporating a income search component in our model in which the return to searching in a particular market is a function of information availability and the income distribution of that market.

When comparing *trends* in income distributions to the mobility decline, no clear picture emerges, although certain individual cities with above average growth in income have shown mobility declines, including several in Texas. Generally, we take this as evidence that it is important to treat each local labor market as its own choice problem. That is, we explicitly account for the local market heterogeneities, whereas Kaplan and Schulhofer-Wohl (2017) use a stylized two-location model.

2.3.4 Other Hypotheses

Finally, we document two potential candidates that are *not* associated with the mobility decline. Figure 7, in the lefthand plot, shows that the decline is not related to market size.[28] There are both large and small declining and non-declining cities. As one example, Los Angeles has slowed but similarly-sized Chicago has not. Neither have these cities disproportionately

---

[27]When discussing information availability, Kaplan and Schulhofer-Wohl (2017) emphasize amenity-based migration, while our emphasis is on labor markets. We do not rule amenity migration, which could be relevant given that many of the declining cities typically rank highly in terms of amenity valuation (e.g. Roback (1982), Chen and Rosenthal (2008), Albouy (2009)). The explanation of Kaplan and Schulhofer-Wohl (2017) is that information has made for better amenity-based matches, meaning fewer move events later need to be reversed. In contrast, we are (1) focused on accounting for widespread labor market heterogeneities rather than a simple proof-of-concept model, and (2) we are trying to rationalize why declining mobility cities have differences in the frequency of moves by natives, not recent arrivals.

[28]Mobility (i.e. population turnover) has no association with population size at the metro level.

Figure 7: Change in Out-Migration Rates to Metro Size and Age
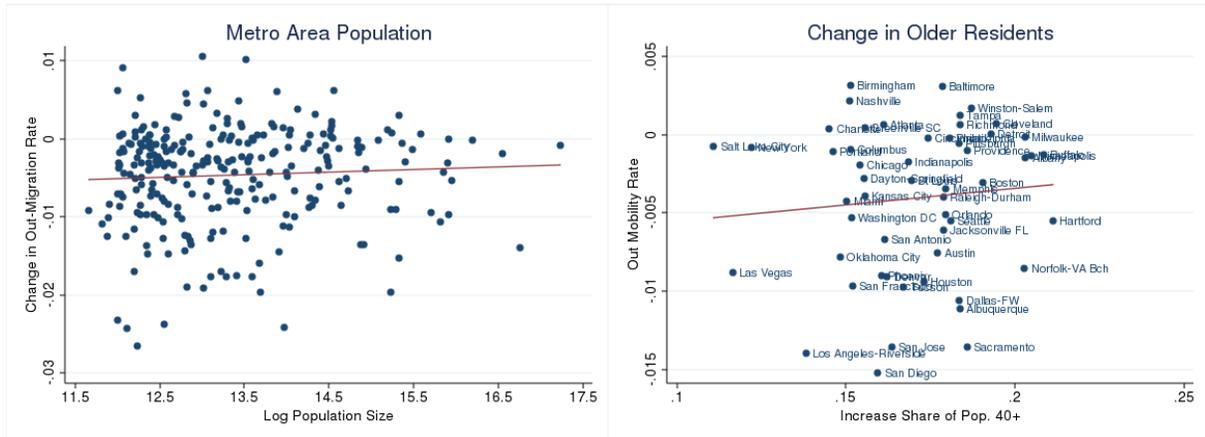
NOTES: The figures plot change in mobility rates (IRS data) to population size (left, Census) and change in share of residents over 40 (right, Census and ACS).

aged. The righthand plot shows the increase in the metro areas' population aged 40 or older. While all cities have aged substantially–the mean is about an 18 percentage point increase–there is no relationship with the size of the decline. Thus, while an aging population may be part of the explanation of the national trend, it clearly is not driving the spatial patterns we document.

3   Model

We now write down a model of income and location search that incorporates the key aspects described above: home utility and its locational heterogeneity via rootedness, local income distributions, and differences in types/states of workers and their life cycles. The model is written in the tradition of Kennan and Walker (2011) with adjustments for our focus on spatial heterogeneities. A dynamic discrete choice model is well suited to the task of explaining costly migration decisions for heterogeneous workers over a set of heterogeneous choice alternatives. We first introduce the model and then discuss how its features will map to the data and affect identification of key parameters.

3.1   Environment

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

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 substages occur. First, workers choose a location. Then, incomes are drawn in a process dependent on the income state with which the worker entered the period.

3.2 Search Across Locations: Migration

Individual workers are endowed with a home location which provides them with a utility flow available nowhere else. The size of this flow utility will be a function of the rootedness of the location endowed at time of birth, which is assumed to be weakly positive and increasing. 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 $u_{j=H}(r)$. The rootedness of a location affects the utility offered to its natives, and not any other workers living in the location but born elsewhere.

Workers also face moving costs to relocating from their origin at the start of the period, whether home or not. While the home premium inclines workers to prefer their birthplace, moving costs introduce frictions. Thus, we draw a distinction between moving costs in the sense of Kennan and Walker (2011) as dynamic frictions to changing location and static preferences for a home location. This will be important for the empirics, because conflating the two concepts could bias estimation of their distinct effects on mobility behavior. Moreover, to account for heterogeneity between locations in mobility rates in ways possibly not captured by home preference, income, or other characteristics, we allow moving costs to be dependent on the pair of locations, representing a generic notion of "distance," and hence we label them $mc_{k,j}$.

We describe the income search process below; for now, it suffices to label the labor market value of a location to be $W^j(k)$ and amenity value to be δ_j, where j indexes destinations/choices and k indexes origins/states. Workers are presented with values for locations according to $\delta + W$, the home utility premium if applicable, and pair specific moving costs. Let v denote the value of making a locational choice,

$$v_j(k) = \delta_j + W^j(k) + u_{j=H}(r) - mc_{k,j} + \beta EV(j) \tag{1}$$

The equation (1) shows the sources of utility, the moving cost wedge, and the continuation value V. Expectation E is taken over future preference shocks.[29] As is common in discrete choice location models, we allow for a temporal idiosyncratic preference shock term to rationalize the gross flows in the data. Calling this ε_j, the worker's choice problem is

---

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

$$V(k) = max_j\{v_j(k) + \varepsilon_j\} \tag{2}$$

Assuming the preference shocks are drawn from a Type I extreme value distribution yields the familiar closed-form solution for choice probabilities of destination j conditional on origin k:

$$\sigma_{kj} = \frac{exp(v_j(k,n))}{\sum_i exp(v_i(k,n))} \tag{3}$$

The shock distribution also provides a convenient closed form for the expected value function which will prove extremely useful in estimation.

$$EV(k) = ln\Big(\sum_i exp(v_i(k))\Big) \tag{4}$$

3.3 Search Within Locations: Income

The worker begins the period at some point in the income distribution, y_n. Let $W(n)$ denote the utility afforded by income at some point n (suppressing location notation). Each period in every location, there is a some probability λ that the worker is contacted with a new offer. If the worker fails to get a contact, occurring with probability $1-\lambda$, 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

$$w_0(n) = \sum_{n'}^N \pi_{n'|n} W(y_{n'}) \tag{5}$$

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,[30] 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

$$w_c(n) = \sum_{n'}^N \pi_{n'|n} E(max\{W(y_n), W(y_{n'})\}) \tag{6}$$

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

$$W(n) = \lambda w_c(n) + (1-\lambda)w_0(n) \tag{7}$$

---

[30]The idiosyncratic shock allows for incomes to fall even when the worker has the option. Income changes in the data may be the result of workers choosing jobs with non pecuniary advantages, shorter commutes, or flexible work hours. Non-income aspects outside the model are subsumed in this idiosyncratic shock.

Thus, the parameter λ 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 $W(n)$ is a function of λ, current income, y_n the income available in the location, $y_{n'}$, and the distribution of income changes.

We allow for the possibility that local and nonlocal searches are not equivalent. Specifically, we allow the probability of drawing points in the income distribution to depend on whether the income search is conducted by an incumbent resident or someone currently residing in a different location. For example, we think it reasonable that search in one's current location is in some sense "easier" than search in another faraway location, as even with increasing availability of information technology, local networks remain important. Proximity provides more face-to-face contacts, unplanned meetings, introduction to mutual acquaintances, etc. Notice that this distinction is based on origin vis-a-vis destination, *not birth location*.

To operationalize this idea, we allow for two distributions on new income shocks: π^{local}, and $\pi^{nonlocal}$.[31] Using data on income dynamics, we can identify differences in income transitions between those who move to new locations and those who do not.

3.4 Locational Heterogeneity in Income Search

The income search component of the model serves several purposes. First, it provides a test of the effect of information availability on the value of a local labor market which could conceivably affect the spatial pattern in mobility. Second, it incorporates heterogeneity in labor market opportunities (by location and age) that might affect the estimate of home premium, which in estimation varies by space and cohort. Third, it allows for direct effects of trends in the labor markets themselves, which could explain some of the variance in the relationship between mobility declines and home attachment observed in, for instance, Figure 6.

All of these concerns relate in some fashion to local labor market heterogeneity, so we briefly describe the ways in which heterogeneity affects choice probabilities. Locations are heterogenous 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.[32]

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

---

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

[32]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.

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, $w_c(n) = \sum_{n'} \pi_{n'|n} ln(exp(w_{n'}) + exp(w_n)) - \gamma$, where γ is Euler's constant (to adjust for the mean of the T1EV distribution), which is increasing in the spread of the distribution of n'.

Finally, the income search value function is clearly affected by the probability of successful search in ways that interact with the income distribution. Writing out the expected value of income search, (7),

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

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

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

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.

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. The quantification of this channel–which accompanies our main story of increasing rootedness and nativity–will depend on the parameters π^{local}, $\pi^{nonlocal}$, and λ, which we calibrate below.

3.5 Worker Types

Up to this point, we have suppressed worker types for exposition. The value function (1) may depend on permanent or exogenously evolving characteristics of the worker. First is "home" location, which is endowed to an individual at the time of her birth. People born in location k will always have some utility premium associated with that place. Relatedly, there is also cohort, as people born to some location k at times t and $t' \neq t$ may be endowed with different attachment via rootedness.

Next, consider age, a deterministic state variable. There are three pertinent features that distinguish workers of different age. One is the horizon of the worker. Incentives for migration may change over the life cycle, as the number of future flow utilities decreases. This would be reflected in the future value term in (1), the exclusion of which could bias moving cost estimates (and by extension, estimates of the home premium). Second, the number past periods of search

increases, so that the worker had more opportunities to ascend the income ladder. This will be reflected in the age-income cell sizes. For instance, if few young workers are at the top of the income ladder, this affects the future value of one location compared to another, and again moving cost estimates could be affected. Third, there may be differences in preferences that change over the life cycle. Specifically, moving costs may increase as the worker ages, evidenced by the decline in move rates over the life cycle. To the extent that this decline is not explained by differences in incentives from horizon effects or selection via the past evolution of state variable,[33] move costs by type will pick up any residual reluctance to migrate.

Lastly, there could be permanent type differences; i.e. those that do not vary after the worker has entered the labor force, like education, though an even richer model could consider a state variable that evolves, like experience within an occupation. In practice, we will separately estimate and simulate the model for workers with and without college degrees, since they face first order differences in income opportunities.

Equation 8 writes out the choice specific values from (1) with all state variables: origin k, age a, and the worker's endowed type τ, which is her birthplace, cohort, and education. k changes endogenously, a deterministically, and τ is fixed over time.

$$v_{j,\tau}(k,n,a) = \delta_{j,\tau} + W^j(n,k,\tau) + u_{j=H|\tau}(r_\tau) - mc_{k,j,\tau} + \beta V_\tau(n',j,a+1) \tag{8}$$

Note the amenity parameter 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.

3.6   Marginal Propensities to Migrate

The discrete choice nature of the model makes explicit the "precariousness" we mentioned above, whereby agents nearer the margin of migration are more greatly affected by changes to the environment. Consider the derivative of a choice probability with respect to choice-specific values in a discrete choice model with Type I Extreme Value errors. The choice probability of location j is

$$\sigma = \frac{exp(v_j)}{exp(v_j) + \sum_{k,k \neq j} exp(v_k)}$$

so the derivative with respect to the value is

---

[33]Bayer and Juessen (2012) study how selection effects can be misattributed to utility-equivalent move costs in models of migration.

$$\frac{d\sigma}{dv} = \frac{exp(v_j)[exp(v_j) + \sum_{k,k \neq j} exp(v_k)] - exp(v_j)^2}{[exp(v_j) + \sum_{k,k \neq j} exp(v_k)]^2} = \sigma - \sigma^2 \qquad (9)$$

For choice probabilities greater than 0.5 (like non-migration propensities in the data), this quadratic is decreasing in σ, the probability of choice. Ex ante higher probability events are less responsive to changes in the choice value, and vice versa. Hence, more mobile agents will be more responsive to changes in their origins. For example, a worker with a home attachment will have a lower elasticity of choice probability with respect to changes in the value of their origin, so that home preference can act as a multiplier, enhancing the staying effects of any other reduction in move propensity. While this particular analytic expression relies on the functional form, the broader point is that ex ante inframarginal agents will be less susceptible to changes in their environments.

4   Demonstrating the Model

The model admits a rich degree of locational and type heterogeneity to account for many possible confounders in estimation and complementary explanations. Before proceeding to a full estimation, we use some simple simulations of the model to demonstrate the model's qualitative features, especially the role of home preference.

4.1   Preference for Home and Endogenous Migration Rates

First, we show how home preference acts as a de-accelerator for migration rates. This begins with a very simple version of the model, an economy of four identical locations (and hence four possible birthplaces) and one agent type. At time $t = 0$ a measure 1 of agents is born across the four locations, distributed equally at 1/4. agents are endowed with a preference for their birthplace, constant across the identical locations. The agents are then given an opportunity to choose a location at each period $t = 1, \cdots$. All moves are idiosyncratic but for the home preference. We ignore aging and cohort effects to emphasize the evolution of the state variables over successive choice opportunities; after sufficient opportunities, the share at-home or away from home reaches a steady state (in practice, we simulate the model for 100 periods). Table 3 reports the economy's moving rate and share living at home under three different regimes: low, medium, and high preferences for living at home.[34]

The table illustrates the direct and indirect ways in which home preference can affect mobility in the economy. First, the direct effect means that there is more staying (less out-mobility) the stronger is home preference. This is seen by comparing move rates in the initial choice

---

[34]For calibrating low, medium, and high preference, we use mean rootedness, and one standard deviation below and above, multiplied by the average preference parameter in Table 4.

Table 3: Simulated Move Rates Under Different Strengths of Preference for Home

Statistics at:	Time 0: All At Home		Steady State		Fixed At Home Share	
Home Preference	Population Share at Home	Migration Rate	Population Share at Home	Migration Rate	Population Share at Home	Migration Rate
Low	100.00	3.32	41.93	9.98	66.99	7.10
Medium	100.00	1.98	66.99	6.67	66.99	6.67
High	100.00	1.17	85.05	3.76	66.99	6.89

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 in percentages (%).

opportunity. The more subtle way is by affecting how many agents would be at-home, and hence have a higher threshold for mobility, in the steady state. Initial differences in home preference accumulate to having more agents in their birthplace even in the long run, so that move rates between the scenarios diverge in the steady state. The last column emphasizes the point by constructing a counterfactual aggregate move rate if the share at-home were that of the average home preference scenario, but moves were made at propensities from the respective simulations. The low home preference scenario would have more out mobility, but the gap is substantially smaller than the one induced by the difference in steady state at-home shares. The high preference scenario would actually have more mobility, as many agents would be "forced" away from home and try to return.

In summary, while the direct effects of home preference are nontrivial, it is their cumulative, indirect effect of placing more people in an away-from-home status that causes migration rates to diverge between scenarios of higher and lower preference for home.

4.2   Home Preference and Regional Shocks

The previous simulation showed how preferences for home can amplify migration rates for a single agent facing multiple decision periods. Next we show how the preference for home mediates a shock to the equilibrium distribution of population in an economy. The simulation that follows has several notable differences from the simple life cycle model. We now use an overlapping generations (OLG) framework. Each agent lives for A periods, and the economy is simulated for $T >> A$ periods. We start with an initial distribution of population for the very first cohort, $t = 0, a = 0$, and simulate their behavior over C periods; they spread out to be at or away from home at a steady state level, like the single cohort simulations above. But when they "die" at A, the distribution of their population forms the rootedness–the strength of the home preference for the next generation at age 0. Thus, the strength of the home preference moves endogenously in the OLG framework. Other generations are born in the intervening periods (e.g. $t = 1, a = 0$), so that at a sufficiently high t, there is a whole distribution of home preferences for A distinct cohorts. We then compare this OLG model, "Roots Preferences," to others in which

there is either no preference for home ("No Home Preference"), or the preference for home is constant ("Home Preference"), and not endogenously generated by prior spatial distributions of population.[35]

We simulate each version of the model until it reaches a steady state in all the relevant home variables–the share living at home, and the strength of that preference for home. At the steady state, all cohorts have the same strength of preference for home, population sizes of the the locations are constant, and all migration is idiosyncratic. Then, we introduce a permanent shock to location attributes in order to cause population reallocation across space. In the simulation of Figure 8, we change the income distributions of the locations in a 2x2 matrix of high or low mean and high or low variance to produce four heterogeneous locations (HH, HL, LH, LL).[36] This particular implementation is not important for generating the qualitative patterns we describe– the dynamics could just as easily be generated by changes in local amenities, for instance. The objective is to see how shocks to the steady state population distribution are mediated over the transition path.

The lefthand plot of Figure 8 shows net population reallocation in the three versions of the simulation, starting at the shock to location incentives after the steady state. When there is no preference for home, the economy reallocates population in response to the shock, with a spike in migration lasting about three periods. When there is a preference for home, the mediation of this shock lasts longer by a factor of about eight. The longer transition path to steady state population comes about because agents are more reluctant to leave home, and more inclined to move back home rather than stay in the better labor markets. Moreover, when the preference for home endogenously varies over the transition path in the "Roots Preference" scenario, there is even more churning in the population because the strength of home preference fluctuates– decreasing, reaching a trough, and then stabilizing.

The righthand plot illustrates the economy's migration rates in the three simulations. Without preference for home, the total migration needed to reallocate the population is relatively low. When there is preference for home, the migration rate is higher (though producing less net reallocation) and the transition path lasts longer. Then, when the preference for home is endogenously generated by previous population shares in, the amplification is even greater. Migration peaks at a higher value and a later date than the simple "Home Preference" since the *size* of the average home preference actually declines as population moves to the better locations. The

---

[35]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.

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

Figure 8: Migration and the Transition Between Steady States
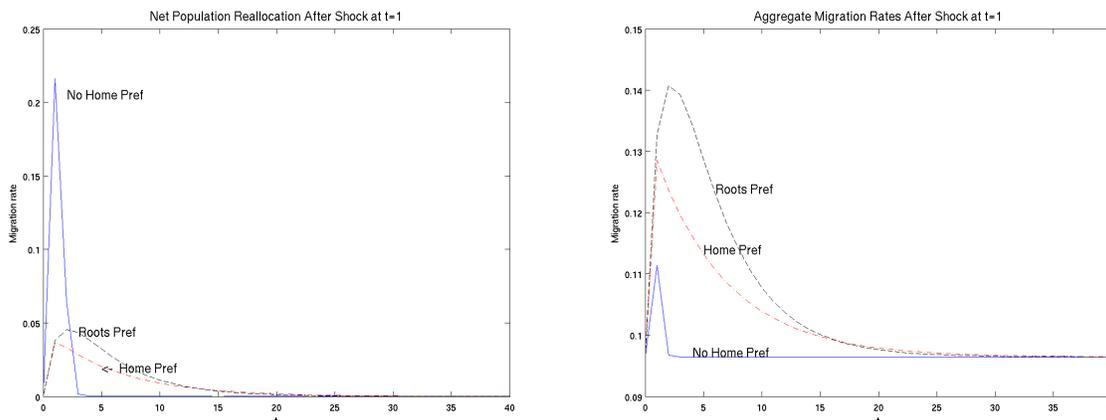
growing populations are less rooted for a time, so that they generate more idiosyncratic out-mobility as they grow. The "Roots Preference" simulation returns to steady state at a similar horizon, but migration rates remain above the constant "Home Preference" simulation for about 15 periods.

In summary, for a given shock to migration incentives, the scenarios with preferences for home have higher gross migration rates, for a longer period of time, along the transitional path. When the strength of preference is endogenous and path-dependent, the churning is even greater. Thus, the model could generate a pattern in gross mobility consistent with the spatial evolution of population in the U.S. and the subsequent heterogeneity in the migration trend, as shown in Section 2. Quantifying this role vis-a-vis other channels will require a formal estimation of the model.

5   Applying the Model to U.S. Data

We now describe how we take the model to data on U.S. cities. Our estimation strategy proceeds more like a location demand model than a dynamic migration model. To study spatial heterogeneity, of course, we need information on location attributes, and longitudinal datasets simply do not have enough observations to identify parameters off of cross sectional differences, such as location rootedness. While our model was written as a dynamic microeconometric model like Kennan and Walker (2011), Bayer and Juessen (2012), or Bishop (2008), the estimation uses aggregate choice probabilities (market shares) instead of, e.g., maximum likelihood estimation on panel microdata. Hence, estimation more closely resembles a location demand model (such as Bayer et al. (2009) or Diamond (2016)). We account for forward-looking behavior by exploiting

28

the structure of the logit choice model, which has a closed form solution for continuation values in a dynamic optimization problem.[37]

We estimate the model on a single cross section of U.S. cities, using the 2007-2011 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 where location features are measured by the 2000 and 1990 census.[38] This allows ceteris paribus simulations to decompose the change in mobility–and its variation across space–explained by the various features of our model, such as trends in at-home status or rootedness, evolving income opportunities, or shifting type shares (e.g. aging population).

5.1 Preliminary Estimation: Income Distributions and Dynamics

Before estimating the model's utility parameters, we obtain measures of income offer distributions and income dynamics. These determine the value of of utility from income as represented in (7) via (5) and (6). There are three sets of parameters to calibrate.

First is the available income distribution of each location. To focus on spatial differences in income opportunities, we construct a measure of the local income distribution having adjusted for differences in the local labor force composition. Specifically, after limiting the data to regularly employed workers, we run a regression of log earnings on controls for sex, race, English proficiency, and household composition in order to strip out compositional differences at national average labor prices.[39] 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.[40] The residualized income for each

---

[37]Bayer et al. (2016) and Davis et al. (2016) 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.

[38]Note that we assume the *primitive* preference parameters are fixed. That is, while home preference can vary with rootedness, or the aggregate share at home may change, the strength of preference of a given rootedness is assumed constant.

[39]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 (and hence captured by the residual) since these can differ materially across labor markets–the kind of spatial variation we want to retain as a city characteristic.

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

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, $w_n = \mu^\omega_{j,\tau} + 0.5 n \sigma^\omega_{j,\tau}$, 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 (5) and (6). 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 the equations (5) and (6) 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, λ, 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; i.e. the distinction between (5) and (6) is conceptual. 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 need actual data on income dynamics, so we turn to the Panel Study of Income Dynamics (PSID, PSID (2014)).[41] Using the post-1997 PSID for workers employed for two consecutive surveys, we measure the change in their income in standard deviations, which becomes the input data to a maximum likelihood estimation. If the step size of discretization is s, the probability of the change to income is

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

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

---

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

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

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, consistent with our conjecture.

The final parameter is the probability of successful search, λ, which gives the worker the preferable option value search (6), instead of the vulnerable random search (5). As noted, we have little sense how to discipline this with 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.

5.2 Main Estimation: Utility Parameters

We next turn to the estimation of the utility parameters. The main parameter of interest is how large is the home premium and how it is affected by rootedness. The parameters of necessity are the move costs and location amenities, which can vary by person type, origin, and destination.

5.2.1 Parameterization

We use a simple linear utility function

$$u = \delta_{j,\tau} + W(n) + I(j = h)(\alpha R_j)$$

where $W(n)$ is the value of income search at point n. (Incomes are in natural logs.) The parameter $\delta_{j,\tau}$ 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-50s, for college and non college educated workers. Attributes of the location, such as amenities or cost of living, will be subsumed in this parameter, but in estimating by type, workers of different age or education can have heterogeneous preferences for these.

Home status and birthplace are then used to identify the home premium, captured by parameter α, multiplied by the birthplace-by-cohort's endowed rootedness.[42] Variation in birthplace

---

[42]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

and cohort provide for identifying variation in rootedness. For example, all young college educated workers may prefer to live and work in San Francisco (captured by δ), 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, heterogeneity in their propensity to choose San Francisco helps to identify this parameter. The logit demand system flexibly allows preferences to depend on attributes of the place, the worker, and interactions between. We allow α to vary by education.

The estimation of home preference will be biased if other frictions are ignored, so we turn to 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. Then, to account for the spatial component of migration probability, we enter the distance in miles between the MSA centroids. Since migration rates fall off with distance, we expect this term to be negative (increasing distance means less moving). We also allow a discrete shift in distance for "neighboring" metro areas, which in practice we define as being in the same state or just across state borders (e.g. Louisville and Cincinnati).

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

Recall that the home preference is identified by two sides of the migration problem: the lower propensity to migrate out when located in one's home and the higher propensity to return home. Both of these are consistent with home providing flow utility. However, in the data, the propensity to return home is an order of magnitude larger than the difference in the propensity of out migration by at-home versus not-at-home. This affects the magnitude of the home premium estimate. Kennan and Walker (2011) show that workers are more likely to return to a location previously visited (even if not "home"), which could be driving this feature of our data. Thus, a more conservative estimate for α includes for a move cost "discount" for returning home. The home premium in utility is still identified through variation across birthplace and cohort in rootedness.[43]

_____

an intercept, for instance, causes Las Vegas' (the lowest rooted city) projected home preference to be negative. While home preference in Las Vegas appears weak, a projected inversion in move rates by at-home status is in conflict with the data.

[43]Essentially, this term forces home preference identification to come from only across-city variation in the

32

All together, the move cost function is

$$mc = \underbrace{\sum_{\hat{\tau}} I(\hat{\tau})mc_\tau}_{types} + \underbrace{D_{j=h}}_{return\,home} + \underbrace{D_{o,j}}_{distance\,mc_d} + \underbrace{\sum_{j}^{J} I(orig = i)mc_i}_{exit\,toll} + \underbrace{\sum_{j}^{J} I(dest = j)mc_i}_{entry\,toll}$$

Table 4 below will compare specifications to demonstrate the importance of each component of the move cost function. We next describe the estimation method. Some details and derivations are relegated to appendix B.

5.2.2 Estimation Method

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

The complete specification contains $(J \times 8) + 2$ parameters in the utility function and $8 + 8 + 2 + (J \times 2)$ in the move cost function. Hence, feasibility is a concern. Fortunately, there is a very tractable way to estimate the model. As a memory-less discrete choice specification, this model is well suited for estimation via conditional choice probabilities (CCPs). CCPs arise in the logistic model because of the mapping between the continuation value and choice probabilities. The advantage is that the model need not be simulated to arrive at parameter estimates. Instead, one needs to derive the model specific equivalence between choice probabilities conditioned on state variables, which are observed in the data, and the model's parameters.

The result admittedly relies on some functional form choices. However, in addition to massive savings in computational burden, using CCPs to approximate the value function has the advantage of putting no structure whatsoever on expectations, which is especially important in our context, since we observe only one choice event but not how choices may change as dynamic states evolve. Whatever workers might believe about the future is subsumed in the CCP term.

Specifically, this model has a simple derivation exploiting finite dependence (Arcidiacono and Miller (2011)).[44] That is, because the choice problem is memory-less by some point s (i.e. does not depend on the whole sequence of choice), two disparate choices in some period t can be

---

return home propensity, not the size of the propensity itself.

[44]Finite dependence and CCP estimation is also lucidly described in Bishop (2008) and Ma (2017).

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.

We now derive this for a simple example of our model. Consider two alternatives available to a person in origin k at income point n (suppressing type indices and using x' to indicate the value of x one period ahead).

$$v_{1,k} = u_1(n, k) + m_{1,k} + \beta E(V|k' = 1, n)$$

$$v_{2,k} = u_2(n, k) + m_{2,k} + \beta E(V|k' = 2, n)$$

Following Hotz and Miller (1993), the expected future value term can be written in terms of the future value of choosing some location 0 in the next period.

$$v_1 = u_1(k) + m_{1,k} + \beta(v_{0,1}' - ln(\sigma_{0,1}'))$$

The next period choice value can then be decomposed again into the flow utility of the location 0, which is affected by the choice of new origin, and two-period ahead value of choosing that reference location 0.

$$v_1 = u_1(n, k) - m_{1,k} - \beta(ln(\sigma_{0,1}') + [u_0(n', 1) - m_{0,1} + v_{0,0}'' - ln\sigma_{00}''])$$

Similarly, the value of alternative 2 can be decomposed as:

$$v_2 = u_2(n, k) - m_{2,k} - \beta(ln(\sigma_{0,2}') + [u_0(n', 2) - m_{0,2} + v_{0,0}'' - ln\sigma_{00}''])$$

The expressions have common two-period-ahead terms. Taking the difference, we have

$$\begin{aligned} v_1 - v_2 = {} & \big(u_1(n, k) - u_2(n, k)\big) \\ & - \big(m_{1,k} - m_{2,k}\big) \\ & + \beta\big(u_0(n', 1) - u_0(n', 2)\big) \\ & - \beta\big(m_{0,1} - m_{0,2}\big) \\ & - \beta\big(ln(\sigma_{0,1}') - ln(\sigma_{0,2}')\big) \end{aligned} \tag{10}$$

Because $ln\frac{\sigma_{1,o}}{\sigma_{2,o}} = v_1 - v_2$, we need only observed move probabilities to estimate the utility parameters in u and the move cost parameters in mc_{kj}. Moreover, the estimation can be written in a linear moments condition, which we now derive.

5.2.3 CCP Method of Moments Estimator

Many studies using CCP estimation use microdata and maximum likelihood estimation. The CCP step is calculated in an initial stage and substituted in for the value function where appropriate. In our context, however, given the available data, we are using aggregated choice probabilities by type-state cell. Therefore, we have one observation per cell but many cells, and we arrive at our estimates using a method of moments estimator.

Rearranging the normalized estimating equation (10), we write the expression for the log odds ratio of choosing some location j versus the normalizing location 0, given type (including a birthplace of h) and origin of o. Letting $u_{j,k}$ denote flow utility offered by destination j from starting in origin k, we have

$$ln(\sigma_{j,k}) - ln(\sigma_{0,k}) + \beta(ln(\sigma'_{0,j}) - ln(\sigma'_{0,0})) = u_{j,k} - u_{0,k} - (m_{j,k} - m_{0,k}) + \beta[u'_{0,j} - u'_{0,0}] - \beta[m_{0,j} - m_{0,0}]$$

The expression above has data on the left and utility functions with parameters on the right. Thus, the parameters can be found by fitting (through, say, least squares) the two sides as best as possible. What is more, given the structure of our utility functions and move costs, these can all be decomposed into a linear-in-parameters expression, so that the parameters can be found through a standard linear estimator like GMM.

The linear representation is straightforward to derive, if a bit tedious. An appendix writes down the full derivation in matrix form. Equation (11) is the resulting estimator, which is measured in differences relative to a normalizing choice (the composite outside option, place 0).

$$ln(\sigma_{j,k}) - ln(\sigma_{0,k}) - \beta(ln(\sigma_{0,j}) - ln(\sigma_{0,0})) = (X_j - X_0)\theta \tag{11}$$

where σ's are vectors of choice probabilities taken from data and the X's are the matrices that represent the structure of the choice problem, which are derived from sets of indicator variables (e.g. moved or did not, interacted by types) and some location attributes (e.g. the rootedness of the home location); the derivation appears in the appendix. The parameters can then be recovered by inversion of the $X_j - X_0$ matrix. In practice, the parameters are overidentified. One may want to weight the conditions by the information contained on the lefthand side–for instance, the inverse of the variance of the cell probabilities.

An advantage of estimating the model, as opposed to say, calibrating to particular targets, is that the estimator naturally embeds controls for the covariances in the data. For example, the city's toll costs will be estimated having controlled for age composition, the share at home and rootedness, and distance relative to other locations in the economy.

We next describe how we form the cell probabilities, the σ's, using the data available.

5.3   Forming the Moment Conditions

The migration choice data come from two sources: the IRS and ACS data discussed above. These complement each other in that the IRS offers detailed place-to-place flows but without information on types, and the ACS data offers choice probabilities by type but with less location information. Also, the ACS and IRS both report one-year move probabilities, so we can use them together without having to infer aggregation over time.[45]

With age group by education by birthplace by origin, there are $4 \times 2 \times J \times J$ cells, with J choices for each. At $J = 55$, there are 24,200 types and 1,331,000 choice probabilities to estimate from the data. In principle, one could form all the moment conditions necessary from ACS, but in practice, the cells simply become too small to reliably estimate a choice probability. There are few individuals once the data are cut to, for example, 40-something college educated workers living in Houston but born in Cleveland–in the full population, and certainly too small in a sample. 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. Hence, a smoothing procedure is in order.

A full description of the choice probability estimation is available in appendix C. In brief, we take directly from the data the cell probabilities after reducing the J-size dimensions in alternating ways, and then impute the full 1.3 million element matrix from a combination of these. For example, our starting point is to use the ACS data to estimate move/not-move probabilities by origin, age, education, by at-home status, combining away-from-home locations and non-reflexive destinations, to arrive at $Pr(move_{\tau,k})$. We then combine origins and estimate return-home and move-elsewhere probabilities, $Pr(choose\, j_\tau | move = 1)$. The full matrix is then derived from the conditional probabilities of these two estimates: $Pr(move_{\tau,k}) \cdot Pr(choose\, j_\tau | move = 1)$. The probabilities are then differenced to form the estimating equation (11).

5.4   Estimates

Table 4 reports the structural parameter estimates for several specifications of the move cost function; our preferred is 6, but the others illustrate the model, so we discuss the progression.

Column 1 uses only age/education intercepts and the distance function. These intercepts reflect the migration pattern with age. Rootedness is strongly associated with staying at home– and returning there if away–for both education groups. At the mean, out-mobility for workers at home is lower by 400 log points than for those not at home. A one standard deviation increase in rootedness decreases out-mobility for natives by about 50 percent. The distance function alone

---

[45]The 1990 and 2000 decennial Censuses, in contrast, report five-year migration rates.Specifically, the migration question in the Census is "In what metropolitan area did you reside five years ago?" Detail on intervening moves is lost. We would have to infer how many moves are reversed by return or onward migration, which are common (see Kennan and Walker (2011)).

seems misspecified, in that longer distance moves are happening at lower cost. Columns 2 and 3 are specification checks that for illustration purposes drop home preference and move costs from the model, respectively. Without home preference in 2, move costs are estimated substantially higher for all age/educations to rationalize the amount of staying in the data. Without move costs in 3, home preferences must be very large to rationalize the degree of non-migration for people at home. We conclude both features are empirically relevant to maintain in the model.

The table proceeds by enriching the specifications. Column 4 adds the moving toll costs. For linear independence we must then drop a move cost category, so an average toll cost is reported. The home preference estimate is very similar, and only the distance coefficients are materially affected. This indicates that rootedness is not simply spuriously correlated with location mobility (as indicated by the informal "placebo check" in Figure 6). Column 5 adds destination fixed effects. Adjusting for average desirability of the location for each age/education group actually slightly increases the home preference estimate in order the rationalize why, for example, there are not more people leaving Albany and Cleveland (low average δ) for Dallas and Phoenix (high average δ). Finally, as discussed before, column 6 adds the return-home move discount to account for the large move-home propensity. The move-home discount is roughly one-third of the average move cost. The home preference parameters decline substantially, but remain statistically and economically significant. At the mean, out-mobility for workers at home is lower than not-at-home by about 200 log points, and a one standard deviation increase in rootedness decreases out-mobility by about 17 percent. We are more comfortable with the conservative estimate of home preference; given the potential for informational advantages in returning home, the specification with move costs seems closer to identifying a true "preference" for home. Column 6 is our preferred specification that we will use in simulations below.

5.5 Model Fit in Baseline Simulation

We next check that the model is able to replicate the main features of the data. Table 5 reports the average moving rates by age and education category and at home status for the data and our baseline simulation of the model. 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 reasonably well the difference between workers at home and not at home, the main qualitative pattern in the data important for our tests of demographic shifts. The preference for home and lower entry costs also enable the model to match the inordinate amount of migration for people not at home back into their birth locations, although it is somewhat over predicted. This will not be a problem for our simulation results, as they focus on out-migration rates. The main reason for including the move-home cost reduction is to avoid over estimating the preference for home via rootedness, which is important in the simulations.

The spatial heterogeneity in migration rates is obviously important for our analysis. Figure 9

Table 4: Parameter Estimates

		1	2	3	4	5	6
Home Preference							
Rootedness	Noncollege	2.422		9.907	2.398	2.799	1.125
		(0.015)		(0.016)	(0.009)	(0.008)	(0.011)
	College	2.347		8.922	2.266	2.504	0.733
		(0.018)		(0.022)	(0.011)	(0.010)	(0.014)
Move Cost							
Noncollege	20s	-7.06	-8.12				
		(0.012)	(0.011)				
	30s	-7.50	-8.54		-0.44	-0.54	-0.47
		(0.012)	(0.010)		(0.002)	(0.012)	(0.012)
	40s	-7.87	-8.91		-0.81	-0.98	-0.90
		(0.012)	(0.010)		(0.002)	(0.012)	(0.012)
	50s	-8.09	-9.14		-1.04	-1.29	-1.19
		(0.012)	(0.010)		(0.002)	(0.012)	(0.012)
College	20s	-7.08	-7.84				
		(0.013)	(0.012)				
	30s	-7.72	-8.46		-0.63	-0.78	-0.71
		(0.013)	(0.012)		(0.002)	(0.016)	(0.016)
	40s	-8.29	-9.04		-1.21	-1.51	-1.41
		(0.013)	(0.012)		(0.002)	(0.016)	(0.016)
	50s	-8.50	-9.25		-1.42	-1.77	-1.66
		(0.013)	(0.012)		(0.003)	(0.016)	(0.016)
Distance	Neighboring MSA	1.499	1.501	-1.42	0.953	0.961	0.960
		(0.004)	(0.004)	(0.007)	(0.002)	(0.002)	(0.002)
	Log(miles)	0.111	0.113	-1.63	-0.16	-0.16	-0.16
		(0.001)	(0.001)	(0.007)	(0.000)	(0.000)	(0.000)
MSA Toll Cost	(suppressed)				yes	yes	yes
Avg Pairwise	Noncollege				-6.19	-5.87	-6.66
Move Cost	College				-6.20	-5.72	-6.39
Return Home	(average)						2.253
MSA Dest FEs	(suppressed)					yes	yes

NOTES: The table reports structural parameter estimates under five versions of the model; standard errors in parentheses. Simulations rely on the estimates from column 6. When no move cost is reported for 20-somethings (columns 4-6), their intercept is absorbed by the MSA toll costs, and older age groups move costs are estimates relative to 20-somethings. Point estimates for metro-area specific move costs and average location quality are available upon request. "Neighboring MSA" refers to metro area pairs in the same U.S. state and/or less than 100 miles apart.

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

		Data				Model			
		Out-Mobility Rate			Move Home	Out-Mobility Rate			Move Home
		All	Natives	Nonnatives	|Move	All	Natives	Nonnatives	|Move
Non-college	20s	6.47	2.81	14.72	42.50	4.55	1.83	10.74	52.11
	30s	4.40	2.01	8.79	33.91	3.00	1.24	6.40	48.00
	40s	2.83	1.14	5.99	34.79	1.99	0.74	4.41	51.76
	50s	2.10	0.86	4.24	31.22	1.53	0.57	3.25	48.99
College	20s	9.11	4.75	14.74	41.06	8.27	3.57	14.23	56.28
	30s	4.68	2.24	7.48	30.67	4.04	1.49	6.80	48.39
	40s	2.34	1.02	3.81	26.10	1.88	0.68	3.10	43.73
	50s	1.99	0.90	3.07	24.01	1.54	0.62	2.42	40.50

NOTES: All figures are in percentages (%).

Figure 9: Model Fit: Predicted and Actual Out-Migration Rates by Metro Area
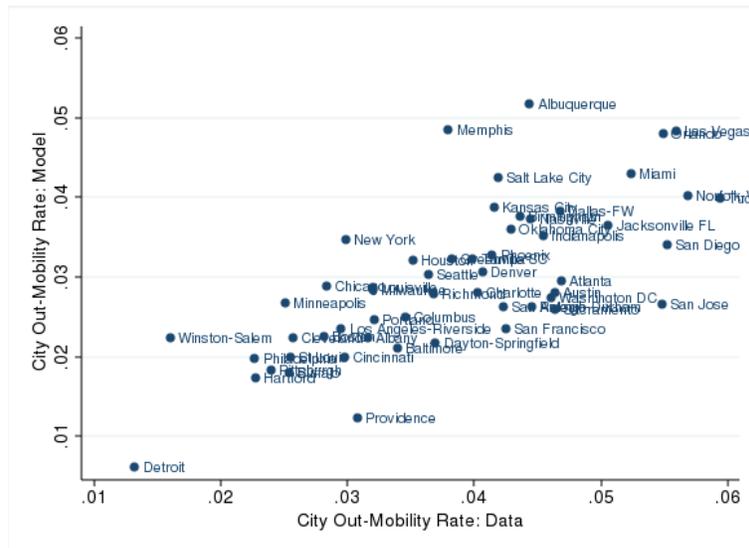
NOTES: The figure plots the model's baseline predicted out-mobility rates by metropolitan area to actual rates (ACS).

plots the actual and predicted out migration rates for each metro area in our analysis. Through a combination of compositional differences, shares of workers away from home, differences in rootedness, and the toll costs, the model is able to match the spatial heterogeneity.

6   Simulations

The purpose of the empirical model is to conduct simulations at alternative population distributions, including home status, and income and rootedness environments to see how the the mobility decline is affected by each channel. A simulation of the model predicts move probabilities for every state (origin, income draw) and agent type (age, education, birthplace) in the economy. To summarize the main findings, we report in the tables below the average mobility rates for the complete set of metro areas as well as a breakdown by more and less mobile locations (as measured by the mobility index), more or less rooted, and higher and lower income dispersion.[46] We then simulate the model for previous years at the same primitive parameter values, but using the environments and weightings as of 2000 and 1990. The environmental factors are the local income distributions and the rootedness of the population. The change in weights consists of shifts in age composition, the spatial distribution of population (i.e., the population share in each location), and the birth locations by cohort. These latter changes are somewhat of a hybrid between strict changes in composition and changes in environment, as they affect the regime of move cost and home-preference incentives faced by a population at a

---

[46]For mobility rate and income dispersion, we split cities at the median. For rootedness, we visually inspected the distribution for a gap, and split into more and less rooted at a rootedness of 0.55, leaving 17 cities as "shallow roots" and 37 as "deep roots."

given cross section. For example, if there are more people in low toll cost locations, aggregate move cost will be lower; or, if there are more people born in less-rooted locations, the aggregate preference for home may be lower.

The categorization of changes allows for a decomposition to measure the relative impact of the various changes to the economy. Most directly, we can alter the weights applied in aggregation. That is, we use the baseline (2010) model's predicted choice probabilities, but summarize the predicted move rates using alternative weightings to reflect compositional changes such as the aging of the population.

The other type of simulations actually alter the features of the model environment: the root-edness and at-home status of the population, the local income distributions, and the information parameter λ. These require new simulations of the model, but summarize the predictions at the baseline weights. We do not change any of the utility parameters in any simulation. Thus, we focus on changes in environment and composition, and do not allow for changing preferences to generate any of our conclusions.

Changing the model environment causes fundamental changes to the choice problem, altering current utilities and future values. We have estimated the utility function but also need a method for deriving new future value terms. In principle we could simulate the model over successive choices (as in section 4), but doing so would require a stance on what agents expect over their entire lifetimes–assumptions we are uncomfortable making, especially in a model with heterogeneous agents and locations. Instead, we again leverage CCPs. We estimate a simple parametric model of the probability of choosing the outside option,[47] which as equation (10) shows, will be sufficient to derive a future value term. The choice model includes environmental features as explanatory variables, so when we change these features in simulations, we can project the new choice probabilities of option 0 and derive the future value term. The advantage of this method is maintaining an agnostic position of how agents form expectations of future values.

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. The model generates a decline in mobility of 0.56 percentage points, about a 17 percent decline, from 1990 to 2010. This is very closely comparable to the actual data (for our subsample of 54 major cities) presented in the lower panel of the table. The model also successfully matches the heterogeneity in the decline, with more mobile cities declining more: 0.90 percentage points, about a 22 percent change, versus 0.43 percentage points, about 14 percent change for less mobile. Similarly, the model also accurately predicts greater declines in shallow rooted locations and those with higher income dispersion.

---

[47]Results available on request.

Table 6: Simulated and Actual Trends in Migration Rates

Cities:	All	Mobility		Roots		Income Dispersion	
		Low	High	Deep	Shallow	Low	High
Model							
1990	3.36	2.90	4.16	3.09	3.93	3.17	3.47
2000	3.14	2.74	3.75	2.91	3.60	3.07	3.17
2010	2.81	2.47	3.25	2.68	3.04	2.76	2.83
Data							
1990	3.34	2.88	4.15	3.00	4.09	3.07	3.46
2000	3.24	2.81	3.92	3.07	3.61	3.19	3.28
2010	2.76	2.51	3.09	2.63	3.00	2.75	2.76

NOTES: All figures are in percentages (%).

6.2 Decomposing the Sources of Decline: Changing Populations

This serves as final result which we can then unpack via *ceteris paribus* breakdowns of the sources of decline. Tables 7 and 8, respectively, report the changes in population weights (which use the baseline simulation probabilities, but aggregate at counterfactual weights), and the changes to the environment (which resimulate the model but use baseline aggregation weights). Tables A1 and A2 in the appendix report the decompositions for each city in our analysis. For brevity we skip 2000 and make comparisons between 2010 and 1990, although the corresponding analysis for the intermediate year of 2000 is available by request.

For reference, columns 0 and 1 of Table 7 display the combined changes to migration by city group (from Table 6). We also include the breakdown by age-education type which shows the decline present *within* person type, just as Molloy et al. (2011) first showed. Columns 2 through 5 then report on counterfactual weightings. Column 2 shifts the composition of age-education within each city origin, but leaves the spatial distribution of population as in the baseline. This shows that aging of the population makes up a nontrivial portion of the aggregate decline in migration rates for all cities, and about half the decline on average.[48] Interestingly, however, it makes up a smaller proportion of the change for more mobile (37 percent) and less deeply rooted locations (29 percent) than low mobility locations (70 percent).

Column 3 then returns to the baseline age/education weights but rebalances the spatial distribution of population across cities to 1990 levels. The projection is for reduced previous migration rates by almost 0.1 percentage point, indicating that the trend in population growth has been to be toward more mobile locations on average, and hence continuing population growth in these locations has actually been a counterweight to the aggregate decline.

Column 4 then returns the spatial distribution and age weights to the baseline, but rebalances the population to reflect 1990 birthplace source. This would reflect, for example, that there

---

[48]There are slight changes in the within-person type because of interactions between age and point in the income distribution, states we left unchanged.

are now more people from Los Angeles and fewer from Boston, a longer run symptom of the population trends exhibited in column 3 and earlier in the paper. This shows that small but still economically meaningful portion of the decline is due the shifting composition of cohort birthplace. In 1990, more people hailing from deeply rooted birthplaces would mean more return migration. Note that under this change alone, the migration rates would be higher for all age/education groups (with the young especially affected). Essentially all of this change is occurring in mobile locations.

To this point, column 5 returns the birthplace location weights to the baseline, but counterfactually projects the share of workers living "at home," in their birth location, to 1990 levels. That is, in contrast to column 4, there is a different proportion of residents not-at-home, but those not-at-home are sourced by birthplace origin (e.g. LA versus Boston) just as in the baseline. Fewer residents at home in 1990 results in a substantial effect of migration rates. The reduction of 0.24 percentage points is comparable in magnitude to the effect of an aging population. It is present in types of cities but largest in mobile and shallow rooted locations.

This effect appears to be the primary driver of the spatial heterogeneity in the decline, and is also pertinent for the aggregate. More residents living in their home location is an increase in the effective home preference for the aggregate population. It is valuable to explain the spatial heterogeneity as well as the aggregate. We attribute at-home status to be a direct mechanism, the consequence of stabilizing population growth rates, because we have estimated a home utility premium separately from move costs.

6.3   Decomposing the Sources of Decline: Changing Environments

Table 8 proceeds to decompose by environmental factors. Column 6 (the numbering continues from Table 7) shows the direct impact of changing rootedness by using the baseline weights (including at-home status) and income distributions, but altering the rootedness of the home locations. The effect here is confined to shallow rooted locations whose roots have deepened since 1990–mainly, California cities and Miami–where by 2010 residents are less likely to move away. Column 7 combines the change in rootedness with the change in the share at home, in which the effects are more widespread and especially large in places where roots are deepening.

Columns 8, 9, and 10 show the effects of the changing income search environment, starting with the information parameters. Here we have to make some decisions about the size of the change in the λ, so specific magnitudes should be interpreted with caution. For this simulation we moved λ from 0.5 in 2010 to 0.125 in 1990, a halving in each decade. By increasing the value to local search, this produces a small but across-the-board decline in mobility. The effect is slightly bigger in locations with more disperse incomes, where the change to local option value is higher. At the utility parameters we estimate, this effect is clearly not large enough to generate by itself the fullness of the observed decline, even at the extreme admissible values for

Table 7: Simulations: Counterfactual Migration Rates at 1990 Weightings

		0	1	2	3	4	5
		All	All	Decompositions: 1990 Values of:			
		Model 2010	Model 1990	Age/Educ.	Spatial	Birthplace	At Home
City Type							
	All	2.80	3.36	3.11	2.72	2.85	3.04
Mobility:	Low	2.46	2.89	2.75	2.42	2.47	2.68
	High	3.25	4.16	3.59	3.21	3.41	3.50
Roots	Deep	2.67	3.07	3.01	2.59	2.69	2.82
	Shallow	3.04	3.93	3.31	2.96	3.20	3.46
Income Disp.	Low, Noncoll.	2.07	2.38	2.42	2.02	2.11	2.12
	Low, Coll.	3.58	4.32	4.17	3.34	3.67	3.90
	High, Noncoll.	2.63	3.50	3.00	2.61	2.72	3.02
	High, Coll.	3.18	3.90	3.70	3.08	3.18	3.37
Person Type:							
Noncollege	20	4.55	5.02	4.53	4.42	4.72	5.15
	30	2.97	3.15	2.96	2.92	3.06	3.13
	40	1.99	2.32	2.00	1.98	2.05	2.24
	50	1.53	1.77	1.56	1.54	1.58	1.68
College	20	8.28	8.91	8.23	7.76	8.38	9.28
	30	4.03	4.07	4.05	3.82	4.09	4.14
	40	1.88	1.96	1.91	1.80	1.91	1.99
	50	1.54	1.60	1.55	1.49	1.56	1.61

NOTES: All figures are in percentages (%).

λ. However, we believe the qualitative pattern to be important, especially if the small changes in migration feedback to larger cumulative effects.

Column 9 returns to baseline λ but uses the 1990 distributions of income. This has a small direct effect on the aggregate move rate. Breaking down the rates by income dispersion shows that the effect is mostly due to college graduates in high dispersion locations, where dispersion is trending down for college educated workers in these markets. While the overall effects of changes to the income distribution are not massive, they are relevant for certain cities showing mobility declines, including several in Texas, North Carolina, and Florida, as well as Washington DC. This is notable result given that many of these decliners continue to grow in population and show little deepening of roots. Hence, different explanations apply to different regions. Moreover, the information parameter shows greater effects in these cities, indicating the interactive effects.

Column 10 combines all environmental factors of rootedness, income distribution, and information parameter. The net effect shows that environmental factors of deepening roots, changing income distributions, and increasing information can explain a meaningful amount of the decline in mobile places, but actually caused an increase in less mobile places. Finally, column 11 combines the environmental changes with the at home share weighting. The offsetting of the two effects accounts for the smaller decline in low mobility places, but combines with environmental

Table 8: Simulations: Counterfactual Migration Rates at 1990 Environmental Changes

		0 All	1 All	6	7	8	9	10	11
					Decompositions: 1990 Values of:				
		Model 2010	Model 1990	Roots	Roots+ At Home	Info	Income	All Env	All Env+ At Home
City Type									
	All	2.80	3.36	2.80	3.03	2.81	2.82	2.82	3.04
Mobility:	Low	2.46	2.89	2.48	2.69	2.47	2.36	2.40	2.59
	High	3.25	4.16	3.22	3.47	3.27	3.42	3.37	3.63
Roots	Deep	2.67	3.07	2.65	2.80	2.68	2.65	2.64	2.77
	Shallow	3.04	3.93	3.08	3.47	3.06	3.14	3.16	3.55
Income Disp.	Low, Noncoll.	2.07	2.38	2.06	2.10	2.08	2.06	2.04	2.08
	Low, Coll.	3.58	4.32	3.56	3.87	3.59	3.54	3.53	3.84
	High, Noncoll.	2.63	3.50	2.64	3.02	2.64	2.66	2.67	3.04
	High, Coll.	3.18	3.90	3.19	3.36	3.20	3.25	3.25	3.43
Person Type:									
Noncollege	20	4.55	5.02	4.56	5.12	4.58	4.57	4.58	5.15
	30	2.97	3.15	2.97	3.13	2.98	2.99	2.99	3.13
	40	1.99	2.32	1.99	2.23	1.99	2.01	2.00	2.23
	50	1.53	1.77	1.53	1.68	1.53	1.52	1.53	1.67
College	20	8.28	8.91	8.24	9.22	8.34	8.39	8.38	9.36
	30	4.03	4.07	4.02	4.14	4.04	4.05	4.05	4.15
	40	1.88	1.96	1.88	1.99	1.89	1.88	1.88	1.98
	50	1.54	1.60	1.54	1.61	1.54	1.54	1.54	1.61

NOTES: All figures are in percentages (%).

factors to magnify the larger decline in mobile places. Notably, these effects are seen within each age-education group.

In summary, the story that emerges from the simulation is that the spatial heterogeneity in the decline can be traced to deepening rootedness and changing income distributions of the more mobile locations, with changes in information availability amplifying the latter. The direct effects of these factors, however, are measured to be relatively small, but when combined with increasing at-home status, these explain nearly half of the decline nationally, but more so in mobile and shallow-rooted locations. As the generic model simulations in section 4 showed, preferences for home can amplify other primitive changes that cause less out mobility. The simulations using the estimated model indicate that the amplification is quantitatively important.

7   Conclusion

This paper has showed that the decline in internal migration in the U.S. has primarily occurred in locations that typically have higher rates of out migration, using the spatial heterogeneity to illuminate the mechanisms in play. We develop, estimate, and simulate a location choice model that simultaneously incorporates multiple mechanisms and rich degree of individual and

spatial heterogeneities. The model illustrates the potential importance of home preference, and then quantifies the effects of various channels. We find a quantitatively important role played by increasing strong and pervasive home preference. Other explanations have smaller direct effects, although we are reluctant to dismiss them, since small primitive changes can be amplified by the existence of home preference.

The paper's contribution is to elaborate on (and mediate between) explanations for the decline in mobility offered in the literature. We find some evidence for both the "long run spatial equilibrium" (Partridge et al. (2012)) and "ease of information" (Kaplan and Schulhofer-Wohl (2017)) explanations, but re-interpreted after recognizing the importance of home preference. People have, on average, a positive preference for home, and this greatly amplifies the effects of primitive changes in net and gross mobility. Home preference affects gross mobility long after net mobility has settled into a long run spatial equilibrium, especially if the strength of the preference depends on the depth of social ties in a region. Home preference will also amplify primitive changes to gross mobility by altering the long-run steady state levels of households living at home or away from home (even when for idiosyncratic reasons).

While we see this paper as significant progress on the economic front by elucidating the mechanism, the implications it leaves for policy are somewhat complicated. In one sense, "all is well" because declining mobility is the result of households optimally choosing to stay in place, not constrained by some new friction. It is simply the natural consequence of population "spreading out" and then "settling in." In another sense, however, home preference could be problematic for the functioning of the labor market if workers are willing to take less productive job matches. Both alarmists and non alarmists may be partly right.

What we hope to offer is a path forward for studying policy prescriptions in light of the importance of home preference. A cost-benefit analysis of migration incentives versus place-based policy may depend on the strength of home preference (which itself depends on the location's history of population adjustment). Policy makers might also consider separate incentives for workers living at home versus away from home, or for workers whose home has a stronger pull. (Specifically what such policies might be is mere speculation at this point.) Broadly, our conclusion is that home preference has import for regional adjustment, and cannot be ignored when making regional policy.

References

Albouy, D. (2009). The unequal burden of federal taxation. *Journal of Political Economy 117*(4), 635–667.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Bayer, C. and F. Juessen (2012). On the dynamics of interstate migration: Migration costs and self-selection. *Review of Economic Dynamics 15*(3), 377–401.

Bayer, P., N. Keohane, and C. Timmins (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management 58*(1), 1 – 14.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

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.

Chen, Y. and S. S. Rosenthal (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics 64*(3), 519 – 537.

Chinitz, B. (1986). The regional transformation of the american economy. *American Economic Review 76*(2), 300–304.

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.

Davis, M. A., J. Gregory, D. A. Hartley, and K. T. Tan (2016). Neighborhood choices and neighborhood effects. Technical report, mimeo, Rutgers University.

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.

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.

Flood, S., M. King, S. Ruggles, and J. R. Warren. (2015). Integrated public use microdata series, current population survey: Version 4.0. [machine-readable database]. minneapolis: University of minnesota. 2015.

Ganong, P. and D. Shoag (2013). Why has regional income convergence in the us declined? $https : //papers.ssrn.com/sol3/papers.cfm?abstract_id = 2081216$.

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

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.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Kambourov, G. and I. Manovskii (2008). Rising occupational and industry mobility in the united states: 1968–97. *International Economic Review 49*(1), 41–79.

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.

Ma, L. (2017). Learning in a hedonic framework: Valuing brownfield remediation. mimeo., University of Kentucky.

Mangum, K. (2016). Coming and going: Metropolitan area heterogeneity in migration rates. Technical Report 2016-1-1, Usery Workplace Research Group.

Miller, E. (1973). Is out-migration affected by economic conditions? *Southern Economic Journal 39*(3), 396–405.

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. S., C. L. Smith, and A. Wozniak (2011). Internal migration in the united states. *Journal of Economic Perspectives 25*(3), 173–196.

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

Morten, M. and J. Oliveira (2016, April). Paving the way to development: Costly migration and labor market integration. Working Paper 22158, National Bureau of Economic Research.

Partridge, M. D., D. S. Rickman, M. R. Olfert, and K. Ali (2012). Dwindling u.s. internal migration: Evidence of spatial equilibrium or structural shifts in local labor markets? *Regional Science and Urban Economics 42*(1-2), 375–388.

PSID (2014). Panel study of income dynamics, public use dataset. Technical report, Institute for Social Research, University of Michigan Survey Research Center.

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

Roback, J. (1982). Wages, rents, and quality of life. *Journal of Political Economy 90*(6), 1257–1278.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek. (2015). Integrated public use microdata series: Version 6.0 [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.

Zabek, M. (2018). Local ties in spatial equilibrium. mimeo., University of Michigan.

Appendices

A   Calculation of Roots

The Census data for 1950, 1970, 1980, 1990, and 2000 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 commuting zone (CZ). We identify a birth cohort by looking at all individuals who are less than 10 years old in a particular Census wave. For example, a twenty-something in 2010 was aged less than 10 in 1990. For the cohort living in each CZ, we calculate the percentage of their parents who were born in a state the CZ includes. 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. We ascribe the cohort-CZ combination to have the rootedness measured by this fraction.

There are a few possible concerns given that we use metro area (CZ) 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 Texas. When comparing across cohorts, if the measurement error is similar, the *change* in rootedness is still accurate. But if we are comparing a CZ in a state with several large cities, such as Texas, to a CZ 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 CZs.

We use the location of residence for children under 10 as home in our cohort matching. While a five-year-old in Dallas is more likely to have been born in Dallas than a 35-year-old, there is still 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.

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. How do we define their 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.

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 considers the relationship between rootedness and education to be the same in every city.

B CCP Method of Moments Estimator

Recall the expression for the log odds ratio of a given choice k:

$$
\begin{aligned}
ln(\sigma_{k,o}) - ln(\sigma_{0,o}) + \beta(ln(\sigma'_{0,k}) - ln(\sigma'_{0,0})) = {} & \underbrace{u_{k,o} - u_{0,o}}_{1} \\
& - \underbrace{(m_{k,o} - m_{0,o})}_{2} \\
& + \underbrace{\beta[u'_{0,k} - u'_{0,0}]}_{3} \\
& - \underbrace{\beta[m_{0,1} - m_{0,0}]}_{4}
\end{aligned}
\tag{12}
$$

(12) has data on the left and utility functions with parameters on the right. The linear representation is straightforward to derive, if a bit tedious. We write it out in four parts, naming various matrices x_i^k for collection in the last step.

1. Current period flow utilities can be decomposed as

$$
\mu_y \underbrace{[w_c^j I(j = o) + w_0^j I(j \neq o)]}_{x_1^1} + \alpha_0 \underbrace{I(j = h)}_{x_2^1} + \alpha_r \underbrace{I(j = h)R_j}_{x_3^1}
$$

where I are indicator functions for, respectively, whether j is the current origin and whether it is home, and R is the location's rootedness for that type.

2. Current period moving costs can be decomposed as

$$
mc_\tau \underbrace{I(j \neq o)I(\tau = \hat{\tau})}_{x_1^2} + mc_d \underbrace{I(j \neq o)D_{j,o}}_{x_2^2} + mc_o \underbrace{I(j \neq o)(o = J)}_{x_3^2} + mc_j \underbrace{I(j \neq o)(j = J)}_{x_4^2}
$$

50

where $I(j \neq o)$ is an indicator for whether j is a different location from origin o (thus toggling the move costs), $I(\tau = \hat{\tau})$ is an indicator for type $\hat{\tau}$ (toggling the type specific shifter), $D_{j,o}$ is a matrix of pairwise distances between locations, and $(o = J)$ or $(j = J)$ is a matrix of indicators for each location in the system (toggling the corresponding toll costs).

3. Next period flow utilities, for the return to location 0 (i.e. finite dependence step), are

$$\mu_y \underbrace{[w_c^0 I(j = 0) + w_0^0 I(j \neq 0)]}_{x_1^3} + \alpha_0 \underbrace{I(0 = h)}_{x_2^3} + \alpha_r \underbrace{I(0 = h)R_0}_{x_3^3}$$

where the indicators are as in 1 above, but for the location being the normalizing place 0.

4. Next period move costs, for the return to location 0 (i.e. finite dependence step), are

$$mc_\tau \underbrace{I(j \neq 0)I(\tau = \hat{\tau} + 1)}_{x_1^4} + mc_d \underbrace{I(j \neq 0)D_{0,j}}_{x_2^4} + mc_o \underbrace{I(j \neq 0)(j = J)}_{x_3^4} + mc_j \underbrace{I(j \neq 0)(0 = J)}_{x_4^4}$$

where the indicators are as in 2 above, but for the location being the normalizing place 0.

The righthand side for a given choice can then be expressed in matrix form as

$$\begin{bmatrix} x_1^1 + x_1^3 & x_2^1 + x_2^3 & x_3^1 + x_3^3 & x_1^2 + x_1^4 & x_2^2 + x_2^4 & x_3^2 + x_3^4 + x_4^2 + x_4^4 \end{bmatrix} * \begin{bmatrix} \mu_y \\ \alpha_0 \\ \alpha_r \\ [mc_\tau] \\ [mc_d] \\ [mc_{0:J}] \end{bmatrix} \quad (13)$$

where mc_τ is a vector of parameters for type-specific move costs, mc_d are the parameters on the pairwise distances, and $mc_{0:J}$ is a vector the location-specific move costs parameters.

Let X_j be the collected matrix of x matrices in (13) for some option j, and let θ be the vector of parameters. Now, the full log odds ratio difference, (11) above, can be written as

$$ln(\sigma_{k,o}) - ln(\sigma_{0,o}) - \beta(ln(\sigma_{0,k}) - ln(\sigma_{0,0})) = (X_j - X_0)\theta$$

C   Forming the Moment Conditions

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

Some states contain multiple MSAs and some MSAs straddle state political boundaries. Note the entire US is partitioned by our geographic areas, so that by definition every state has at least one MSA (the outside option, location 0) 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 MSAs, so even in this easy case, we have overlapping boundaries. Hence, we need to map between state of birth and MSA of birth.

At one extreme, we could assign someone in their birth state to be fully at-home. At the other extreme, we could fully partition the observation by population in his year of birth. To be more specific, consider an examples. Say we observe someone living in Houston whose state of birth is Texas. We could simply assume he was born in Houston, but actually he could have been born in Houston, Dallas, San Antonio, Austin, some smaller MSA like El Paso or Abilene, or in a rural county. We could also use 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 this drastically understates the at-home share because it effectively assumes full mobility within state. In practice, we settle in between. For instate births, we first assign the fraction of out of state residents as a probability of the person migrating in to Houston. We then assign birthplace as one minus the in-migration probability times the population of the sending locations (Dallas, Austin, etc.) in their birth year. On the other hand, say we observe two people–one born in New Jersey and another born in New York. By the same method as the Texas example, each of them will be by some fraction assigned to be a New York City MSA native.

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

We then proceed to calculate choice probabilities. This is where we need to make some decisions about when and how to combine cells in order to get reliable CCP estimates without

losing the kind of detail in the stylized facts presented above. We create three tables of move probabilities from the ACS and one form the IRS. These are:

1. the probability of migrating (to anywhere), by age, education, origin, and whether the origin is home. This will capture the main differences in move 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.

2. the probability of returning home (i.e. moving from an away-location back to one's birth-place) 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 non natives 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.

3. the probability of moving into a location (as a destination) by age and education, com-bining over birthplace and origin for precision. This captures differences in preferences for destinations for workers with different skills, and helps to identify the income component of utility. Call this p_3.

4. probability of moving into a location by origin, from the IRS data. This helps to identify distance and toll move costs, and also provides some information on the utility from income. These data are already aggregated over types. Call this p_4.

Note that each p_n is computed by age and education, although we drop subscripts for expo-sition. The full matrix of cell-specific choice probabilities is then a formed from the product of these for the corresponding cases, which are as follows. For the reflexive entry (i.e., "stayers," the diagonal in the matrix) is simply the probability p_1. Moves home are a conditional proba-bility, $(1 - p_1)p_2$. Moves elsewhere for people living away from their birthplace, the conditional probability is $(1 - p_1)(1 - p_2)(\omega p_3 + (1 - \omega)p_4)$, where the ω is a weight assigned to average the information from the ACS destinations and IRS flow matrix. We use a distance weight, $\omega = 1/d$, so that proximate moves are more heavily represented by the IRS (geographically detailed) data and distant moves by the ACS (demographically detailed), although results are not very sensitive to the choice of weighting. Finally, for people living in their birthplace, the probability is $(1 - p_1)(\omega p_3 + (1 - \omega)p_4)$.

All together, this forms a full $J \times J$ matrix (origin to destination) of flow probabilities for each type of worker, which is placed in the lefthand side of (11) above.

D   Additional Simulation Results

Table A1: Simulations: Baseline and Counterfactual Migration Rates

		0 All Model 2010	1 All Model 1990	2 Age/Educ.	3 Spatial	4 Birthplace	5 At Home
				Decompositions: 1990 Values of:			
Albany	16	2.24	2.95	2.48	2.24	2.53	2.54
Albuquerque	20	5.17	6.40	5.83	5.17	5.47	5.72
Atlanta	52	2.96	3.43	3.44	2.96	3.05	2.93
Austin	64	2.81	4.02	3.13	2.81	3.18	2.88
Baltimore	72	2.11	2.57	2.46	2.11	2.07	2.18
Birmingham	100	3.76	3.87	4.09	3.76	3.86	3.64
Boston	112	2.26	2.71	2.78	2.26	2.23	2.27
Buffalo	128	1.81	2.02	2.00	1.81	1.95	2.11
Charlotte	152	2.80	2.97	3.16	2.80	2.85	2.46
Chicago	160	2.88	3.39	3.28	2.88	2.76	3.11
Cincinnati	164	1.99	2.40	2.31	1.99	2.00	2.14
Cleveland	168	2.24	2.96	2.60	2.24	2.28	2.85
Columbus	184	2.50	2.93	2.65	2.50	2.74	2.84
Dallas-FW	192	3.83	5.18	4.39	3.83	3.99	4.17
Dayton-Springfield	200	2.17	2.53	2.27	2.17	2.26	2.74
Denver	208	3.06	3.57	3.33	3.06	3.17	3.17
Detroit	216	0.61	0.68	0.71	0.61	0.59	0.71
Winston-Salem	312	2.23	2.14	2.78	2.23	2.26	1.96
Greenville SC	316	3.23	3.03	3.55	3.23	3.13	2.91
Hartford	328	1.74	2.22	2.08	1.74	1.80	1.96
Houston	336	3.21	4.25	3.58	3.21	3.30	3.62
Indianapolis	348	3.52	3.97	3.83	3.52	3.96	3.94
Jacksonville FL	359	3.65	4.55	4.04	3.65	4.02	3.85
Kansas City	376	3.88	4.39	4.34	3.88	4.19	4.06
Las Vegas	412	4.83	4.61	4.59	4.83	4.94	4.92
Los Angeles-Riverside	448	2.36	3.05	2.46	2.36	2.49	3.05
Louisville	452	2.38	2.23	2.60	2.38	2.55	2.17
Memphis	492	4.85	5.55	5.11	4.85	4.87	5.00
Miami	500	4.29	5.43	4.52	4.29	4.60	5.08
Milwaukee	508	2.84	3.29	3.19	2.84	2.84	3.01
Minneapolis	512	2.67	3.32	3.06	2.67	2.50	2.78
Nashville	536	3.72	3.68	3.94	3.72	3.86	3.41
New York	560	3.48	4.14	4.05	3.48	3.26	3.61
Norfolk-VA Bch	572	4.03	4.58	4.25	4.03	4.20	4.12
Oklahoma City	588	3.60	4.18	3.94	3.60	3.78	3.77
Orlando	596	4.81	5.35	5.15	4.81	5.10	5.07
Philadelphia	616	1.99	2.37	2.18	1.99	1.90	2.21
Phoenix	620	3.28	3.83	3.62	3.28	3.36	3.42
Pittsburgh	628	1.83	1.70	1.97	1.83	1.81	1.76
Portland	644	2.47	2.65	2.80	2.47	2.51	2.37
Providence	648	1.23	1.95	1.43	1.23	1.31	1.45
Raleigh-Durham	664	2.63	2.79	3.05	2.63	2.70	2.29
Richmond	676	2.79	2.97	3.00	2.79	3.22	2.83
Sacramento	692	2.59	4.45	2.83	2.59	2.97	3.52
St Louis	704	2.00	2.31	2.17	2.00	1.97	2.19
Salt Lake City	716	4.25	4.40	4.33	4.25	4.72	4.43
San Antonio	724	2.62	3.22	2.88	2.62	3.04	2.86
San Diego	732	3.40	5.17	3.57	3.40	3.72	4.27
San Francisco	736	2.36	3.90	2.67	2.36	2.42	2.93
San Jose	740	2.67	4.96	2.89	2.67	2.76	3.85
Seattle	760	3.04	3.73	3.34	3.04	3.08	3.07
Tampa	828	3.22	4.13	3.61	3.22	3.37	3.46
Tucson	852	3.98	5.03	4.48	3.98	4.32	4.53
Washington DC	884	2.74	3.92	3.19	2.74	2.83	2.97

Table A2: Simulations: Baseline and Counterfactual Migration Rates

		0 All	1 All	6	7	8	9	10	11
				Decompositions: 1990 Values of:					
		Model 2010	Model 1990	Roots	Roots+ At Home	Info	Income	All Env	All Env+ At Home
Albany	16	2.24	2.95	2.25	2.54	2.24	2.18	2.19	2.47
Albuquerque	20	5.17	6.40	5.10	5.65	5.19	5.17	5.12	5.67
Atlanta	52	2.96	3.43	2.88	2.85	2.97	3.01	2.94	2.91
Austin	64	2.81	4.02	2.74	2.81	2.82	3.53	3.29	3.39
Baltimore	72	2.11	2.57	2.12	2.19	2.11	2.10	2.11	2.19
Birmingham	100	3.76	3.87	3.73	3.61	3.77	3.67	3.67	3.57
Boston	112	2.26	2.71	2.23	2.24	2.26	2.22	2.21	2.22
Buffalo	128	1.81	2.02	1.81	2.11	1.81	1.52	1.57	1.84
Charlotte	152	2.80	2.97	2.76	2.41	2.81	3.07	2.98	2.61
Chicago	160	2.88	3.39	2.88	3.11	2.90	2.72	2.76	2.97
Cincinnati	164	1.99	2.40	1.98	2.13	1.99	1.91	1.92	2.07
Cleveland	168	2.24	2.96	2.28	2.89	2.24	1.93	2.03	2.57
Columbus	184	2.50	2.93	2.49	2.82	2.51	2.38	2.39	2.71
Dallas-FW	192	3.83	5.18	3.73	4.09	3.85	4.11	3.97	4.35
Dayton-Springfield	200	2.17	2.53	2.16	2.73	2.17	1.82	1.88	2.37
Denver	208	3.06	3.57	3.02	3.13	3.08	3.25	3.19	3.30
Detroit	216	0.61	0.68	0.64	0.74	0.61	0.47	0.52	0.60
Winston-Salem	312	2.23	2.14	2.21	1.94	2.24	1.92	1.97	1.72
Greenville SC	316	3.23	3.03	3.21	2.89	3.24	2.99	3.03	2.71
Hartford	328	1.74	2.22	1.73	1.95	1.75	1.67	1.68	1.89
Houston	336	3.21	4.25	3.14	3.56	3.23	3.40	3.32	3.75
Indianapolis	348	3.52	3.97	3.51	3.93	3.53	3.28	3.31	3.71
Jacksonville FL	359	3.65	4.55	3.67	3.86	3.66	3.95	3.91	4.11
Kansas City	376	3.88	4.39	3.81	3.99	3.89	3.83	3.78	3.96
Las Vegas	412	4.83	4.61	4.70	4.80	4.84	4.93	4.79	4.88
Los Angeles-Riverside	448	2.36	3.05	2.56	3.20	2.37	1.98	2.23	2.80
Louisville	452	2.38	2.23	2.36	2.15	2.39	2.26	2.27	2.05
Memphis	492	4.85	5.55	4.82	4.97	4.87	5.15	5.07	5.22
Miami	500	4.29	5.43	4.36	5.11	4.31	4.13	4.25	5.00
Milwaukee	508	2.84	3.29	2.81	2.98	2.85	2.93	2.88	3.04
Minneapolis	512	2.67	3.32	2.64	2.75	2.69	2.84	2.79	2.91
Nashville	536	3.72	3.68	3.64	3.33	3.75	3.93	3.80	3.46
New York	560	3.48	4.14	3.51	3.65	3.51	3.39	3.45	3.57
Norfolk-VA Bch	572	4.03	4.58	4.03	4.12	4.03	4.13	4.10	4.18
Oklahoma City	588	3.60	4.18	3.54	3.72	3.61	3.62	3.56	3.74
Orlando	596	4.81	5.35	4.77	5.04	4.82	4.65	4.65	4.92
Philadelphia	616	1.99	2.37	1.99	2.21	1.99	1.91	1.93	2.14
Phoenix	620	3.28	3.83	3.23	3.37	3.29	3.37	3.32	3.46
Pittsburgh	628	1.83	1.70	1.82	1.76	1.83	1.60	1.64	1.57
Portland	644	2.47	2.65	2.43	2.34	2.47	2.52	2.48	2.39
Providence	648	1.23	1.95	1.23	1.46	1.23	1.23	1.23	1.45
Raleigh-Durham	664	2.63	2.79	2.58	2.23	2.63	2.86	2.76	2.39
Richmond	676	2.79	2.97	2.77	2.80	2.80	2.75	2.73	2.75
Sacramento	692	2.59	4.45	2.71	3.59	2.60	2.70	2.80	3.72
St Louis	704	2.00	2.31	1.99	2.18	2.00	1.95	1.95	2.14
Salt Lake City	716	4.25	4.40	4.13	4.30	4.26	4.16	4.07	4.24
San Antonio	724	2.62	3.22	2.59	2.83	2.63	2.66	2.64	2.88
San Diego	732	3.40	5.17	3.50	4.32	3.42	3.63	3.70	4.56
San Francisco	736	2.36	3.90	2.44	2.99	2.38	2.69	2.72	3.32
San Jose	740	2.67	4.96	2.71	3.86	2.69	3.00	2.99	4.24
Seattle	760	3.04	3.73	3.00	3.04	3.05	3.48	3.35	3.39
Tampa	828	3.22	4.13	3.20	3.44	3.23	3.40	3.35	3.60
Tucson	852	3.98	5.03	3.93	4.46	4.00	3.91	3.87	4.40
Washington DC	884	2.74	3.92	2.72	2.95	2.75	3.18	3.08	3.33