Labor Market Frictions and Moving Costs of the Employed and Unemployed*

Tyler Ransom†

February 1, 2019

Abstract

This paper examines the role of labor market frictions and moving costs in explaining the migration behavior of US workers by employment status. Using data on low-skilled workers from the Survey of Income and Program Participation (SIPP), I estimate a dynamic model of individual labor supply and migration decisions. The model incorporates a reduced-form search model and allows for migration for non-market reasons. My estimates show that moving costs are substantial and that labor market frictions primarily inhibit migration of the employed. I use the model to study migration responses to local labor market shocks and to a moving subsidy. Workers' preferences for non-market amenities, coupled with substantial moving costs and employment frictions, grant market power to incumbent employers. Large moving costs also likely affect employers' recruiting behavior.

JEL Classification: C35, E32, J22, J61, J64, R23

Keywords: Migration, Job search, Dynamic discrete choice

*A previous version of this paper was circulated under the title, "The Effect of Business Cycle Fluctuations on Migration Decisions." I would like to thank Peter Arcidiacono, Patrick Bayer, V. Joseph Hotz, and Arnaud Maurel for their helpful comments and encouragement. Martha Stinson and Gary Benedetto provided expert knowledge and invaluable assistance with the SIPP data and methodology, and Christopher Timmins generously provided locational price data. Funding from NSF grant SES-11-31897 and the OU College of Arts & Sciences is gratefully acknowledged. I would also like to thank participants at the Princeton Conference on Monopsony in the Labor Market, and in particular Evan Starr and my discussant Ted To. This paper also benefited from conversations with Jared Ashworth, Esteban Aucejo, Patrick Coate, Christos Makridis, Kyle Mangum, Ekaterina Jardim, Michael Ransom, Seth Sanders, Juan Carlos Suárez Serrato, and various other conference and seminar participants. All errors are my own. Any opinions and conclusions expressed herein are my own and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

†University of Oklahoma and IZA. Contact: Department of Economics, University of Oklahoma, 308 Cate Center Dr, Room 158, Norman, OK 73072. E-mail: ransom@ou.edu

1 Introduction

Migration is widely considered to be a key indicator of labor market health, for two reasons. First, it is understood to be the primary way by which local labor markets adjust to shocks (Topel, 1986; Blanchard and Katz, 1992; Yagan, 2014). Second, lower levels of migration may indicate a less competitive labor market: when workers are unable or unwilling to move, employers can compensate them below their market value (Ransom, 1993; Fox, 2010), or recruit only within the local area.[1]

In this paper, I develop and estimate a dynamic structural model. In the model, workers face frictions in obtaining employment and costs in moving across locations or entering or exiting the labor force. These frictions and costs depend on the worker's current employment status. I use the model to compute moving costs and to examine workers' relocation behavior in response to local labor market shocks or to a moving subsidy (e.g. Moretti, 2012; H.R. 2755, 2015). Workers' willingness to move is informative of the extent of employers' monopsony power through wage-setting or selective recruiting strategies.

I study individual migration, employment, and labor force transitions across U.S. metropolitan areas over the period 2004–2013. My primary data source is confidential panel data collected by the Survey of Income and Program Participation (SIPP). My sample consists of prime-age white men who are not college educated. The large coverage of the SIPP allows me to observe many moves and to accurately observe the conditions of many local labor markets. The SIPP also contains detailed information on demographic characteristics and labor market experience.

The econometric model characterizes locations in three dimensions which enter workers' utility functions and govern their decision making. The three dimensions are market and non-market amenities, expected earnings, and expected employment. Each worker has common preferences for a location's market amenities (e.g. climate), but workers may value non-market amenities differently. Earnings and employment differ across workers based on differences in their observable characteristics. Workers are also classified into two unobservable types which differ in terms of wages, employment, and switching costs. Using estimates of worker preferences and productivity, I calculate each type of worker's willingness to move under alternative scenarios such as a local labor market shock or a government-provided moving subsidy. Migration responses to changes in the local labor market are similar to a spatial labor supply elasticity.

A key component of my analysis is that search frictions differ based on current employment and residence status. That is, in the style of Burdett and Mortensen (1998), the employed and non-employed face different search processes. This paper builds on their framework by also allowing the search process to differ based on whether a worker has recently moved from another labor

[1] Geography is an important part of monopsony power. See Bhaskar and To (1999); Bhaskar, Manning, and To (2002); Bhaskar and To (2003); and Staiger, Spetz, and Phibbs (2010) who examine multi-firm monopsony power through the lens of the canonical spatial models of Hotelling (1929) and Salop (1979).

market. In this sense, cross-location search frictions can also function as a type of moving cost. Descriptive statistics show that these dimensions are important to migration and job search. I show that the non-employed are much more likely than the employed to move.[2] I also show that employed movers are much less likely to receive a job offer than employed stayers. On the other hand, job offer arrival rates are about the same for non-employed movers and stayers.

This paper models workers' locational choice and labor supply as a discrete choice dynamic programming problem.[3] Search frictions enter the model in a reduced form. Each period, an individual chooses whether or not to enter the labor force in one of 55 locations. If he chooses to supply labor, he pre-commits to working, enters a lottery, and becomes employed (or continues to be employed) with some probability (Harris and Todaro, 1970). The employment probability depends on local labor market conditions as well as the worker's previous location decision and observable characteristics.[4] The model assumes that workers have symmetric information about future labor market shocks in all locations, as well as their own taste shocks. When making their decision, forward-looking individuals forecast these shocks and also take into account that it may be costly to change their decision in future periods.

In order to estimate the model, I utilize recent developments in the estimation of large-state-space dynamic discrete choice models. By making use of conditional choice probabilities (CCPs) and the property of finite dependence, I tractably estimate a model that includes many alternative choices and uncertainty in choice outcomes.[5]

The estimated parameters imply that moving costs are substantial, and that labor market frictions are especially burdensome for the employed. I estimate the moving cost to the average person to be on the order of $400,000. The primary reason for this large magnitude is that there is a weak empirical relationship between expected earnings and observed moves. With regard to search frictions, the descriptive finding of reduced job offer arrivals for employed movers continues to hold in the structural model after allowing for employment to depend on unobserved worker ability. Thus, search frictions act as an additional hindrance to migration for the employed. In contrast, the non-employed are equally likely to receive an offer whether or not they move, so search frictions are less binding to their migration behavior.

[2]See also Schmutz and Sidibé (2017) and Schlottmann and Herzog (1981) who find similar results in other contexts (France and USA, respectively).

[3]See also Gould (2007); Gemici (2011); Kennan and Walker (2011); Baum-Snow and Pavan (2012); Bishop (2012); Coate (2013); Adda, Dustmann, and Görlach (2015); Mangum (2015); Schmutz and Sidibé (2017); and Bartik (2018) who estimate dynamic models of migration.

[4]For other papers examining migration as adjustment to labor market shocks, see Notowidigdo (2011); Sastry and Gregory (2014); Yagan (2014); Gardner and Hendrickson (2017); Huttunen, Møen, and Salvanes (2018); Monras (2018); and Foote, Grosz, and Stevens (2019).

[5]The literature on estimation of dynamic models using CCPs originated with Hotz and Miller (1993). Altuğ and Miller (1998) expanded the set of models that can be estimated in this fashion by introducing the concept of finite dependence. Arcidiacono and Miller (2011) show how CCPs can be used in models with unobserved heterogeneity. I employ methods from all three of these papers while showing how finite dependence can be achieved even in the presence of a stochastic choice set (Arcidiacono and Miller, Forthcoming).

I use the structural model estimates to study migration responses to local labor market shocks, and to a government move subsidy. Employed workers are more likely to *stay* in a place experiencing a local economic downturn, but less so if the economic downturn is nationwide. The opposite is true for the unemployed, who are more likely to move in response to a local economic downturn. If the government were to offer a $10,000 moving subsidy to the unemployed (e.g. as discussed in Moretti, 2012, and the American Worker Mobility Act of 2015 (H.R. 2755, 2015)), my model predicts that there would be low take-up rates ($\approx 3\% - 5\%$).[6] The model also predicts that subsidy recipients would move to locations with higher amenities, near their birth location, and with better employment prospects.

Taken together, I find that high moving costs and labor market frictions tend to keep workers in their current location, even if workers are offered a moving subsidy. Response to the subsidy differs by the conditions of the local labor market and the desirability of the location. Combined with the tendency of workers who do take the moving subsidy to choose locations with high market- or non-market amenities, these findings imply that spatial labor supply elasticity is low (Balasubramanian et al., 2018; Dube et al., 2018; Hirsch et al., 2018). The findings can be taken to imply that firms have monopsony power due to workers' high moving costs and the employed's lower likelihood of finding a job after a move. High moving costs may also reduce the radius of employer recruiting (Krueger and Ashenfelter, 2018; Naidu and Posner, 2018; Webber, 2018).

2 Data and Stylized Facts about Migration and Unemployment

This section introduces the main data sources used in the paper and presents stylized facts about migration and unemployment upon which the structural model builds.

2.1 Data

This subsection describes the data and key variables used to estimate the model. The main data source is the 2004 and 2008 panels of the Survey of Income and Program Participation (SIPP). I supplement the SIPP with data on location characteristics and local labor market conditions.

2.1.1 The SIPP

The SIPP is a longitudinal survey of a stratified random sample of residents of the United States, administered by the United States Census Bureau. Respondents are interviewed every four months over a four- or five-year span. Each four-month period is referred to as a wave. Survey respondents

[6]See also Marinescu and Rathelot (2018) who find that relocating job seekers to minimize unemployment would have only modest effects on the aggregate unemployment rate. Caliendo, Künn, and Mahlstedt (2017) find that mobility assistance programs in Germany increase geographical mobility of the unemployed.

are asked questions regarding their living arrangements, labor force participation, earnings, assets, government program participation, migration, and education, among many other topics. Within each wave, respondents provide additional information on many of these activities at the monthly level.

In order to preserve confidentiality, the data used here—which make use of detailed residence location and earnings that are not top-coded—are not released publicly by the SIPP and are only available through the Census Research Data Center (RDC) Network.[7] Furthermore, the confidential version of the SIPP is linked via the respondent's social security number to Internal Revenue Service (IRS) and Social Security Administration (SSA) administrative data on annual earnings, employment history, government program participation, and social security benefits receipts. I make use of this link to create work experience profiles based on the administrative data that are less vulnerable to survey recall error.

The SIPP is essential to studying migration and labor supply behavior during the Great Recession.[8] It is the only national panel survey that covers the time period of the Great Recession on a multi-cohort basis.[9] Furthermore, the SIPP is a survey and thus can separately measure unemployment and labor force detachment, two effects that are conflated in studies that use administrative data such as tax records (Schmutz and Sidibé, 2017; Yagan, 2014, Forthcoming). Finally, the SIPP is a much larger sample than any other longitudinal surveys that have primarily been used in the literature (e.g. the National Longitudinal Survey of Youth (NLSY) cohorts or the Panel Study of Income and Dynamics (PSID)). This feature allows me to estimate location-specific parameters with greater precision.[10]

The main disadvantages of the SIPP are two-fold. First, its panels are relatively short—four to five years in length. Second, attrition rates in the SIPP are higher than in other longitudinal surveys. However, there is evidence that high attrition rates do not bias labor market outcomes (Zabel, 1998).[11]

2.1.2 Individual variables

With the data in hand, I now introduce the outcome and explanatory variables used in the analysis. The three outcomes of interest are location, labor force status, and earnings. Within the labor force, the two outcomes are employment or unemployment. For the employed, monthly salary is

[7] For more information regarding the SIPP, see http://www.census.gov/sipp/. For more information about conducting research using confidential data in an RDC, see http://www.census.gov/ces/rdcresearch/.

[8] Other migration studies using the SIPP include Kaplan and Schulhofer-Wohl (2017), Oswald (2018), and Balgova (2018).

[9] In contrast, the National Longitudinal Survey of Youth (NLSY) cohorts, which are heavily used in the literature, only follow a certain age group of individuals during the Great Recession.

[10] For example, Kennan and Walker (2011) and Bishop (2012) are both required to estimate location specific earnings effects in the CPS because of small cell sizes in the NLSY79.

[11] Zabel compares the attrition behavior of the SIPP and PSID.

an additional outcome.

Labor force participation and unemployment are defined as follows:

labor force participants are those who have a full-time job or are seeking a full-time job. Those who are self-employed or who voluntarily work part-time are excluded from the labor force.

unemployment closely resembles the U-6 unemployment definition reported by the BLS.[12] Unemployment is defined here as labor force participation that is not full-time employment. Full-time employment is defined as working 35 or more hours per week for all weeks in the survey month.

I focus on full-time employment for three reasons. First, full-time employees are most likely to be employed throughout the year, which more closely matches the time horizon of the model. Second, the SIPP does not measure hours worked at the monthly level—only at the wave level. Thus, measuring earnings at the hourly level is more difficult. I focus on full-time jobs because these jobs are most likely to be salaried, and an hourly earnings measure does not appropriately capture marginal labor productivity for salaried workers. Finally, there is evidence that part-time employment and unemployment are highly correlated, making it innocuous to consider involuntary part-time workers as unemployed workers.[13]

I define monthly earnings as the sum of earnings across all jobs in the survey month. I deflate earnings by cost of living in the location as described later in this section. All monetary figures throughout this paper are expressed in constant 2000 dollars unless otherwise noted.

The primary explanatory variables are work experience, age, and birth location. I indirectly use additional demographic variables such as education level, gender, and race/ethnicity to determine the estimation subsample. I create work experience from IRS records as an annualized measure of the sum of all quarters worked. I similarly construct age from the SSA data by comparing the calendar year and month with the birth year and month. Respondents report their state or country of birth in Wave 2 of each SIPP panel.

2.1.3 Geographical variables

I define locations as cities (Core Based Statistical Areas or CBSAs).[14] In order to maintain tractability, I restrict to the 35 cities that are most frequently observed in the SIPP. I construct

[12]U-6 is defined as total unemployed, plus all marginally attached workers, plus total involuntary part-time workers, as a percent of the civilian labor force plus all marginally attached workers.

[13]For instance, the trend in the overall ratio of part-time workers to full-time workers closely follows the trend in the U-6 unemployment rate. Figure A2 shows this relationship.

[14]My definition of city is the Core Based Statistical Area (CBSA) as defined in 2009 by the U.S. Office of Management and Budget (OMB). CBSAs include one or more counties and are defined according to commuting ties. As such, they are a reasonable measure of whether or not a county belongs to a city.Using the 2009 definition, there are a total of 942 CBSAs—366 Metropolitan Statistical Areas (MSAs) and 576 Micropolitan Statistical Areas (μSAs). Because it is infeasible to estimate this many locations, the choice set is aggregated.

an additional 20 residual synthetic locations to ensure that the choice set is geographically exhaustive. These synthetic locations are grouped into two population bins (small and medium) based on population. Table A4 contains a complete list of all 55 locations. A map of the 35 cities can be found in Figure A1.

Modeling a large number of locations is essential to capturing the actual locational choice alternatives that individuals face. I focus on cities rather than states because business cycle effects are heterogeneous across cities, even within the same state.[15] Furthermore, because many cities cross state boundaries, focusing on cities more closely characterizes the actual local labor market. Modeling the largest cities is a parsimonious way of categorizing the choice set: 43% of the US population resides in the 30 largest cities (CBSAs).[16] Finally, the residual locations are divided into population categories because there is evidence in the urban economics literature that a variety of labor market outcomes differ systematically by city size due to agglomeration economies, thick market effects, human capital externalities, and labor market competition (Glaeser and Maré, 2001; Wheeler, 2006; Yankow, 2006; Gould, 2007; Baum-Snow and Pavan, 2012; Bleakley and Lin, 2012; Hirsch et al., 2018). Breaking out the residual categories by city size is a parsimonious way of capturing these effects.

Beyond the geographical definition of location, I also make use of the population, unemployment rate, and price level of the worker's city. Population is defined as the 2000 Census population level in the county of residence, aggregated to the CBSA level. It is used to divide locations that are smaller than the top 35 cities. The unemployment rate is taken at the county level from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics data series and aggregated to the CBSA level, weighting by county population.[17] This variable is used in the model to inform individuals about their employment prospects in each location. I merge these city characteristics using a crosswalk that maps counties to CBSAs. Further details on data sources can be found in Table A3.

Following a number of papers in the literature, I spatially deflate earnings using the American Chamber of Commerce Research Association's Cost of Living Index (ACCRA-COLI).[18] This is done primarily because the model presented in the next section abstracts from the housing market. I calculate the local cost of living index in a like manner as Baum-Snow and Pavan (2012), who use a utility indifference condition to derive the price index at location j relative to base location 0 as a share-weighted geometric mean of prices in the consumer's basket of goods. For housing prices, I follow Winters (2009) and use quality-adjusted gross rents instead of the housing prices listed in ACCRA. Winters finds that spatially deflating earnings in this way yields an elasticity of

[15] See, e.g., Moretti (2012) who contrasts the labor market trajectories of different areas within California.

[16] See https://en.wikipedia.org/wiki/List_of_core-based_statistical_areas.

[17] For the 20 residual locations, unemployment is aggregated to the location level.

[18] Studies using this data include: Glaeser and Maré (2001); Yankow (2006); Kennan and Walker (2011); and Baum-Snow and Pavan (2012), among others.

1 between nominal earnings and nominal prices, as predicted by economic theory. Further details on the construction of this index can be found in Appendix A.6.

2.1.4 Estimation subsample

I estimate the model using non-Hispanic white men of prime working age (i.e. ages 18–55 at the beginning of the survey) who have completed school and do not have a bachelor's degree. I focus on males of a particular education level, race, and ethnicity in order to form a homogeneous sample.[19] The final estimation subsample comprises 16,648 males each averaging 3.03 annual observations.

Tables 1 and 2 list descriptive statistics for the estimation subsample. The average individual in the sample is 42 years old and has 23 years of work experience. Living in the same location of birth is common, with almost 75% of the sample residing in their state of birth. Table 2 lists the migration statistics in the sample, which contains 568 movers who make 653 moves.[20]

I use four annual observations for the 2004 panel—the interview month of waves 2, 5, 8, and 11—to measure location, labor market outcomes, and individual characteristics. The 2008 Panel is slightly longer, so I use the same waves in addition to wave 14. The entire dataset spans the years 2004–2013, but any given individual can only appear in at most five of those years. The average individual is observed for between two and three years, but over half the sample has at least three observations. For more details on sample selection and construction of key variables, see Appendix A.5 and Table A1.

2.2 Stylized Facts about Migration and Unemployment

With the data in hand, I now present three stylized facts about migration and unemployment. First, the non-employed are more geographically mobile than the employed. Second, employed workers who move are much less likely to become employed after the move than employed workers who stay. This phenomenon is restricted to employed movers; non-employed movers are just as likely to have a job as non-employed stayers. Third, the employment prospects of non-employed workers are relatively worse during local economic downturns, compared to employed workers. For expositional reasons, I illustrate these facts using publicly available SIPP data on all workers in the United States. In all cases, employment is defined as described previously.

Figure 1 shows that, across multiple distances, the non-employed move more frequently than the employed. The difference amounts to about a 50% higher mobility rate for across-state moves,

[19] See Kennan and Walker (2011), Bishop (2012), and Bartik (2018) for other studies that focus on a similar demographic group.

[20] While the observed rate of repeat and return migration is lower than in other studies, this is due to the short panel of the SIPP. While not the focus of the current analysis, the dynamic model used in this paper allows for repeat and return migration.

and about a 30% higher mobility rate for within-state, across-county moves.[21]

To see if movers and stayers tend to have different employment outcomes, I estimate a simple linear probability model. The left-hand side variable is an indicator for full-time employment in the current period. The right-hand side variables include race-cross-gender dummies, a quadratic in experience, the previous-period unemployment rate in the current state of residence, and an indicator for having made a move since the previous period. Here, I define a move as changing counties or states, i.e. moving to a different local labor market. I estimate this linear probability model separately by previous employment status.

The results of these descriptive regressions are shown in Table 3. For employed workers, the mover dummy coefficient is -12 percentage points, indicating a large penalty for employed movers. For the non-employed, there is actually a slight gain to moving—non-employed movers are about 5 percentage points more likely to be employed than non-employed stayers.[22] Finally, an increase in the state unemployment rate reduces the employment probability of the non-employed by more than it does the employed. In this sense, the employed are more insulated from local economic downturns than those who are not employed.

While the above facts are illustrative, they are likely biased due to mismeasurement of the local labor market, endogeneity of migration, and unobserved worker heterogeneity. Additionally, the incentives to move faced by the employed and non-employed are myriad and require a more careful unpacking. In the next section, I introduce a structural model in which forward-looking individuals choose where to live and whether to supply labor. Individuals take into account that moving is costly and that employment is uncertain and is affected by local labor market conditions. I then estimate the model on the estimation subsample using restricted-access SIPP data that allows me to more precisely observe local labor markets.

3 A Model of Search Frictions, Labor Supply, and Migration

3.1 Overview

I now introduce the model that I will estimate and use to quantify moving costs, labor market frictions, and to examine counterfactual scenarios that will shed light on workers' spatial responsiveness to changes in their local labor market.

In each period, individuals choose whether or not to supply labor in one of 55 locations. The choice set is exhaustive in that in covers every possible location in the United States, and every possible labor market status. Search frictions are a key element of the model. That is, while an

[21]While not the focus of this paper, Figure 1 exhibits declining migration rates over time. This is consistent with the findings of Molloy, Smith, and Wozniak (2011) and Kaplan and Schulhofer-Wohl (2017).

[22]These findings can be replicated in other survey data from the US, such as the National Longitudinal Survey of Youth or the American Community Survey.

individual may control his labor supply decision, he cannot control his employment *outcome*. For example, a non-employed worker may exogenously receive a job offer or an employed worker may exogenously be laid off. Furthermore, these job offer and destruction rates are allowed to vary by location, migration status, and calendar time, thus capturing heterogeneity in local business cycles and spatial frictions in job search. Allowing individuals to choose to supply labor is essential to the model, because the employment probabilities are conditional on labor force participation.[23] I specify the job search parameters in a reduced form, but the underlying search process relates to a Burdett and Mortensen (1998) approach where workers can move locations and enter or exit the labor force.

Individuals are forward-looking and in each period choose the alternative that maximizes their present discounted value of utility. Thus, individuals take into account local labor market conditions when choosing where to locate—in addition to amenities and earnings prospects, which have been traditionally modeled in the migration literature.[24] Individuals also understand that there are costs associated with changing locations or labor force status. These costs motivate individuals to be forward-looking when considering their decision in each period.

The model is the first to examine locational choice and labor supply in a dynamic setting with time-varying search frictions and are tied to local business cycles. I now present each feature of the model in more detail, beginning with earnings and flow utilities for each alternative.[25]

3.2 Earnings

Monthly earnings for individual i in location ℓ of unobserved type r and calendar year t are a function of location-year fixed effects $\psi_{0\ell t}$, work experience x_{it}, a person-specific unobserved type indicator $\tau_i r$, and measurement error $\eta_{i\ell t}$. The log earnings equation is

$$\ln w_{i\ell rt} = \psi_{0\ell t} + \psi_1 G(x_{it}) + \psi_2 \tau_{ir} + \eta_{i\ell t} \tag{3.1}$$

where $G(\cdot)$ is a quadratic polynomial. Human capital accumulation is accounted for by including work experience as a determinant of earnings. Person-specific unobserved heterogeneity in earnings is captured by the discrete type indicator τ_{ir}. Cross-location heterogeneity and nonstationarity in earnings is accounted for by the location-year fixed effects, which allow for business cycle effects to be different across locations. Importantly, individuals observe the time dummies in calendar year t but must form expectations about their evolution in future periods. Measurement

[23] Additionally, Amior and Manning (2018) argue that local labor supply ratios are key indicators for individuals' economic opportunity.

[24] The model relates to Molloy and Wozniak (2011), who examine migration across the business cycle. In my model, individuals are assumed to know what each location's labor market conditions are, as well as their trends and persistence. See Ilut and Schneider (2014) and Jeon (2017) for models which relax this assumption.

[25] A cross-referenced notation glossary for all Greek symbols is available in Table A9.

error $\eta_{i\ell t}$ is assumed to be distributed $N\left(0,\sigma_\eta^2\right)$ and independent over time, locations, and all other state variables.

3.2.1 Forecasting earnings

Individuals are uncertain about future realizations of the $\psi_{0\ell t}$'s and the $\eta_{i\ell t}$'s. They forecast future earnings according to an AR(1) process (with drift) on the location-year fixed effects $\psi_{0\ell t}$:

$$\psi_{0\ell t} = \rho_{0\ell} + \rho_1 \psi_{0\ell t-1} + \zeta_{\ell t} \tag{3.2}$$

where $\zeta_{\ell t}$ is distributed $N\left(0,\sigma_{\zeta_\ell}^2\right)$. In other words, individuals know the drift and autocorrelation coefficients ($\rho_{0\ell}$ and ρ_1) and shock variances ($\sigma_{\zeta_\ell}^2$) for each location, and integrate over future realizations of $\zeta_{\ell t}$ using information about the distribution from which $\zeta_{\ell t}$ is drawn.[26]

Individuals also forecast future earnings innovations $\eta_{i\ell t}$ but do not need to integrate over any distribution because the $\eta_{i\ell t}$'s are measurement error and assumed to be mean-zero and independent over time, and because they are assumed to not affect decisions.

3.3 Employment probabilities

Individuals who choose to supply labor obtain employment with probability $\pi_{i\ell rt}$, which depends on their level of work experience x_{it}, their unobserved type τ_i, their current location ℓ, their previous location and employment status, and the previous unemployment rate in each location.

$$\pi_{i\ell rt}\left(x_{it},\tau_{ir},UR_{\ell t-1}\right) = \begin{cases} \left(1-\delta_{i\ell rt}\right) & \text{if employed in } \ell \text{ in } t-1 \\ \lambda_{i\ell rt} & \text{if not employed in } \ell \text{ in } t-1 \\ \lambda_{i\ell rt}^e & \text{if employed in } \ell' \neq \ell \text{ in } t-1 \\ \lambda_{i\ell rt}^u & \text{if not employed in } \ell' \neq \ell \text{ in } t-1 \end{cases} \tag{3.3}$$

Equation (3.3) operates as follows. Individuals arrive in a location and, if choosing to supply labor, are entered into a lottery that assigns them employment with probability $\pi_{i\ell rt}$. Individuals pre-commit to working if they are assigned to employment.[27] Employed individuals are thus laid off with probability $\delta_{i\ell rt}$ and unemployed individuals receive a job offer with probability $\lambda_{i\ell rt}$. Individuals coming from employment in another location receive an offer with probability $\lambda_{i\ell rt}^e$. Individuals coming from non-employment in another location receive an offer with probability

[26] The assumption that ρ_1 is not location-specific has also been made by Kaplan and Schulhofer-Wohl (2017).

[27] In this sense, individuals choose locations not by the availability of job offers, but by the likelihood of finding a job once in the destination location. Thus, workers search for a job in the location upon arrival. This motivates the sample selection discussed in the next section, since this group of people are more likely to move before finding a job (Balgova, 2018).

$\lambda^u_{i\ell rt}$. In particular, these employment probability parameters are indexed by location and time, which allows for heterogeneous business cycles across locations—a trend seen in the data. As with other non-stationary parameters in the model, individuals observe these probabilities in calendar year t but must form expectations about their evolution in future periods.

The model introduces search frictions in a reduced form primarily for computational reasons. Adding non-market migration and labor force participation to an equilibrium search model would be computationally prohibitive.[28]

In practice, each of the employment probabilities is parameterized as a predicted logistic probability, where the logistic regression has right hand side variables as follows: location fixed effects, lagged unemployment rate, an unobserved type indicator, and a mover dummy. Non-stationarity enters the employment probabilities through movements in the lagged unemployment rate, which I now detail.

3.3.1 Forecasting employment probabilities

Individuals forecast future employment probabilities according to an AR(1) process (with drift) on the local unemployment rate $\mathrm{UR}_{\ell t}$:

$$\mathrm{UR}_{\ell t} = \phi_{0\ell} + \phi_{1\ell}\mathrm{UR}_{\ell t-1} + \xi_{\ell t} \tag{3.4}$$

where, again, individuals know the shock variance and autocorrelation of the $\mathrm{UR}_{\ell t-1}$'s in each location, but integrate over possible realizations of the $\xi_{\ell t}$'s given realizations of $\mathrm{UR}_{\ell t-1}$. $\xi_{\ell t}$ is assumed to be drawn from a distribution that is $N\left(0, \sigma^2_{\xi_\ell}\right)$.

This implies that, for each t, the time-t forecast of the future employment probability is

$$\mathbb{E}_t \pi_{i\ell t+1} = \int_{\xi_\ell} \Pr\left\{\mu_2\left(\phi_{0\ell} + \phi_{1\ell}\mathrm{UR}_{\ell t} + \xi_{\ell t+1}\right) + z_{i\ell r}\mu > 0\right\} dF\left(\xi_{\ell t+1}\right) \tag{3.5}$$

where the argument inside the probability is a latent linear index comprised of μ_2, a parameter governing the relationship between the lagged unemployment rate in location ℓ and the employment probability, and $z_{i\ell r}\mu$, which represents other right-hand side variables and parameters in the probability that are stationary (e.g. location, work experience, unobserved type). In practice, μ_2 is allowed to vary by employment status to capture the fact that unemployed workers are more vulnerable to economic downturns. More details on the form and estimation of these probabilities are given in section 4.2.

[28] For a treatment of migration in the style of a classical job search model, see Schmutz and Sidibé (2017). For recent work on extending classical job search models to include preference heterogeneity over non-wage amenities, see Sullivan and To (2014); Taber and Vejlin (2016); Hall and Mueller (2018); Sorkin (2018).

3.4 State variables, flow utilities, and stochastic employment

Denote by $d_{it} = (j, \ell)$ the choice for individual i in calendar year t, where $j \in \{0,1\}$ indexes labor force status, $\ell \in \{1,\ldots,L\}$ indexes locations, and $\mathscr{J} = \{0,1\} \times \{1,\ldots,L\}$ denotes the entire choice set. Labor force participation is denoted by $j = 1$ while $j = 0$ indicates out of the labor force. As mentioned before, individuals control their labor supply decision, but not their employment outcome. To differentiate between the two, let $y_{it} \in \{e,u,n\} \times \{1,\ldots,L\}$ be the choice *outcome*, where e denotes employment, u unemployment, and n non-participation. The set $\{1,\ldots,L\}$ covers all possible locations in the United States. A complete list of these locations can be found in Table A4.

Let Z_{it} denote the state variables for individual i in calendar year t. Z_{it} contains work experience, age, calendar time t, previous decision d_{it-1}, previous employment outcome y_{it-1}, and unobserved type τ_i. The flow utility associated with making choice (j, ℓ) is a function of the observed state variables and unobservables $\varepsilon_{ij\ell t}$:

$$U_{ij\ell t}\left(Z_{it}, \varepsilon_{ij\ell t}\right) = u_{ij\ell t}\left(Z_{it}\right) + \varepsilon_{ij\ell t} \tag{3.6}$$

The flow payoffs associated with choices j and ℓ are an expectation of the employment outcomes:

$$u_{i1\ell t}\left(Z_{it}\right) = \pi_{i\ell rt}\left(Z_{it}\right) u^{e}_{i\ell t}\left(Z_{it}\right) + \left(1 - \pi_{i\ell rt}\left(Z_{it}\right)\right) u^{u}_{i\ell t}\left(Z_{it}\right) \tag{3.7}$$

$$u_{i0\ell t}\left(Z_{it}\right) = u^{n}_{i\ell t}\left(Z_{it}\right) \tag{3.8}$$

The flow utility corresponding to each employment outcome is given by

$$u^{e}_{i\ell t}\left(Z_{it}\right) = \alpha_{\ell} + \Delta_{\ell}\left(Z_{it}\right) + \Xi_{1}\left(Z_{it}\right) + \gamma_{3} b^{ST}_{i\ell} + \gamma_{4} b^{DIV}_{i\ell} + \gamma_{0} \ln \widetilde{w}_{i\ell t}\left(Z_{it}\right) \tag{3.9}$$

$$u^{u}_{i\ell t}\left(Z_{it}\right) = \alpha_{\ell} + \Delta_{\ell}\left(Z_{it}\right) + \Xi_{1}\left(Z_{it}\right) + \gamma_{3} b^{ST}_{i\ell} + \gamma_{4} b^{DIV}_{i\ell} + \gamma_{1} + \gamma_{2} \tag{3.10}$$

$$u^{n}_{i\ell t}\left(Z_{it}\right) = \alpha_{\ell} + \Delta_{\ell}\left(Z_{it}\right) + \Xi_{0}\left(Z_{it}\right) + \gamma_{3} b^{ST}_{i\ell} + \gamma_{4} b^{DIV}_{i\ell} + \gamma_{1} \tag{3.11}$$

The first term α_ℓ is a location fixed effect measuring the net value of all amenities in location ℓ. The variables $b^{ST}_{i\ell}$ and $b^{DIV}_{i\ell}$ are dummies that indicate if the individual was born in any of the states (ST) or Census divisions (DIV) contained in location ℓ.[29] $\ln \widetilde{w}_{i\ell t}\left(Z_{it}\right)$ is the deterministic component of log earnings for an individual in location ℓ with states Z_{it}, γ_1 is a home production benefit, and γ_2 is a net cost of searching for employment.[30] Home production benefits and search costs are constant

[29] For example, an individual born in Indiana would have this dummy turned on for the following locations: Indianapolis, Chicago, and the two East North Central Census division synthetic locations. There are two cities in the model that straddle Census divisions: New York City and St. Louis, MO.

[30] While γ_1 is labeled as a benefit and γ_2 a cost, the sign of each is freely estimated. γ_2 can be thought of as the additional cost to searching off-the-job relative to on-the-job.

across locations. $\Delta_\ell(Z_{it})$ are costs of moving to ℓ which are incurred if the previous location is different from ℓ. Likewise, $\Xi(Z_{it})$ are labor supply switching costs which are incurred whenever a person enters or exits the labor force. The unobserved part of the flow utility $\varepsilon_{ij\ell t}$ is a preference shock, but can equivalently be thought of as a shock to moving costs or switching costs. These preference shocks are assumed to be drawn from a standard Type I extreme value distribution, independently across i, j, ℓ, and t. The $\varepsilon_{ij\ell t}$'s are also independent of the other variables in the model.

3.5 Moving costs and switching costs

3.5.1 Moving costs

Let h indicate the previous location, which is embedded in Z_{it}, and define $D(\ell,h)$ as the great circle distance between the two locations. The moving costs are a function of a fixed cost, distance, age, previous employment status, and unobserved type, and are specified as

$$\begin{aligned}\Delta_\ell(Z_{it}) = \big(&\theta_0 + \theta_1 D(\ell,h) + \theta_2 D^2(\ell,h) + \theta_3 \mathrm{age}_{it} + \theta_4 \mathrm{age}_{it}^2 + \\ &\theta_5 \mathrm{employed}_{it-1} + \theta_6 \mathrm{unemployed}_{it-1} + \theta_7 \tau_i\big)\, 1\{\ell \neq h\}\end{aligned} \tag{3.12}$$

where $1\{A\}$ is an indicator meaning that A is true. Moving costs are specified in a reduced-form manner to flexibly capture trends in the data. This specification is in line with Kennan and Walker (2011) and Bishop (2012).[31] Additionally, I allow the moving cost to differ by previous employment status to capture the possibility that the non-employed are financially constrained, or the employed are more tied to a location. This dimension of moving costs has not been explored previously in the literature.

The intercept θ_0 corresponds to the fixed cost of moving, while θ_1 and θ_2 capture the fact that moving costs increase with distance, but at a potentially decreasing rate. θ_3 and θ_4 capture a similar idea with age. θ_5 and θ_6 are included to capture differences in psychic costs and financial costs for those who were previously employed compared with those who were previously unemployed.[32] θ_7 captures the fact that some moving costs may be unobserved to the researcher. The signs on the θ's are written here as being positive but are allowed to be freely estimated. The fixed cost and linear terms of distance and age are expected to have a negative sign, while the signs on the two quadratic terms are expected to be positive. The expected sign of the previous employment status dummies is ambiguous because, for example, while employed individuals might have more financial capital to facilitate moving, they also might face steeper psychic costs to moving locations relative to unemployed persons due to their greater attachment to the current location.

[31] Davies, Greenwood, and Li (2001) also include moving costs in a conditional logit analysis of migration.

[32] Those who were previously out of the labor force serve as the reference group.

3.5.2 Labor force switching costs

Labor force switching costs are modeled in a similar way as moving costs. Let k index the previous labor supply choice (0 for out of the labor force, 1 for in the labor force), which is embedded in Z_{it}. The labor force switching costs are allowed to vary by entry or exit and have the following form for each:

$$\Xi_1(Z_{it}) = \left(\theta_8 + \theta_9 \text{age}_{it} + \theta_{10}\text{age}_{it}^2 + \theta_{11}\tau_{ir}\right) 1\{k=0\} \tag{3.13}$$

$$\Xi_0(Z_{it}) = \left(\theta_{12} + \theta_{13} \text{age}_{it} + \theta_{14}\text{age}_{it}^2 + \theta_{15}\tau_{ir}\right) 1\{k=1\} \tag{3.14}$$

As with the moving costs, the intercepts θ_8 and θ_{12} are fixed costs of switching employment status, while the other parameters capture the fact that switching costs increase with age (but potentially at a decreasing rate) and that switching costs may be unobserved. As with the moving costs, the signs of the θ's are freely estimated. The fixed costs and linear age terms are expected to have a negative sign, while the θ's associated with the quadratic age terms are expected to be positive.[33]

3.6 The optimization problem

Individuals sequentially choose d_{it} to maximize the sum of their present discounted utility according to the following expression:

$$\max_{d_{it}} \mathbb{E}\left[\sum_{t=0}^{T} \beta^t \sum_j \sum_\ell \left(u_{ij\ell t}(Z_{it}) + \varepsilon_{ij\ell t}\right) 1\{d_{it} = (j,\ell)\}\right] \tag{3.15}$$

with discount factor β. Individuals observe current-period preference shocks before making a decision, but do not observe future shocks and must take expectations accordingly.

I now introduce notation to empirically represent the optimization problem. Define by $V_{it}(Z_{it})$ the ex ante value function for individual i in period t just before ε_{it} is revealed.

$$V_{it}(Z_{it}) = \mathbb{E}_\varepsilon \max_{j,\ell} \left\{ u_{ij\ell t}(Z_{it}) + \varepsilon_{ij\ell t} + \beta \int V_{it+1}(Z_{it+1})\, dF(Z_{it+1}|Z_{it}) \right\} \tag{3.16}$$

Under mild regularity conditions, (3.16) follows Bellman's optimality principle.[34] Now define the choice-specific value function $v_{ij\ell t}$ as the flow payoff of choosing (j,ℓ) minus $\varepsilon_{ij\ell t}$ plus future

[33] Previous employment status is not included in the switching costs so as to maintain clear interpretation of the search cost parameter γ_2 in (3.10).

[34] These conditions include additive separability of the flow utility covariates and preference shocks, and conditional independence of the state variables and preference shocks.

utility assuming that the optimal decision is made in every period from $t+1$ on.

$$v_{ij\ell t}(Z_{it}) = u_{ij\ell t}(Z_{it}) + \beta \int V_{it+1}(Z_{it+1})\, dF(Z_{it+1}|Z_{it}) \tag{3.17}$$

$$= u_{ij\ell t}(Z_{it}) + \beta \int \mathbb{E}_{\varepsilon} \max_{k,m} \{v_{ikmt+1}(Z_{it+1}) + \varepsilon_{ikmt+1}\}\, dF(Z_{it+1}|Z_{it}) \tag{3.18}$$

Expressing the value functions as choice-specific value functions simplifies the connection between the theory and empirics, as will be shown in greater length later on.

In summary, equations (3.15)-(3.18) establish the mathematical framework through which individuals make forward-looking decisions. Specifically, individuals integrate over unknown future preference shock realizations $\varepsilon_{ij\ell t}$ using the value function.

4 Identification and Estimation

This section informally discusses identification of the model and provides further details on the estimation procedure. Estimation is greatly simplified by two key standard assumptions: (i) the unobserved type dummy entering each of the components of the model accurately captures unobserved productivity and unobserved preferences; and (ii) the measurement error in earnings does not affect the flow utility of employment in (3.9); that is, individuals make their location decision based on expected earnings, and not a specific job offer. The first assumption allows me to make use of the Expectation-Maximization (EM) algorithm in estimation. The second assumption greatly simplifies the estimation of the flow utility parameters.[35]

4.1 Identification

This section informally discusses identification of various parameters of the model.[36] As in all dynamic discrete choice models, only differences in utility are identified. I choose to normalize all parameters relative to labor force participation in the first location $(d_{it} = (1,1))$. The scale of utility is normalized by assuming payoff shocks are drawn from a Type I extreme value distribution. The discount factor β is not identified, so I set it equal to 0.9.

I now present a more detailed discussion of the identification of each of the model's parameters: the parameters of the earnings equation, employment probabilities, and flow utilities, and the population parameters of unobserved productivity and preferences.

The vector of earnings parameters is identified from variation in earnings across locations, time, and levels of work experience. These parameters are consistently estimated using OLS as part of the EM algorithm. Within-person serial correlation in earnings identifies the unobserved type

[35] See Kennan and Walker (2011) and Arcidiacono et al. (2016) who also make use of both of these assumptions.

[36] See Magnac and Thesmar (2002) for a formal discussion of identification in dynamic discrete choice models.

coefficient. That is, those who have earnings that are persistently higher than would be predicted by their observable characteristics are labeled as "high" earnings types.

Job destruction and job offer probabilities ($\pi_{i\ell rt}$) are identified non-parametrically from transitions between employment states. This is possible because of the assumption that employment happens according to a lottery with pre-commitment. As with earnings, within-person serial correlation in employment identifies the unobserved type dummy.

4.1.1 Identification of utility parameters

The search cost parameter γ_2 is identified from the share of labor force participants that are unemployed. This share is in turn identified through the employment probabilities $\pi_{i\ell rt}$, which are identified from transitions between employment states. It is not possible to identify separate values of γ_2 for each type, because γ_2 depends on the employment probabilities, which themselves are type-specific.

Parameters in the moving cost equation (Δ_ℓ) are identified from variation between the observed characteristics of movers and the probability of moving, along with the assumption that moving costs are symmetric (i.e. a move from Boston to Chicago has the same cost as a move from Chicago to Boston). Specifically, variation in the origin and destination of moves identifies the distance parameters, and variation in the ages of movers identifies the age parameters. Variation in the previous employment status of movers identifies the employment parameters. As with earnings and employment probabilities, serial correlation in moves—compared to what would be predicted given observables—identifies the type dummy coefficient in the moving cost equation.

Switching cost parameters Ξ_j are identified from the individual's observed characteristics and serial correlation in labor force entry and exit. However, these switching costs cannot be separately identified from home production benefits and local amenities because the set of all three is linearly dependent. Thus, identification is only possible under either a symmetry assumption or by taking the difference in the costs. I choose the latter because there is no theoretical reason for why the entry and exit costs should be symmetric. The results presented hereafter represent the cost of labor force entry net of labor force exit, because the utility of labor force participation is the baseline alternative.

Identification of the expected earnings coefficient in the flow utility of employment requires variation in earnings that is excluded from the flow utility equation. I make use of two such exclusion restrictions. The first is variation in work experience, and the second is variation in mean earnings across time periods within each location. These exclusion restrictions allow me to distinguish between expected earnings and amenities.

I now discuss the implications of the exclusion restrictions for identification of the expected earnings coefficient. The work experience exclusion restriction hinges on the assumption that work experience is uncorrelated with time-varying amenities. This is a reasonable assumption because most time-varying amenities in a given location (e.g. crime, air pollution) are indeed uncorrelated with the level of work experience in that location. The time dummy exclusion restriction implies that amenities are fixed over time. In the short run, this is likely to hold, as amenities that vary over time within a location (e.g. crime or economic development) are much less volatile than local labor market conditions.[37] Given that this model focuses on a 10-year period, this is a reasonable exclusion restriction.

Finally, the population proportion of each unobserved type, denoted $\overline{\pi}_r$, is directly identified from the frequency of individuals with each particular type label.

4.2 Estimation of earnings, employment, and utility parameters

As discussed in the preceding section, the earnings equation parameters are consistently estimated by OLS as part of the EM algorithm. The employment probabilities are estimated using transitions between employment and unemployment. While they could in principle be estimated non-parametrically, I estimate them as predicted probabilities from a pair of logit regressions in order to allow for a common effect of moving from a different location in each employment state, and to avoid imprecision due to sparse cells. The logit models are estimated separately by previous employment status. Right hand side variables include current location fixed effects, lagged unemployment rate in current location, a quadratic in experience, a mover dummy, and an unobserved type dummy. For futher details, see Appendix A.1.

Estimation of the flow utility parameters follows Hotz and Miller (1993) and Arcidiacono and Miller (2011) by making use of two separate simplification tools that are closely related: (i) conditional choice probabilities (CCPs); and (ii) finite dependence. CCPs make use of a function mapping future value terms from the individual's dynamic programming problem into the probability of making a discrete choice. These probabilities are called CCPs and are found in the data. Finite dependence allows the researcher to formulate the future value terms into a finite sequence of future payoffs. Together, the two strategies yield substantial computational savings by eliminating the need to solve the dynamic programming problem using backwards recursion. I additionally use insights from Arcidiacono and Miller (Forthcoming) to handle finite dependence when choice outcomes are stochastic, as is the case with my search model.

Estimation reduces to a multi-stage static problem, which can be estimated using an alternative-specific conditional logit with an adjustment term that captures the future value associated with each alternative. I provide full details on how CCPs and finite dependence are used in the estimation algorithm in Appendix A.2. Full details on the specification of the likelihood function and computational steps for estimation of the flow utility parameters are included in Appendix A.3.

[37]For instance, over the 11-year period from 2000-2010, annual crime rates in Washington, D.C. for a variety of crimes remained mostly stable. See http://www.dccrimepolicy.org/Briefs/images/Volatility-Brief-3-10-11_1.pdf for more details.

5 Empirical Results

This section discusses in detail each of the empirical results of the model. The model estimates confirm the stylized facts presented in Section 2.2. Even after accounting for non-random moves and person-specific unobserved heterogeneity, the employment penalty for employed movers persists. Additionally, when using higher-quality data that allow me to identify local labor markets, I find that the shielding effect of employment is even stronger than initially presented. The combination of these two effects turns out to be an important factor in explaining the differential migration response to labor market shocks by employment status. Another major finding is that, even though non-employed workers are more mobile, they actually face higher moving costs. This latter finding highlights the usefulness of the structural model in distinguishing between search frictions and moving costs.

Results from the model estimation are found in Tables 4 through 6. These results comprise estimates from each component of the model. Additional tables of results are included in the Appendix (see Tables A6 through A8). I also discuss post-estimation results contained in Table 7.

5.1 Model estimates

I begin by discussing the estimated employment probabilities and their evolution over the business cycle, as reported in Table 4. This table lists the estimates of separate binary logits that predict the probability of being employed conditional on previous employment status. I present the estimates under two different scenarios: no unobserved heterogeneity, and $R = 2$ unobserved types. The results confirm the findings in Section 2.2: The employed are more shielded from local economic downturns, but employed movers face an employment penalty in the new location. An additional finding from the structural model is that there is comparative advantage in job-finding based on employment status. That is, type 1 workers are much more likely than type 2 workers to stay employed, but are much worse at being hired.

Table 5 presents estimates of the structural earnings equation. The main takeaway is that type 1 workers are more productive when employed, as they earn a 67% wage premium over type 2 workers.

Table 6 presents the flow utility parameter estimates. The highlight here is that both employed and type 1 workers have lower moving costs. The former finding is somewhat surprising, given that employed workers are less mobile than unemployed workers. Combined with the earnings equation estimates, the latter finding is consistent with other studies that have found that cognitive and non-cognitive abilities are correlated with migration—that is, those who are more productive in the labor market also have lower moving costs (Bütikofer and Peri, 2017). In other aspects, the flow utility parameter estimates conform to economic theory and the previous literature.[38]

[38]For example, the positive coefficient on expected log earnings indicates that cross-location differences in earnings

5.2 Moving costs, amenity values, and location-specific parameters

With estimates on the elasticity of income and moving costs in Table 6, I can calculate the monetary value of moving costs and amenity values. The expected earnings parameter converts utility to money and can thus be used to express the structural parameter estimates in monetary units. It is also important to note that these moving cost estimates represent the moving costs faced by the *average* individual, not the *marginal* individual (i.e. not the person who is just indifferent between staying and moving). Table 7 displays sample moving costs by previous employment status and unobserved type. The fixed cost of moving ranges from -\$126,505 for an employed type 1 person to -\$162,030 for an unemployed type 2 person. The moving cost evaluated at the average person's characteristics and for the average move path range from -\$375,274 to -\$417,830. These figures are similar in magnitude to those reported in Kennan and Walker (2011), Bishop (2012), and Bartik (2018).[39] Importantly, the monetary value of the moving cost reflects psychological costs of moving (e.g. acclimating to a new location or leaving behind friends and family) in addition to monetary costs (e.g. costs to procure a moving truck or close on a mortgage). For more details on the computation of the moving costs, see Appendix A.4.

As discussed in Kennan and Walker (2011) and Bishop (2012), the moving cost represents the cost faced to the average individual if he were forced to move to an arbitrary location in an arbitrary time period. Allowing the individual to choose the best available location would substantially reduce this cost. Kennan and Walker also show that the moving cost for actual moves is much lower than for the average mover. A similar line of logic applies to the current model, but I omit the discussion here for expositional purposes.

In addition to moving costs, I analyze amenity values and find them to be economically significant, but not nearly as large as moving costs. This is not surprising given that amenities are consumed in each period, whereas moving costs are incurred only once for each move. To put the amenity values in perspective, a one-standard-deviation increase in local amenities is valued at just under \$4,000, while moving from the bottom to the top of the amenity distribution would be worth about \$17,000 per period. Such a move would compensate for the average fixed cost of moving in just over nine years. Preferences for birth state proximity—which are also received each period—are in between these two values at \$10,000.

Due to disclosure rules, I am not allowed to present location-specific parameter estimates. In order to gauge the plausibility of my parameter estimates, I regress various location-specific

matter to migration decisions, as found by Kennan and Walker (2011) and others. Individuals value locations that are in their state of birth, more than for locations in their Census division of birth (Diamond, 2016). Fixed costs of moving are substantial, but also steeply increase with distance and age (Bishop, 2012).

[39] Other papers estimating moving costs include Bayer, Keohane, and Timmins (2009); Morten and Oliveira (2016); Diamond (2016) and Shenoy (2016). Exact values of moving costs depend on assumptions of the underlying model, including whether the model is static or dynamic.

parameters from the model on a set of location characteristics.[40] The results are reported in the Appendix in Table A8.

One might wonder why the estimated moving costs are so large. The primary reason is that there is a weak relationship between expected earnings and observed moves. Expected earnings is just one of a list of many potential reasons for moving, and while the elasticity of earnings is positive (as predicted by economic theory), the moves observed in the data on average are not strongly related to increases in expected earnings.

6 Model Fit and Counterfactual Simulations

In this section, I verify that the structural model fits the data well, and then discuss the results obtained from counterfactual simulations of the model. The results of these simulations comprise the main empirical conclusions of the paper.

6.1 Model Fit

Tables 8 and 9 show migration probabilities and employment transitions in the model and in the data. Panel (a) of Table 8 shows how migration varies by previous employment status and calendar time. The model matches these differences well. Migration probabilities over previous employment crossed with age and distance are shown in panels (b) and (c) of this table. The model and data also match up well along these dimensions.

Table 9 compares employment transitions in the data and model conditional on migrating or staying. Panel (a) compares employment transition rates conditional on migrating. These match up very closely with the exception of remaining out of the labor force for non-participants. This is likely due to the fact that, in the data, there are relatively few non-participant movers who remained out of the labor force after moving. Panel (b) compares these transitions conditional on staying in a location. Again, the data and model match up well.

6.2 Counterfactual Simulations

Now that I have established that the model fits the data well, I discuss counterfactual simulations of the model that make up my primary empirical results. Specifically, I simulate five different counterfactual policies, separately by employment status and unobserved type for six different locations. The six locations correspond to three pairs of artificial cities, each possessing characteristics at specific points in the respective distribution of city characteristics for local amenities, earnings, and employment probabilities. For example, I calculate the difference in the probability

[40] Schmutz and Sidibé (2017) conduct a similar analysis using the parameters of their model with 200 locations.

of out-migration with and without the policy in a city at the 75th percentile of the amenities distribution versus a city at the 25th percentile of the amenities distribution. All other city characteristics are identical across the two cities.[41] This process is repeated for earnings and employment probabilities. Constructing the counterfactuals in this way allows me to hold fixed city characteristics, which turn out to be important determinants of migration behavior.

The four main findings of the counterfactual exercises are as follows:

1. The employment penalty for employed movers makes the employed less likely to move in response to negative local economic shocks.

2. Correlated labor market shocks dampen out-migration from impacted areas. This result is strongest for unemployed workers facing a local unemployment shock, where out-migration is anywhere from 8-12 times larger for this group in response to a local unemployment shock. This result is driven in part by the fact that non-employed movers and stayers are equally likely to be employed.

3. City characteristics affect migration response. For example, out-migration response to a moving subsidy is strongest in a low-amenity city. Similarly, out-migration response to an adverse unemployment shock is strongest in a high-unemployment city.

4. Unemployed workers who have been given a moving subsidy favor locations closer to their place of birth, locations with higher amenities, and locations with higher employment certainty—they do not favor locations with higher earnings once accounting for these other attributes. Workers favor employment certainty over higher earnings because off-the-job search is costly.

Each counterfactual policy is purely temporary. That is, each policy is in effect for only one calendar year. This is because analyzing policy innovations with a longer time horizon would require conditional choice probabilities (CCPs) derived from the new policy, and generating these counterfactual CCPs would require fully solving the model via backwards recursion. This is infeasible given the size of the model's state space. By restricting the policy horizon to a short term, I can calculate the appropriate CCPs without solving the model by making use of the finite dependence structure outlined in Section A.2.2. I emphasize that, because of the autocorrelated structure of some components of the model, the *effect* of each counterfactual policy may not be temporary. However, each policy innovation is purely transitory.

I focus my discussion on the impact of the policies on out-migration of young unemployed workers who were not born in the impacted location, because these are the workers who are most responsive to such policies. The results of the simulations are reported in Figures 2 (type 1 workers) and 3 (type 2 workers).

[41] In all cases, the artificial city is set to be in the same geographical location. The exact geographical location of the artificial city makes little difference to the final results.

Predicted out-migration rates for each city and employment group are listed just above the horizontal axis. These migration rates are heterogeneous across cities, employment status, and type. In particular, predicted out-migration is highest for the city with the lowest amenities, and for those who are type 1. In contrast, out-migration is smallest for the city with high amenities and those who are type 2. These results point to the importance of considering amenities when forming policy that is intended to affect migration behavior. The baseline migration rates also differ markedly by employment status. The rate of out-migration for unemployed workers is 1.2 to 1.5 times the rate for employed workers, consistent with the stylized fact presented in Figure 1. However, there is substantial heterogeneity across cities and unobserved worker types.

In order to weigh the mechanisms influencing migration, I compare adverse earnings and unemployment shocks that are either completely localized, or correlated across all locations (but originating in the current location). The first four bars in each panel of Figures 2 and 3 respectively report the simulated response to independent and correlated adverse shocks to earnings and employment in each location. The earnings shock corresponds to the 70th percentile of the cross-location distribution in earnings AR(1) shock deviations. The unemployment shock corresponds to the 2008–2009 increase in the local unemployment rate for the average location in the data.

The key result from this exercise is the difference in behavior between employed and unemployed workers when faced with unemployment shocks.[42] This difference stems from the difference in job offer arrival rates that these groups face when moving. Employed workers are more likely to stay in their current location when faced with either a localized or correlated shock, whereas the opposite is true for unemployed workers. Comparing the third and fourth bars of each panel of the table shows that unemployed workers would have been anywhere from 8 to 12 times more likely to out-migrate in response to a localized unemployment shock.

There is also substantial heterogeneity in migration responses to local economic shocks with respect to unobserved type. For example, employed type 1 workers are more likely than type 2 workers to stay in response to either of the four shocks. This is because type 2 workers have a comparative advantage in job finding, and employed movers are much less likely to find a job upon arrival in a new location. The comparative advantage of type 2 workers also explains why unemployed type 2 workers are more likely than unemployed type 1 workers to leave in response to an unemployment shock. This is true even though type 2 individuals have larger moving costs.

I also study the role of moving costs in explaining migration behavior. The last bar of each panel of Figures 2 and 3 reports the simulated impact of a moving cost subsidy equal to 10% of the fixed cost of moving for employed type 1 workers. For all cities and employment statuses, out-migration rates increase. The increase in migration rates under a moving subsidy is largest in areas with low amenities and low job-finding rates. Unemployed workers are more responsive

[42] These findings contrast with those of Gardner and Hendrickson (2017), who show that labor markets with higher variance in unemployment rates have lower out-migration rates, all else equal. My approach underscores that moving incentives differ drastically by employment status.

to each subsidy, even though they have larger moving costs. The reason ties back to the fact that unemployed movers do not face an employment penalty. Finally, type 2 workers are more likely to stay because their moving costs are higher.

Finally, I emphasize that the above migration responses are upper bounds on the population-level average response. This is because the counterfactual simulations have focused on the group of people that is most mobile to begin with. Repeating the exercise for older workers, or workers born in the origin location, would result in much lower responses because these other groups are more tied to their current location.

6.3 Where do moving subsidy recipients relocate?

I now use the model to analyze what happens when workers are given a moving subsidy. A similar policy has been proposed by Moretti (2012), and was introduced in the US House of Representatives as a bill sponsored by Rep. Tony Cardenas (H.R. 2755, 2015).[43] Because the moving subsidy is not tied to a particular destination location, I analyze where workers who accept the subsidy would choose to relocate. Table 10 shows estimates of a regression of the net migration probability (multiplied by 100) on a vector of location characteristics (amenities, earnings, job offer probability, and birth location proximity). The origin location is excluded from this regression. These regressions are run separately for each of the three origin cities and each of the two unobserved worker types.

Table 10 predicts that, all else equal, migrants will choose locations that are near their birth location, close to their origin location, and that have higher employment certainty and higher amenities.[44] This finding is consistent with Monras (2018), who finds that out-migration from heavily shocked areas was constant during the Great Recession, but that in-migration into heavily shocked areas decreased markedly. Interestingly, migrants value employment certainty much more than earnings, regardless of the origin city. The reason for this is that unemployment risk enters the flow utility of labor force participation twice (multiplied by the earnings and multiplied by the search cost and home production benefit), but earnings enters the flow utility once (multiplied by the unemployment risk). Hence, workers are more sensitive to unemployment uncertainty because it is costly to find a job when unemployed.

In summary, I emphasize that the response to each counterfactual shock is heterogeneous across

[43] These proposed policies focus on the fact that unemployment is an externality that should be internalized through subsidized migration. The model presented here does not include such externalities. However, such proposals abstract from preferences for amenities, which my model shows are important determinants of migration.

[44] These results echo findings by Deryugina, Kawano, and Levitt (2018) who conclude that individuals displaced by Hurricane Katrina migrated to areas offering better economic opportunities, resulting in immediate wage gains. However, they also conclude that these wage gains were likely nominal (i.e. there were no utility gains), because housing prices for these people also increased by the same amount. My results show that migrants tend to choose places with higher amenities and that are closer to family (as proxied by birth location). This suggests that there can, in fact, be utility gains for displaced workers provided these workers are not native to the shocked location.

locations. Specifically, cities with low amenities and low job-finding rates see the largest out-migration response to adverse shocks and favorable subsidies. Furthermore, areas with higher amenities and higher job-finding rates are the prime destinations for out-migrants. This heterogeneity underscores the difficulty in implementing migration policy that would have the intended consequence of inducing migration from high-unemployment areas to low-unemployment areas, because workers value amenities about the same as employment certainty and much more strongly than earnings.

6.4 Implications for Monopsony

High moving costs and labor market frictions keep workers in their current location, even if workers are offered a moving subsidy. Workers who do take the moving subsidy are more likely to choose locations with high market- or non-market amenities, or locations with higher likelihood of employment. These findings together imply that spatial labor supply elasticity is low. Thus, low labor supply elasticity leads to monopsony power (Manning, 2003). While the model here can only ascertain the reduced-form labor supply curve facing an individual firm, the findings can be taken to imply that firms have monopsony power due to workers' high moving costs and the employed's lower likelihood of finding a job after a move.[45] High moving costs may also induce firms to reduce their geographical recruitment radius.

7 Conclusion

This paper studies the role of labor market frictions and moving costs in determining migration and labor supply decisions of American workers. To quantify the impacts of these mechanisms, I develop and tractably estimate a rich dynamic structural model that incorporates search frictions. I estimate the model using two panels of the Survey of Income and Program Participation (SIPP).

I find that moving costs are substantial and that employed movers see a steep reduction in job offers after moving. Moving costs are large because migration decisions observed in the data are only loosely related to cross-location earnings differences. That is, workers have sizable preferences for market and non-market amenities which weaken the role of earnings on the migration decision. Labor market frictions are also important. Even though the employed have lower moving costs, counterfactual simulations of the model show that they are *less* likely to move in response to a shock to the local unemployment rate. This is because they face a steep decline in employment likelihood if they move locations.

[45] In the Burdett and Mortensen (1998) model, firms have monopsony power due to the presence of search frictions. In my model, the composition of the local labor force is affected by moving costs and labor force switching costs. Thus, a firm's labor supply curve depends on the composition of the local labor force in addition to the job separation and offer arrival rates.

I use the model to simulate the effect of a moving subsidy offered to both employed and unemployed workers. Owing to large moving costs, the subsidy has low take-up rates ($\approx 3\% - 5\%$). The unemployed are more likely to take the subsidy. On average, subsidy recipients go to locations with higher amenities and higher employment certainty.

Taken together, this paper finds evidence of a low spatial elasticity of labor supply. Workers face costs to moving locations and uncertainty about gaining employment after a move. These factors combine to result in relatively low migration rates. Moving costs, in addition to search frictions, grant employers monopsony power. Employers may also strategically adjust their recruiting behavior as a result of worker moving costs.

References

Adda, Jérôme, Christian Dustmann, and Joseph-Simon Görlach. 2015. “The Dynamics of Return Migration, Human Capital Accumulation, and Wage Assimilation.” Working paper, Bocconi University.

Altuğ, Sumru and Robert A. Miller. 1998. “The Effect of Work Experience on Female Wages and Labour Supply.” *Review of Economic Studies* 65 (1):45–85.

Amior, Michael and Alan Manning. 2018. “The Persistence of Local Joblessness.” *American Economic Review* 108 (7):1942–1970.

Arcidiacono, Peter. 2004. “Ability Sorting and the Returns to College Major.” *Journal of Econometrics* 121 (1):343–375.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom. 2016. “College Attrition and the Dynamics of Information Revelation.” Working Paper 22325, National Bureau of Economic Research.

Arcidiacono, Peter and John B. Jones. 2003. “Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm.” *Econometrica* 71 (3):933–946.

Arcidiacono, Peter and Robert A. Miller. 2011. “Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity.” *Econometrica* 79 (6):1823–1867.

———. Forthcoming. “Nonstationary Dynamic Models with Finite Dependence.” *Quantitative Economics* .

Balasubramanian, Natarajan, Jin Woo Chang, Mariko Sakakibara, Jagadeesh Sivadasan, and Evan Starr. 2018. “Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers.” Working paper, Syracuse University.

Balgova, Maria. 2018. “Why Don’t Less Educated Workers Move? The Role of Job Search in Migration Decisions.” Working paper, Oxford University.

Bartik, Alexander W. 2018. “Moving Costs and Worker Adjustment to Changes in Labor Demand: Evidence from Longitudinal Census Data.” Working paper, University of Illinois at Urbana-Champaign.

Baum-Snow, Nathaniel and Ronni Pavan. 2012. “Understanding the City Size Wage Gap.” *Review of Economic Studies* 79 (1):88–127.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. “Migration and Hedonic Valuation: The Case of Air quality.” *Journal of Environmental Economics and Management* 58 (1):1–14.

Bhaskar, V., Alan Manning, and Ted To. 2002. “Oligopsony and Monopsonistic Competition in Labor Markets.” *Journal of Economic Perspectives* 16 (2):155–174.

Bhaskar, V. and Ted To. 1999. “Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition.” *Economic Journal* 109 (455):190–203.

———. 2003. “Oligopsony and the Distribution of Wages.” *European Economic Review* 47 (2):371–399.

Bishop, Kelly. 2012. “A Dynamic Model of Location Choice and Hedonic Valuation.” Working paper, Washington University in St. Louis.

Blanchard, Olivier and Lawrence F. Katz. 1992. “Regional Evolutions.” *Brookings Papers on Economic Activity* 1992 (1):1–75.

Bleakley, Hoyt and Jeffrey Lin. 2012. “Thick-Market Effects and Churning in the Labor Market: Evidence from US Cities.” *Journal of Urban Economics* 72 (2):87–103.

Burdett, Kenneth and Dale T. Mortensen. 1998. “Wage Differentials, Employer Size, and Unemployment.” *International Economic Review* 39 (2):257–273.

Bütikofer, Aline and Giovanni Peri. 2017. “The Effects of Cognitive and Noncognitive Skills on Migration Decisions.” Working Paper 23877, National Bureau of Economic Research.

Caliendo, Marco, Steffen Künn, and Robert Mahlstedt. 2017. “Mobility Assistance Programmes for Unemployed Workers, Job Search Behaviour and Labour Market Outcomes.” Discussion Paper 11169, IZA.

Coate, Patrick. 2013. “Parental Influence on Labor Market Outcomes and Location Decisions of Young Workers.” Working paper, Duke Univeristy.

Davies, Paul S., Michael J. Greenwood, and Haizheng Li. 2001. “A Conditional Logit Approach to U.S. State-to-State Migration.” *Journal of Regional Science* 41 (2):337–360.

Deryugina, Tatyana, Laura Kawano, and Steven Levitt. 2018. “The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns.” *American Economic Journal: Applied Economics* 10 (2).

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

Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. 2018. “Monopsony in Online Labor Markets.” Working paper, University of Massachusetts Amherst.

DuMond, J. Michael, Barry T. Hirsch, and David A. Macpherson. 1999. “Wage Differentials across Labor Markets and Workers: Does Cost of Living Matter?” *Economic Inquiry* 37 (4):577–598.

Foote, Andrew, Michel Grosz, and Ann Stevens. 2019. “Locate Your Nearest Exit: Mass Layoffs and Local Labor Market Response.” *ILR Review* 72 (1):101–126.

Fox, Jeremy T. 2010. “Estimating the Employer Switching Costs and Wage Responses of Forward-Looking Engineers.” *Journal of Labor Economics* 28 (2):357–412.

Gardner, John and Joshua R. Hendrickson. 2017. “If I Leave Here Tomorrow: An Option View of Migration When Labor Market Quality Declines.” Working paper, University of Mississippi.

Gayle, Wayne-Roy. 2018. “CCP Estimation of Dynamic Discrete/Continuous Choice Models with Generalized Finite Dependence and Correlated Unobserved Heterogeneity.” Working paper, University of Massachusetts.

Gemici, Ahu. 2011. “Family Migration and Labor Market Outcomes.” Working paper, Royal Holloway, University of London.

Glaeser, Edward and David Maré. 2001. “Cities and Skills.” *Journal of Labor Economics* 19 (2):316–342.

Gould, Eric. 2007. “Cities, Workers and Wages: A Structural Analysis of the Urban Wage Premium.” *Review of Economic Studies* 74 (2):477–506.

Hall, Robert E. and Andreas I. Mueller. 2018. “Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values.” *Journal of Political Economy* 126 (4):1594–1637.

Harris, John R. and Michael P. Todaro. 1970. “Migration, Unemployment and Development: a Two-Sector Analysis.” *American Economic Review* 60 (1):126–142.

Hirsch, Boris, Elke J. Jahn, Alan Manning, and Michael Oberfichtner. 2018. “The Urban Wage Premium in Imperfect Labour Markets.” Working paper, Leuphana University of Lüneburg.

Hotelling, Harold. 1929. “Stability in Competition.” *Economic Journal* 39 (153):41–57.

Hotz, V. Joseph and Robert A. Miller. 1993. “Conditional Choice Probabilities and the Estimation of Dynamic Models.” *Review of Economic Studies* 60 (3):497–529.

H.R. 2755. 2015. “American Worker Mobility Act of 2015.” 114th Congress.

Humphries, John Eric. 2018. “The Causes and Consequences of Self-Employment over the Life Cycle.” Working paper, Yale University.

Huttunen, Kristiina, Jarle Møen, and Kjell G. Salvanes. 2018. “Job Loss and Regional Mobility.” *Journal of Labor Economics* 36 (2):479–509.

Ilut, Cosmin L. and Martin Schneider. 2014. “Ambiguous Business Cycles.” *American Economic Review* 104 (8):2368–2399.

James, Jonathan. 2012. “Learning and Occupational Sorting.” *Federal Reserve Bank of Cleveland Working Paper 12-25* .

Jeon, Jihye. 2017. “Learning and Investment under Demand Uncertainty in Container Shipping.” Working paper, Boston University.

Kaplan, Greg and Sam Schulhofer-Wohl. 2017. “Understanding the Long-Run Decline in Interstate Migration.” *International Economic Review* 58 (1):57–94.

Keane, Michael P. and Kenneth I. Wolpin. 1994. “The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence.” *Review of Economics and Statistics* 76 (4):648–672.

Kennan, John and James R. Walker. 2011. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica* 79 (1):211–251.

Krueger, Alan B and Orley Ashenfelter. 2018. “Theory and Evidence on Employer Collusion in the Franchise Sector.” Working Paper 24831, National Bureau of Economic Research.

Magnac, Thierry and David Thesmar. 2002. “Identifying Dynamic Discrete Decision Processes.” *Econometrica* 70 (2):801–816.

Mangum, Kyle. 2015. “Cities and Labor Market Dynamics.” Working Paper 2015-2-3, Georgia State University.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Marinescu, Ioana and Roland Rathelot. 2018. “Mismatch Unemployment and the Geography of Job Search.” *American Economic Journal: Macroeconomics* 10 (3):42–70.

McFadden, Daniel. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontiers in Econometrics*, edited by Paul Zarembka. Academic Press, New York, 105–142.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. “Internal Migration in the United States.” *Journal of Economic Perspectives* 25 (3):173–196.

Molloy, Raven S. and Abigail Wozniak. 2011. “Labor Reallocation over the Business Cycle: New Evidence from Internal Migration.” *Journal of Labor Economics* 29 (4):697–739.

Monras, Joan. 2018. “Economic Shocks and Internal Migration.” Discussion Paper 12977, Centre for Economic Policy Research.

Moretti, Enrico. 2012. *The New Geography of Jobs*. New York: Houghton Mifflin Harcourt.

Morten, Melanie and Jaqueline Oliveira. 2016. “Paving the Way to Development: Costly Migration and Labor Market Integration.” Working Paper 22158, National Bureau of Economic Research.

Naidu, Suresh and Eric A. Posner. 2018. “Labor Monopsony and the Limits of the Law.” Working paper, Columbia University.

Notowidigdo, Matthew J. 2011. “The Incidence of Local Labor Demand Shocks.” Working Paper 17167, National Bureau of Economic Research.

Oswald, Florian. 2018. “The Effect of Homeownership on the Option Value of Regional Migration.” Working paper, Sciences Po.

Ransom, Michael R. 1993. “Seniority and Monopsony in the Academic Labor Market.” *American Economic Review* 83 (1):221–233. URL https://www.jstor.org/stable/2117505.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]*. Minneapolis: University of Minnesota.

Rust, John. 1987. “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher.” *Econometrica* 55 (5):999–1033.

Salop, Steven C. 1979. “Monopolistic Competition with Outside Goods.” *Bell Journal of Economics* 10 (1):141–156.

Sastry, Narayan and Jesse Gregory. 2014. “The Location of Displaced New Orleans Residents in the Year After Hurricane Katrina.” *Demography* 51 (3):753–775.

Schlottmann, Alan M. and Henry W. Herzog, Jr. 1981. “Employment Status and the Decision to Migrate.” *Review of Economics and Statistics* 63 (4):590–598.

Schmutz, Benoît and Modibo Sidibé. 2017. “Frictional Labor Mobility.” ERID Working Paper 234, Duke University.

Shenoy, Ashish. 2016. “Migration Decisions and Persistent Earnings Differentials: Evidence from Thailand.” Working paper, University of California at Davis.

Sorkin, Isaac. 2018. “Ranking Firms Using Revealed Preference.” *Quarterly Journal of Economics* 133 (3):1331–1393.

Staiger, Douglas O., Joanne Spetz, and Ciaran S. Phibbs. 2010. “Is There Monopsony in the Labor Market? Evidence from a Natural Experiment.” *Journal of Labor Economics* 28 (2):211–236.

Sullivan, Paul and Ted To. 2014. “Search and Nonwage Job Characteristics.” *Journal of Human Resources* 49 (2):472–507.

Taber, Chris and Rune Vejlin. 2016. “Estimation of a Roy/Search/Compensating Differential Model of the Labor Market.” Discussion Paper 9975, IZA.

Topel, Robert H. 1986. “Local Labor Markets.” *Journal of Political Economy* 94 (3):S111–S143.

Webber, Douglas A. 2018. “Employment Adjustment Over the Business Cycle: The Impact of Competition in the Labor Market.” Working paper, Temple University.

Wheeler, Christopher H. 2006. “Cities and the Growth of Wages Among Young Workers: Evidence from the NLSY.” *Journal of Urban Economics* 60 (2):162–184.

Winters, John V. 2009. “Wages and Prices: Are Workers Fully Compensated for Cost of Living Differences?” *Regional Science and Urban Economics* 39 (5):632–643.

Yagan, Danny. 2014. “Moving to Opportunity? Migratory Insurance over the Great Recession.” Working paper, UC Berkeley.

———. Forthcoming. “Employment Hysteresis from the Great Recession.” *Journal of Political Economy* .

Yankow, Jeffrey J. 2006. "Why Do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium." *Journal of Urban Economics* 60 (2):139–161.

Zabel, Jeffrey E. 1998. "An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior." *Journal of Human Resources* 33 (2):479–506.

Figures and Tables

Table 1: Descriptive statistics of the estimation subsample of the SIPP, 2004-2013

Variable	Mean	Std Dev
Log monthly earnings (2000 dollars)[a]	7.96	0.52
Work experience (years)	22.60	9.49
Age (years)	42.29	9.76
Lives in location in birth state	0.74	0.44
Lives in location in birth Census division	0.75	0.43
Number of persons	16,648	
Number of observations	50,415	

Notes: For complete sample selection rules, see Table A1.

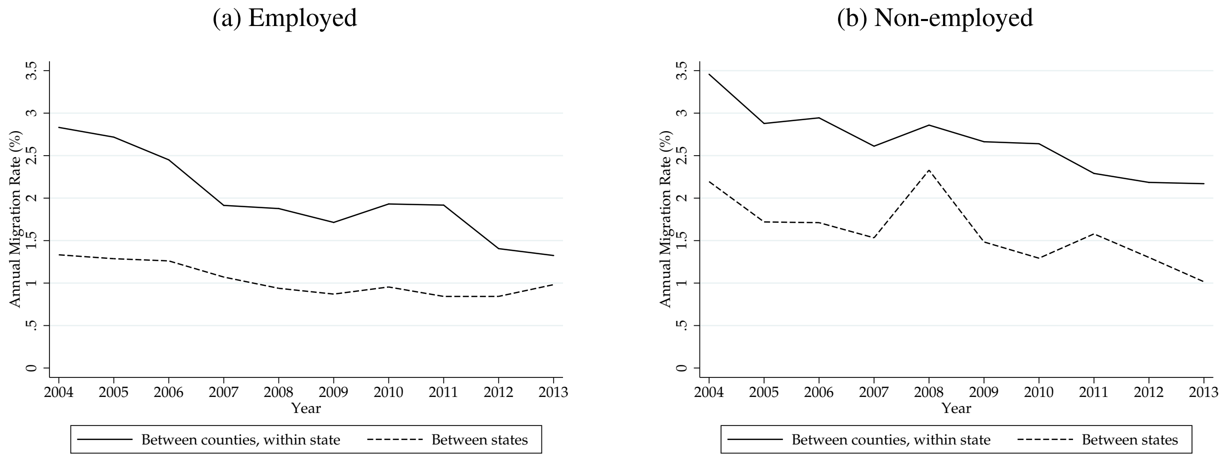
[a] Conditional on being employed full-time with monthly earnings between \$400 and \$22,000. This variable has 29,238 person-year observations. The earnings variable is spatially deflated to account for differences in cost of living according to the procedure outlined in Appendix A.6

Table 2: Migration in the SIPP, 2004-2013

Number of persons (age 18-55)	16,648
Movers	568
Movers (%)	3.41
Moves	653
Moves per mover	1.15
Repeat moves (% of all moves)	13.38
Return moves (% of all moves)	8.98

Note: Moves are defined as changing locations as defined in the model.

Figure 1: Annual migration rates by lagged employment status and migration distance



Source: 2004 and 2008 Panels of the Survey of Income and Program Participation. Figures include all non-college graduates aged 18-55 who have completed their schooling. Employment is defined as full-time employment.

Table 3: Linear probability models of employment, by lagged employment status

	Prev. employed		Prev. non-employed	
Variable	Coeff	Std Err	Coeff	Std Err
Constant	0.7243***	0.0071	0.1976***	0.0059
Experience	0.0123***	0.0005	0.0077***	0.0004
Experience^2/100	-0.0200***	0.0012	-0.0142***	0.0011
Lagged state unempl. rate	-0.0038***	0.0006	-0.0060***	0.0006
Mover dummy	-0.1219***	0.0080	0.0468***	0.0076
Race $\times$ gender dummies	✓		✓	
Observations	83,324		78,057	

Notes: Dependent variable is an indicator for being employed full-time in the current period. Sample includes all non-college graduates aged 18-55 in the 2004 and 2008 panels of the SIPP who have completed their schooling. *** p<0.01; ** p<0.05; * p<0.10.

Table 4: Structural employment probability equation estimates

	1 type				2 types			
	Prev. employed		Prev. non-employed		Prev. employed		Prev. non-employed	
Variable	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
Constant	1.3056***	0.2220	0.2566	0.2237	0.9812***	0.2241	0.5035***	0.2264
Experience	0.0858***	0.0091	0.0359***	0.0086	0.0847***	0.0092	0.0366***	0.0087
Experience^2/100	-0.1228***	0.0208	-0.0285	0.0219	-0.1224***	0.0211	-0.0310	0.0221
Lagged local unempl. rate	-0.0314***	0.0104	-0.0922***	0.0110	-0.0342***	0.0105	-0.0937***	0.0111
Mover dummy	-0.9257***	0.1280	0.1929	0.1557	-1.0483***	0.1300	0.2359	0.1572
Unobserved type 1					0.7958***	0.0400	-0.5751***	0.0429
Location fixed effects	✓		✓		✓		✓	
Observations	30,898		9,949		30,898		9,949	
Persons	12,013		6,087		12,013		6,087	

Notes: Reported numbers are coefficients from logit regressions conditional on previous employment status. *** p<0.01; ** p<0.05; * p<0.10.

Table 5: Structural earnings equation estimates

	1 type		2 types	
Parameter	Coeff	Std Err	Coeff	Std Err
Constant	7.5708***	0.0673	7.2074***	0.0470
Experience	0.0432***	0.0015	0.0411***	0.0010
Experience^2/100	-0.0595***	0.0033	-0.0575***	0.0023
Unobserved type 1			0.6773***	0.0039
Location-time fixed effects	✓		✓	
Persons	11,404		11,404	
Observations	29,238		29,238	

Notes: Reported numbers are coefficients from an OLS log earnings regression conditional on full-time employment and observing earnings. See footnote (a) of Table 1 for complete details on this subsample. *** p<0.01; ** p<0.05; * p<0.10

Table 6: Structural choice equation estimates

Parameter	Symbol	1 type Coeff	1 type Std Err	2 types Coeff	2 types Std Err
Job & location preferences					
Expected log earnings	(γ_0)	0.916**	0.397	1.001**	0.412
Home production benefit	(γ_1)	-0.902	3.477	11.333***	3.453
Search cost	(γ_2)	-1.195***	0.069	-1.008***	0.070
Birth state bonus	(γ_3)	0.207***	0.072	0.210***	0.072
Birth division bonus	(γ_4)	-0.002	0.073	-0.003	0.073
Switching costs					
Fixed cost	$(\theta_{12}-\theta_8)$	0.335**	0.127	0.910***	0.126
Age	$(\theta_{13}-\theta_9)$	-0.095***	0.006	-0.106***	0.006
$\text{Age}^2/100$	$(\theta_{14}-\theta_{10})$	0.109***	0.008	0.121***	0.008
Unobserved type 1	$(\theta_{15}-\theta_{11})$			-0.746***	0.019
Moving costs					
Fixed cost	(θ_0)	-3.148***	0.361	-3.165***	0.362
Distance (1000 miles)	(θ_1)	-2.063***	0.078	-2.066***	0.078
Distance^2	(θ_2)	0.369***	0.025	0.369***	0.025
Age	(θ_3)	-0.094***	0.018	-0.101***	0.018
$\text{Age}^2/100$	(θ_4)	0.056**	0.023	0.063***	0.023
Employed_{t-1}	(θ_5)	0.197*	0.110	0.252**	0.110
Unemployed_{t-1}	(θ_6)	-0.230*	0.128	-0.239*	0.129
Unobserved type 1	(θ_7)			0.256***	0.045
$\Pr(\text{type}=1)$	$(\overline{\pi}_r)$	N/A		0.4926	
Observations		50,415		50,415	
Persons		16,648		16,648	
Discount factor	(β)	0.9		0.9	

Notes: Reported numbers are flow utility parameter estimates from the dynamic choice model described in Sections 3 and 4. Estimates of location-specific amenities (the α_ℓ's) are not reported due to Census Bureau rules regarding disclosure risk. *** p<0.01; ** p<0.05; * p<0.10

Table 7: Sample moving costs and amenity values (2014 dollars)

	Monetary value			
	Type 1		Type 2	
Utility component	Employed	Unemployed	Employed	Unemployed
Moving costs				
Fixed cost of moving	-$126,505	-$149,849	-$138,687	-$162,030
Average mover, 500-mile move	-$375,274	-$403,238	-$389,867	-$417,830
Average mover, NY to LA	-$483,867	-$511,831	-$498,460	-$526,423
Young mover, NY to LA	-$285,720	-$304,068	-$295,295	-$313,644
Std. Dev. of local amenities	$3,807			
Range of local amenities	$17,114			
Birth state bonus	$9,980			

Notes: Values are evaluated at the parameter estimates in Table 6 and at the sample average of monthly earnings (log earnings equal to 7.96). The average mover is age 39, and a young mover is age 25. The great-circle distance from New York to Los Angeles is 2,446 miles. For more details on how these values are calculated, see Section A.4.

Table 8: Model fit: observed vs. predicted migration probabilities

(a) Migration probabilities by calendar time and $t-1$ employment status

	2004-2008		2009-2013		All	
$t-1$ Employment status	Data	Model	Data	Model	Data	Model
Employed	1.30%	1.33%	1.28%	1.24%	1.29%	1.29%
Unemployed	1.21%	1.25%	1.15%	1.10%	1.19%	1.19%
Out of labor force	1.88%	1.73%	1.52%	1.66%	1.69%	1.70%
Overall	1.14%	1.27%	1.38%	1.22%	1.25%	1.25%

(b) Migration probabilities by age and $t-1$ employment status

	Employed		Unemployed		Out of LF		All	
Age range	Data	Model	Data	Model	Data	Model	Data	Model
18-25	2.31%	2.11%	2.54%	3.62%	3.37%	3.31%	2.52%	2.84%
26-35	1.90%	1.65%	2.31%	2.32%	1.97%	2.13%	2.00%	1.86%
36-45	1.00%	1.20%	1.57%	1.42%	1.23%	1.30%	1.13%	1.25%
46-55	0.80%	0.82%	1.09%	0.85%	0.88%	0.84%	0.86%	0.83%

(c) Migration probabilities by distance migrated and $t-1$ employment status

	Employed		Unemployed		Out of LF		All	
Distance (miles)	Data	Model	Data	Model	Data	Model	Data	Model
0-500	0.72%	0.70%	0.68%	0.65%	0.92%	0.94%	0.72%	0.70%
501-1,000	0.31%	0.35%	0.29%	0.33%	0.41%	0.45%	0.31%	0.35%
1,001-1,500	0.13%	0.13%	0.10%	0.12%	0.20%	0.17%	0.13%	0.13%
1,501-2,000	0.07%	0.05%	0.06%	0.05%	0.11%	0.07%	0.07%	0.05%
2,001+	0.06%	0.05%	0.07%	0.04%	0.05%	0.06%	0.06%	0.05%

Notes: All numbers in this table correspond to migration probabilities (multiplied by 100 and expressed as percentages). Data probabilities consist of conditional means of an indicator for migration. Model probabilities consist of conditional means of the predicted probability of leaving the current location.

Table 9: Model fit: employment transitions by migration status

(a) Employment transitions conditional on migrating

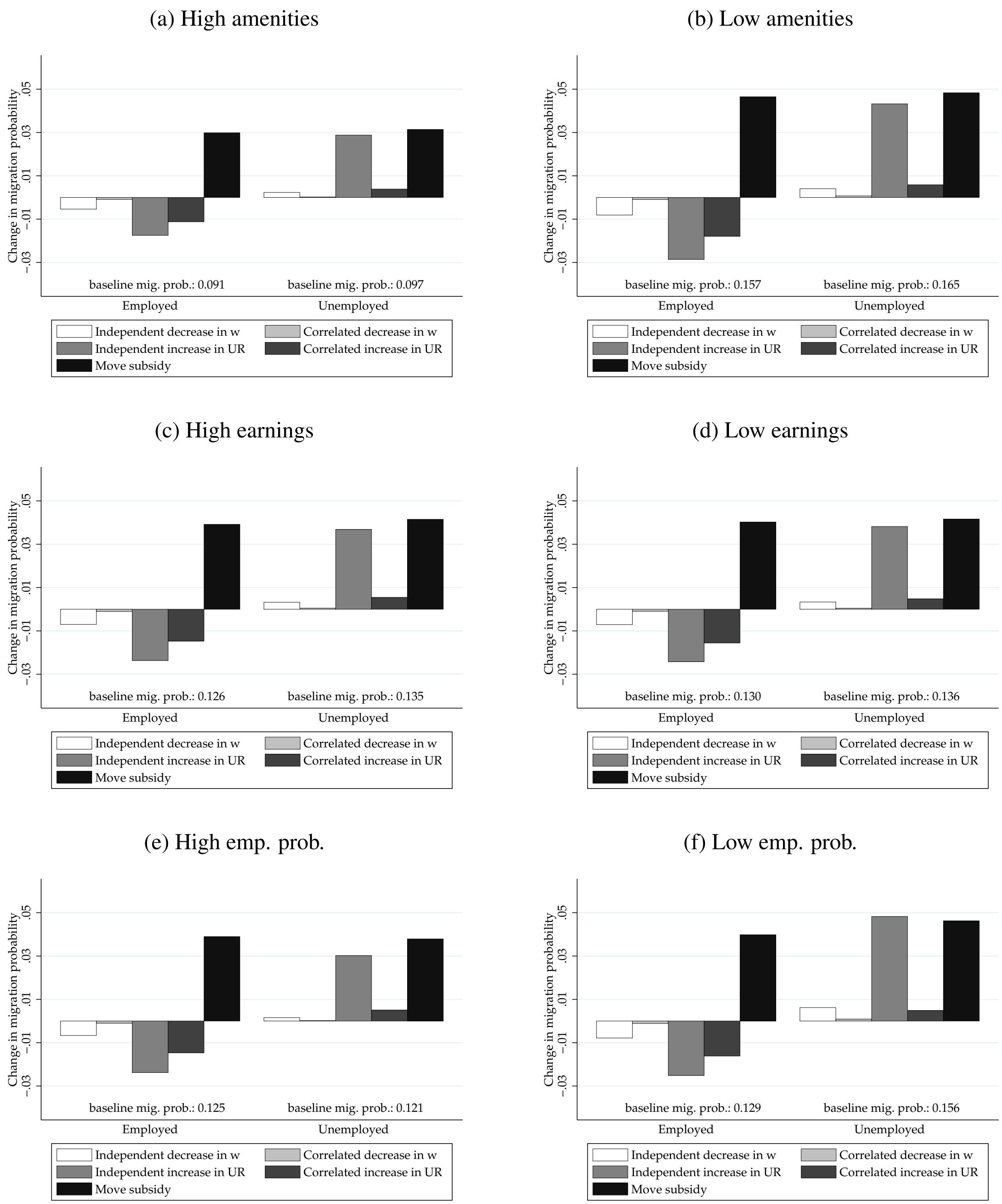
	Period t					
	Data			Model		
Period $t-1$	E	U	N	E	U	N
Employed (E)	70.98%	22.69%	6.33%	71.99%	22.26%	5.75%
Unemployed (U)	41.40%	46.50%	12.10%	45.06%	44.77%	10.17%
Out of labor force (N)	16.52%	17.39%	66.09%	13.88%	12.58%	73.54%

(b) Employment transitions conditional on staying

	Period t					
	Data			Model		
Period $t-1$	E	U	N	E	U	N
Employed (E)	86.92%	9.86%	3.23%	86.45%	9.83%	3.71%
Unemployed (U)	36.33%	49.75%	13.93%	38.09%	49.60%	12.31%
Out of labor force (N)	10.81%	10.41%	78.78%	10.56%	12.22%	77.21%

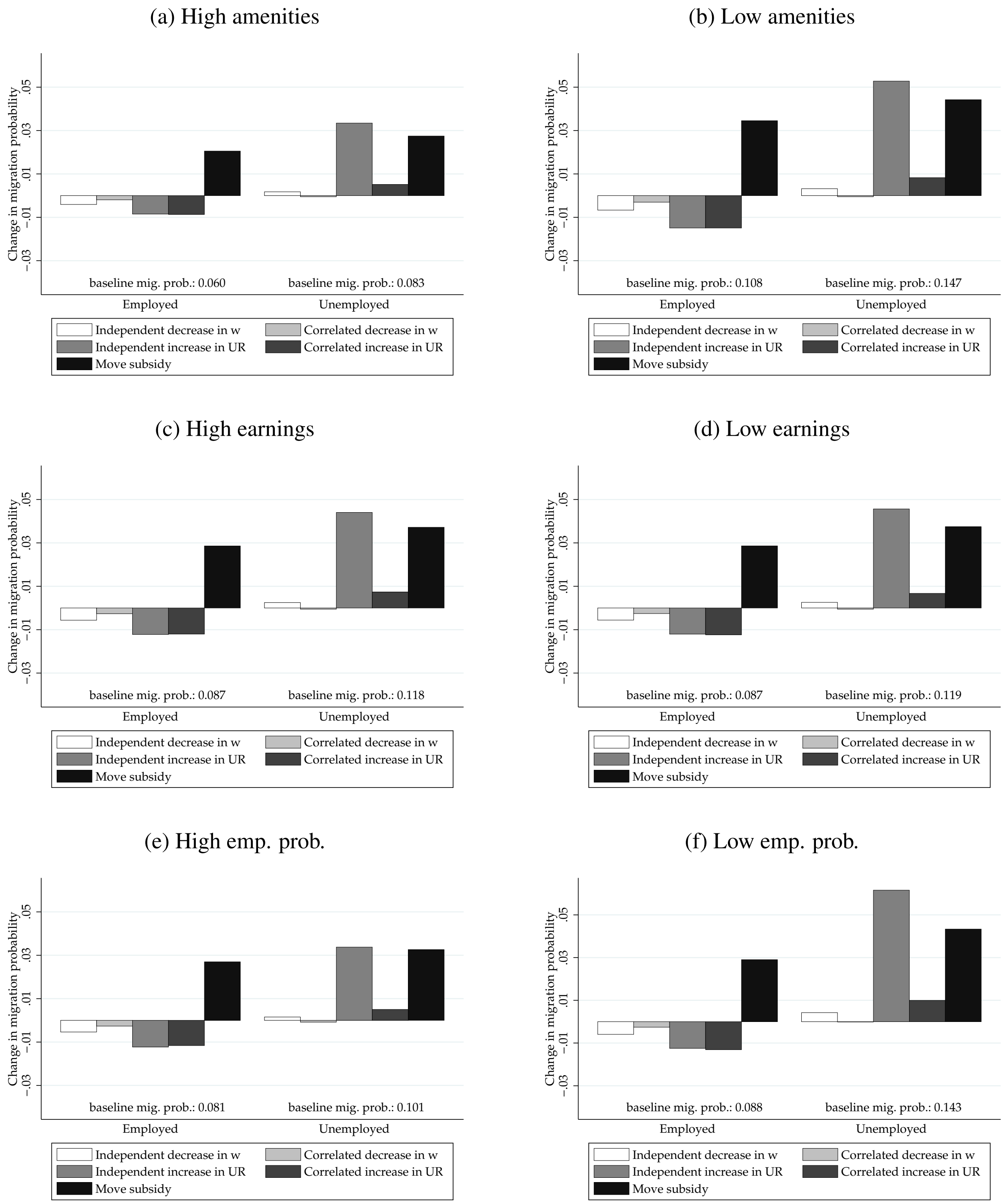
Notes: All numbers in this table correspond to employment transition probabilities (multiplied by 100 and expressed as percentages). Data probabilities consist of conditional means of employment transition by migration status. Model probabilities consist of conditional means (by employment status) of the predicted conditional probability of making an employment transition (conditional on leaving or staying).

Figure 2: Counterfactual change in migration for Type 1 individuals, by origin city and prior employment status



Notes: Each panel corresponds to a different origin city. Bar heights refer to the change in the out-migration rate from the specified location in response to the listed counterfactual. All figures are for 25-year-olds who were not born in the origin location. "high" refers to a location in the 75th percentile of the given distribution; "low" refers to the 25th percentile. All characteristics not set to "high" or "low" are set to the median. The earnings shock ($\downarrow w$) corresponds to the 70th percentile of the cross-location distribution in earnings AR(1) shock deviations. The unemployment shock corresponds to the jump from 2008 to 2009 for the average location in the data. To focus the results, each candidate location has median AR(1) parameters for both earnings and employment. Birth location is held fixed in all counterfactuals. Individual characteristics are set to the average for all 25-year-olds, conditional on employment status.

Figure 3: Counterfactual change in migration for Type 2 individuals, by origin city and prior employment status



Notes: See notes to Figure 2.

Table 10: Characteristics of destination location given moving cost subsidy to unemployed workers in various origin cities, year 2007

(a) Amenities

	Type 1 workers				Type 2 workers			
	high amenity city		low amenity city		high amenity city		low amenity city	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	0.3405	0.1790	0.4587	0.2750	0.3000	0.1643	0.4205	0.2643
amenities	0.0284*	0.0086	0.0436*	0.0132	0.0261*	0.0079	0.0419*	0.0127
earnings (conditional on working)	-0.0005	0.0067	-0.0008	0.0104	-0.0006	0.0062	-0.0009	0.0100
employment probability	0.0229*	0.0078	0.0352*	0.0120	0.0221*	0.0072	0.0357*	0.0116
ln(distance)	-0.0456*	0.0143	-0.0701*	0.0220	-0.0395*	0.0131	-0.0635*	0.0211
state of birth	0.2724*	0.0277	0.4195*	0.0426	0.2396*	0.0254	0.3859*	0.0410
region of birth	0.0149	0.0225	0.0230	0.0347	0.0129	0.0207	0.0208	0.0333

(b) Earnings level

	Type 1 workers				Type 2 workers			
	high earnings city		low earnings city		high earnings city		low earnings city	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	0.4162	0.2352	0.4151	0.2378	0.3745	0.2211	0.3749	0.2246
amenities	0.0372*	0.0112	0.0375*	0.0114	0.0349*	0.0106	0.0354*	0.0107
earnings (conditional on working)	-0.0007	0.0089	-0.0006	0.0090	-0.0008	0.0084	-0.0008	0.0085
employment probability	0.0298*	0.0103	0.0307*	0.0104	0.0296*	0.0097	0.0306*	0.0098
ln(distance)	-0.0603*	0.0188	-0.0603*	0.0190	-0.0535*	0.0176	-0.0537*	0.0179
state of birth	0.3611*	0.0364	0.3607*	0.0368	0.3255*	0.0343	0.3261*	0.0348
region of birth	0.0196	0.0296	0.0199	0.0300	0.0173	0.0279	0.0177	0.0283

(c) Employment probability level

	Type 1 workers				Type 2 workers			
	high emp. prob. city		low emp. prob. city		high emp. prob. city		low emp. prob. city	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	0.3823	0.2129	0.4558	0.2672	0.3304	0.1917	0.4290	0.2639
amenities	0.0337*	0.0102	0.0422*	0.0128	0.0303*	0.0092	0.0415*	0.0126
earnings (conditional on working)	-0.0006	0.0080	-0.0007	0.0101	-0.0007	0.0072	-0.0008	0.0100
employment probability	0.0263*	0.0093	0.0357*	0.0117	0.0251*	0.0084	0.0370*	0.0115
ln(distance)	-0.0553*	0.0170	-0.0666*	0.0213	-0.0471*	0.0153	-0.0618*	0.0211
state of birth	0.3313*	0.0330	0.3972*	0.0414	0.2868*	0.0297	0.3740*	0.0409
region of birth	0.0176	0.0268	0.0227	0.0337	0.0149	0.0242	0.0211	0.0333

Notes: Dependent variable is predicted migration rate to location ℓ (in percentage points). Covariates are locational characteristics of the candidate destination locations. The amenities, earnings, and employment probability variables are each standardized to have mean-zero, unit variance. All locations (including synthetic locations) are included in the regression. Controls also included for local earnings drift, earnings volatility, unemployment drift, unemployment persistence, and unemployment volatility. * p<0.05

A Online Appendix

A.1 Details for estimation of employment probabilities

This subsection details the functional form and specification for the logit models from which the employment probabilities are derived. Each of the estimated employment probabilities can be summarized in words. $\hat{\lambda}_{i\ell rt}$ measures the probability that an individual was not employed in the previous period and is employed in the current period in the same location. Likewise, $\hat{\delta}_{i\ell rt}$ measures job destruction, i.e. the probability that an individual is not employed in the current period but was employed in the current location in the previous period. Finally, $\hat{\lambda}^{e}_{i\ell rt}$ and $\hat{\lambda}^{u}_{i\ell rt}$ measure the probability of employment in a new location given previous employment status e or u.

The logit equation for $\hat{\delta}_{i\ell rt}$ and $\hat{\lambda}^{e}_{i\ell rt}$ is estimated conditional on choosing to supply labor in the current period and having been employed in the previous period.

$$\Pr\left[y_{it} = (e,\ell) | \, d_{it} = (1,\ell), y_{it-1} = (e,\cdot)\right] = \frac{\exp\left(\Theta_1\right)}{1+\exp\left(\Theta_1\right)} \tag{A.1}$$

where

$$\Theta_1 = \mu^{e}_{1\ell} + \mu^{e}_{2}\text{UR}_{\ell t-1} + \mu^{e}_{3} G\left(x_{it}\right) + \mu^{e}_{4}\tau_{ir} + \mu^{e}_{5} 1\left\{y_{it-1} = \left(e,\ell'\right)\right\}$$

and where $\text{UR}_{\ell t-1}$ is the lagged unemployment rate in location ℓ and $G(\cdot)$ is a quadratic polynomial. The excluded category is $1\{y_{it-1} = (e,\ell)\}$. A similar regression can be estimated conditional on non-employment in the previous period.

$$\Pr\left[y_{it} = (e,\ell) | \, d_{it} = (1,\ell), y_{it-1} = (\{u,n\},\cdot)\right] = \frac{\exp\left(\Theta_2\right)}{1+\exp\left(\Theta_2\right)} \tag{A.2}$$

where

$$\Theta_2 = \mu^{u}_{1\ell} + \mu^{u}_{2}\text{UR}_{\ell t-1} + \mu^{u}_{3} G\left(x_{it}\right) + \mu^{u}_{4}\tau_{ir} + \mu^{u}_{5} 1\left\{y_{it-1} = \left(\{u,n\},\ell'\right)\right\}$$

The excluded category in (A.2) is $1\{y_{it-1} = (\{u,n\},\ell)\}$. Non-stationarity in the $\pi_{i\ell rt}$'s is accounted for through the evolution of the local unemployment rate.

Conditioning (A.1) and (A.2) on labor force participants is crucial, because in the model individuals do not exogenously transition between labor force participation states—they only exogenously transition between employment states (conditional on participating in the labor force).

It then follows that

$$
\begin{aligned}
1-\hat{\delta}_{i\ell rt} &= \frac{\exp\left(\hat{\mu}^e_{1\ell}+\hat{\mu}^e_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^e_3 G\left(x_{it}+\hat{\mu}^e_4\tau_{ir}\right)\right)}{1+\exp\left(\hat{\mu}^e_{1\ell}+\hat{\mu}^e_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^e_3 G\left(x_{it}+\hat{\mu}^e_4\tau_{ir}\right)\right)} && \text{(A.3)}\\
\hat{\lambda}^e_{i\ell rt} &= \frac{\exp\left(\hat{\mu}^e_{1\ell}+\hat{\mu}^e_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^e_3 G\left(x_{it}\right)+\hat{\mu}^e_4\tau_{ir}+\hat{\mu}^e_5\right)}{1+\exp\left(\hat{\mu}^e_{1\ell}+\hat{\mu}^e_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^e_3 G\left(x_{it}\right)+\hat{\mu}^e_4\tau_{ir}+\hat{\mu}^e_5\right)} && \text{(A.4)}\\
\hat{\lambda}_{i\ell rt} &= \frac{\exp\left(\hat{\mu}^u_{1\ell}+\hat{\mu}^u_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^u_3 G\left(x_{it}\right)+\hat{\mu}^u_4\tau_{ir}\right)}{1+\exp\left(\hat{\mu}^u_{1\ell}+\hat{\mu}^u_2\mathrm{UR}_{\ell t-1}\right)+\hat{\mu}^u_3 G\left(x_{it}+\hat{\mu}^u_4\tau_{ir}\right)} && \text{(A.5)}\\
\hat{\lambda}^u_{i\ell rt} &= \frac{\exp\left(\hat{\mu}^u_{1\ell}+\hat{\mu}^u_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^u_3 G\left(x_{it}\right)+\hat{\mu}^u_4\tau_{ir}+\hat{\mu}^u_5\right)}{1+\exp\left(\hat{\mu}^u_{1\ell}+\hat{\mu}^u_2\mathrm{UR}_{\ell t-1}+\hat{\mu}^u_3 G\left(x_{it}\right)+\hat{\mu}^u_4\tau_{ir}+\hat{\mu}^u_5\right)} && \text{(A.6)}
\end{aligned}
$$

A.2 Details for estimation of the flow utility parameters

This subsection provides additional details for how the flow utility parameters are estimated. The main idea is that the recursive Bellman equation can be reduced to a static problem by making use of conditional choice probabilities (CCPs) and the property of finite dependence. Throughout this section, I suppress the individual subscript i and unobserved type subscript r for expositional purposes.

A.2.1 Conditional choice probabilities

I first describe how CCPs are employed. Recall equations (3.17) and (3.18), which show that, by definition, the value function $V_{t+1}(Z_{t+1})$ is equivalent to the $\mathbb{E}$max of the conditional value functions in period $t+1$ plus the ε_{t+1}'s.

When the ε's are assumed to be Type I extreme value, equation (3.18) simplifies to

$$
v_{j\ell t}(Z_t)=u_{j\ell t}(Z_t)+\beta\int\ln\left(\sum_k\sum_m\exp\left(v_{kmt+1}(Z_{t+1})\right)\right)dF\left(Z_{t+1}|Z_t\right)+\beta\overline{\gamma} \tag{A.7}
$$

where $\overline{\gamma}$ is Euler's constant, the mean of a standard Type I extreme value distribution (McFadden, 1974; Rust, 1987). Thus, the $\mathbb{E}$max is the natural log of the sum of the exponentiated conditional value functions, plus Euler's constant.[46]

I will now show how (A.7) can be manipulated to admit CCPs. First, multiply and divide by the exponentiated conditional value function associated with a given choice alternative (e.g. (j', ℓ')), $\exp\left(v_{j'\ell' t+1}(Z_{t+1})\right)$, to get

$$\int V_{t+1}(Z_{t+1})\, dF(Z_{t+1}|Z_t) = \int \ln\left(\frac{\exp\left(v_{j'\ell' t+1}(Z_{t+1})\right)}{\exp\left(v_{j'\ell' t+1}(Z_{t+1})\right)} \times \sum_k \sum_m \exp\left(v_{kmt+1}(Z_{t+1})\right)\right) dF(Z_{t+1}|Z_t) + \overline{\gamma} \tag{A.8}$$

$$= \int \left[v_{j'\ell' t+1}(Z_{t+1}) + \ln\left(\frac{\sum_k \sum_m \exp\left(v_{kmt+1}(Z_{t+1})\right)}{\exp\left(v_{j'\ell' t+1}(Z_{t+1})\right)}\right)\right] dF(Z_{t+1}|Z_t) + \overline{\gamma} \tag{A.9}$$

$$= \int \left[v_{j'\ell' t+1}(Z_{t+1}) - \ln p_{j'\ell' t+1}(Z_{t+1})\right] dF(Z_{t+1}|Z_t) + \overline{\gamma} \tag{A.10}$$

Comparing (3.17) with (A.10) shows that, for any choice alternative (j', ℓ'), the future value function is equal to the conditional value function $v_{j'\ell' t+1}$ minus the log probability of choosing (j', ℓ'). This log probability is the conditional choice probability, and can in principle be recovered non-parametrically from the data. The CCP method pares down the number of future-period conditional value functions from $2L$ to 1.

While it is helpful that the number of conditional value functions has decreased, the value function as currently expressed still has a recursive structure. In mathematical terms, $v_{j'\ell' t+1}(Z_{t+1})$ in (A.10) is a function of V_{t+2}, which is a function of V_{t+3}, etc. In order to eliminate this recursive structure and the need to use backward recursion to solve the model, I make use of the property of finite dependence.

[46] This follows from the fact that the Type I extreme value distribution has a closed-form CDF. The mean of the distribution is Euler's constant. If the ε_t's were assumed to be normally distributed, the $\mathbb{E}$max term would not have a closed form, for the same reason that the Normal CDF does not have a closed form. The CCP method works for any distribution of ε, but requires numerical integration or simulation methods for distributions that are outside of the Generalized Extreme Value family.

A.2.2 Finite dependence

Finite dependence is based on the fact that in discrete choice models only differences in utility (or, in dynamic models, differences in the present value of utility) matter in estimation, e.g. $v_{j'\ell't} - v_{0\ell t}$. Hence, it is possible to express the value function for choosing (j', ℓ') in period t in terms of a sequence of decisions up to N periods ahead, then create a corresponding sequence of decisions for choosing the base alternative $(0, \ell)$ in period t such that after N periods the value functions are the same and can cancel out. The key insight is that this sequence of decisions need not be optimal.[47]

In the case where the choice outcomes correspond to the choice alternatives, the following sequences could be used for all (j', ℓ') to create a cancellation in period $t+3$:

- $v_{j'\ell't}$ path: choose $d_t = (j', \ell')$; $d_{t+1} = (0, \ell')$; $d_{t+2} = (0, \ell)$
- $v_{0\ell t}$ path: choose $d_t = (0, \ell)$; $d_{t+1} = (j', \ell)$; $d_{t+2} = (0, \ell)$

where ℓ is the location in period $t-1$. In both cases, the states in period $t+3$ are one additional year of work experience, three additional years of age, and previous decision equal to non-participation in location ℓ. The value function $V_{t+3}(Z_{t+3})$ is thus the same for both and vanishes when the standard utility normalization is applied.

In the case where labor market outcomes are stochastic, however, inducing the cancellation of the future value terms is not as straightforward. To illustrate how the setup proceeds in this case, recall equation (A.10), rewritten below in conserved notation:

$$\mathbb{E}_t V_{t+1}(Z_{t+1}) = \mathbb{E}_t \left[v_{j'\ell't+1}(Z_{t+1}) - \ln\left(p_{j'\ell't+1}(Z_{t+1})\right)\right] + \overline{\gamma}$$

The key idea is that this equality holds for a weighted sum of $v_{j'\ell't+1}$'s such that the weights

[47] For other studies using finite dependence to aid estimation, see Altuğ and Miller, 1998; Arcidiacono and Miller, 2011; Bishop, 2012; Coate, 2013; Arcidiacono, Aucejo, Maurel, and Ransom, 2016; Arcidiacono and Miller, Forthcoming; Gayle, 2018; and Humphries, 2018.

add up to unity:

$$\begin{aligned}
\mathbb{E}_t V_{t+1}(Z_{t+1}) &= \mathbb{E}_t\left[v_{j'\ell' t+1}(Z_{t+1}) - \ln\left(p_{j'\ell' t+1}(Z_{t+1})\right)\right] + \overline{\gamma} \\
&= \sum_{(k,m)\in\mathscr{J}} \omega_{(k,m)} \left\{\mathbb{E}_t\left[v_{kmt+1}(Z_{t+1}) - \ln\left(p_{kmt+1}(Z_{t+1})\right)\right] + \overline{\gamma}\right\} \\
\text{s.t.} &\sum_{(k,m)\in\mathscr{J}} \omega_{(k,m)} = 1
\end{aligned} \tag{A.11}$$

In the application below, the $\omega_{(k,m)}$'s are functions of the current and future employment probabilities.

Figure A3 shows how the finite dependence structure works in the case of stochastic choice outcomes. It depicts the choice sequences for $v_{j'\ell' t}$ and $v_{0\ell t}$ conditional on the previous choice outcome y_{t-1}. Because of the random nature of employment outcomes, the individual must take expectations over all possible outcomes. Thus, each period of labor force participation induces two outcomes, which are depicted in tree form in Figure A3 (recall that $\pi_{\ell t} = 0$ for the non-participation decision). Decisions are depicted by boxes, and outcomes are depicted by nodes. Probabilities are written next to edges connecting the nodes.

The top branches of each sub-tree in the diagram have the same state variables in $t+3$. However, because the individual must take expectations over the future employment outcomes, cancellation of these terms is not possible except for the case of degenerate employment probabilities or the case where $\pi_{\ell' t} = \pi_{\ell t+1}$. This equality does not hold in general.

In order to induce the cancellation, I make use of the insights provided by equation (A.11). The diagram for this case is provided in Figure A4. The difference is that now the $t+1$ decision in the expression for $v_{0\ell t}$ is a weighted sum of $d_{t+1} = (1,\ell)$ and $d_{t+1} = (0,\ell)$. The ω's are pushed through to the $t+3$ states as with the other probabilities in the tree.

Cancellation is possible by solving for the ω's that make the top branches of each tree equal:

$$\pi_{\ell' t} V_{t+3} = \omega_{(1,\ell)} \pi_{\ell t+1} V_{t+3} \tag{A.12}$$

Solving (A.12) for ω gives

$$\omega_{(1,\ell)} = \frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \tag{A.13}$$

A similar solution strategy can be used for the bottom branches of each tree, with the same value of ω being true for both cases.

Putting everything together, the final equation for the differenced conditional value function expression is then (suppressing i and r subscripts and assuming $j' = 1$):

$$\begin{aligned}
v_{j'\ell' t} - v_{0\ell t} = {} & \pi_{\ell' t} u^{e}_{\ell' t}(Z_t) + (1-\pi_{\ell' t}) u^{u}_{\ell' t}(Z_t) - u^{n}_{\ell t}(Z_t) \\
& + \beta \left[\pi_{\ell' t} u^{n}_{\ell t+1}\left(Z^{1}_{t+1}\right) - \pi_{\ell' t} \ln p_{0\ell' t+1}\left(Z^{1}_{t+1}\right) \right. \\
& + (1-\pi_{\ell' t}) u^{n}_{\ell t+1}\left(Z^{2}_{t+1}\right) - (1-\pi_{\ell' t}) \ln p_{0\ell' t+1}\left(Z^{2}_{t+1}\right) \\
& - \pi_{\ell' t} u^{e}_{\ell t+1}\left(Z^{3}_{t+1}\right) \\
& - \left(\frac{\pi_{\ell' t}(1-\pi_{\ell t+1})}{\pi_{\ell t+1}} \right) u^{u}_{\ell t+1}\left(Z^{3}_{t+1}\right) - \left(1 - \frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \right) u^{n}_{\ell t+1}\left(Z^{3}_{t+1}\right) \\
& \left. + \left(\frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \right) \ln p_{1\ell t+1}\left(Z^{3}_{t+1}\right) + \left(1 - \frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \right) \ln p_{0\ell t+1}\left(Z^{3}_{t+1}\right) \right] \\
& + \beta^2 \left[\pi_{\ell' t} u^{n}_{\ell t+2}\left(Z^{4}_{t+2}\right) - \pi_{\ell' t} \ln p_{0\ell t+2}\left(Z^{4}_{t+2}\right) \right. \\
& + (1-\pi_{\ell' t}) u^{n}_{\ell t+2}\left(Z^{5}_{t+2}\right) - (1-\pi_{\ell' t}) \ln p_{0\ell t+2}\left(Z^{5}_{t+2}\right) \\
& - \pi_{\ell' t} u^{n}_{\ell t+2}\left(Z^{6}_{t+2}\right) + \pi_{\ell' t} \ln p_{0\ell t+2}\left(Z^{6}_{t+2}\right) \\
& - \left(\frac{\pi_{\ell' t}(1-\pi_{\ell t+1})}{\pi_{\ell t+1}} \right) u^{n}_{\ell t+2}\left(Z^{7}_{t+2}\right) + \left(\frac{\pi_{\ell' t}(1-\pi_{\ell t+1})}{\pi_{\ell t+1}} \right) \ln p_{0\ell t+2}\left(Z^{7}_{t+2}\right) \\
& \left. - \left(1 - \frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \right) u^{n}_{\ell t+2}\left(Z^{8}_{t+2}\right) + \left(1 - \frac{\pi_{\ell' t}}{\pi_{\ell t+1}} \right) \ln p_{0\ell t+2}\left(Z^{8}_{t+2}\right) \right]
\end{aligned} \tag{A.14}$$

where the integrals over the future state variables have been suppressed for notational simplicity. Superscripts on the state variables Z denote different sets of states.[48]

Equation (A.14) is a complex formula that includes employment probabilities, flow utility parameters, and log CCPs. However, it is a linear function of all structural parameters which greatly simplifies the estimation. Most importantly, there is no need to use backward recursion in the

[48] Additionally, the state dependence of the employment probabilities $\pi_{\ell' t}$ and $\pi_{\ell t+1}$ is also suppressed for simplicity. $\pi_{\ell' t}$ is always evaluated at Z_t while $\pi_{\ell t+1}$ is always evaluated at Z^{3}_{t+1}.

estimation procedure.

For non-employment alternatives, the finite dependence formula written in equation (A.14) is much simpler (because $pi_{\ell t} = 0$ when $j = 0$ for all ℓ and t):

$$\begin{aligned}
v_{j'\ell' t}(Z_t) - v_{0\ell t}(Z_t) = & u^n_{\ell' t}(Z_t) - u^n_{\ell t}(Z_t) + \\
& \beta \left[\left(u^n_{\ell' t+1}\left(\left\{0, \ell'\right\}, x_t\right) - \ln p_{0\ell' t+1}\left(\left\{0, \ell'\right\}, x_t\right)\right) \right. \\
& \left. - u^n_{\ell t+1}\left(\left\{0, \ell\right\}, x_t\right) + \ln p_{0\ell t+1}\left(\left\{0, \ell\right\}, x_t\right)\right] + \\
& \beta^2 \left[u^n_{\ell t+2}\left(\left\{0, \ell'\right\}, x_t\right) - \ln p_{0\ell t+2}\left(\left\{0, \ell'\right\}, x_t\right) \right. \\
& \left. - u^n_{\ell t+2}\left(\left\{0, \ell\right\}, x_t\right) + \ln p_{0\ell t+2}\left(\left\{0, \ell\right\}, x_t\right)\right]
\end{aligned} \tag{A.15}$$

Figures A3 and A4 have illustrated how finite dependence can be used even in models where choice outcomes are not included in the choice set. This method can be used in a variety of other discrete choice applications where stochastic choice outcomes might not be aligned with deterministic choices.

Using CCPs and finite dependence, the optimization problem has been reduced from a backward recursion problem to a simple multi-stage static estimation problem with an adjustment term comprised of CCPs, current and future flow utilities, and employment probabilities, resulting in impressive computational gains that make possible the estimation of the model.

A.2.3 Integrating out local labor market shocks

When making decisions about the future, agents need to form expectations over the evolution of the labor market conditions in each location. This is outlined in equations (3.2) and (3.4). However, the evolution of these labor market conditions also enters the future value term associated with each alternative. Because this future value term is non-linear, the future labor market shocks need to be integrated out of the value function. Furthermore, because the shock in *each* location enters the choice probability associated with any given location, the dimension of this integral is on the order of double the number of locations.[49] With many locations, the only way to compute the integral is using Monte Carlo techniques.

[49] L of the $2L$ dimensions correspond to the earnings AR(1) shocks ζ_t and the other L dimensions correspond to the unemployment rate AR(1) shocks ξ_t.

The structure of the forecasting problem further underscores the advantages in using CCPs and finite dependence to estimate the flow utility parameters. If estimating the parameters using the full solution (backwards recursion) method, the researcher would be required to evaluate the value function at each realization of the labor market shocks and integrate accordingly. To make the backwards recursion tractable, interpolation methods (Keane and Wolpin, 1994) or simplification of the state space (Kennan and Walker, 2011) would have to be used.

In my case, I can use the finite dependence assumption to exactly rewrite the value function in terms of one- and two-period ahead CCPs and flow payoffs. This only requires integration of the relevant CCPs and employment probabilities, of which there are only nine for each choice alternative (see equation A.14).

Formally, an example of the time-t expectation of one of the log CCPs (choosing alternative $(0,\ell')$) is written as follows:

$$E_t\left[\ln\left(p_{0\ell' t+1}\left(\xi_{t+1},\zeta_{t+1}\right)\right)|Z_t\right]=\int \ln p_{0\ell' t+1}\left(\xi_{t+1},\zeta_{t+1}\right)dF\left(\xi,\zeta\right) \tag{A.16}$$

where ξ_{t+1} and ζ_{t+1} are respectively L-dimensional vectors of earnings and employment shocks in period $t+1$. f is the density of a multivariate normal distribution with mean 0 and covariance Ψ.[50]

The integral in (A.16) is of dimension $2L$ and thus needs to be estimated using Monte Carlo methods. This is done by drawing D draws from the $N(0,\Psi)$ density, plugging them into the CCPs, and averaging over the draws as written below:

$$\int \ln p_{0\ell' t+1}\left(\xi_{t+1},\zeta_{t+1}\right)dF\left(\xi,\zeta\right)\approx\frac{1}{D}\sum_{d=1}^{D}\ln p_{0\ell' t+1}\left(\xi_d,\zeta_d\right) \tag{A.17}$$

where (ξ_d,ζ_d) is the dth draw from f.

For integration of the two-period-ahead CCPs, the variance of f is modified to account for uncertainty in the one-period-ahead outcomes. In this case, the variance matrix of f is

$$\Psi+\Psi\odot RR' \tag{A.18}$$

[50] Ψ is estimated by computing the covariance of the AR(1) residuals for all equations in both AR(1) systems.

where $\odot$ is the element-wise (Hadamard) product and R is a $2L \times 1$ vector of autocorrelation parameters corresponding to earnings or employment forecasting (ρ_1 or $\phi_{1\ell}$). The result in (A.18) comes about because the forecasting shocks are assumed to be normally distributed and independent over time.

A.3 Details on joint likelihood function and the EM algorithm

This subsection outlines the exact functional form of each of the components of the likelihood function. They main idea is that each of the components of the model introduced in Section 3 contains an intercept for unobserved type, which means that estimation is not separable across components. I explain how to use the EM algorithm to obtain parameter estimates of the model. The EM algorithm is a sequential algorithm that allows me to break the dependence of any model component on the rest of the components. Estimation reduces to an iterative procedure where each component of the model can be estimated separately.[51]

The overall log likelihood of the model in the presence of unobserved heterogeneity is

$$L = \sum_i \ln \left(\sum_{r=1}^{R} \overline{\pi}_r \mathscr{L}_{d,i|r} \mathscr{L}_{w,i|r} \mathscr{L}_{\pi,i|r} \right) \tag{A.19}$$

where $\mathscr{L}_{d,i|r}$ denotes the likelihood contribution of the choice parameters conditional on being unobserved type r, $\mathscr{L}_{w,i|r}$ the earnings likelihood contribution, and $\mathscr{L}_{\pi,i|r}$ the employment probability contribution.

$$L = \sum_i \ln \left(\sum_{r=1}^{R} \overline{\pi}_r \prod_j \prod_\ell \prod_t \left\{ \left[P_{ij\ell rt} \Lambda_{i\ell rt} h_{i\ell rt} \right]^{d_{ij\ell t}} \right\}^{1[j=1]} \Big] \left[P_{ij\ell rt}^{d_{ij\ell t}} \right]^{1[j=0]} \right) \tag{A.20}$$

where the following hold:

[51] For further details on the EM algorithm, see Arcidiacono and Jones (2003) and Arcidiacono and Miller (2011). The algorithm is used in Arcidiacono (2004); James (2012); Arcidiacono, Aucejo, Maurel, and Ransom (2016); Humphries (2018); and others.

- P is a multinomial logit function:

$$P_{ij\ell rt} = \frac{\exp(\cdot)}{\sum_{k=1}^{J\times L}\exp(\cdot)} \tag{A.21}$$

- Λ is the likelihood associated with an individual's employment probability, which contributes to the likelihood only when $j = 1$ (i.e. labor is supplied), and which depends on the prior labor market outcome. Λ is a product of two binary logit likelihoods.

$$\Lambda_{i\ell rt} = \left[\pi_{i\ell rt}^{1[y_{i\ell t}=e]}\left(1-\pi_{i\ell rt}\right)^{1[y_{i\ell t}\neq e]}\right]^{1[y_{i\ell t-1}=e]}\left[\pi_{i\ell rt}^{1[y_{i\ell t}=e]}\left(1-\pi_{i\ell rt}\right)^{1[y_{i\ell t}\neq e]}\right]^{1[y_{i\ell t-1}\neq e]} \tag{A.22}$$

where $\pi_{i\ell rt}$ is as defined in (3.3) and where y denotes employment outcome.

- h is an individual's wage likelihood, which contributes to the likelihood only when $j = 1$ and the individual is employed with valid earnings. Under the assumption that the measurement error in wages is normally distributed (see Equation 3.1), the likelihood is given by

$$h_{i\ell rt} = \frac{1}{\sigma_\eta}\phi^h\left(\frac{\ln w_{i\ell t} - \psi_{0\ell t} - \psi_1 G(x_{it}) - \psi_2\tau_{ir}}{\sigma_\eta}\right) \tag{A.23}$$

where $\phi^h(\cdot)$ is the standard normal density function.

Using Bayes' rule, the probability that i is of unobserved type r is defined as follows:

$$q_{ir} = \frac{\overline{\pi}_r \mathscr{L}_{d,i|r}\mathscr{L}_{w,i|r}\mathscr{L}_{\pi,i|r}}{\sum_{r'=1}^{R}\overline{\pi}_{r'}\mathscr{L}_{d,i|r'}\mathscr{L}_{w,i|r'}\mathscr{L}_{\pi,i|r'}} \tag{A.24}$$

It then follows that the population unobserved type probabilities are given by

$$\overline{\pi}_r = \frac{1}{N}\sum_{i}^{N} q_{ir} \tag{A.25}$$

The key insight of (A.24) is that (A.19) can be rewritten as

$$\widetilde{L} = \sum_i \sum_{r=1}^{R}\left\{q_{ir}\ln\mathscr{L}_{d,i|r} + q_{ir}\ln\mathscr{L}_{w,i|r} + q_{ir}\ln\mathscr{L}_{\pi,i|r}\right\} \tag{A.26}$$

where the first order conditions of $\widetilde{L}$ are equal to the first order conditions of L in (A.19). This equivalence at the first order conditions enables implementation of the EM algorithm, which iterates on the first order conditions until convergence. The convergence point is the solution to both (A.19) and (A.26). Most importantly, (A.26) is additively separable in each likelihood, so each likelihood can be estimated separately.

Each likelihood contribution differs by type through the inclusion of a type dummy in each of the separate likelihoods. The algorithm iterates on the following steps:

1. Given estimates of the earnings parameters, employment probabilities, structural flow utilities, update the q_{ir}'s according to equation (A.24) and $\overline{\pi}_r$'s according to (A.25).

2. Estimate the earnings parameters and employment probabilities by weighted OLS and a pair of weighted binary logits, respectively, where the q_{ir}'s are used as weights.

3. Estimate the CCPs using a weighted flexible multinomial logit model, where the q_{ir}'s are used as weights.

4. Calculate the expected future value terms along the finite dependence paths, using the estimated earnings parameters, the employment probabilities, and the CCPs as inputs. Integrate over future local labor market shocks.

5. Estimate the flow utility parameters in the structural choice model. This amounts to estimating a weighted multinomial logit with an offset term containing the future value terms computed in Step 4, where the q_{ir}'s are used as weights.

6. Repeat until convergence

Step 1 is referred to as the Expectation- or E-step because the researcher integrates over the type-conditional likelihoods, and Steps 2–5 constitute the Maximization- or M-step because the researcher maximizes the type-conditional likelihoods.

A.4 Calculating moving costs and amenity values

With expected earnings included in the utility function, I can use the parameter γ_0 in equation (3.9) to convert from units of utility to money and assign a monetary value to the cost of moving and to the amenities. Because earnings do not enter linearly, however, the moving cost needs to be evaluated at some value (e.g. the average earnings in the data). The fixed cost of moving for a previously employed person ($\theta_0+\theta_5$) given in equation (3.13) is calculated as follows:

$$\begin{aligned}
\text{Moving Cost} &= \frac{\theta_0+\theta_5}{\gamma_0}\left(\frac{12\exp\left(\overline{w}\right)\,(2000\,\text{dollars})}{\text{year}}\right) && \text{(A.27)}\\
&= \frac{12\theta_0\exp\left(\overline{w}\right)}{\gamma_0}\,(2000\,\text{dollars}) && \text{(A.28)}\\
&= \frac{12\,(-3.165+0.252)\exp\left(7.96\right)}{1.001}\cdot\frac{1.38\,(2014\,\text{dollars})}{(2000\,\text{dollars})}\\
&= -\$138,687\;(2014\,\text{dollars})
\end{aligned}$$

where I multiply by 12 in order to convert monthly earnings to an annual measure and then multiply by 1.38 to inflate from 2000 to 2014 dollars using the CPI.

The other moving costs (e.g. moving costs at a particular age or between two particular locations) are obtained by calculating the predicted value from the moving cost equation at the relevant characteristics of the mover and origin and destination locations. This predicted value is then substituted for $\theta_0+\theta_5$. A similar method is used to calculate amenity values. However, due to Census Bureau policies surrounding disclosure risk, I am unable to report the estimates of the amenity values α_ℓ.

A.5 Estimation subsample

The estimation subsample is restricted to non-Hispanic white males aged 18-55 who have completed schooling by the time of the first SIPP interview and who do not hold a bachelor's degree. The final estimation subsample comprises 16,648 males each averaging 3.03 annual observations. Earnings are computed as total monthly earnings across all jobs in the interview month. Observations with monthly earnings higher than \$22,000 or lower than \$400 are excluded from earnings estimates. The small percentage of workers with survey data containing missing or imputed monthly earnings are assigned a monthly estimate of annual earnings reported on their W-2 tax form. For complete details on sample selection, see Table A1.

A.6 Population and Prices

I gather locational characteristics from a variety of sources. Using the Missouri Census Data Center's MABLE/Geocorr12 program, I form a crosswalk that maps every county to its Core Based Statistical Area (CBSA) as of 2009. The locational characteristics used in this analysis are population (in 2000) and prices (varying by year). Population is calculated by summing the population of each component county. If the individual does not live in a CBSA, his county population is used instead.

Locational prices come from the ACCRA-COLI data. This data, generously provided by Christopher Timmins, contains quarterly information from 1990-2008 on six different categories of goods (groceries, housing, utilities, medical, transportation and miscellaneous) across a wide range of surveyed locations, both metropolitan and rural. I average prices over quarters and CBSA (since some large CBSAs have multiple price listings) to form an annual price index for each CBSA. For locations that are not included in a particular year, I assign each location to one of five population categories and then impute the price by assigning the average price of all other locations in the same state and population category. If the location still has no price information, I repeat the process but aggregate at the level of census region instead of state.

Multiple studies have found that housing prices listed in ACCRA are not good measures of true housing costs (e.g. DuMond, Hirsch, and Macpherson, 1999; Winters, 2009; Baum-Snow and Pavan, 2012). As a result, I follow Winters (2009) and use quality-adjusted gross rents from the 2005 American Community Survey (ACS) compiled by Ruggles et al. (2010). This consists of regressing log gross rents on a vector of housing characteristics and CBSA fixed effects. The housing price level of a given city is then the predicted average gross rents for that city evaluated at the mean housing characteristics for the entire sample. This price level is then included in place of the ACCRA housing price level when forming the price index in (A.29) below. For more details regarding the specific housing characteristics included in the analysis, see p. 636 of Winters (2009). It is also important to note that the ACS does not include location information for low populated areas. For locations that are not identifiable in the ACS, I use states instead of CBSAs. I exclude houses that are in an identifiable CBSA and repeat the process outlined above, assigning rural housing prices as state fixed effects plus average sample characteristics.

With location-specific prices in hand, I compute the price index according to Baum-Snow and Pavan (2012):

$$\text{INDEX}_j = \prod_g \left(\frac{p_g^j}{p_g^0} \right)^{s_g} \tag{A.29}$$

where g indexes goods in the consumer's basket, p_g^j is the price of good g in location j, and s_g is the share of income on good g. In practice, g corresponds to the six categories of goods included in the ACCRA data: groceries, housing, utilities, transportation, health care and all other goods. I use the income shares provided by ACCRA which were computed using the Consumer Expenditure Survey (CEX).

Once this is accomplished, I temporally deflate the indices using the CPI-U in 2000 and spatially deflate using the population-weighted average location in 2000. I then deflate earnings by dividing monthly earnings by this index.

Equation (A.29) is derived from an indifference relationship for identical workers in location j with utility function U over a vector of goods z (which is allowed to differ in price across locations):

$$\bar{v} = \max_z U(z) + \lambda \left[w_j - \sum_g p_g^j z_g \right]. \tag{A.30}$$

Log-linearizing (A.30) around a mean location (indexed by 0) yields an equilibrium relationship in earnings adjusted for cost of living between locations j and 0, with s_g indicating the share of income spent on good z_g:

$$\ln(w_0) = \ln(w_j) - \sum_g s_g \left[\ln\left(p_g^j\right) - \ln\left(p_g^0\right) \right] \tag{A.31}$$

Taking the exponential of both sides and rearranging terms yields equation (A.29).

A.7 SIPP Sample Design

The SIPP is a two-stage stratified random sample. The sampling frame is the Master Address File (MAF), which is a database maintained by the Census Bureau and used in other surveys such as the American Community Survey (ACS) and Decennial Censuses. The primary sampling unit

(PSU) is one or more bordering counties. Within the PSU, addresses are divided into two groups: those with lower incomes and those with higher incomes. Addresses in the lower-income group are sampled at a higher rate.

Online Appendix Figures and Tables

Table A1: Sample selection

	Remaining Persons	Remaining Person-years
Non-Hispanic, non-college graduate white males in wave 1 of 2004 or 2008 SIPP panel	37,499	124,719
Drop those enrolled in school at any point of survey	30,410	102,740
Drop those outside of 18-55 age range at start of survey	20,153	65,836
Drop those who attrited from survey	20,148	58,320
Drop those missing link to administrative data	16,648	50,415
Final estimation sample	16,648	50,415

Table A2: Distribution of person-years

Years per person	Persons	Person-years
1	3,576	3,576
2	2,117	4,234
3	4,641	13,923
4	2,888	11,552
5	3,426	17,130
Final estimation sample	16,648	50,415

Table A3: Data sources

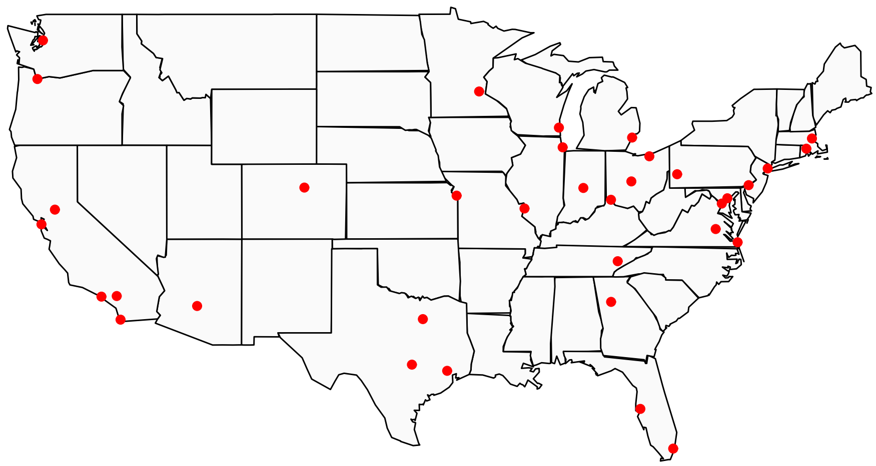
Data	Source	Years
Earnings and location & employment transitions	SIPP, 2004 and 2008 Panels	2004–2013
CBSA population	Census Bureau	2000
County unemployment rate	Bureau of Labor Statistics (BLS)	1990–2013
Local price level	American Chamber of Commerce Researchers Assoc. (ACCRA)	1990–2008

Table A4: Locations in the model

Location	Location
Atlanta , GA	San Diego , CA
Austin , TX	San Francisco , CA
Baltimore , MD	Seattle , WA
Boston , MA	St. Louis , MO
Chicago , IL	Tampa , FL
Cincinnati , OH	Virginia Beach, VA
Cleveland , OH	Washington , DC
Columbus , OH	New England Division small
Dallas , TX	New England Division medium
Denver , CO	Mid Atlantic Division small
Detroit , MI	Mid Atlantic Division medium
Houston , TX	E N Central Division small
Indianapolis , IN	E N Central Division medium
Kansas City , MO	W N Central Division small
Knoxville , TN	W N Central Division medium
Los Angeles , CA	S Atlantic Division small
Miami , FL	S Atlantic Division medium
Milwaukee , WI	E S Central Division small
Minneapolis , MN	E S Central Division medium
New York , NY	W S Central Division small
Philadelphia , PA	W S Central Division medium
Phoenix , AZ	Mountain Division small
Pittsburgh , PA	Mountain Division medium
Portland , OR	Pacific Division small
Providence , RI	Pacific Division medium
Richmond , VA	Alaska
Riverside , CA	Hawaii
Sacramento , CA	

Notes: The cutoff between small and medium is defined by CBSA population of 193,000. This number corresponds to the first tercile of the observed city population distribution in the SIPP. Rural areas (i.e. areas not in any CBSA) are included with small CBSAs.

Figure A1: Map of cities in the model

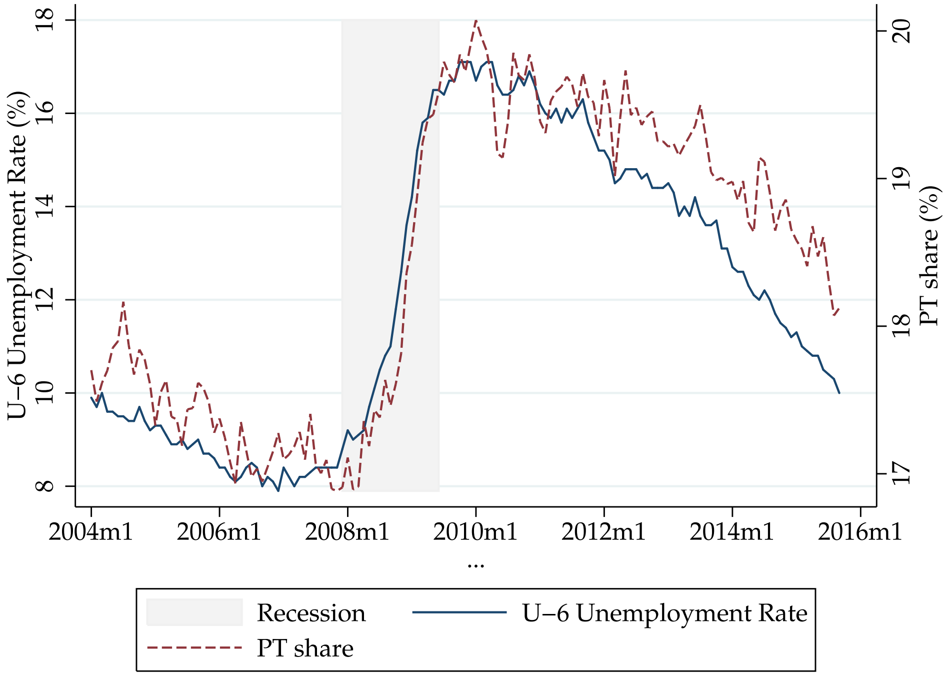


Note: Dots correspond to CBSA centroids of cities that are included in the model.

Table A5: Census divisions and their component states

Census Division Name	States Included
New England	CT, RI, MA, VT, NH, ME
Middle Atlantic	NY, NJ, PA
South Atlantic	DE, MD, DC, VA, WV, NC, SC, GA, FL
East South Central	KY, TN, MS, AL
East North Central	OH, IN, IL, WI, MI
West North Central	MN, IA, MO, KS, NE, SD, ND
West South Central	AR, LA, OK, TX
Mountain	MT, WY, CO, NM, AZ, UT, NV, ID
Pacific	CA, OR, WA, AK, HI

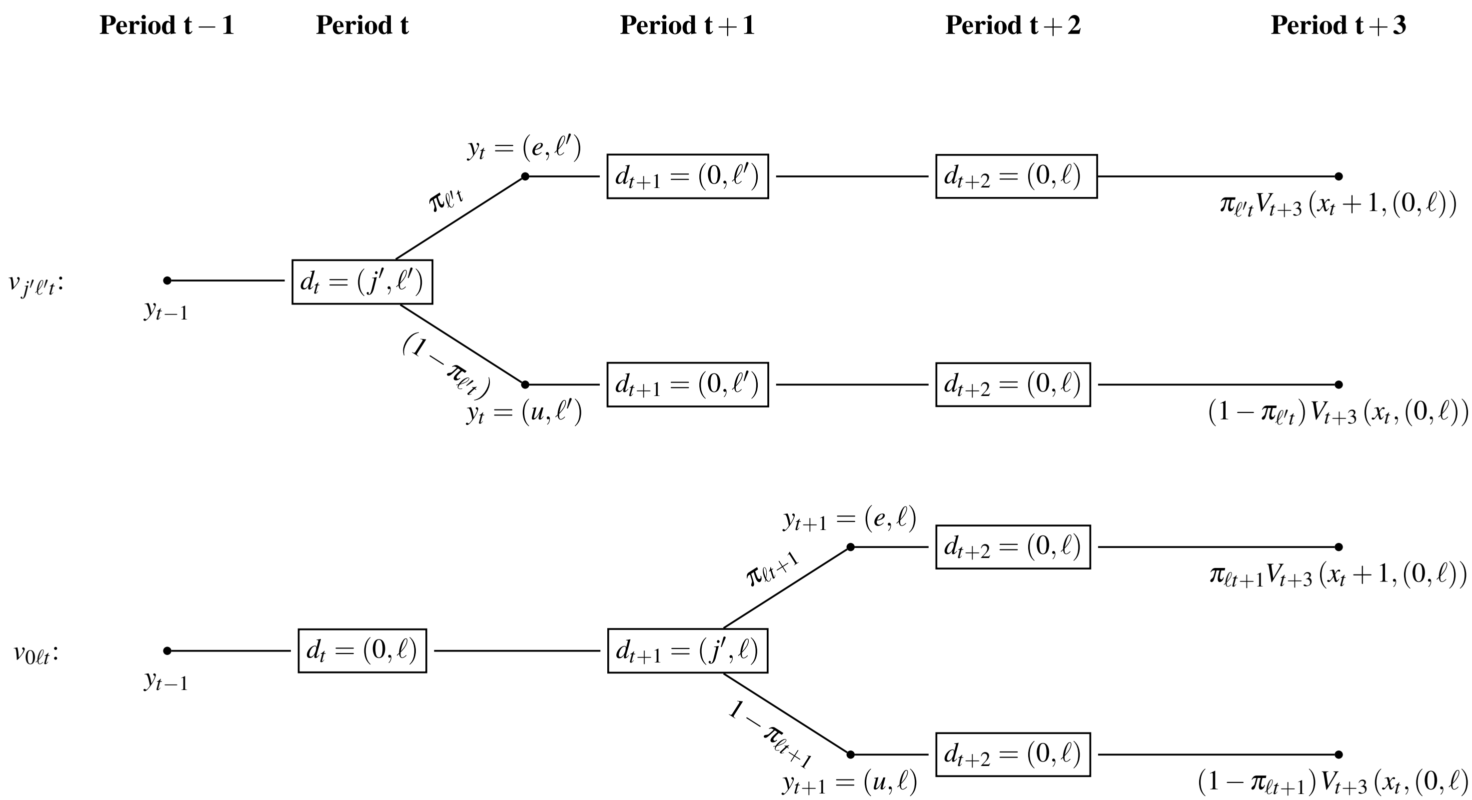
Figure A2: Trends in U-6 Unemployment Rate and Part-time Employment Share



Source: FRED Economic Data, Federal Reserve Bank of St. Louis.

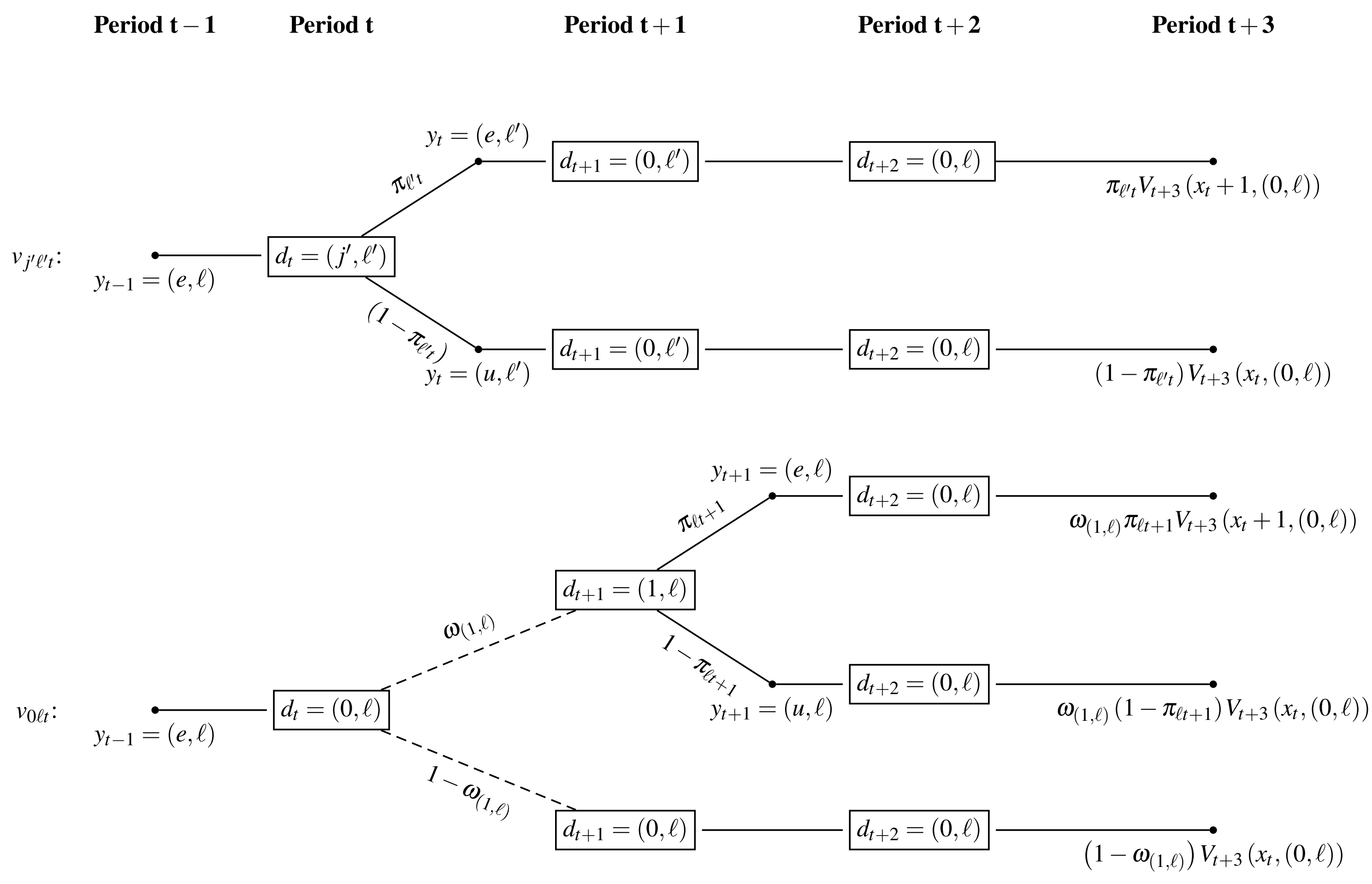
Note: PT share corresponds to the fraction of employed persons who work part-time.

Figure A3: Finite dependence paths conditional on y_{t-1}



Note: This figure depicts the evolution of the state space given the finite dependence paths described in Section A.2.2. Cancellation of the future value terms does not occur unless $\pi_{\ell' t} = \pi_{\ell t+1}$. This equality does not hold in general.

Figure A4: Expanded finite dependence paths conditional on y_{t-1}



Note: This figure depicts the evolution of the state space given the finite dependence paths described in Section A.2.2. Cancellation of the future value terms occurs when $\omega_{(1,\ell)} = \frac{\pi_{\ell' t}}{\pi_{\ell t+1}}$, as described in Equations (A.12) and (A.13).

Table A6: Estimates of unemployment rate forecasting equations

Parameter	Symbol	Mean	Std Dev
Drift	(ϕ_0)	0.0171	0.0052
Autocorrelation	(ϕ_1)	0.7670	0.0840
SD of shock	(σ_ξ)	0.0137	0.0035

Notes: Reported numbers are distributional moments of parameters from L separate AR(1) regressions.

Table A7: Earnings forecasting estimates

Parameter	Symbol	Mean	Std Dev
Earnings			
Drift	(ρ_0)	-0.0797	0.0352
Autocorrelation	(ρ_1)	0.7415	—
SD of shock	(σ_ζ)	0.0792	0.0416

Notes: Reported numbers are distributional moments of parameters from a pooled AR(1) regression with location-specific drift and shock variance, but common autocorrelation coefficient. The standard error of ρ_1 is 0.0359, which both rejects that the process is a unit root, and rejects that the process is white noise.

Table A8: Determinants of local labor market attributes

(a) Amenities, earnings, and employment levels

	Amenities (α_ℓ)		Earnings (w_ℓ)		Job destruction (δ_ℓ)		Job offer (λ_ℓ)	
Variable	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	0.0422	0.3531	7.3423***	0.1919	0.5860***	0.0376	0.0019	0.1060
ln(population)	0.0023	0.0238	-0.0272**	0.0129	0.0150***	0.0025	0.0236***	0.0071
New England	-0.0017	0.0566	-0.1923***	0.0308	-0.0197***	0.0060	-0.0265	0.0170
Mid Atlantic	-0.0403	0.0553	-0.0057	0.0301	0.0239***	0.0059	-0.0123	0.0166
E N Central	-0.0572	0.0386	0.1036***	0.0210	0.0140***	0.0041	-0.0253***	0.0116
W N Central	0.0225	0.0531	0.1017***	0.0289	0.0442***	0.0057	-0.0525***	0.0159
S Atlantic	-0.0338	0.0398	0.0480**	0.0217	0.0268***	0.0042	0.0107	0.0120
E S Central	-0.0744	0.0659	0.1061***	0.0358	-0.0064	0.0070	0.1038***	0.0198
W S Central	-0.1033*	0.0551	0.2983***	0.0299	0.0591***	0.0059	0.0192	0.0165
Mountain	0.0274	0.0643	0.1071***	0.0350	-0.0046	0.0069	0.0771***	0.0193
R^2	0.3848		0.3368		0.4507		0.2230	

Notes: Each column is a separate regression with 350 observations (35 cities, 10 time periods) of the corresponding model parameter on the log population of the location and Census division dummies. Amenities and AR(1) shock standard deviations do not vary over time, so these regressions have 35 observations. *** p<0.01; ** p<0.05; * p<0.10

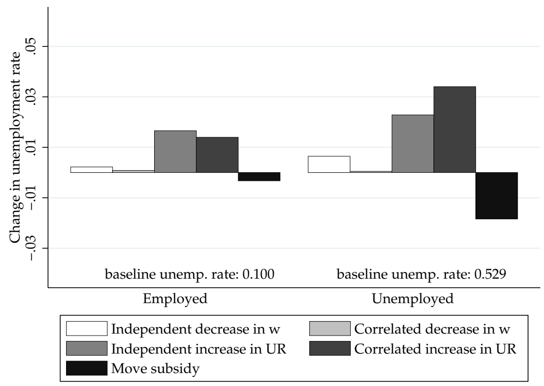
(b) Earnings and unemployment drift, persistence, and volatility

	Earnings drift ($\rho_{0\ell}$)		UR drift ($\phi_{0\ell}$)		UR persistence ($\phi_{1\ell}$)		Earnings volatility (σ_{ζ_ℓ})		UR volatility (σ_{ξ_ℓ})	
Variable	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	-0.0630	0.1235	0.0190	0.0197	0.4310	0.2896	0.3938	0.3048	-0.0091	0.0137
ln(population)	-0.0018	0.0083	-0.0001	0.0013	0.0237	0.0195	-0.0192	0.0205	0.0018*	0.0009
New England	-0.0585***	0.0198	-0.0015	0.0032	-0.0132	0.0464	-0.0280	0.0489	-0.0034	0.0022
Mid Atlantic	-0.0113	0.0193	-0.0051*	0.0031	0.0505	0.0453	0.0172	0.0477	-0.0052***	0.0021
E N Central	0.0312***	0.0135	0.0028	0.0022	-0.0454	0.0317	-0.0188	0.0333	-0.0019	0.0015
W N Central	0.0136	0.0186	0.0011	0.0030	-0.0660	0.0436	-0.0123	0.0458	-0.0027	0.0021
S Atlantic	0.0008	0.0139	-0.0047**	0.0022	0.0394	0.0327	-0.0436	0.0344	-0.0024	0.0015
E S Central	0.0236	0.0231	-0.0002	0.0037	-0.0163	0.0541	-0.0284	0.0569	-0.0006	0.0026
W S Central	0.0672***	0.0193	0.0010	0.0031	-0.1074***	0.0452	-0.0787	0.0475	-0.0060***	0.0021
Mountain	0.0276	0.0225	-0.0026	0.0036	-0.0043	0.0527	-0.0393	0.0555	-0.0011	0.0025
R^2	0.5989		0.4347		0.5004		0.3534		0.3953	

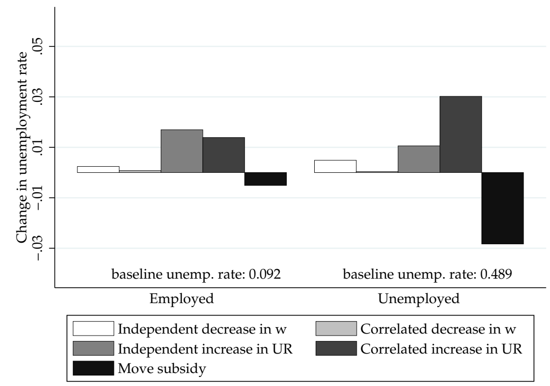
Notes: "UR" denotes unemployment rate. Each column is a separate regression with 350 observations (35 cities, 10 time periods) of the corresponding model parameter on the log population of the location and Census division dummies. Amenities and AR(1) shock standard deviations do not vary over time, so these regressions have 35 observations. *** p<0.01; ** p<0.05; * p<0.10

Figure A5: Counterfactual change in unemployment rate for Type 1 individuals, by origin city and prior employment status

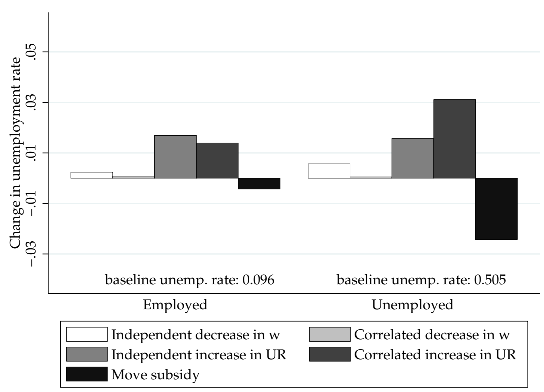
(a) High amenities



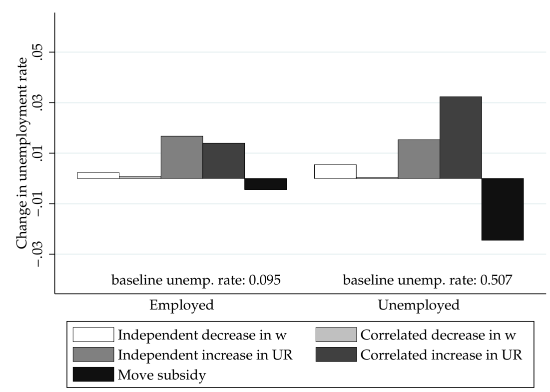
(b) Low amenities



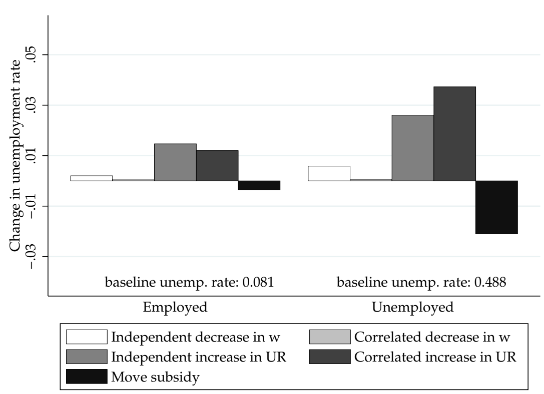
(c) High earnings



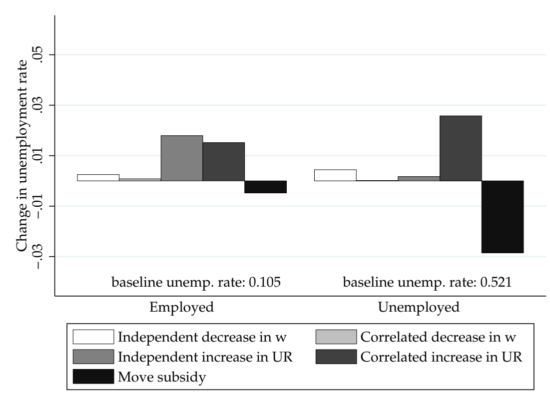
(d) Low earnings



(e) High emp. prob.

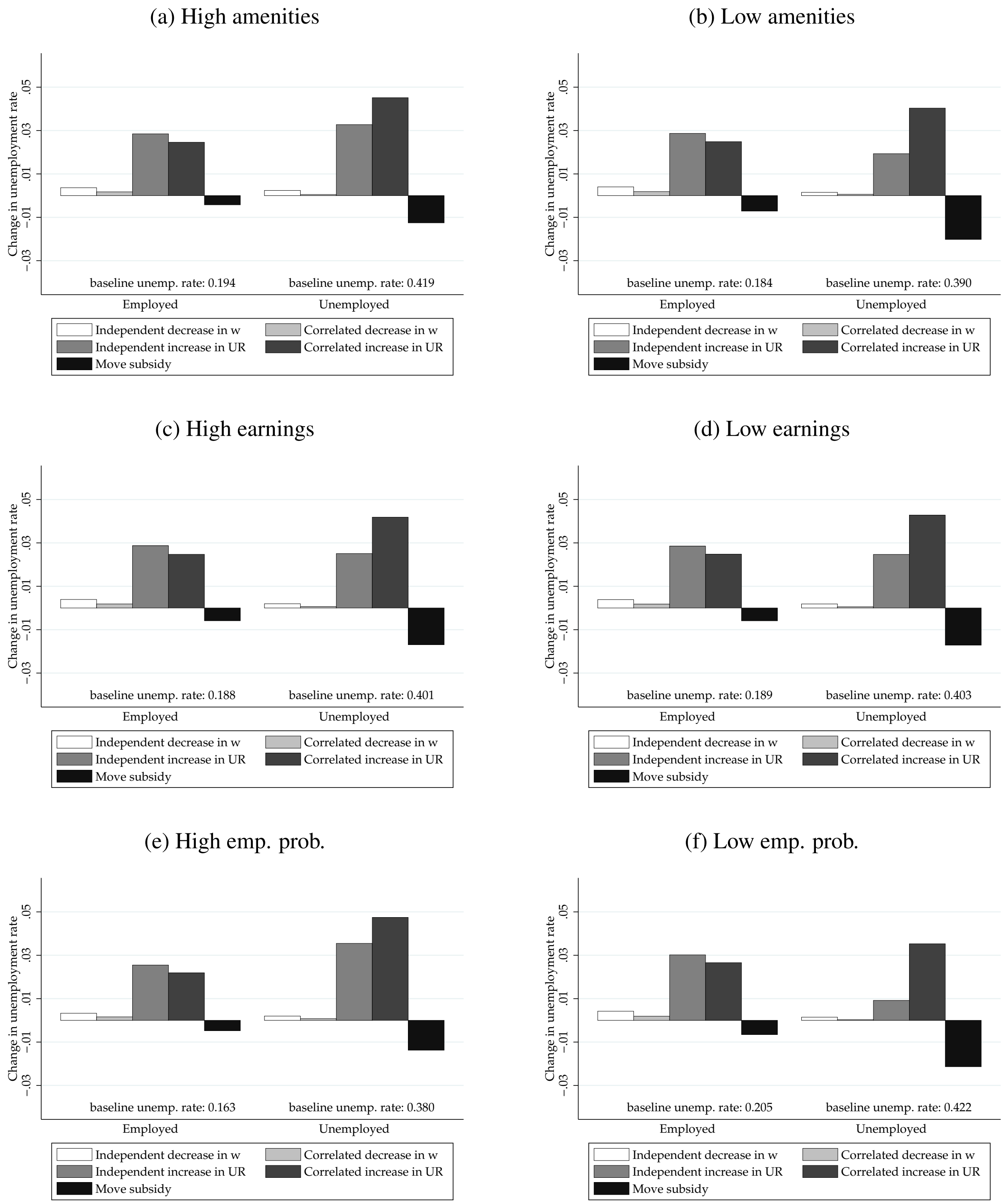


(f) Low emp. prob.



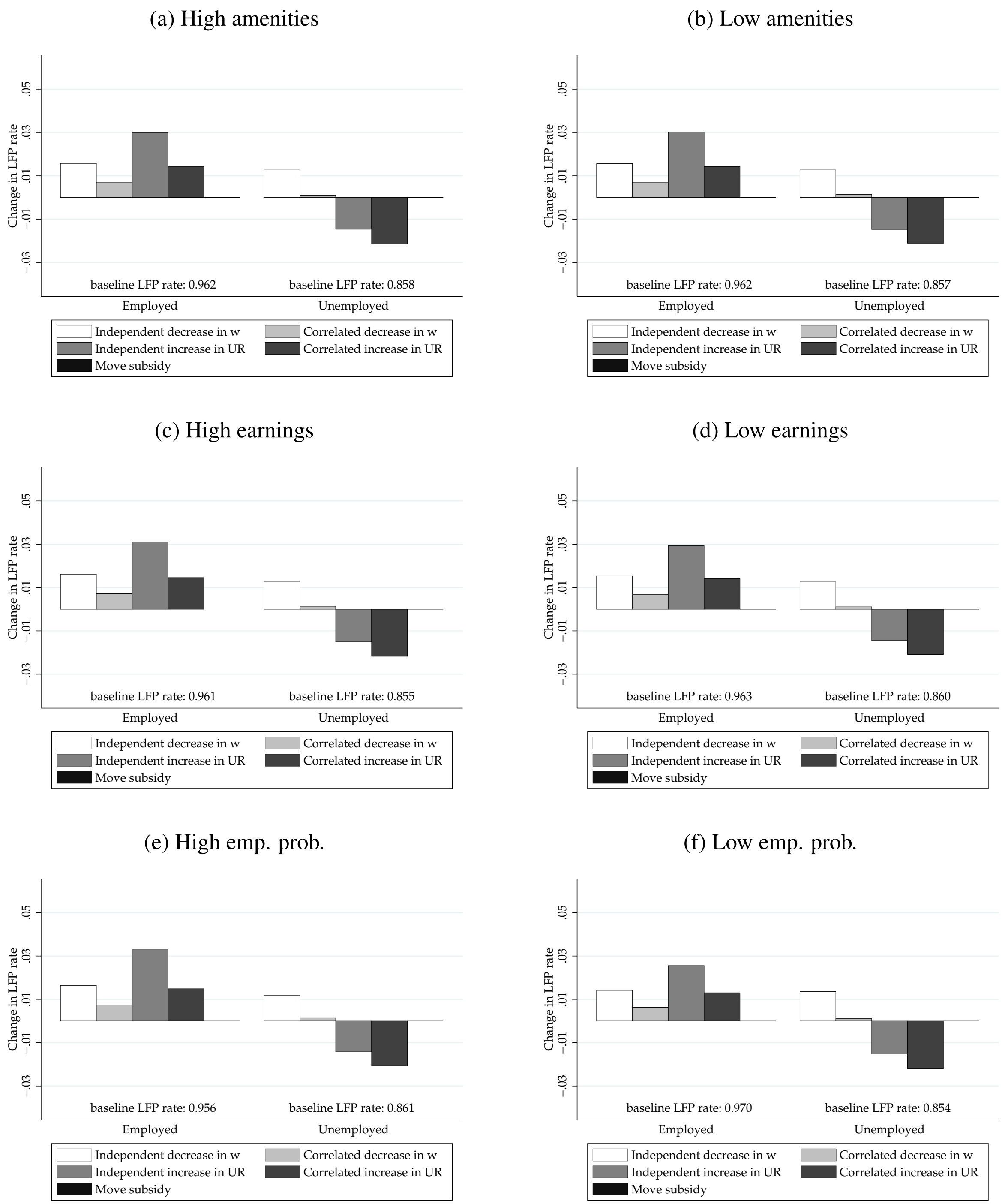
Notes: Each panel corresponds to a different origin city. Bar heights refer to the change in the out-migration rate from the specified location in response to the listed counterfactual. All figures are for 25-year-olds who were not born in the origin location. "high" refers to a location in the 75th percentile of the given distribution; "low" refers to the 25th percentile. All characteristics not set to "high" or "low" are set to the median. The earnings shock ($\downarrow w$) corresponds to the 70th percentile of the cross-location distribution in earnings AR(1) shock deviations. The unemployment shock corresponds to the jump from 2008 to 2009 for the average location in the data. To focus the results, each candidate location has median AR(1) parameters for both earnings and employment. Birth location is held fixed in all counterfactuals. Individual characteristics are set to the average for all 25-year-olds, conditional on employment status.

Figure A6: Counterfactual change in unemployment rate for Type 2 individuals, by origin city and prior employment status



Notes: See notes to Figure 2.

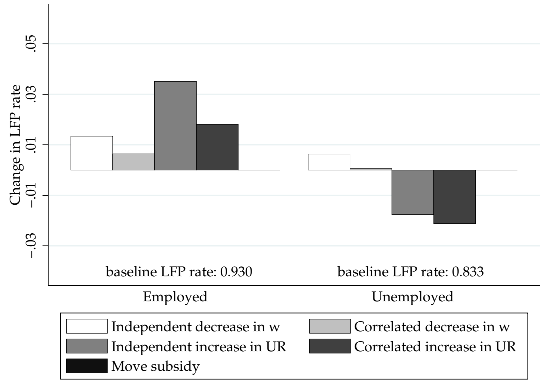
Figure A7: Counterfactual change in labor force participation rate for Type 1 individuals, by origin city and prior employment status



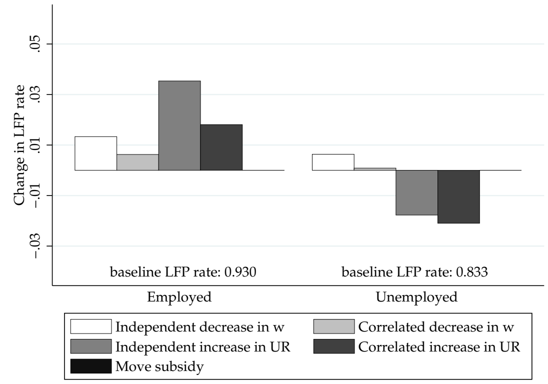
Notes: Each panel corresponds to a different origin city. Bar heights refer to the change in the out-migration rate from the specified location in response to the listed counterfactual. All figures are for 25-year-olds who were not born in the origin location. "high" refers to a location in the 75th percentile of the given distribution; "low" refers to the 25th percentile. All characteristics not set to "high" or "low" are set to the median. The earnings shock ($\downarrow w$) corresponds to the 70th percentile of the cross-location distribution in earnings AR(1) shock deviations. The unemployment shock corresponds to the jump from 2008 to 2009 for the average location in the data. To focus the results, each candidate location has median AR(1) parameters for both earnings and employment. Birth location is held fixed in all counterfactuals. Individual characteristics are set to the average for all 25-year-olds, conditional on employment status.

Figure A8: Counterfactual change in labor force participation rate for Type 2 individuals, by origin city and prior employment status

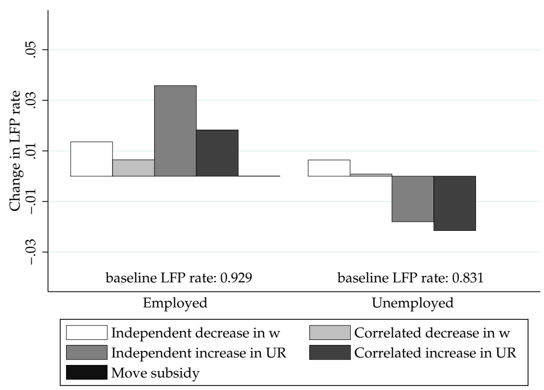
(a) High amenities



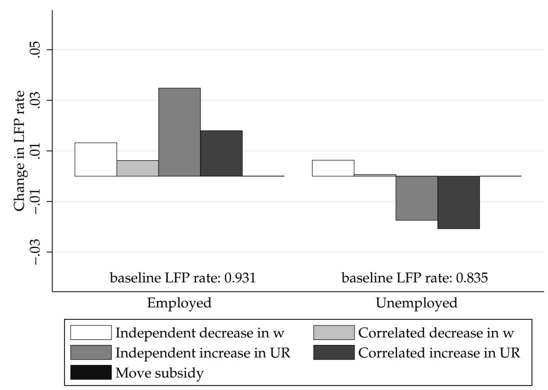
(b) Low amenities



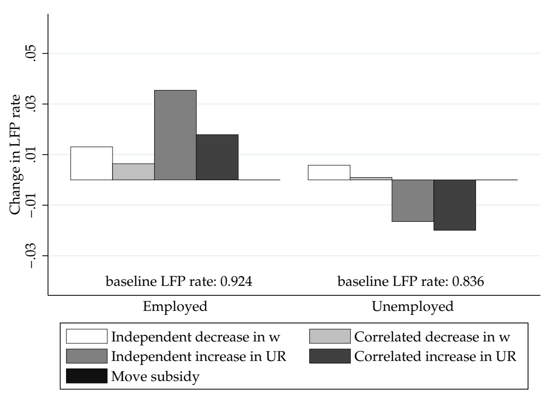
(c) High earnings



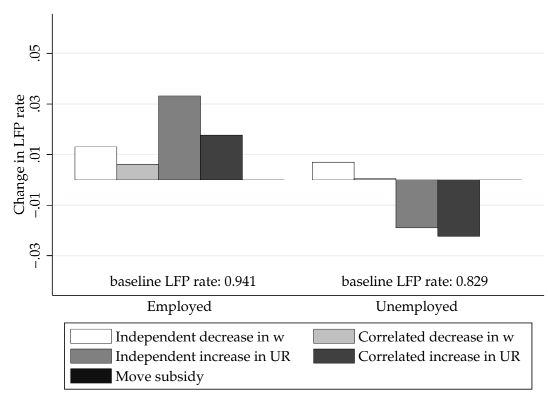
(d) Low earnings



(e) High emp. prob.



(f) Low emp. prob.



Notes: See notes to Figure 2.

Table A9: Greek symbol notation glossary

Greek symbol	Equation of first reference	Description
α	(3.9)	Local amenities
β	(3.15)	Discount factor
γ	(3.9)	Flow utility parameters
$\overline{\gamma}$	(A.7)	Euler's constant
δ	(3.3)	Job destruction probability
ε	(3.6)	Preference shocks
ζ	(3.2)	Shocks to evolution of earnings parameters
η	(3.1)	Earnings measurement error
θ	(3.13)	Moving and switching cost parameters
λ	(3.3)	Job offer probability
μ	(A.1)	Parameters in estimation of employment probabilities
ξ	(3.4)	Shocks to evolution of unemployment rate
π	(3.7)	Employment probabilities
$\overline{\pi}$	(A.19)	Population unobserved type probabilities
ρ	(3.2)	Parameters governing evolution of earnings parameters
σ_ζ	(3.2)	Std deviation of shocks to evolution of earnings parameters
σ_η	(3.1)	Std deviation of earnings measurement error
σ_ξ	(3.4)	Std deviation of shocks to evolution of unemployment rate
ϕ	(3.4)	Parameters governing evolution of employment probabilities
ϕ^h	(A.23)	Density of wage equation errors
ψ	(3.1)	Earnings parameters
ω	(A.11)	Value function weights
τ	(3.1)	Unobserved type (discrete)
Δ	(3.9)	Moving cost
Λ	(A.21)	Likelihood contribution of employment probabilities
Θ	(A.1)	Employment probability determinants
Ξ	(3.9)	Switching cost
Ψ	(A.16)	Covariance of local labor market shocks