NBER WORKING PAPER SERIES

THE DETERMINANTS OF DECLINING INTERNAL MIGRATION

William W. Olney
Owen Thompson

Working Paper 32123
http://www.nber.org/papers/w32123

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2024

We thank numerous colleagues and seminar participants for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Determinants of Declining Internal Migration
William W. Olney and Owen Thompson
NBER Working Paper No. 32123
February 2024
JEL No. J31,J61,R23,R31

ABSTRACT

Internal migration in the United States has declined substantially over the past several decades, which has important implications for individual welfare, macroeconomic adjustments, and other key outcomes. This paper studies the determinants of internal migration and how they have changed over time. We use administrative data from the IRS covering the universe of bilateral moves between every Commuting Zone (CZ) in the country over a 23 year period. This data is linked to information on local wage levels and home prices, and we estimate bilateral migration determinants in rich regression specifications that contain CZ-pair fixed effects. Consistent with theoretical predictions, results show that migration is decreasing with origin wages and destination home prices, and is increasing with destination wages and origin home prices. We then examine the contributions of earnings and home prices to the noted overall decline in internal migration. These analyses show that wages on their own would have led to an increase in migration rates, primarily because migrants are increasingly responsive to high earnings levels in potential destination CZs. However, these wage effects have been more than offset by housing related factors, which have increasingly impeded internal mobility. In particular, migration has become much less responsive to housing prices in the origin CZ, such that many households that would have left in response to high home prices several decades ago now choose to stay.

William W. Olney
Department of Economics
Williams College
24 Hopkins Hall Dr
Williamstown, ma 01267
wwo1@williams.edu

Owen Thompson
Department of Economics
Williams College
24 Hopkins Hall Drive
Williamstown, MA 01267
and NBER
ot3@williams.edu

1 Introduction

The United States has historically been a country with high levels of internal migration and the willingness of Americans to move has contributed to its reputation for having a dynamic national economy.[1] Prominent historical examples of internal migration in the U.S. include westward expansion, rural to urban migration in the period of industrialization, the Dust Bowl epoch, and the Great Migration(s) of African Americans out of the South, among many others. These moves were often motivated by economic conditions and internal migration has been found to be a powerful tool for reducing poverty and increasing economic opportunity, both within and between generations. Moving from socioeconomically under-performing locations to more prosperous ones, commonly referred to as "moving to opportunity," is found to increase a variety of future outcomes [Chetty and Hendren, 2018a,b].[2]

In light of these historical migration episodes and its centrality to popular narratives about American dynamism, an important and surprising development over the past several decades is that internal migration rates in the U.S. have steadily *declined*. Figure 1 demonstrates this development by plotting cross-state migration rates in Current Population Survey (CPS) data, as well as cross-Commuting-Zone (CZ) migration rates using Internal Revenue Service (IRS) data.[3] Similar declines have been documented using alternative migration measures, including cross-region, cross-MSA, or cross-county moves and the share of individuals living outside their state of birth, as well as alternative data sets, including the ACS and Decennial Censuses [Molloy et al., 2011, Jia et al., 2022].

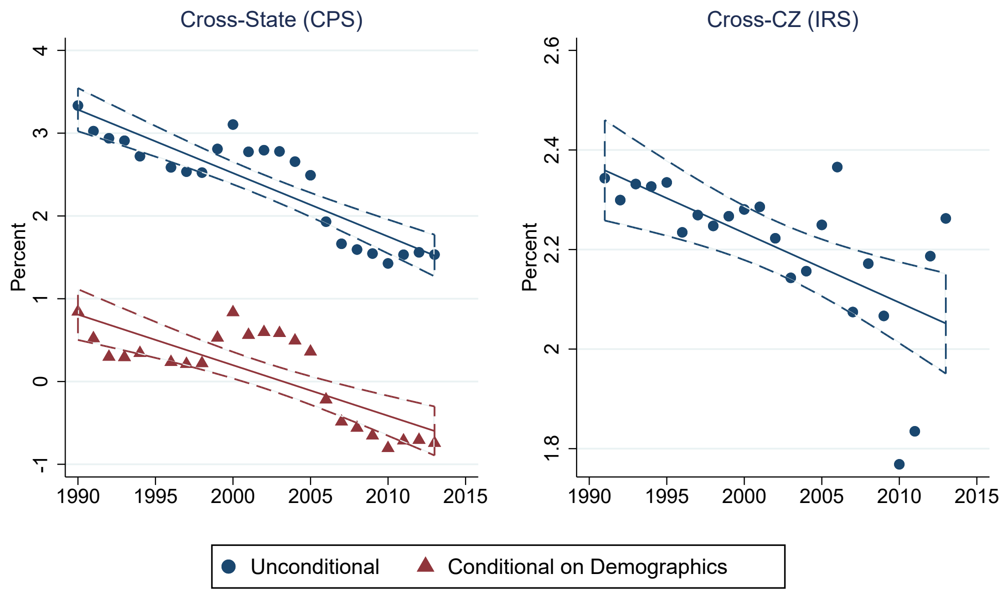
While there is broad agreement that internal migration is declining and that this decline has far reaching economic implications, much less is known about exactly *why* Americans are moving less frequently. Many obvious potential explanations, especially those related to changing population characteristics, can readily be ruled out by simply conditioning on demographic measures. As a demonstration of this, Figure 1 re-estimates the cross-state migration trend in the CPS while controlling for age, home ownership, education, race/ethnicity, marital status, employment status, and (international) immigrant status. While the *level* of migration falls substantially after conditioning on these characteristics, indicating that they do indeed affect migration behavior, the *slopes* of the unconditional and conditional lines are approximately equal, indicating that the declines in migration over time are not attributable to these demographic characteristics. Similarly, migration has fallen *within* all of these demographic and socioeconomic groups [Molloy et al., 2011]. This suggests that declining internal migration is due to changes in structural aspects of migration choice, which affect individuals and families of many different backgrounds.

[1] See Bentivogli and Pagano [1999], Ferrie [2003], Magrini [2004].

[2] Although Derenoncourt [2022] shows that in some cases these positive outcomes are more nuanced.

[3] The unconditional fitted line on the left fell from approximately 3.30 in 1990 to 1.50 in 2013, while the fitted line on the right fell from approximately 2.35 in 1991 to 2.05 in 2013.

Figure 1
Internal Migration Rates



Notes: The left panel uses CPS data to measure annual internal cross-state migration rates. The conditional migration rate controls for home ownership, age, education, sex, race/ethnicity, marital status, employment status, and immigrant status. The right panel uses IRS data to measure annual internal cross-commuting-zone (CZ) migration rates.

Better understanding these broad structural detriments of declining internal migration is the focus of the current paper. We first use the canonical Rosen-Roback model of migration choice [Roback, 1982, Rosen, 1979] to identify key migration determinants and to inform our empirical approach. This model highlights how migration depends on wages and housing prices in both the origin and destination locations; amenities in each location such as weather and recreational opportunities; and moving costs. Our data and methodology allow us to isolate the influences of wages and home prices specifically, and to account for changes in these factors on a *bilateral* basis in both the origin location and in potential destination locations.

Our primary migration data source is administrative tax data from the IRS. Using information on the tax filers location in consecutive years, this data identifies the universe of bilateral migration flows between every two-way pairing of CZs in the United States for every year from 1991 to 2013.[4] The ability to observe bilateral gross migrant flows (i.e. all moves from commuting zone A to B as well as from B to A) is strongly preferable to net flows for the current application. This is both because bilateral data allows us to disentangle push and pull factors influencing migration decisions, and because it is the decline in *gross* migration that we are interested in explaining. We link the IRS migration data to CZ-level wage

[4]Our main analysis goes through 2013 because the IRS changed how it reports migration data in 2014. However, we show in the appendix that our results are similar if we use data through 2019 and make adjustments to account for the changes in data structure.

and housing price data, from the IRS and Census respectively, to create a data set with 740 CZs that form approximately 550,000 bilateral CZ pairs, and are observed over a 23 year period.

Our preferred empirical specification uses Poisson Pseudo-Maximum Likelihood (PPML) to account for CZ-pairs with zero migrant flows in a given year. Critically, our analysis contains CZ-*pair* fixed effects, as well as year fixed effects, and controls for a variety of potentially confounding factors that vary at the CZ*year level. The CZ-pair fixed effects in particular account for an extensive set of otherwise unobservable migration determinants, including time-invariant amenities (like weather) and geographic distance (a strong proxy for moving costs), and allow us to isolate plausibly causal impacts of wages and housing prices.

Our baseline estimates show that people respond to both wages and housing prices when making migration decisions, and do so in ways that are intuitive and consistent with theoretical predictions. We start by using wage and housing price *gaps* between the destination and origin CZs, and the results indicate that migration increases with larger wage gaps but decreases with larger housing price gaps. We then estimate a more flexible specification that allows for an asymmetric response of migration to economic conditions in the origin and destination CZ. This distinction proves to be important. A 10% increase in the origin wage is associated with a 3.5% decline in migration, while a 10% increase in the destination wage is associated with a 7.8% increase in migration. For housing, a 10% increase in origin home prices is associated with a 1.4% increase in migration, but a 10% increase in destination home prices is associated with a 2.6% decline in migration. These effects are precisely estimated and are robust to a wide variety of data construction and modeling choices. The finding that migrants differentially respond to "pull" and "push" factors is consistent with previous work [Monras, 2020, Yagan, 2019, Howard, 2020, Autor et al., 2013, Wilson, 2021, 2022] and highlights the importance of not simply considering origin location conditions when examining migration decisions, which is common in studies of how local labor market shocks affect migration.

The second component of our analysis examines whether our estimates can help explain the overall decline in internal migration seen in Figure 1. To do so, we estimate models where the elasticity of migration with respect to wages and home prices in both origin and destination are allowed to vary by year, and then use these estimates along with the observed annual levels of wages and housing prices, again in both origin and destination locations, to construct *predicted* migration flows in each year. We find that these predicted migration levels track actual migration flows well: by simply taking the products of the time-varying elasticity estimates and the levels of the four key independent variables and then summing, we can predict about a third of observed migration *levels*, and more importantly can (more than) explain the overall *decline* in domestic migration rates.

The additive nature of our predicted migration calculation allows us to examine whether the decline in internal migration is driven by 1) wages or housing prices, 2) origin or destination conditions, and 3) migrant's changing responsiveness to economic conditions or changes in the levels of wages and housing prices across CZs. Our findings indicate that wages alone would have actually led to an increase in internal migration, because migration decisions have become more sensitive to wages in potential destination CZs. In contrast, housing is an impediment to moving and can effectively explain all of the aggregate decline in internal migration. Most importantly, migration decisions have become less sensitive to housing, and particularly to housing prices in the *origin* location. In all cases the key change is in household's *responsiveness* to wages and home prices, not differential changes in the *levels* of these variables themselves across different CZs. We conclude that household's decision making about how to respond to changing home prices in their current locations is key to understanding the decline in internal migration over the last three decades, more so than wage related factors or the spacial dispersion of migration determinants.

Understanding why people are becoming less responsive to origin housing prices is an inherently challenging question, given that the details of the relevant decision making processes are seldom observed. Nevertheless we do try to provide some evidence on this issue, first by investigating *where* home price elasticities have declined the most, and then by incorporating data from the CPS to identify which *demographic* groups these changes were most concentrated in. These analyses indicate that declines in sensitivity to origin home prices are most acute in large urban states with high housing prices, such as California, New York and New Jersey, and also among older, less-educated individuals who own their homes.

Our paper contributes to three broad literatures. First are studies that examine how migration affects the earnings of workers. This literature use longitudinal data to observe earnings before and after migration [Glaeser and Maré, 2001, Card et al., 2023], while others have used natural disasters, military spouses, refugee resettlement, or similar factors to identify the causal effect of moving on earnings [Deryugina et al., 2018, Burke and Miller, 2018, Nakamura et al., 2022]. These studies are important background for our own work because they suggest that moving to a higher wage area does indeed causally affect the earnings of migrants, rather than spacial differences in wage levels simply reflecting sorting or other endogenous processes. The recent study by Card et al. [2023] is particularly relevant for current purposes because they find that individual earnings increase when people move to higher wage areas, but that these location premiums are similar for workers from different industries and education levels, with little evidence of quantitatively meaningful "match effects." This suggests that even in aggregate data like ours, which does not contain information on the individual's industry or education, the mean wage in an origin CZ and a potential destination CZ are indicative of the earnings change the average individual could expect from a move.[5] More generally, the finding that locational choice affects a variety of key outcomes underscores the importance of understanding why internal migration is declining, which is our focus.

Second, we contribute to work that examines the consequences of labor market shocks on internal migration [Blanchard and Katz, 1992, Dao et al., 2017, Foote et al., 2019, Partridge et al., 2012].[6] For

[5]Card et al. [2023] also find that the earnings premium associated with a given CZ is more than offset by higher housing costs, which is broadly consistent with our finding that housing is a first-order determinant of migration decisions.

[6]A subset of these papers focus specifically on the U.S. locational choices of international migrants, for example Basso and Peri [2020] and Cadena and Kovak [2016].

example, Chinese imports adversely affected particular U.S. labor markets but, surprisingly, most evidence suggests that this has not led people to migrate out of these under-performing areas [Autor et al., 2013, Faber et al., 2022]. Our results provide a potential explanation for this finding by emphasizing the need to account for the bilateral nature of the migration decision, by considering both origin *and* destination wages. In particular, deteriorating labor market opportunities in the origin location may not induce out-migration if the individual's opportunities are also adversely affected in the most germane potential destination locations, which are often geographically proximate and subject to correlated shocks, a point emphasized by Borusyak et al. [2022].[7]

Third, we relate to papers on how housing markets can influence mobility [Ferreira et al., 2010, Plantinga et al., 2013, Coulson and Grieco, 2013, Barkema and Bayoumi, 2019, Notowidigdo, 2020]. This literature is somewhat less developed than that on wages, but has addressed a variety of important topics including how home ownership, interest rates, and the share of people with negative equity can influence migration choice. While all of these factors do appear to affect migration, it does not seem likely that the trends in any of them over the last thirty years are large enough to account for the steady decline in internal migration rates over this period. We contribute to this literature by showing that housing market considerations are in fact the key to understanding internal migration declines. However, not for the reasons often mentioned (i.e. homeownership rates and interest rates), but rather because people are becoming less responsive to changing origin home prices.

Finally, and most directly, our work builds on an existing literature that is concerned specifically with the decline in internal migration [Molloy et al., 2011, Kaplan and Schulhofer, 2017, Basso and Peri, 2020, Jia et al., 2022]. This literature has carefully documented the decline in domestic mobility, and has ruled out some potential explanations like demographic shifts or increasing remote work, but has not come to any general conclusions about the root causes of declining migration. We contribute to this literature by utilizing a comprehensive administrative data set covering the universe of bilateral moves between every pair of commuting zones in the country over a long period, and by exploiting the size and features of this data to more carefully identify how migration responds to labor market and housing conditions in both the origin and destination locations.

The remainder of the paper is structured as follows. In the next section, we provide a brief theoretical framework for analyzing the determinants of domestic migration. Section 3 outlines the data sources used in our analysis and provides some descriptive evidence. Section 4 describes our empirical strategy and our baseline results are presented in Section 5. We then use these estimates to help explain the decline in domestic migration over the last three decades in Section 6. Finally, Section 7 concludes.

[7]Other explanations for the lack of outmigration include Greenland et al. [2019] who find that migration does respond to trade shocks but over a much longer time horizon (i.e. 7-10 years) and Monras [2020] who finds more of a response of migration to "pull" factors than "push" factors. While our results do suggest a potential explanation for the low origin wage migration elasticities found in the shocks literature, we emphasize that the goal of the current paper is not to narrowly focus on trade-induced wage changes, but to examine more broadly why internal migration is declining and how this relates to labor *and* housing market conditions.

2 Theoretical Framework

This section presents a basic Rosen-Roback theoretical framework. While this is not the main contribution of our paper, it does help fix ideas and guides our empirical analysis.

Assume there are two locations: the origin commuting zone and the destination commuting zone. Individual i derives utility from locating in the origin commuting zone "o" or in the destination commuting zone "d" in year t. Utility in both locations depends on local wages, housing prices, and amenities in the following way:[8]

$$U_{iot} = Wage_{ot} - Housing_{ot} + Amenity_o + \varepsilon_{iot} \tag{1}$$

$$U_{idt} = Wage_{dt} - Housing_{dt} + Amenity_d + \varepsilon_{idt} \tag{2}$$

In this static framework, an individual's moving decision depends on the relative utilities they can obtain in the two locations. Specifically, an individual chooses to migrate if the difference between the utility in the destination and origin location is greater than the cost of moving:

$$Migrate_{iodt} = (U_{idt} - U_{iot}) - MovingCosts_t > 0 \tag{3}$$

Plugging Equations 1 and 2 into Equation 3, we find that the internal migration decision depends on conditions in the origin and destination locations in the following way:

$$Migrate_{iodt} = Wage_{dt} - Wage_{ot} - Housing_{dt} + Housing_{ot} \tag{4}$$

$$+Amenity_d - Amenity_o - MovingCosts_t + \varepsilon_{idt} - \varepsilon_{iot} \tag{5}$$

Anticipating our subsequent empirical approach, note that all observed and unobserved time-invariant factors, including time-invariant amenities and the distance between the two CZs in a given pair, can be accounted for using CZ-pair fixed effects (γ_{od}) and all location-invariant factors that change over time can be controlled for with year fixed effects (γ_t). Since we believe that amenities are predominantly time-invariant and that moving costs are determined by some combination of geographic distance and year specific factors like labor and fuel prices, amenities and moving costs can be accounted for with CZ-pair and year fixed effects, which suggests the following empirical migration equation:

$$Migrate_{iodt} = Wage_{dt} - Wage_{ot} - Housing_{dt} + Housing_{ot} + \gamma_{od} + \gamma_t + \varepsilon_{iodt} \tag{6}$$

[8]For the time being we assume amenities are time-invariant, which is a reasonable assumption for many important amenities such as weather, beaches, mountains, or distance to other population centers. However, later we empirically account for the possibility that amenities could be time-varying at the state or even CZ-level by controlling for state*year fixed effects and total migration into or out of a CZ.

Overall, we see that this basic theoretical framework generates some clear and testable predictions. Specifically, rising wages in the destination location and rising housing prices in the origin location should both increase incentives to migrate, while rising wages in the origin location and rising housing prices in the destination location should both reduce migration. We test whether migration actually responds to these economic factors in this anticipated way in the subsequent empirical analysis.

3 Data

Our main analysis combines data from three distinct sources that provide information on bilateral migration flows between CZs, local earnings levels, and local housing prices. Here we describe each of these data sources in more detail and provide descriptive statistics.

3.1 Migration Data

Annual county-to-county migration flows from 1991 through 2013 come from the IRS, which constructs this information using the mailing addresses on the same tax filer's returns in consecutive years. For every possible combination of origin and destination counties, this data reports the total number of tax exemptions that were filed in a given origin county in the previous year but filed in a different destination county in the current year.[9]

The IRS migration data is currently available through 2019. However, the IRS maintains minimum reporting thresholds for confidentiality purposes, and these thresholds were changed in 2014. Specifically, prior to 2014 all county pairs with 10 or more migrating tax returns were included in the dataset, whereas beginning in 2014 this threshold was increased to 20. Since the number of migrants from one specific county to another specific county in a single year is often quite small, this reporting threshold change roughly doubled the number of suppressed county-to-county flows. To maintain consistent migration measures over time we therefore restrict the sample in our main analyses to 1991-2013. However, in appendix Table A2 we show that our key findings are similar when using data extended through 2019, while setting migration flows between 10 and 20 to zero in years prior to 2014.[10]

The IRS migration data has the highly desirable features of coming from administrative records that incorporate a large share of the overall population, of providing high levels of geographic specificity on an annual basis, and of reporting gross rather than net migration flows. These are all significant advantages over survey based migration data sources like the American Community Survey (ACS) and Current

[9]The IRS migration data is available for both the number of returns filed, which approximate the number of migrating households, and for the number of personal exemptions, which approximates the number of migrating individuals. Since we are interested in understanding trends in total migrants, we use the exemptions-based measure in our main estimates, but also demonstrate that our findings are robust to using the returns-based measure. Households are identified across tax years using the Social Security number of the primary filer.

[10]There are also previously documented anomalies in more recent years of the IRS migration data [DeWaard et al., 2020], and in our own construction of the data we observe large outliers in migration rates in 2015 and 2017. This provides another reason to focus on the 1991-2013 period for our main analyses.

Population Survey (CPS). That being said, there are also some shortcomings of the IRS data, that are worth acknowledging. First, there could be some incompleteness due to non-filers, those that use the address of a non-local tax preparer or business on their return, those that file returns after September when the migration data for the previous calendar year is typically released, and households that form or dissolve between tax years. Despite these coverage limitations, the IRS migration data still captures a large share of U.S. households, and the noted measurement issues are generally consistent from year to year, which is important given that we are primarily interested in changes in migration over time. The other limitation of the IRS migration data is that it is aggregated, and thus we are generally unable to observe characteristics like age, race, occupation or industry of the individual migrant.[11]

Because we are primarily interested in how local labor and housing market conditions impact migration choices, we aggregate the migration data (as well as our other data sources) to the Commuting Zone level. Commuting Zones are collections of counties that are designed to approximate local labor markets, and that have been widely used in studies of local economic activity, labor market shocks, and migration. To aggregate we use the 1990 county-to-CZ crosswalks from Autor et al. [2013] and sum the number of migrants and non-migrants over all non-suppressed counties within each CZ.

3.2 Wage Data and Housing Price Data

Our second key data set is the Statistics of Income (SOI) County Level Income file, also published by the IRS and based on tax returns over the same range of years. We calculate mean household labor market earnings by taking the sum of the wage and salary income reported on all of the returns filed in each CZ-year and dividing that sum by the total number of returns filed in each CZ. We calculate the natural log of mean CZ earnings, in order to facilitate the interpretation of the corresponding coefficients as elasticities.

Our key data source for local housing prices is the county-level price index provided and described by Bogin et al. [2019]. These data use repeated sales of the same home, with adjustments for changes in observable home characteristics, to construct housing price indices that account for a large set of observable and unobservable home characteristics. This index is available at the county level from 1975 onward.[12] Relative to other commonly used indices, for instance the S&P/Case-Shiller index or the Zillow Value index, these data have the advantage of providing home value information at the local level, covering all housing markets in the U.S. (rather than only a set of medium and large cities), being available annually for our full sample period, and being free and publicly available. We aggregate this index to the CZ-level by taking its median value across all counties in each CZ. Finally, to express home

[11]We do utilize IRS information on the average Adjusted Gross Income of the bilateral migrants (see Section 5.2), and we complement our main analysis with CPS data that has demographic information of cross-state migrants (see Section 6.4.

[12]Indices are calculated using over 97 million transactions from the FHFA's "all transactions" sample and are available for all areas with at least 100 repeated sales. To limit the influence of major renovations and home-flipping, transaction pairs with annual appreciation over 30% and cases where a home was sold twice within 12 months are excluded. See Bogin et al. [2019] for details.

prices in the same units as wages we convert the index values to nominal dollars using median home values from the 2000 Census.

3.3 Other Data Sources

We supplement our data on migration, earnings levels and home values with two additional data sources.

First, we compile basic demographic characteristics at the CZ-year level from the Survey of Epidemiology and End Results (SEER) to use as covariates in some specifications. We specifically calculate, for each CZ-year, the share of the population in eight age categories (19 or younger, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80 or older), three race categories (white, Black, and other race), two ethnicity categories (Hispanic and non-Hispanic) and two gender categories (male and female).

Second, in our heterogeneity analyses we utilize data on the geographic distance between every CZ-pair. Specifically, the county distance file provided by NBER contains great circle distances between the "internal points" of each U.S. county (typically their centroids).[13] We convert these county based distance measures to the CZ-level by assigning each CZ the geographic internal point of its most populous county.

Third, we complement our baseline findings with an analysis (in Section 6.4) that utilizes CPS's Annual Social and Economic Supplement to measure state-level migration. While the more aggregate geographic unit of observation and the relatively small sample size of the CPS is less appealing than our administrative IRS data, the CPS dataset has the advantage of providing demographic information of migrants.

3.4 Descriptive Statistics

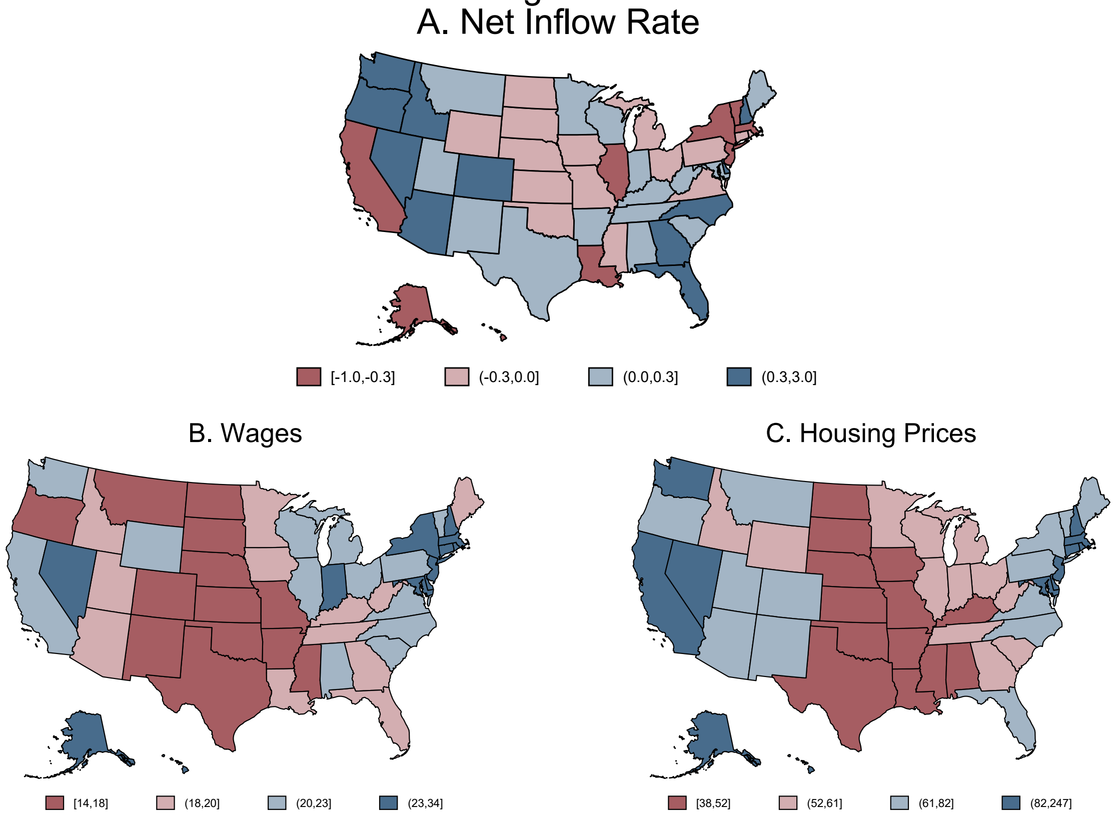
Combining these various data sources generates a data set spanning 740 Commuting Zones (CZs), 550,000 two-way pairs of CZs, and twenty three years (1991-2013). Figures 2 and 3 provide some descriptive statistics for this working data set.

To give a sense for the geography of the prevailing migration patterns, Figure 2a shows the *net* migration rate (inflows minus outflows divided by population) at the state-level. We see that on net people are typically migrating out of the Northeast, Midwest, and California and moving to the Southeast, the Southwest, and the Northwest. While net outflows from "Rust Belt" states like Michigan and Ohio are observed, it is notable that we see even stronger negative net-migration out of states like New York, New Jersey, Massachusetts, Illinois, and California. Figures 2b and 2c show mean wages and housing prices by state. Notably, many of the states experiencing negative net migration have relatively high wages, which immediately suggests that earnings levels alone are not sufficient for explaining migrant outflows. Interestingly, these states also have relatively high housing prices. On the other hand, the set of states

[13] https://www.nber.org/research/data/county-distance-database

with large net-positive migration rates are often "Sunbelt" states, like Texas, Georgia, North Carolina, and Tennessee, that tend to have lower wages and housing prices.

Figure 2



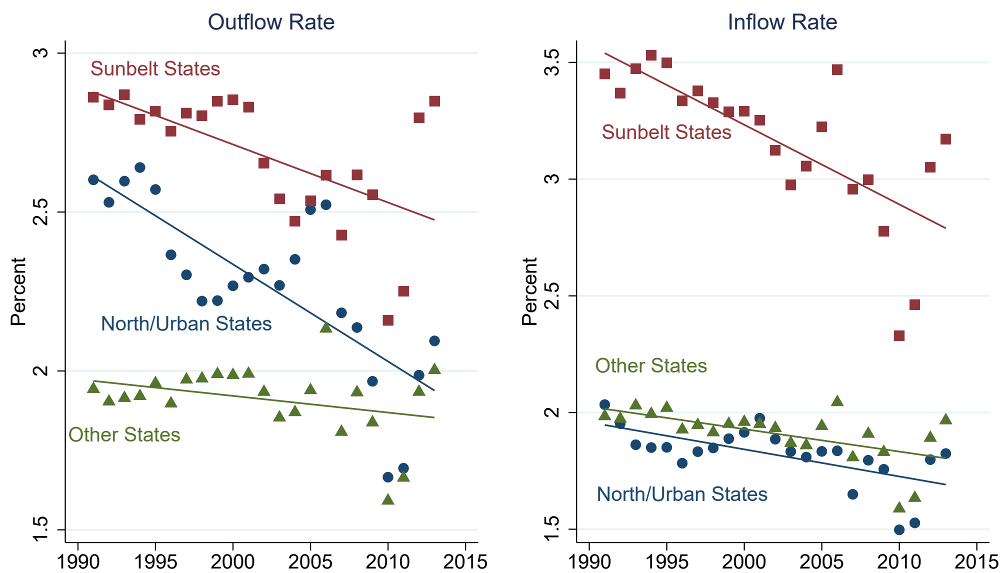
Notes: Net Inflow Rate is defined as inflows minus outflows divided by the population over the 1991-2013 period (in panel A). Average wage and salary income (thousands of $) and average median house price (thousands of $) over the 1991-2013 period is reported in panels B and C.

While the net flows in Figure 2a are informative, our primary interest is in understanding declines in gross migration flows. To characterize the geography of gross flows, Figures 3a and 3b respectively depict trends in inflow rates and outflow rates for the set of populous Northern and urban states that have relatively high wages and housing prices (NY, NJ, MA, IL, and CA), the set of Sunbelt states that tend to have lower wages and housing prices (TX, FL, NV, AZ, NM, GA, SC, NC, and TN), and all other states.[14] Figure 3a shows that while migration outflows have declined for all three regions, these declines were strongest in the set of northern and urban states. With respect to inflows, Figure 3b indicates that it was predominantly Sunbelt states that experienced declining migration inflows. Together these figures suggest that the key geographic feature of the overall declines in internal migration since the early 1990s

[14]We calculate outflow and inflow rates as the share of the population in each set of states that out-migrated/in-migrated in each year. These regional groupings are examined in more detail in Section 6.4.

has been a slow down in the pace of moves out of north/urban states like California, New York, and Illinois and into Sunbelt states like Texas, Nevada, and Florida.[15]

Figure 3
Regional Outflow and Inflow Rates



Notes: Outflow Rate is the regional outflows to all CZs divided by the regional population. The Inflow Rate is the regional inflows from all CZs divid by the regional population. 'Northern/Urban States' include NY, NJ, MA, IL, and CA; 'Sunbelt States' include TX, FL, NV, AZ, NM, GA, SC, NC, and TN; and 'Other' includes the remaining states.

4 Empirical Strategy

Motivated by the Rosen-Roback theoretical framework described above, we pursue two empirical specifications. We begin with an intuitively appealing "gap" specification, where migration flows depend on the *difference* between wages and home prices in the destination and origin CZ, and then we turn to a more flexible "origin-destination" specification that allows for asymmetric push and pull factors.

4.1 Gap Specification

In the gap specification we regress the migration rate between origin Commuting Zone o and destination Commuting Zone d in calendar year t onto the gap in mean wage earnings and the gap in mean home values between the destination and origin CZ. Specifically, we estimate the following equation:

[15]Of course, our subsequent empirical analysis will focus on migrant flows at the CZ-level, the largest of which are listed in Figure A1. For instance San Diego to LA is the second most common move in our data, while LA to San Diego is the third most common move. If we focused on net migration, these two flows would effectively cancel out, and thus Figure A1 highlights the importance of measuring changes in *gross* migration, which is our primary interest.

$$MigRate_{odt} = \alpha_1 * (lnWage_{dt-1} - lnWage_{ot-1}) + \alpha_2 * (lnHouse_{dt-1} - lnHouse_{ot-1}) + \gamma_{od} + \gamma_t + \epsilon_{odt} \quad (7)$$

The dependent variable, migration rate, is the ratio of migrating tax exemptions to the total number of tax exemptions in the origin CZ. The wage and housing price gaps are lagged one year to account for the time it takes for migration decisions to respond to changing economic conditions, and are expressed in logs. We cluster our standard errors at the bilateral CZ-pair level. Based on the theoretical insights from Equation 6, we anticipate that a larger wage gap between the destination and origin location will increase migration ($\alpha_1 > 0$) and a larger housing price gap between the destination and origin location will reduce migration ($\alpha_2 < 0$).

Equation 7 also includes CZ-pair fixed effects (γ_{od}) and year fixed effects (γ_t). The CZ-pair fixed effects will account for unobserved time-invariant conditions in the two commuting zones that could affect migration decisions. For example, they will control for the distance between the two CZs, natural amenities such as weather, mountains and oceans, as well as the historical industrial or ethnic composition of each CZ. Likewise, the year fixed effects are included to account for unobserved time-varying factors that affect overall migration or economic conditions in all CZs. For instance, these control for aspects such as skill-biased technical change, increasing trade exposure, the prevalence of remote work, and changes in family structure.

One feature of our data structure that has implications for estimating Equation 7 is that, since we include all possible migration flows between every commuting zone in the country, the migration rate between two particular CZs in a given year is frequently zero.[16] Expressing the dependent variable in logs and estimating our main specification via OLS would therefore exclude many observations. To avoid this we follow the now standard approach from the gravity equation literature and estimate Equation 7 using Poisson Pseudo-Maximum Likelihood (PPML) [Santos Silva and Tenreyro, 2006]. However, we demonstrate in Table A2 that similar results are obtained if we estimate Equation 7 using Ordinary Least Squares (OLS) and define the dependent variable as log of migrants plus one.

4.2 Origin and Destination Specification

The "gap" specification outlined above is intuitive, since bilateral migration decisions often entail comparing wages and housing prices in the origin and destination locations, and entering these variables as gaps allows each of these comparisons to be captured by a single parameter. However, a less appealing feature of this empirical specification is that it implicitly assumes that changing conditions in the destination and origin locations have a symmetric effect on migration decisions. For instance, the α_1 estimate assumes that the migration "pull" factors associated with rising wages in the destination CZ

[16] Or more precisely, given the suppression rules in the IRS migration data, it is possible that none of the counties within a CZ in a given year has 10 or more migrating tax units to other counties in a different CZ.

are equivalent in magnitude to the "push" factors associated with declining wages in the origin CZ. This need not be the case [Monras, 2020], and so our preferred specification allows for asymmetric origin and destination effects by estimating the following equation:

$$MigRate_{odt} = \beta_1 lnWage_{dt-1} + \beta_2 lnWage_{ot-1} + \beta_3 lnHouse_{dt-1} + \beta_4 lnHouse_{ot-1} + \gamma_{od} + \gamma_t + \epsilon_{odt}. \quad (8)$$

This model is similar in most respects to Equation 7, but now we identify separate β estimates for wages and housing in the destination and origin locations. Based on our theoretical predictions, we anticipate that migration flows are increasing with the destination wage ($\beta_1 > 0$), declining with the origin wage ($\beta_2 < 0$), decreasing with the destination home prices ($\beta_3 < 0$), and increasing with the origin home prices ($\beta_4 > 0$).

Equation 8 is a two-way fixed effects specification with a continuous "treatment" and thus causal inference requires a common trends assumption, specifically that the counterfactual trends in migration flows between CZ-pairs where wages and housing prices changed more are well characterized by the migration flows that occurred between CZ pairs where there were smaller or no change in wages and housing prices. While the included fixed effects absorb many potential confounders, the identifying assumption could still be violated if there are unobserved time-varying CZ-level factors that are correlated with both changes in wages or housing prices and with changes in migration flows. For instance, an improvement in school funding or quality that both increased local earnings and independently attracted migrants, would be a potential issue for our empirical approach.

While it is hard to rule out these types of time-varying CZ level confounders entirely, we do estimate and report three extensions of Equation 8 that will help to account for potential time-varying CZ-specific omitted factors.

First and perhaps more importantly, we show that our main findings are similar if we replace the year fixed effects in Equation 8 with origin-state-by-year fixed effects and destination-state-by-year fixed effects. These state-by-year interactions will account for factors that vary over time at the state level, including state level policies like minimum wage laws, education spending/quality, and changes in the generosity of social safety net programs like Medicaid or SNAP, among many other characteristics that differ over time across states.

Second, we control for the *total* number of migrants entering and leaving the destination and origin CZ in each year. This is intended to address potential reverse causality; specifically the possibility that migration flows into a particular CZ could themselves affect the wages or home values of that CZ.[17] While reverse causality is a possibility, note that our dependent variable is the *bilateral* migration rate between two CZs, not total migration, which likely mitigates this concern. For example, while the *total* number of migrants entering San Francisco likely affects wages and home values in San Francisco, the number of

[17]For instance, Moretti [2012] has emphasized how influxes of more educated migrants can positively influence the economic trajectories of cities.

migrants entering San Francisco *from Houston* (or any other CZ) is much less likely to materially impact local wage levels and home values. Still, estimating models that control for the *total* number of migrants entering and leaving the destination and origin CZ in each year will help account for any remaining reverse causality concerns. Extending the previous example, estimating the determinants of migration from Houston to San Francisco conditional on the total number of migrants to San Francisco (from all origins) and the total number of migrants from Houston (to all destinations) in each year would directly account for any effect that overall migration may have on local wage and home values.

Third, we estimate models that directly control for the demographic characteristics of each CZ in each year. To the extent that the age, race, ethnicity and gender composition of a CZ (or factors well proxied by these demographics) influence the conditional relationship between wages, housing prices, and migration, the inclusion of these controls will mitigate these concerns.

Finally, we note that our primary interest is in understanding how wages and home prices have contributed to *changes* in migration rates over time. As such, bias in the estimates of our baseline specification are primarily a concern only to the extent that they differ by time period.[18] Variation over time in the level of bias is of course plausible, but would also require more narrow and specific processes that we believe are less likely.

5 Results on Migration Determinants

We present three sets of results in this section, including findings from our "gap" specification, findings from our "origin and destination" specification, and results addressing threats to idenfitication.

5.1 Gap Results

The estimated impact of wage and housing price gaps on bilateral migration flows are reported in Table 1. Column 1 starts by regressing the migration rate on the gap in wages between the destination and origin location, after accounting for CZ-pair fixed effects and year fixed effects.[19] The point estimate of 0.356 is highly statistically significant and indicates that a 10% increase in the gap between the wages of a potential destination CZ relative to the current CZ of residence increases the migration rate by 3.56%. Column 2 shows that bilateral migration rates are significantly negatively related to housing price gaps, with a 10% increase in the gap in home prices between a destination CZ and an origin CZ reducing the migration rate by an estimated 0.85%.

Column 3 then incorporates both wage and housing price gaps in the same analysis, which generates estimates that are larger in magnitude and statistically significant at conventional levels. This highlights the importance of accounting for both labor market and housing market determinants when considering

[18] This is in contrast to studies of migratory responses to local economic shocks, where a strongly causal interpretation of migration elasticity estimates is of first order importance and shift-share instruments are typically employed.

[19] Note that over 12 million observations are included in all our PPML estimations, but by default only the number of observations with a positive value for the dependent variable are reported in the sample size.

Table 1: Migration Determinants

	Migration Rate		
	(1)	(2)	(3)
Wage_Gap	0.356***		0.524***
	(0.043)		(0.049)
Home_Price_Gap		-0.085***	-0.193***
		(0.017)	(0.018)
CZ-Pair FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	856,543	856,543	856,543

Notes: PPML estimation. The dependent variable is the migration rate between the origin and destination CZ. All independent variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level in parentheses. Data source: IRS. * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

migration decisions, which is not always done in the existing literature. Furthermore, the estimates are consistent with theory and intuition: larger wages gaps between the destination and origin CZ are associated with more migration, while larger housing prices gaps are associated with less migration.

5.2 Origin and Destination Results

Results from our origin and destination specification (Equation 8), which allows for asymmetric "push" and "pull" factors, are reported in Table 2. Overall, we find that our fully specified model generates results that strongly conform to the predictions from our theoretical framework, and these findings also show that less comprehensive specifications can be misleading.

Column 1 of Table 2 reports a naive specification that examines how migration rates respond to wages in the origin CZ, ignoring housing prices in either location as well as wages in the destination CZ. Variations of this specification are often used in existing studies and they typically find that labor market conditions in the origin location do not lead to out-migration. Consistent with this existing evidence, we find in column 1 that the estimated coefficient on origin wage levels is statistically insignificant.

We examine two potential explanations for this surprising finding. The first is a standard omitted variable bias story, where wages may be correlated with other local conditions that could counteract the wage effect on migration. For instance, when wages fall perhaps housing prices also decline causing people not to leave. Second, this naive specification may be misleading because it does not take into account the bilateral nature of the migration decision [Borusyak et al., 2022]. In particular migrants likely consider not only origin CZ conditions (i.e. wages and housing prices) but also economic conditions

Table 2: Migration Determinants

	Migration Rate		
	(1)	(2)	(3)
Orig_Wage	-0.051	-0.067	-0.353***
	(0.047)	(0.054)	(0.057)
Origin_Home_Price		0.016	0.141***
		(0.018)	(0.021)
Dest_Wage			0.777***
			(0.058)
Dest_Home_Price			-0.262***
			(0.021)
CZ-Pair FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	856,543	856,543	856,543

Notes: PPML estimation. The dependent variable is the migration rate between the origin and destination CZ. All independent variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level in parentheses. Data source: IRS. * p< 0.1, ** p< 0.05, *** p< 0.01.

in potential destination CZs when deciding whether to move. We investigate these potential explanations in the remainder of Table 2.

In Column 2 we add origin housing prices to the empirical specification. Consistent with theoretical predictions, migration is decreasing with origin wages and increasing with origin housing prices, but neither coefficient is statistically significant or of an economically meaningful magnitude. Furthermore, adding housing prices to the specification does not alter the wage coefficient substantially, which falls from -0.051 to -0.067. We conclude that accounting for local housing conditions alone does not change the puzzling finding that migration appears to be unresponsive to local origin wages.

Next we exploit the bilateral nature of our data by also incorporating labor market and housing conditions in potential *destination* CZs. In Column 3, which is our preferred specification, we add wage levels and home values in the destination CZ. Accounting for these characteristics substantially alters the estimates. In particular, all four coefficients of interest in this specification have the expected signs and are now statistically significant at the 1% level. The magnitudes of the estimated coefficients also grow larger, and are economically meaningful. For instance, a 10% increase in the origin CZ wage is associated with a 3.5% decrease in the migration rate and a 10% increase in the destination CZ wage is associated with a 7.8% increase in the migration rate. With respect to home values, the estimates in Column 3 indicate that a 10% increase in the origin home price is associated with a 1.4% increase in the migration rate, while a 10% increase in the destination home price is associated with a 2.6% reduction in migration. Comparing Column 3 to the first two columns indicates that it is essential to account for the bilateral nature of the migration decisions. Considering only labor and housing market conditions in the origin location can generate highly misleading estimates.

In addition to producing large, statistically significant coefficients with the expected signs, an important finding in this preferred specification in column 3 is that origin and destination effects are *asymmetric*, and in particular that pull factors are more important than push factors. Indeed the magnitude of the coefficient on destination wages is almost 70% larger than the coefficient on origin wages, while the magnitude of the coefficient on destination home values is 35% larger than origin home values.[20]

In Appendix Table A1 we report a heterogeneity analysis that shows these baseline estimates are qualitatively similar for CZ pairs where migrant's Adjusted Gross Incomes (AGIs) are above versus below the median (AGI is one of the few characteristics the IRS data provides); for CZ pairs that are above versus below median geographic distance; and for CZ pairs with more versus fewer families with children (measured using the difference between total exemptions and total returns). Appendix Table A2 demonstrates the robustness of the baseline findings to a variety of alternative modeling and variable definition choices. These include transforming the dependent variable as the natural log of (migrants + 1) and using OLS rather than PPML; extending the sample through 2019 while imposing a consistent minimum migrant flow of 20 rather than 10; applying sampling weights equal to total annual migrants or to total migrants in 1991; defining migration rates using the number of tax returns filed rather than the number of tax exemptions; excluding CZ-pairs that are within 100 miles of each other, which drops observations where commuting across CZs may be viable; and excluding Alaska and Hawaii, which are outliers in terms of the distance and difficulty of migration. In all cases the results do not change substantively, with economically large and statistically significant estimates of the key parameters that are of the expected sign.

We draw three main conclusions from the results in Table 2 and the corresponding heterogeneity and robustness analyses in the appendix. First, it is important to account for both labor market and housing market conditions when considering internal mobility. Second, it is essential to simultaneously account for both origin and destination characteristics when studying migration decisions. And third, migration push and pull factors are not, generally speaking, symmetric. These findings constitute three key contributions of the current study, and it is notable that when fully specified, the results of a transparent migration model like Equation 8 strongly conform to intuition and theoretical predictions.

5.3 Threats to Identification

Given the features of our baseline specification, and in particular the inclusion of CZ-pair fixed effects, we are cautiously optimistic that our estimates primarily reflect a causal relationship. This section reports

[20] Chi-Square tests (not shown) indicate that both of these differences in the magnitudes of the coefficients are statistically significant at beyond the 1% level.

results that address some of the most obvious threats to identification, which were briefly discussed above.

We begin by evaluating the extent to which omitted time-varying characteristics of origin and destination locations may be affecting our results. Specifically, we are interested in whether policy changes (like minimum wage laws, educational investments, or social safety net programs), location specific productivity shocks (like trade penetration for a locally important industry), or changing amenities (like warming temperatures or increasing numbers of college graduates) may affect both migration flows and labor and housing markets. We attempt to account for these potentially confounding factors by including, not only the usual CZ-pair fixed effects, but also origin-state*year fixed effects and destination-state*year fixed effect, will account for all unobserved factors that vary over time within states. This empirical approach draws on insights from the gravity trade literature, where the inclusion of both bilateral-pair fixed effects and origin*year and destination*year fixed effects is considered the optimal empirical specification [Baldwin and Taglioni, 2006].[21]

The estimates with origin-state*year and destination-state*year fixed effects are reported in Column 2 of Table 3, while for comparison purposes Column 1 reproduces our baseline results with only CZ-pair and year fixed effects. The coefficients on our four key independent variables of interest are all of the same sign and remain statistically significant at the one percent level. The magnitude of these point estimates change slightly, with two coefficients becoming a bit smaller in magnitude and two becoming a bit larger, but overall there is little meaningful change in the relationship between bilateral migration rates and origin and destination conditions. We conclude that unobserved time-varying factors within states are not driving our key results.

A related concern is that perhaps wages and housing prices are changing in response to migration flows themselves, a form of reverse causality. As discussed above, we feel the bilateral analysis is unlikely to suffer from this type of reverse causality, since the migration flows to or from any particular CZ are unlikely to be large enough to drive overall CZ-level wages and housing prices. Nonetheless, to help address this potential concern we also estimate models that control for *total* migrant inflows into each CZ and *total* migrant outflows from each CZ, which will account for the potential impacts of aggregate migration in each CZ.

The results are reported in Column 3 of Table 3. Not surprisingly, we see that bilateral migration is increasing with the total migrant inflows into the destination CZ and with total migrant outflows from the origin CZ. Returning to the San Francisco-Houston example, these coefficients mean that if more people overall are moving to San Francisco and more people overall are moving away from Houston, then all else equal more people are moving from Houston to San Francisco. What is noteworthy is that while controlling for these overall CZ level migration flows does change the magnitudes for some of the coefficients of interest, the key estimates are qualitatively similar to baseline in that they retain the same sign and are still statistically significant. We conclude that there is little evidence that our findings are

[21] In this context it is not possible to include origin-CZ*year FE and destination-CZ*year FE because these fixed effects would subsume the independent variables of interest, but most of the factors we are interested in accounting for vary at the state level.

Table 3: Addressing Threats to Identification

	Migration Rate			
	(1)	(2)	(3)	(4)
	Baseline	StatexYear_FE	Total_Migrants	Demographics
Orig_Wage	-0.353***	-0.265***	-0.597***	-0.506***
	(0.057)	(0.056)	(0.049)	(0.056)
Origin_Home_Price	0.141***	0.220***	0.043**	0.189***
	(0.021)	(0.030)	(0.020)	(0.020)
Dest_Wage	0.777***	0.893***	0.301***	0.758***
	(0.058)	(0.058)	(0.046)	(0.056)
Dest_Home_Price	-0.262***	-0.179***	-0.179***	-0.217***
	(0.021)	(0.032)	(0.019)	(0.019)
Orig_Total_Out_Migration			0.770***	
			(0.023)	
Dest_Total_In_Migration			0.598***	
			(0.015)	
CZ-Pair FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Orig_State*Year FE	No	Yes	No	No
Dest_State*Year FE	No	Yes	No	No
Orig Demographics	No	No	No	Yes
Dest Demographics	No	No	No	Yes
Observations	856,543	856,470	855,766	849,270

Notes: PPML estimation. The dependent variable is the migration rate between the origin and destination CZ. Column 1 reports the baseline results, Column 2 includes state*year fixed effects, Column 3 controls for total out-migration from the origin CZ and total in-migration into the destination CZ, and Column 4 controls for demographic characteristics in the origin and destination CZs. The wage and housing variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level in parentheses. Data source: IRS. * p< 0.1, ** p< 0.05, *** p< 0.01.

driven by reverse causality.

In the final column of Table 3 we include year-specific demographic characteristics in the origin and destination CZs. In particular we control for the share of the population that is in eight age categories (19 or younger, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80 or older), three race categories (white, Black, and other race), two ethnicity categories (Hispanic and non-Hispanic) and two gender categories (male and female). As discussed above, if the coefficients on wages and housing partially reflect demographic changes in particular CZs, or any factors that are well proxied by demographic changes, the inclusion of these controls will reduce any such omitted variable bias. Reassuringly, Column 4 shows that the estimates are very similar with these demographic controls included.

6 Explaining the Decline in Migration

Our results so far show that migration rates strongly and consistently respond to origin and destination wages and home prices in the anticipated directions. We now turn to using our estimates to better understand the declines in internal migration that were documented in Figure 1 and in several influential previous studies.

Our overall approach is to use the coefficients from our regression models and the mean levels of the key independent variables observed in our data to calculate various *predicted* migration rates. By plugging in alternative values for the relevant elasticities and means, we can construct transparent counterfactual estimates of how much migration would have changed over time depending on the evolution of the key housing and wage related parameters.

As a starting point, for each year t in our sample period we construct a predicted migration rate by multiplying *year-specific* elasticities with the annual mean values of wages and home prices, then summing over these four products:

$$PredM\hat{i}gRate_t = \hat{\beta_{1t}}\overline{lnWage_{dt-1}} + \hat{\beta_{2t}}\overline{lnWage_{ot-1}} + \hat{\beta_{3t}}\overline{lnHousing_{dt-1}} + \hat{\beta_{4t}}\overline{lnHousing_{ot-1}} \quad (9)$$

While generally straightforward, there are two important aspects of this approach that are worth highlighting.

First, note that the β terms contain t subscripts, indicating that the *responsiveness* of migration to wages and home prices is time-varying. We estimate these time varying β's by interacting our wage and housing variables with year fixed effects and then summing the coefficients on the main effects and these interaction terms.

Second, if Equation 9 were to use the simple means of the independent variables calculated in the full sample, the mean origin wage would be equal to the mean destination wage by construction (and the same applies for home prices), simply because the bilateral structure of our data means that each CZ appears an equal number of times as an origin and as a destination location. Since the mean values of wages and home prices in origin and destination CZs would be mechanically equal, variation in the predicted migration rate would come exclusively from asymmetries in the origin and destination coefficients, and the calculation performed in Equation 9 would be less informative.[22]

As discussed in more detail in Section 6.3, we believe that the symmetry of mean wages and home values in origins and destinations is actually an insightful feature of migration determination. However, this only occurs when all CZ-pairs receive equal weight, but in practice some CZ-pairs and some migration directions are surely much more important than others for determining aggregate migration. For example,

[22]This point is especially easy to see when considering the specification where wages and home values are expressed as *gaps* between origin and destination and therefore not allowed to be asymmetric. In this case the *means* of these gaps would be equal to zero in each year in a full bilateral data set, which mechanically causes overall predicted migration to be zero.

consider Bakersfield, CA and Dallas, TX, which both appear in our data as origins and destinations. An increase in wages in Dallas would attract migrants *from* Bakersfield (since Dallas is their destination), but would also suppress migrants *to* Bakersfield (since Dallas is their origin). However, given that the dominant flow of migrants is from Bakersfield to Dallas (this annual flow is more than twice as large on average in our data), and not the reverse, we would not generally expect these two effects to be equal and opposite. Furthermore, the CZ-pair of San Francisco and Los Angeles appears in our data the same number of times as Bakersfield and Dallas, but surely wage and home price changes in the San Francisco and Los Angeles pair have a larger impact on gross migration, given that Los Angeles and San Francisco are highly populous and geographically proximate.

Given these considerations and our goal of explaining declines in aggregate migration (Figure 1), in Equation 9 we calculate and use annual mean values of wages and housing prices that are weighted by the *total* number of migrants in each bilateral CZ-pair (over the full sample period).[23] This accounts for the fact that some CZ-pairs are simply much more important than others for aggregate migrant flows, and accordingly gives more weight to the CZ-pairs where actual migration occurs.

Bearing these points in mind, we calculate and plot in Figure 4 the predicted migration rates for each year. For reference we also plot the actual observed migration rate by year (the same series that was shown in Figure 1). Figure 4 illustrates two important points.

First, in general, wages and housing prices can account for a substantial share of overall migration. For instance, the mean of the actual migration rate across all years is 2.17, while the mean of predicted migration as defined in Equation 9 is 0.70. That is, while people migrate for myriad and idiosyncratic reasons, simply observing the time-varying components of home prices and wage, along with our regression coefficients, can explain about one third of observed migration flows in a given year.[24]

Second, and more directly relevant for current purposes, we observe a negative *trend* in predicted migration that tracks the negative trend in observed migration. In fact, it substantially over-predicts the fall in migration rates. More specifically, the fit line for total migration in Figure 4 falls by 0.31 percentage points over the study period, while the fit line for predicted migration falls by approximately 1 percentage point. Our framework may miss other possible determinants of internal migration, but we see in Figure 4 that these omitted factors, in anything, seem to have increased migration rates over time.

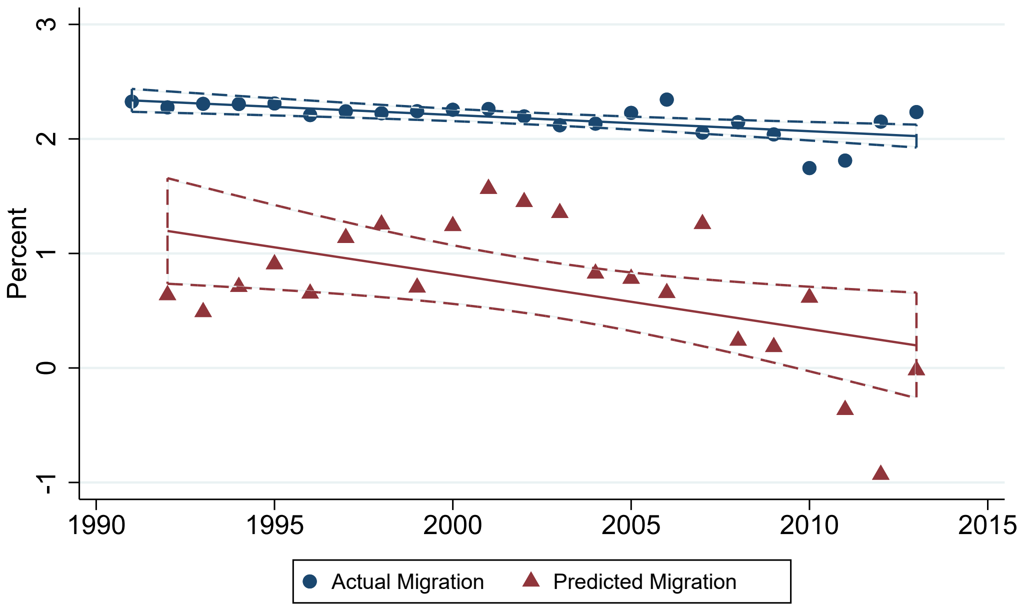
On balance, we conclude from Figure 4 that wages and housing are major contributors to aggregate domestic migration in general, and have also played a key role in the decline in internal migration over the last three decades. We next extend our approach in three directions, all of which turn out to be important in practice and in our view contribute significantly to understanding why internal migration has fallen in recent decades. First, we differentiate the relative importance of wages versus home prices in predicting

[23] Weighting by the total number of migrants in the first year of the study period is arguably more exogenous, but we prefer weighting by total migrants because in any given year many CZ-pairs have no reported migrants, so that averaging over the full sample period preserves a large number of CZ-pairs in the weighting. However, in practice the two approaches yield virtually identical results.

[24] The time-invariant components of wages and home prices, as well as all other time invariant origin and destination CZ factors, are captured by the bilateral pair-fixed effects, which are not included in the predicted migration calculation.

migration trends. Second, we examine the importance of origin versus destination CZ factors in driving migration declines. Finally, we consider the importance of changes in the *levels* of wages and home values relative to the importance of changes in how *responsive* migration is to wages and home values, similar to the "levels versus prices" distinction in a traditional Kitigawa-Blinder-Oaxaca decomposition.

Figure 4
Actual and Predicted Migration

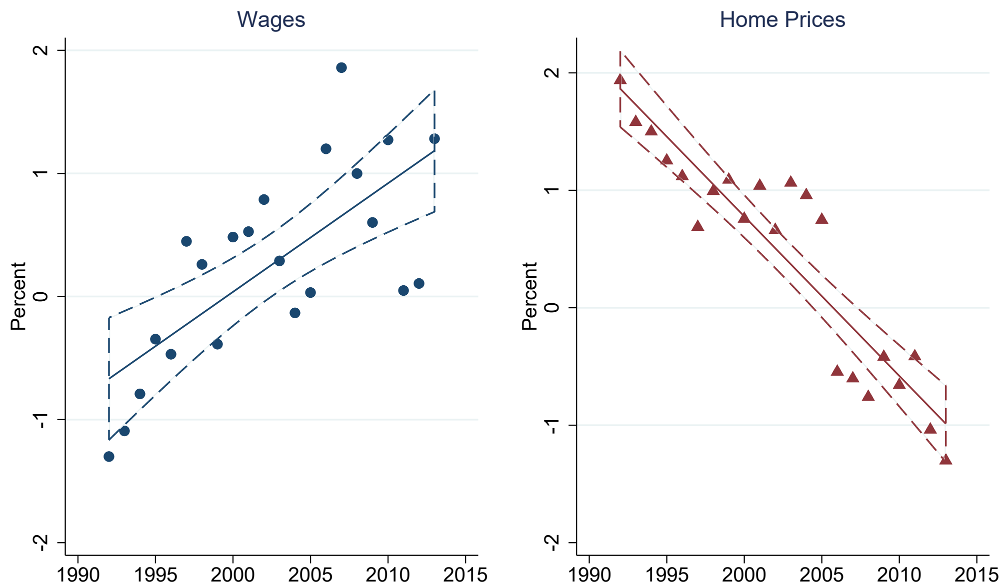


Notes: Predicted migration rate is calculated using time-varying elasticities and the mean values of the independent variables.
Data source: IRS.

6.1 Wages versus Housing

The additive nature of our predicted migration measure allows us to relatively easily decompose predicted internal migration rates into individual components. We start by examining whether the decline in predicted migration is predominantly driven by wages or housing prices. Specifically, we construct two different predicted migration rates, one due to wages using the first two terms on the right side of Equation 9 and the other due to housing prices using the the second two terms. We plot these two predicted migration rates and the lines of best fit in Figure 5.

Figure 5
Predicted Migration: Wages vs Housing



Notes: Predicted migration rates are calculated using time-varying elasticities and the mean values of the independent variables.
Data source: IRS.

Wages alone predict that internal migration would have risen over our sample period, as we see on the left side of Figure 5. Specifically, wages are associated with an *increase* in the migration rate of 2.6 percentage points. While some previous studies have argued that labor market changes are an explanation for declines in internal migration [Kaplan and Schulhofer, 2017], Figure 5 indicates that changing features of the labor market, on their own, would have actually led to an *increase* in migration.

Figure 5 also clearly demonstrates that predicted migration rates due to housing have steadily declined over the sample period. Specifically, housing prices are associated with a 3.3 percentage point *decrease* in the internal migration rate. Observed changes in housing prices and the evolving responsiveness of migrants to housing price changes has led to a significant decline in internal migration. We conclude that housing is key to understanding why migration rates have declined.

6.2 Origin versus Destination

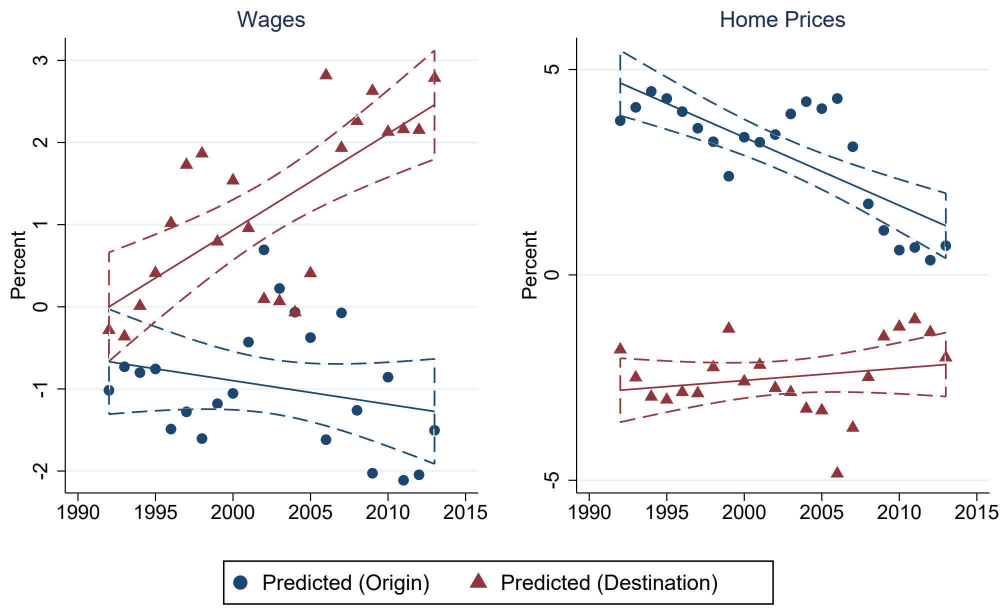
We next evaluate the relative importance of push and pull factors in driving migration trends. Specifically, we construct four different predicted migration rates driven by origin wages, destination wages, origin housing prices, and destination housing prices respectively (i.e. a separate predicted migration rate using each of the four terms on the right side of Equation 9). These predicted migration rates and their lines of best fit are reported in Figure 6.

In the left panel of Figure 6 we see that wages in the destination commuting zone leads to an increase in internal migration rates, while there is little significant impact of origin wages on migration rates. This is an important distinction, because as discussed above much of the existing literature focuses on shocks in the origin location [Autor et al., 2013] and attempts to understand why deteriorating conditions do not cause more people to leave. Figure 6 suggests that migrants are much more responsive to wages in the destination location, and it is these pull factors that have changed the most over time. This is consistent with papers studying how wages have an asymmetric effect on *in-migration* relative to *out-migration* [Bartik and Rinz, 2018, Yagan, 2019]), and with papers finding significant migration responses to positive shocks like the fracking boom [Wilson, 2022]

The results for housing are quite different. On the right panel of Figure 6, we see that it is origin housing prices that are associated with a decline in internal migration, while destination housing prices are essentially unrelated to trends in migration rates. This allows us to further refine the finding from the previous sub-section that housing is responsible for declines in internal mobility, by specifically noting that it is changes in *origin* housing prices that are most germane, rather than destination home prices.

On balance, the patterns in Figure 6 indicate that over the past 20 to 25 years, Americans have become increasingly influenced by wage levels in prospective destinations, which on its own would have led to an increase in the domestic migration rate of more than two percentage points. However, over the same period Americans have become less influenced by home prices in their current locations, to the point where in recent years this elasticity is close to zero and changes in origin home prices have virtually no effect on migration choice. This latter effect has reduced migration by over three percentage points, and led to the observed decrease in domestic migration.

Figure 6
Predicted Migration: Origin vs Destination

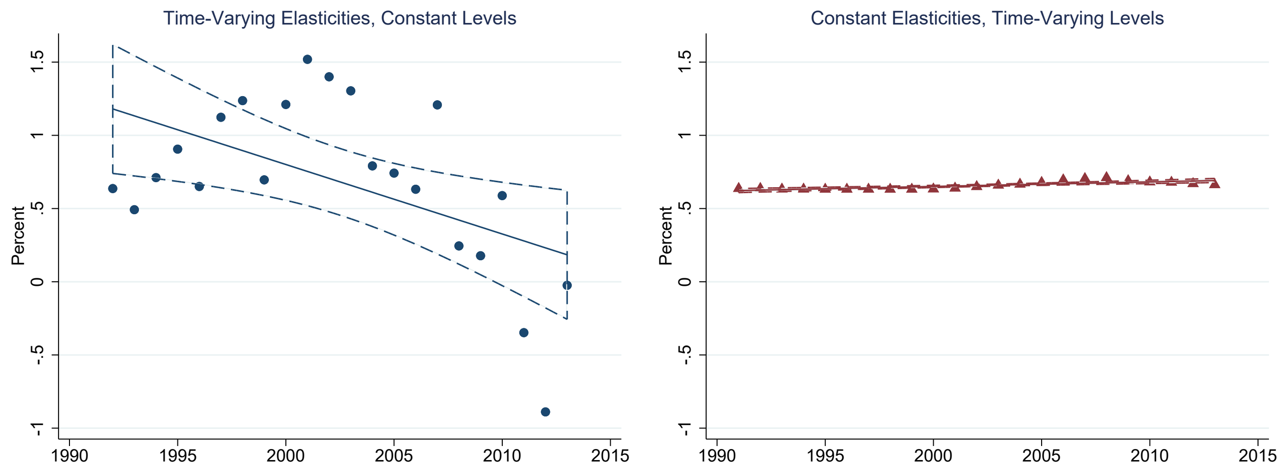


Notes: Predicted migration rates are calculated using time-varying elasticities and the mean values of the independent variables.
Data source: IRS.

6.3 Levels versus Elasticities

Finally, in the spirit of a Kitigawa-Blinder-Oaxaca decomposition, we disaggregate the overall trend in predicted migration from Figure 4 into a portion that is attributable to changes in the *levels* of the independent variables and a portion that is attributable to changes in the *elasticities*. Our predicted migration approach can be easily adapted to investigate this issue. In particular, we first calculate predicted migration using the observed time-varying levels of the independent variables, but with the elasticity estimates from the first year of our sample. This identifies what predicted migration rates would have been if the level of these variables changed but migrant's responsiveness to these changes remained constant. We then calculate predicted migration using time-varying elasticities, but fix our wage and housing variables at the levels observed in the first year of our sample. This calculates what annual predicted migration rates would have been if the responsiveness of migration to wages and home values changed, but the levels of these variables did not.

Figure 7
Predicted Migration: Elasticities vs Levels



Notes: The left panel calculates predicted migration using the time-varying elasticities and the mean values of the independent variables from the pre-sample year. The right panel constructes predicted migration using the elasticities from the presample year and the time-varying mean values of the independent variables. Data source: IRS.

Figure 7 plots the annual predicted migration rates that are specifically attributable to changing elasticities (left side) and to changing levels of wage and home values (right side). The results suggest that trends in predicted migration rates are almost wholly attributable to changing elasticities: had wages and housing prices remained at their presample values and the estimated β's evolved as they actually did, then there would be approximately a 1 percentage point decline in the migration rate. However, if the β's stayed fixed at their initial values while wage and home values evolved as they actually did, then there would have been essentially no change in migration.

This result may seem surprising, but it largely follows from the simple fact that wages and home values have evolved in broadly similar ways in the places that people tend to move from and the places that people tend to move to. Given this, the predictive power of wages and home values must be due to changes in the responsiveness of migration to these factors, not changes in their levels.

More specifically, it is important to remember that the means in Equation 9 are weighted by aggregate migration flows. In the absence of these weights, origin means would exactly equal destination means. Even with weights applied, the differences in the mean values of origins and destinations are similar in part because CZs frequently appear as both origins and destinations. For instance suppose that in aggregate over the full sample period 100 people moved from Boston to Philadelphia and 200 people moved from Philadelphia to Boston. In this case Boston would get twice as much weight as Philadelphia in the destination mean calculation, and vice versa for the origin mean calculation. But both locations would still get significant weight as both an origin and as a destination, which would partially offset each other and drive mean differences between origins and destinations towards zero.

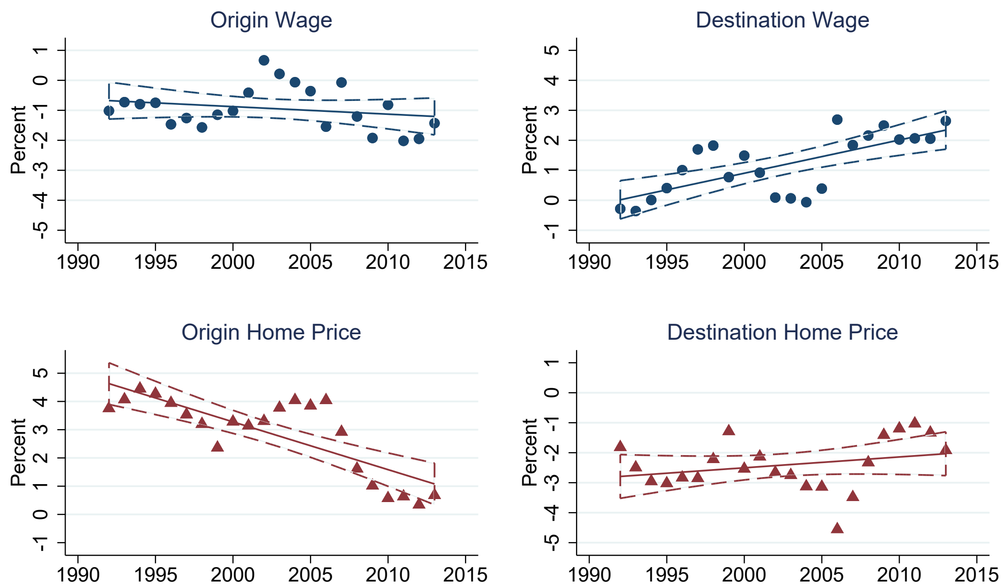
While in some sense this reflects how the structure of bilateral data mechanically suppresses mean differences, we believe that the modest contributions of the levels shown in Figure 7 actually reflect a substantive feature of migration choice. In particular, every location genuinely does simultaneously operate as both an origin *and* as a destination. This implies that if, for example, wages in Boston rise while wages in Philadelphia are unchanged, there will be an increase in the incentive to migrate from Philadelphia to Boston. However, there will *also* be a decrease in the incentive to migrate from Boston to Philadelphia, and the impact on total migration is ambiguous. The fact that migration incentives are fundamentally two-sided makes it less likely that the *levels* of wages and home prices have exerted large influences on total migration trends, simply because a change in levels in any given location generally exerts these two countervailing influences.[25]

We extend the analysis reported in Figure 7 by separately reporting the time-varying elasticities for our four key independent variables of interest. Specifically, Figure 8 reports predicted migration due to the time-varying elasticities, while holding the variable levels constant at their presample values. As we saw previously (Figure 6) origin housing prices are key to understanding the decline in internal migration, and furthermore the results in Figure 8 show that this is predominantly due to the declining responsiveness of migrants to origin housing prices.

On balance, Figures 5-8 suggest a rather specific explanation for declining internal migration over the last thirty years: Changes in migrants' *responsiveness* (elasticities) to *home prices* in their *origin* locations. In the next section we evaluate possible explanations for why migration decisions have become less responsive to origin housing prices.

[25]It is sometimes argued that more geographic dispersion in wage levels, measured for instance by the standard deviation of mean wages across states or CZs, will tend to increase total migration. Our argument is that this is not necessarily true, since larger wage differences across space may increase migration in one direction but reduce it in the other direction. Ultimately the effect of dispersion on total migration will depend on whether the places where wages are rising are more common as destinations or as origins. And even in cases where the net effect of wage dispersion on migration is positive, the magnitude will be attenuated by these countervailing impacts.

Figure 8
Predicted Migration - Time-Varying Elasticities and Constant Levels



Notes: Predicted migration calculated using the time-varying elasticities and the mean values of the independent variables from the pre-sample yea
Data source: IRS.

6.4 The Declining Migration Response to Origin Housing Prices

We have documented that declining domestic migration is primarily attributable to falling responsiveness to origin home prices, but this still leaves open the question of *why* origin home price elasticities have fallen. This is an inherently challenging question, since it is ultimately related to individual preferences and behaviors, but in this section we provide some suggestive evidence on the reasons for this decline. Specifically, we first evaluate whether there is regional variation in the extent to which origin home price elasticities have fallen, and then we examine the characteristics of the individuals who have disproportionately become less responsive to origin home prices.

With respect to geographic heterogeneity, in Table 5 we investigate *where* home price elasticities have fallen the most. In Column 1 we use the full sample and estimate a modified version of our baseline model that includes an interaction between origin home prices and a continuous year variable.[26] The coefficient on the interaction term is -0.008 and is statistically significant, indicating that for each 10 year period over our study period the origin home price elasticity fell by 0.08. These findings are similar to the trends observed in Figures 6 and 8, but this approach has the benefit of more succinctly depicting

[26]For consistency, the other three independent variables are also allowed to vary linearly with year in Table 4 (although due to space constraints these results are not reported). The results are similar if only origin home prices are interacted with year.

the time-varying effect of origin home prices.

We next estimate this same model for different groups of states. In Column 2 we restrict the sample to CZ-pairs whose origin is in large urban states that had relatively high baseline home prices and experienced declining out migration over our sample period, specifically CA, IL, NY, NJ, and MA (see Figure 3). Within these states, the trend in home price elasticities was -0.023, approximately three times the national trend. In Column 3 we restrict our analysis to observations whose origin CZ was in a "Sunbelt" states that had relatively low baseline home prices, specifically NV, AZ, NM, TX, FL, GA, SC, NC, and TN. In these states the trend in origin home price elasticities is just -0.005. For completeness Column 4 of Table 5 focuses on outflows from the remaining states, which also have an estimated origin home price elasticity trend of -0.005.[27]

Table 4: Regional Variation in Origin Home Price Elasticities

	Migration Rate			
	(1)	(2)	(3)	(4)
	All_States	North_Urban	Sunbelt	Other
Origin_Home_Price_x_Year	-0.008***	-0.023***	-0.005**	-0.005***
	(0.001)	(0.002)	(0.002)	(0.001)
CZ-Pair FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	856,543	120,911	289,823	445,809

Notes: PPML estimation. The dependent variable is the migration rate between the origin and destination CZ. Regressions also include Origin Wage, Dest Wage, and Dest Home Price, as well as the interaction of these variables with year. All independent variables are logged and lagged one year. Column 1 includes all states, column 2 restricts the sample to outflows from North/Urban states (CA, IL, NY, NJ, and MA), column 3 restricts the sample to outflows from Sunbelt states (NV, AZ, NM, TX, FL, GA, SC, NC, and TN), and column 4 includes outflows from all other states. Robust standard errors clustered at the CZ-pair level in parentheses. Data source: IRS. * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$.

The results in Table 4 indicate that while declines in origin home price elasticities were broad-based, they were much stronger in states like California and New York, where individuals have become less willing to migrate away from rising home prices over time. This matches the regional trends in gross mobility shown in Figure 3, and is consistent with declining sensitivity to origin home prices being a main driver of the fall in domestic migration.

We next estimate differences in home price elasticity trends across different demographic groups. To do so we draw samples from the CPS's Annual Social and Economic Supplement over our sample period.

[27]The groupings of states used in Table 4 are somewhat arbitrary, but were chosen to correspond to the broad spacial patterns of migration over this period, as shown in Figures 2 and 3. Similar results are obtained when using alternative groupings.

This data contains current state of residence, as well as information on whether each respondent has moved in the past year and, if so, what their state of residence was one year ago. We use these questions to construct an individual-level binary variable identifying moves from each origin state and to each destination state in the past year, and then we link this CPS data to state-level versions of the income and home value measures used in our main analysis.

Relative to the CZ-level IRS migration data constructed from the universe of tax filers, the CPS has the disadvantage of being available only at the state-level and is also a relatively small sample. For these reasons we strongly prefer the baseline estimates from the IRS data. However, the CPS data has the key advantage of containing migrant demographic and socioeconomic information, which allows us to observe the types of people whose responsiveness to origin house prices have changed the most. We specifically use the CPS data to estimate the following specification:

$$Migrated_{iodt} = \beta_1 lnWage_{dt-1} + \beta_2 lnWage_{ot-1} + \beta_3 lnHouse_{dt-1} + \beta_4 lnHouse_{ot-1} + \gamma_o + \gamma_d + \gamma_t + X_i\alpha + \epsilon_{ist} \tag{10}$$

where $Migrated_{iodt}$ is a binary variable indicating whether individual i moved in the past year t, and if so what their origin state o (state of prior residence) was and what their destination state d (state of current residence) is.[28] Our main independent variables are the origin and destination home price and wage levels, constructed as they were in the main analysis but at the state-level. The γ terms are fixed effects for origin state, destination state, and year, and X_i is a vector of individual-level controls including home ownership, age, education, sex, race, ethnicity, marital status, employment status, children, and immigrant status.[29]

In Table 5 we report time-invariant estimates from this specification. For reference, the first column of Table 5 reproduces the baseline estimates from our preferred CZ-level specification using the IRS data (Column 3 of Table 2). The second column of Table 5 reports the analogous estimates using the state-level CPS data. Despite differences in the structure of the estimating equations, the different units of analyses, and the use of wholly separate data sets, the results in Column 2 are highly consistent with the baseline IRS estimates from Column 1. For instance the coefficients on destination wages and home prices are 0.777 and -0.262 in Column 1, and are 1.103 and -0.337 in Column 2. The estimates for origin wages and home values are more different, but they are of the same sign, statistically significant, and are of generally consistent magnitudes.

The key benefit of the CPS data set is that we can include a large set of individual level covariates

[28] For non-movers, origin and destination state are the same.

[29] Note that while origin, destination, and year fixed effects are included in our specification, it is not possible to also include origin-destination *pair* fixed effects because there is no variation in the migration dummy within a state-pair (i.e. migrated=1 for every observation within the CA-OR pair and migrated=0 for every observation within the CA-CA pair). Given our interest in individual characteristics, as well as the smaller CPS sample size, we chose to estimate Equation 10 using individual-level microdata, and not to collapse the data and construct migration *rates* between states, which could then have facilitated state-pair fixed effects.

Table 5: Migration Determinants and Demographic Controls

	(1) Cross-CZ Migration (IRS)	(2) Cross-State Migration (CPS)	(3) Cross-State Migration (CPS)
Orig_Wage	-0.353***	-1.103***	-1.082***
	(0.057)	(0.201)	(0.198)
Origin_Home_Price	0.141***	0.337***	0.335***
	(0.021)	(0.078)	(0.075)
Dest_Wage	0.777***	1.095***	1.076***
	(0.058)	(0.206)	(0.203)
Dest_Home_Price	-0.262***	-0.337***	-0.334***
	(0.021)	(0.080)	(0.077)
Own_House			-0.038***
			(0.004)
Age_55_Plus			-0.015***
			(0.001)
BA_Degree			0.013***
			(0.001)
Male			0.001***
			(0.000)
White (Non-Hispanic)			0.007***
			(0.001)
Married			0.002***
			(0.001)
Employed			-0.006***
			(0.001)
Children			-0.007***
			(0.001)
Foreign_Born			-0.001**
			(0.001)
Observations	856,543	2,918,163	2,918,163

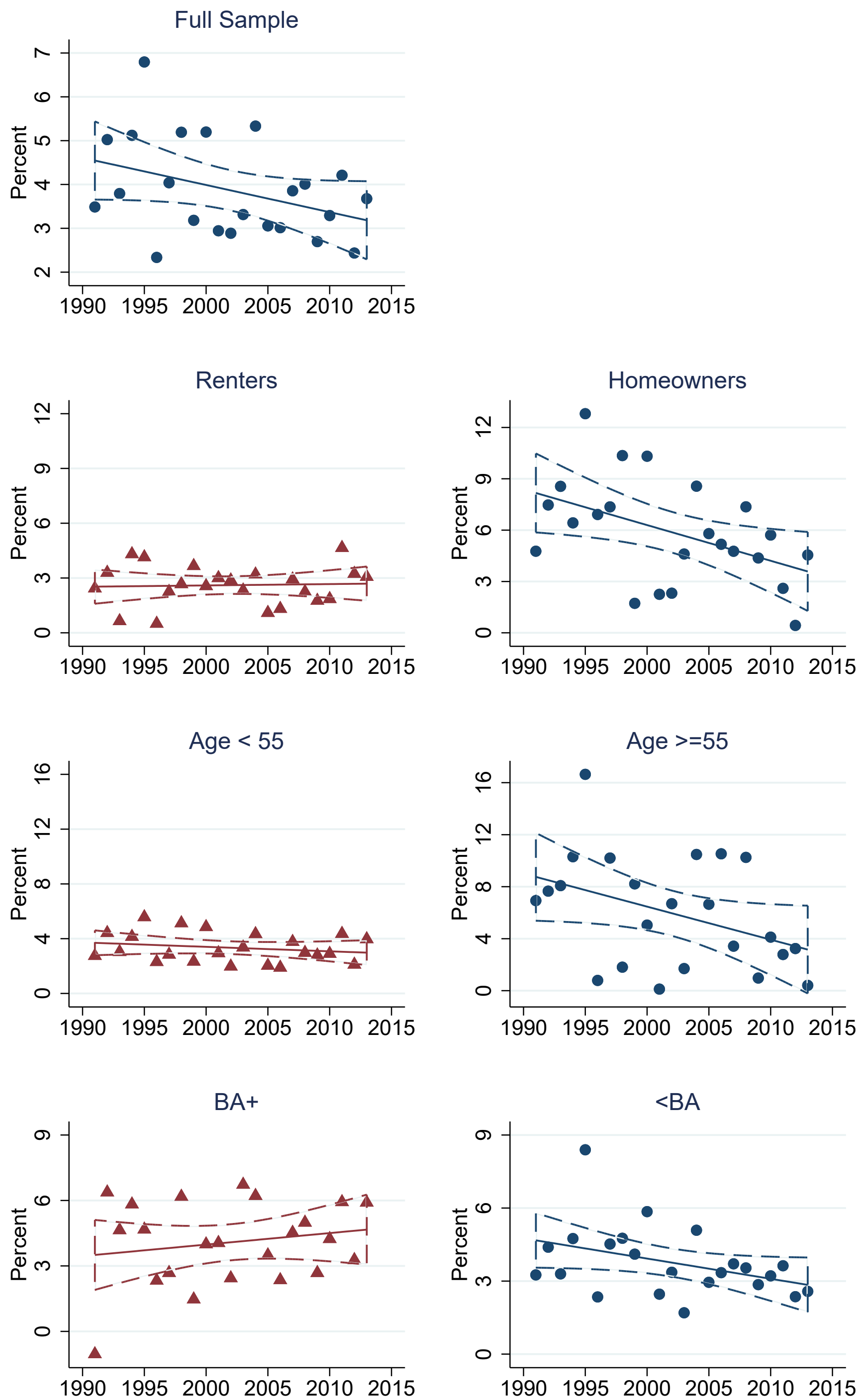
Notes: Column 1 uses the cross-CZ migration rate (IRS) as the dependent variable, uses PPML estimation, and includes CZ-pair and year fixed effects. Columns 2 and 3 use the cross-state migration dummy (CPS) as the dependent variable, use OLS estimation, and include origin-state, destination-state, and year fixed effects. The CZ- and state-level wage and housing price variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level (Col 1) and at the origin-state-level (Col 2 and 3) in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01.

in Column 3 of Table 5. The coefficients on these demographic controls have sensible signs and most are statistically significant. For instance migration is significantly less common among homeowners, older individuals, parents, and those that are employed, while it is significantly more common among the more highly educated. However, adding these covariates has virtually no impact on the estimated effects of origin and destination wages and home values.

While the consistency of the baseline results across data sets and specifications is reassuring, at present we are primarily interested in investigating the time-varying effects of origin home prices. As a starting place, the top panel of Figure 9 uses CPS data and predicts migration due to changes in origin home price elasticities over time, following the same approach utilized with the IRS data in Section 6.3.[30] Figure 9 shows that origin home prices predict a decline in migration rates, and the magnitude of this decline is similar to results obtained using the IRS data (bottom left panel of Figure 8).

[30]Specifically, we regress the cross-state migration dummy on origin home prices, origin wages, destination home prices, and destination wages all interacted with year, as well as the demographic controls, origin-state fixed effects, destination-state fixed effects, and year fixed effects included in column 3 of Table 5. Predicted migration due to origin home prices is then calculated using the time-varying elasticities and the origin home prices from the presample year.

Figure 9
Predicted Migration - Origin Home Price Elasticities (CPS)



Notes: Predicted migration due to origin-home-prices is calculated using the time-varying elasticities and the mean home price from the pre-sample year. Data source: CPS.

To investigate the determinants of this decline in home price elasticities, we utilize the information on demographic characteristics available in the CPS data. Specifically, we separately estimate our full empirical specification for different groups of migrants, including homeowners and renters, younger and older individuals, and more and less educated people. The time-varying origin home price elasticities of interest are reported in the bottom panels of Figure 9 and show significant variation across these groups. Specifically homeowners, older individuals, and those without a college degree are becoming less responsive to origin home prices over time, while trends in the sensitivity to origin home prices are flat or increasing for renters, younger individuals, and college graduates.

Additional unreported results show that origin home price elasticity trends to not substantially differ by marital status, employment status, children, or immigration status. Thus there is little support for explanations for the housing-induced decline in migration that focus on the rise of two income households, changes in labor force participation, the role that children play in family decision making, or changes in international immigration. Instead it is older, less-educated, home owners that are becoming less responsive to origin housing prices.

Notably, these three characterises which are correlated with declining sensitivity to housing prices are frequently overlapping in practice: home-ownership is more common later in the life-cycle, and the prevalence of college completion has increased substantially over time, such that older Americans are less likely to hold a college degree. In conjunction with the geographic heterogeneity in origin home price elasticities documented in Table 4, we can heuristically imagine an older homeowner in New Jersey or California. Thirty years ago, when such an individual experienced an increase in home values in their current location, they typically responded by moving to a Sunbelt location, perhaps to retire, but today they are much less likely to do so.[31]

While it is not obviously precisely *why* the preferences of these geographic and demographic groups have shifted, we believe that highlighting the specific *mechanism* that is the leading cause of declining mobility (origin home price elasticities) and also specifying the specific *types* of individuals and families for whom this mechanism appears most salient (older homeowners in urban states like California, Illinois and New York) substantially improves our understanding of why internal mobility has declined in recent decades.

[31]Our findings are consistent with evidence that newly aging individuals are reluctant to sell their homes and downsize in retirement. For instance, a survey of baby-boomer homeowners found that the majority of them (52%) said they will never move from their current house (https://media.chase.com/news/downsize-or-right-size-baby-boomers-are-looking-to-renovations-in-order-to-age-in-place)

7 Conclusion

Geographic mobility has historically been a powerful tool for increasing economic opportunity both within and between generations in the United States. However, internal migration has steadily *declined* over the last three decades, Understanding why Americans are moving less has important welfare and policy implications.

In this paper we use administrative tax records on the universe of cross-CZ moves to show that migration decisions respond in intuitive ways to wages and housing prices. We emphasize the need to consider the bilateral nature of migration decisions, by accounting for labor and housing market conditions in both the origin and destination location. Transparent predicted values from these regressions can explain all of the decline in mobility observed in the United States over the last thirty years.

Additional results showed that wages alone would have actually led to an increase in internal migration, and instead it is housing that has been an impediment to mobility in the U.S. Furthermore, it is not a rise in the geographic dispersion of wages or housing prices that are driving migration trends, but rather responsiveness to these factors. Even more specifically, the decline in internal migration appears to be primarily driven by the migration decisions of older homeowners in expensive urban states becoming less responsive to local home prices. Researchers and policy makers interested in internal mobility should focus on why Americans are increasingly willing to tolerate higher housing prices in their current location, whereas in the past they frequently responded by moving to more affordable areas.

References

D. H. Autor, D. Dorn, and G. H. Hanson. Import competition in the united states. *American Economic Review*, 103(6):2121–68, 2013.

R. Baldwin and D. Taglioni. Gravity for dummies and dummies for gravity equations, 2006.

J. Barkema and T. Bayoumi. Stranded! how rising inequality surpressed us migration and hurt those 'left behind.'. *IMF Working Paper*, 19(122), 2019.

A. W. Bartik and K. Rinz. Moving costs and worker adjustment to changes in labor demand: Evidence from longitudinal census data. *Manuscript, University of Illinois at Urbana-Champaign*, 2018.

G. Basso and G. Peri. Internal mobility: The greater responsiveness of foreign-born to economic conditions. *Journal of Economic Perspectives*, 34(3):77–98, Summer 2020.

C. Bentivogli and P. Pagano. Regional disparities and labour mobility: The euro-11 versus the usa. *Labour*, 13(3):737–760, 1999.

O. J. Blanchard and L. F. Katz. Regional evolutions. *Brookings Papers on Economic Activity*, 23(1): 1–76, 1992.

A. Bogin, W. Doerner, and W. Larson. Local house price dynamics: New indices and stylized facts. *Real Estate Economics*, 47(2):365–398, 2019.

K. Borusyak, R. Dix-Carneiro, and B. K. Kovak. Understanding migration responses to local shocks. *Working Paper*, 2022.

J. Burke and A. R. Miller. The effects of job relocation on spousal careers: Evidence from military change of station moves. *Economic Inquiry*, 56(2):1261–1277, 2018.

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

D. Card, J. Rothstein, and M. Yi. Location, location, location. *NBER Working Paper 31587*, 2023.

R. Chetty and N. Hendren. The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162, 2018a.

R. Chetty and N. Hendren. The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228, 2018b.

N. E. Coulson and P. L. Grieco. Mobility and mortgages: Evidence from the psid. *Regional Science and Urban Economics*, 43(1):1–7, 2013.

M. Dao, D. Furceri, and P. Loungani. Regional labor market adjustment in the united states: Trend and cycle. *American Economic Review*, 99(2):243–57, 2017.

E. Derenoncourt. Can you move to opportunity? evidence from the great migration. *American Economic Review*, 112(2):369–408, 2022.

T. Deryugina, L. Kawano, and S. Levitt. The economic impact of hurricane katrina on its victims: Evidence from individual tax returns. *American Economic Journal: Applied Economics*, 10(2):202–33, 2018.

J. DeWaard, M. Hauer, E. Fussell, K. J. Curtis, S. Whitaker, K. McConnell, K. Price, and D. Egan-Robertson. User beware: Concerning findings from recent us internal revenue service migration data. *Medium: PDF Publisher: Minneapolis, MN: Minnesota Population Center, Minnesota Population Center Working Paper Series. Visited on*, 8(16):2020, 2020.

M. Faber, A. P. Sarto, and M. Tabellini. Local shocks and internal migration: The disparate effects of robots and chinese imports in the US. *NBER Working Paper 30048*, 2022.

F. Ferreira, J. Gyourko, and J. Tracy. Housing busts and household mobility. *Journal of urban Economics*, 68(1):34–45, 2010.

J. P. Ferrie. Internal migration. In S. B. Carter, S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright, editors, *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, pages 489–494. Cambridge University Press, New York, 2003.

A. Foote, M. Grosz, and A. Stevens. Locate your nearest exit: Mass layoffs and local labor market response. *Industrial and Labor Relations Review*, 72(1):101–26, 2019.

E. L. Glaeser and D. C. Maré. Cities and skills. *Journal of Labor Economics*, 19(2):316–342, 2001.

A. Greenland, J. Lopresi, and P. McHenry. Import competition and internal migration. *The Review of Economics and Statistics*, 101(1):44–59, 2019.

G. Howard. The migration accelerator: Labor mobility, housing, and demand. *American Economic Journal: Macroeconomics*, 12(4):147–79, 2020.

N. Jia, R. Molloy, C. Smith, and A. Wozniak. The economics of internal migration: Advances and policy questions. *Journal of Economic Literature*, Forthcoming, 2022.

G. Kaplan and S. Schulhofer. Understanding the long-run decline in interstate migration. *International Economic Review*, 58(1):57–94, 2017.

S. Magrini. Regional (di) convergence. In *Handbook of regional and urban economics*, volume 4, pages 2741–2796. Elsevier, 2004.

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

J. Monras. Economic shocks and internal migration. *CEPR Discussion Paper 12977*, 2020.

E. Moretti. *The new geography of jobs.* Houghton Mifflin Harcourt, 2012.

E. Nakamura, J. Sigurdsson, and J. Steinsson. The gift of moving: Intergenerational consequences of a mobility shock. *Review of Economic Studies*, 89(3):1557–92, 2022.

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

M. D. Partridge, D. S. Rickman, M. R. Olfert, and K. Ali. Dwindling u.s. internal migration: Evidence of spatial equilibrium or structural shifts in local labor markets? *Regional Science and Urban Economics*, 42(1-2):375–88, 2012.

A. J. Plantinga, C. Detang-Dessendre, G. L. Hunt, and V. Piguet. Housing prices and inter-urban migration. *Regional Science and Urban Economics*, 43(2):296–306, 2013.

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

S. Rosen. Wage-based indexes of urban quality of life. *In P. Mieszkowski and M. Straszheim, Editors, Current Issues in Urban Economics*, Johns Hopkins University Press, Baltimore:74–104, 1979.

J. Santos Silva and S. Tenreyro. The log of gravity. *The Review of Economics and Statistics*, 8(4): 641–658, 2006.

R. Wilson. Moving to jobs: The role of information in migration decisions. *Journal of Labor Economics*, 39(4):1083–1128, 2021.

R. Wilson. Moving to economic opportunity: The migration response to the fracking boom. *Journal of Human Resources*, 57(3):918–55, 2022.

D. Yagan. Employment hysteresis from the great recession. *Journal of Political Economy*, 127(5):2505–58, 2019.

A ONLINE APPENDIX

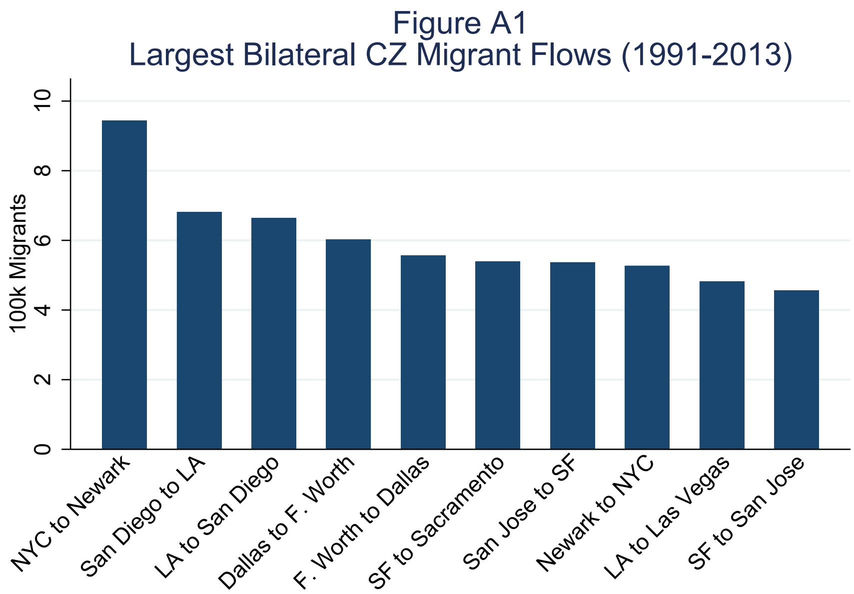


Table A1: Heterogeneity

	Migration Rate						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Poor_Mig	Rich_Mig	Short_Dist	Long_Dist	Non_Families	Families
Orig_Wage	-0.451***	-0.453***	-0.343***	-0.463***	-0.194**	-0.374***	-0.537***
	(0.051)	(0.059)	(0.099)	(0.057)	(0.096)	(0.067)	(0.076)
Origin_Home_Price	0.162***	0.101***	0.309***	0.162***	0.218***	0.052**	0.275***
	(0.020)	(0.026)	(0.030)	(0.027)	(0.033)	(0.023)	(0.033)
Dest_Wage	0.760***	0.675***	1.163***	0.741***	1.154***	0.811***	0.688***
	(0.051)	(0.060)	(0.103)	(0.059)	(0.090)	(0.064)	(0.079)
Dest_Home_Price	-0.248***	-0.232***	-0.236***	-0.263***	-0.131***	-0.211***	-0.288***
	(0.020)	(0.026)	(0.035)	(0.028)	(0.029)	(0.023)	(0.033)
CZ-Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	846,400	423,154	423,154	423,177	423,223	423,200	423,200

Notes: PPML estimation. The dependent variable is the migration rate between the origin and destination CZ. Column 1 reports the baseline results. Column 2 and 3 split the bilateral CZ pairs based on the average per-capita adjusted gross income of migrants. Columns 4 and 5 split the sample based on the distance between the bilateral CZ pairs. Finally, Columns 6 and 7 split the bilateral CZ pairs based on the average share of dependents migrating (which is identified by comparing tax returns to exemptions). All independent variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01.

Table A2: Robustness

	Migration Rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	OLS	Years_91to19	Weighted	Weighted_1991	Returns	Moves_More100Miles	Not_HI_AK
Orig_Wage	-0.451***	-0.042***	-0.220***	-0.304**	-0.248**	-0.376***	-0.477***	-0.484***
	(0.051)	(0.005)	(0.062)	(0.125)	(0.124)	(0.048)	(0.064)	(0.051)
Origin_Home_Price	0.162***	0.064***	0.085***	0.183**	0.202***	0.106***	0.208***	0.162***
	(0.020)	(0.003)	(0.030)	(0.076)	(0.074)	(0.019)	(0.021)	(0.020)
Dest_Wage	0.760***	0.174***	0.499***	0.305**	0.350***	0.723***	1.213***	0.739***
	(0.051)	(0.005)	(0.066)	(0.128)	(0.126)	(0.047)	(0.070)	(0.051)
Dest_Home_Price	-0.248***	-0.020***	-0.219***	-0.285***	-0.315***	-0.190***	-0.268***	-0.258***
	(0.020)	(0.003)	(0.028)	(0.062)	(0.061)	(0.018)	(0.024)	(0.020)
CZ-Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	846,400	6,399,888	593,949	846,400	455,262	846,400	782,000	829,725

Notes: The dependent variable is the migration rate between the origin and destination CZ. Column 1 is the baseline PPML estimation, while Column 2 uses OLS and the log of migrants plus one as the dependent variable. Column 3 uses the years 2014-2019 but drops 10-20 person migrant flows in earlier years to be consistent with the later bottom coding. Column 4 and 5 reported regressions weighted by the mean bilateral migration flow and by the 1991 bilateral migration flow respectively. Column 6 uses tax returns rather than tax exemptions to construct the migration rate. Column 7 drops bilateral CZ pairs that are less than 100 miles apart. Column 8 excludes any migration flows involving Alaska and Hawaii. All independent variables are logged and lagged one year. Robust standard errors clustered at the CZ-pair level in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01.