

Online Appendix for
Labor Market Frictions and Moving Costs of the
Employed and Unemployed

Tyler Ransom

A Online Appendix

This appendix is organized as follows. Section A.1 provides complete details of each component of the structural model. Section A.2 discusses identification in further detail, and Section A.3 provides complete detail on how I estimate the model. Section A.4 explains how I calculate moving costs and amenity values, and Section A.5 evaluates destination location choice in the counterfactual simulation. Section A.6 explains details on selecting the estimation subsample, while Section A.7 explains how I account for cost of living. Section A.8 explains the sampling design of the SIPP. All appendix tables are included after Section A.8.

A.1 Detailed definition of the structural model

This section of the Online Appendix presents each element of the model in complete detail.³⁷

A.1.1 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 τ_{ir} , and measurement error $\eta_{i\ell t}$. The log earnings equation is

$$\ln w_{i\ell r t} = \psi_{0\ell t} + \psi_1 G(x_{it}) + \psi_2 \tau_{ir} + \eta_{i\ell t} \quad (\text{A.1})$$

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

³⁷A cross-referenced notation glossary for all Greek symbols is available in Online Appendix Table A14.

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 $\psi_{0\ell t}$:

$$\psi_{0\ell t} = \rho_{0\ell} + \rho_1 \psi_{0\ell t-1} + \zeta_{\ell t} \quad (\text{A.2})$$

where $\zeta_{\ell t}$ is distributed $N(0, \sigma_{\zeta_{\ell t}}^2)$. In other words, individuals know the drift and autocorrelation coefficients ($\rho_{0\ell}$ and ρ_1) and shock variances ($\sigma_{\zeta_{\ell t}}^2$) for each location, and integrate over future realizations of $\zeta_{\ell t}$ using information about the distribution from which $\zeta_{\ell t}$ is drawn.³⁸

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.

A.1.2 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.

$$\pi_{i\ell rt}(x_{it}, \tau_{ir}, UR_{\ell t-1}) = \begin{cases} (1 - \delta_{i\ell rt}) & \text{if employed in } \ell \text{ in } t-1 \\ \lambda_{i\ell rt} & \text{if not employed in } \ell \text{ in } t-1 \\ \lambda_{i\ell' rt}^e & \text{if employed in } \ell' \neq \ell \text{ in } t-1 \\ \lambda_{i\ell' rt}^u & \text{if not employed in } \ell' \neq \ell \text{ in } t-1 \end{cases} \quad (\text{A.3})$$

Equation (A.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.³⁹ Employed individuals are thus laid off with probability $\delta_{i\ell rt}$ and unemployed individuals receive a job offer with probability $\lambda_{i\ell rt}$.

³⁸The assumption that ρ_1 is not location-specific has also been made by [Kaplan and Schulhofer-Wohl \(2017\)](#).

³⁹In 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 previously, since this group of people is more likely to move before finding a job ([Balgova, 2018](#)).

Individuals coming from employment in another location receive an offer with probability $\lambda_{i\ell r t}^e$. Individuals coming from non-employment in another location receive an offer with probability $\lambda_{i\ell r t}^u$. 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.⁴⁰

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.

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

$$UR_{\ell t} = \phi_{0\ell} + \phi_{1\ell} UR_{\ell t-1} + \xi_{\ell t} \quad (\text{A.4})$$

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

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

$$\mathbb{E}_t \pi_{i\ell t+1} = \int_{\xi_{\ell t}} \Pr\{\mu_2(\phi_{0\ell} + \phi_{1\ell} UR_{\ell t} + \xi_{\ell t+1}) + z_{i\ell r} \mu > 0\} dF(\xi_{\ell t+1}) \quad (\text{A.5})$$

where the argument inside the probability is a latent linear index comprised of μ_2 , a parameter

⁴⁰For a treatment of migration in the style of a classical job search model, see [Schluter and Willems \(2018\)](#) and [Schmutz and Sidibé \(2019\)](#). For recent work on extending classical job search models to include preference heterogeneity over non-wage amenities, see [Sullivan and To \(2014\)](#); [Taber and Vejin \(2020\)](#); [Hall and Mueller \(2018\)](#); [Sorkin \(2018\)](#).

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.

A.1.3 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, $\ell \in \{1, \dots, L\}$ indexes locations, and $\mathcal{J} = \{0, 1\} \times \{1, \dots, L\}$ 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 $y_{it} \in \{e, u, n\} \times \{1, \dots, L\}$ be the choice *outcome*, where e denotes employment, u unemployment, and n non-participation. The set $\{1, \dots, L\}$ covers all possible locations in the United States. A complete list of these locations can be found in Online Appendix 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 $\varepsilon_{ij\ell t}$:

$$U_{ij\ell t}(Z_{it}, \varepsilon_{ij\ell t}) = u_{ij\ell t}(Z_{it}) + \varepsilon_{ij\ell t} \quad (\text{A.6})$$

The flow payoffs associated with choices j and ℓ are an expectation of the employment outcomes because individuals pre-commit to working if they get a job offer:

$$u_{i1\ell t}(Z_{it}) = \pi_{i\ell r t}(Z_{it}) u_{i\ell t}^e(Z_{it}) + (1 - \pi_{i\ell r t}(Z_{it})) u_{i\ell t}^u(Z_{it}) \quad (\text{A.7})$$

$$u_{i0\ell t}(Z_{it}) = u_{i\ell t}^n(Z_{it}) \quad (\text{A.8})$$

The flow utility corresponding to each employment outcome is given by

$$u_{i\ell t}^e(Z_{it}) = \alpha_\ell + \Delta_\ell(Z_{it}) + \Xi_1(Z_{it}) + \gamma_3 b_{i\ell}^{ST} + \gamma_4 b_{i\ell}^{DIV} + \gamma_0 \ln \tilde{w}_{i\ell t}(Z_{it}) \quad (\text{A.9})$$

$$u_{i\ell t}^u(Z_{it}) = \alpha_\ell + \Delta_\ell(Z_{it}) + \Xi_1(Z_{it}) + \gamma_3 b_{i\ell}^{ST} + \gamma_4 b_{i\ell}^{DIV} + \gamma_1 + \gamma_2 \quad (\text{A.10})$$

$$u_{i\ell t}^n(Z_{it}) = \alpha_\ell + \Delta_\ell(Z_{it}) + \Xi_0(Z_{it}) + \gamma_3 b_{i\ell}^{ST} + \gamma_4 b_{i\ell}^{DIV} + \gamma_1 \quad (\text{A.11})$$

The first term α_ℓ is a location fixed effect measuring the net value of all amenities in location ℓ . The variables $b_{i\ell}^{ST}$ and $b_{i\ell}^{DIV}$ are dummies that indicate if the individual was born in any of the states (*ST*) or Census divisions (*DIV*) contained in location ℓ .⁴¹ $\ln \tilde{w}_{i\ell t}(Z_{it})$ 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.⁴² Home production benefits and search costs are constant 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.

A.1.4 Moving costs and switching costs

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

$$\Delta_\ell(Z_{it}) = (\theta_0 + \theta_1 D(\ell, h) + \theta_2 D^2(\ell, h) + \theta_3 \text{age}_{it} + \theta_4 \text{age}_{it}^2 + \theta_5 \text{employed}_{it-1} + \theta_6 \text{unemployed}_{it-1} + \theta_7 \tau_i) 1\{\ell \neq h\} \quad (\text{A.12})$$

⁴¹For 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.

⁴²While 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.

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 \(2012\)](#).⁴³ 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.⁴⁴ 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.

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:

$$\Xi_1(Z_{it}) = (\theta_8 + \theta_9 \text{age}_{it} + \theta_{10} \text{age}_{it}^2 + \theta_{11} \tau_{ir}) 1\{k = 0\} \quad (\text{A.13})$$

$$\Xi_0(Z_{it}) = (\theta_{12} + \theta_{13} \text{age}_{it} + \theta_{14} \text{age}_{it}^2 + \theta_{15} \tau_{ir}) 1\{k = 1\} \quad (\text{A.14})$$

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

⁴³[Davies, Greenwood, and Li \(2001\)](#) also include moving costs in a conditional logit analysis of migration.

⁴⁴Those who were previously out of the labor force serve as the reference group.

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.⁴⁵

A.2 Further details on identification of the model's parameters

This section informally discusses identification of various parameters of the model.⁴⁶ 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. As with any causal analysis using observational data, identification relies on a set of exclusion restrictions.

A.2.1 Identification of earnings and employment parameters

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 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.

⁴⁵Previous employment status is not included in the switching costs so as to maintain clear interpretation of the search cost parameter γ_2 in (A.10).

⁴⁶See [Magnac and Thesmar \(2002\)](#) for a formal discussion of identification in dynamic discrete choice models.

A.2.2 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 there is no reason to think that time-varying amenities in a given location (e.g. crime, air pollution) would be correlated 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.⁴⁷ 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 $\bar{\pi}_r$, is directly identified from the frequency of individuals with each particular type label.

A.3 Further details on estimation of the model

This section explains in detail the various steps to estimate the model's parameters.

A.3.1 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, $\hat{\lambda}_{i\ell rt}^e$ and $\hat{\lambda}_{i\ell rt}^u$ 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 $\hat{\lambda}_{i\ell rt}^e$ is estimated conditional on choosing to supply labor in the current period and having been employed in the previous period.

$$\Pr[y_{it} = (e, \ell) | d_{it} = (1, \ell), y_{it-1} = (e, \cdot)] = \frac{\exp(\Theta_1)}{1 + \exp(\Theta_1)} \quad (\text{A.15})$$

⁴⁷For 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.

where

$$\Theta_1 = \mu_{1\ell}^e + \mu_2^e \text{UR}_{\ell t-1} + \mu_3^e G(x_{it}) + \mu_4^e \tau_{ir} + \mu_5^e 1 \{y_{it-1} = (e, \ell')\}$$

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.

$$\Pr[y_{it} = (e, \ell) | d_{it} = (1, \ell), y_{it-1} = (\{u, n\}, \cdot)] = \frac{\exp(\Theta_2)}{1 + \exp(\Theta_2)} \quad (\text{A.16})$$

where

$$\Theta_2 = \mu_{1\ell}^u + \mu_2^u \text{UR}_{\ell t-1} + \mu_3^u G(x_{it}) + \mu_4^u \tau_{ir} + \mu_5^u 1 \{y_{it-1} = (\{u, n\}, \ell')\}$$

The excluded category in (A.16) 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.15) and (A.16) 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

$$1 - \hat{\delta}_{i\ell rt} = \frac{\exp(\hat{\mu}_{1\ell}^e + \hat{\mu}_2^e \text{UR}_{\ell t-1} + \hat{\mu}_3^e G(x_{it} + \hat{\mu}_4^e \tau_{ir}))}{1 + \exp(\hat{\mu}_{1\ell}^e + \hat{\mu}_2^e \text{UR}_{\ell t-1} + \hat{\mu}_3^e G(x_{it} + \hat{\mu}_4^e \tau_{ir}))} \quad (\text{A.17})$$

$$\hat{\lambda}_{i\ell rt}^e = \frac{\exp(\hat{\mu}_{1\ell}^e + \hat{\mu}_2^e \text{UR}_{\ell t-1} + \hat{\mu}_3^e G(x_{it}) + \hat{\mu}_4^e \tau_{ir} + \hat{\mu}_5^e)}{1 + \exp(\hat{\mu}_{1\ell}^e + \hat{\mu}_2^e \text{UR}_{\ell t-1} + \hat{\mu}_3^e G(x_{it}) + \hat{\mu}_4^e \tau_{ir} + \hat{\mu}_5^e)} \quad (\text{A.18})$$

$$\hat{\lambda}_{i\ell rt}^u = \frac{\exp(\hat{\mu}_{1\ell}^u + \hat{\mu}_2^u \text{UR}_{\ell t-1} + \hat{\mu}_3^u G(x_{it}) + \hat{\mu}_4^u \tau_{ir})}{1 + \exp(\hat{\mu}_{1\ell}^u + \hat{\mu}_2^u \text{UR}_{\ell t-1} + \hat{\mu}_3^u G(x_{it}) + \hat{\mu}_4^u \tau_{ir})} \quad (\text{A.19})$$

$$\hat{\lambda}_{i\ell rt}^u = \frac{\exp(\hat{\mu}_{1\ell}^u + \hat{\mu}_2^u \text{UR}_{\ell t-1} + \hat{\mu}_3^u G(x_{it}) + \hat{\mu}_4^u \tau_{ir} + \hat{\mu}_5^u)}{1 + \exp(\hat{\mu}_{1\ell}^u + \hat{\mu}_2^u \text{UR}_{\ell t-1} + \hat{\mu}_3^u G(x_{it}) + \hat{\mu}_4^u \tau_{ir} + \hat{\mu}_5^u)} \quad (\text{A.20})$$

A.3.2 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.

Conditional choice probabilities I first describe how CCPs are employed. To do so, I define the following equation, which is a rewritten form of (3.2). The choice-specific value function $v_{ij\ell t}$ is defined as the flow payoff of choosing (j, ℓ) minus $\varepsilon_{ij\ell t}$ plus future utility assuming that the optimal decision is made in every period from $t + 1$ on.

$$\begin{aligned} v_{ij\ell t}(Z_{it}) &= u_{ij\ell t}(Z_{it}) + \beta \int V_{it+1}(Z_{it+1}) dF(Z_{it+1}|Z_{it}) \\ &= u_{ij\ell t}(Z_{it}) + \beta \int \mathbb{E}_{\varepsilon} \max_{k,m} \{v_{ikm+1}(Z_{it+1}) + \varepsilon_{ikm+1}\} dF(Z_{it+1}|Z_{it}) \end{aligned} \quad (\text{A.21})$$

Equation (A.21) shows 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 (A.21) simplifies to

$$v_{j\ell t}(Z_t) = u_{j\ell t}(Z_t) + \beta \int \ln \left(\sum_k \sum_m \exp(v_{kmt+1}(Z_{t+1})) \right) dF(Z_{t+1}|Z_t) + \beta \bar{\gamma} \quad (\text{A.22})$$

where $\bar{\gamma}$ 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.⁴⁸

I will now show how (A.22) can be manipulated to admit CCPs. First, multiply and divide by the exponentiated conditional value function associated with a given choice alternative (e.g.

⁴⁸This 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.

(j', ℓ') , $\exp(v_{j'\ell'_{t+1}}(Z_{t+1}))$, to get

$$\int V_{t+1}(Z_{t+1}) dF(Z_{t+1}|Z_t) = \int \ln \left(\frac{\exp(v_{j'\ell'_{t+1}}(Z_{t+1}))}{\exp(v_{j'\ell'_{t+1}}(Z_{t+1}))} \right. \\ \left. \times \sum_k \sum_m \exp(v_{kmt+1}(Z_{t+1})) \right) dF(Z_{t+1}|Z_t) + \bar{\gamma} \quad (\text{A.23})$$

$$= \int \left[v_{j'\ell'_{t+1}}(Z_{t+1}) \right. \\ \left. + \ln \left(\frac{\sum_k \sum_m \exp(v_{kmt+1}(Z_{t+1}))}{\exp(v_{j'\ell'_{t+1}}(Z_{t+1}))} \right) \right] dF(Z_{t+1}|Z_t) + \bar{\gamma} \quad (\text{A.24})$$

$$= \int \left[v_{j'\ell'_{t+1}}(Z_{t+1}) - \ln p_{j'\ell'_{t+1}}(Z_{t+1}) \right] dF(Z_{t+1}|Z_t) + \bar{\gamma} \quad (\text{A.25})$$

Comparing (A.21) with (A.25) 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.25) 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.

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.⁴⁹

⁴⁹For 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, 2019](#); [Gayle, 2018](#); and [Humphries, 2018](#).

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.25), rewritten below in conserved notation:

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

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

$$\begin{aligned} \mathbb{E}_t V_{t+1}(Z_{t+1}) &= \mathbb{E}_t [v_{j'\ell'_{t+1}}(Z_{t+1}) - \ln(p_{j'\ell'_{t+1}}(Z_{t+1}))] + \bar{\gamma} \\ &= \sum_{(k,m) \in \mathcal{J}} \omega_{(k,m)} \{ \mathbb{E}_t [v_{kmt+1}(Z_{t+1}) - \ln(p_{kmt+1}(Z_{t+1}))] + \bar{\gamma} \} \\ \text{s.t. } \sum_{(k,m) \in \mathcal{J}} \omega_{(k,m)} &= 1 \end{aligned} \quad (\text{A.26})$$

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

Figure A4 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 A4 (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.26). The diagram for this case is provided in Figure A5. 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:

$$\pi_{\ell_t} V_{t+3} = \omega_{(1,\ell)} \pi_{\ell_{t+1}} V_{t+3} \quad (\text{A.27})$$

Solving (A.27) for ω gives

$$\omega_{(1,\ell)} = \frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \quad (\text{A.28})$$

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 $j' = 1$):

$$\begin{aligned}
v^{j'\ell_t} - v_{0\ell_t} = & \pi_{\ell_t} u_{\ell_t}^e(Z_t) + (1 - \pi_{\ell_t}) u_{\ell_t}^u(Z_t) - u_{\ell_t}^n(Z_t) \\
& + \beta \left[\pi_{\ell_t} u_{\ell_{t+1}}^n(Z_{t+1}^1) - \pi_{\ell_t} \ln p_{0\ell_{t+1}}(Z_{t+1}^1) \right. \\
& + (1 - \pi_{\ell_t}) u_{\ell_{t+1}}^n(Z_{t+1}^2) - (1 - \pi_{\ell_t}) \ln p_{0\ell_{t+1}}(Z_{t+1}^2) \\
& - \pi_{\ell_t} u_{\ell_{t+1}}^e(Z_{t+1}^3) \\
& - \left. \left(\frac{\pi_{\ell_t} (1 - \pi_{\ell_{t+1}})}{\pi_{\ell_{t+1}}} \right) u_{\ell_{t+1}}^u(Z_{t+1}^3) - \left(1 - \frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \right) u_{\ell_{t+1}}^n(Z_{t+1}^3) \right. \\
& + \left. \left(\frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \right) \ln p_{1\ell_{t+1}}(Z_{t+1}^3) + \left(1 - \frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \right) \ln p_{0\ell_{t+1}}(Z_{t+1}^3) \right] \\
& + \beta^2 \left[\pi_{\ell_t} u_{\ell_{t+2}}^n(Z_{t+2}^4) - \pi_{\ell_t} \ln p_{0\ell_{t+2}}(Z_{t+2}^4) \right. \\
& + (1 - \pi_{\ell_t}) u_{\ell_{t+2}}^n(Z_{t+2}^5) - (1 - \pi_{\ell_t}) \ln p_{0\ell_{t+2}}(Z_{t+2}^5) \\
& - \pi_{\ell_t} u_{\ell_{t+2}}^n(Z_{t+2}^6) + \pi_{\ell_t} \ln p_{0\ell_{t+2}}(Z_{t+2}^6) \\
& - \left. \left(\frac{\pi_{\ell_t} (1 - \pi_{\ell_{t+1}})}{\pi_{\ell_{t+1}}} \right) u_{\ell_{t+2}}^n(Z_{t+2}^7) + \left(\frac{\pi_{\ell_t} (1 - \pi_{\ell_{t+1}})}{\pi_{\ell_{t+1}}} \right) \ln p_{0\ell_{t+2}}(Z_{t+2}^7) \right. \\
& - \left. \left(1 - \frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \right) u_{\ell_{t+2}}^n(Z_{t+2}^8) + \left(1 - \frac{\pi_{\ell_t}}{\pi_{\ell_{t+1}}} \right) \ln p_{0\ell_{t+2}}(Z_{t+2}^8) \right]
\end{aligned} \tag{A.29}$$

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.⁵⁰

Equation (A.29) 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 estimation procedure.

For non-employment alternatives, the finite dependence formula written in equation (A.29) is

⁵⁰ Additionally, the state dependence of the employment probabilities π_{ℓ_t} and $\pi_{\ell_{t+1}}$ is also suppressed for simplicity. π_{ℓ_t} is always evaluated at Z_t while $\pi_{\ell_{t+1}}$ is always evaluated at Z_{t+1}^3 .

much simpler (because $pi_{\ell t} = 0$ when $j = 0$ for all ℓ and t):

$$\begin{aligned}
v_{j\ell t}(Z_t) - v_{0\ell t}(Z_t) = & u_{\ell t}^n(Z_t) - u_{\ell t}^n(Z_t) + \\
& \beta \left[(u_{\ell t+1}^n(\{0, \ell'\}, x_t) - \ln p_{0\ell t+1}(\{0, \ell'\}, x_t)) \right. \\
& \left. - u_{\ell t+1}^n(\{0, \ell\}, x_t) + \ln p_{0\ell t+1}(\{0, \ell\}, x_t) \right] + \\
& \beta^2 \left[u_{\ell t+2}^n(\{0, \ell'\}, x_t) - \ln p_{0\ell t+2}(\{0, \ell'\}, x_t) \right. \\
& \left. - u_{\ell t+2}^n(\{0, \ell\}, x_t) + \ln p_{0\ell t+2}(\{0, \ell\}, x_t) \right]
\end{aligned} \tag{A.30}$$

Figures A4 and A5 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.

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 (A.2) and (A.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.⁵¹ With many locations, the only way to compute the integral is using Monte Carlo techniques.

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

⁵¹ 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 .

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.29).

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

$$E_t [\ln(p_{0\ell't+1}(\xi_{t+1}, \zeta_{t+1})) | Z_t] = \int \ln p_{0\ell't+1}(\xi_{t+1}, \zeta_{t+1}) dF(\xi, \zeta) \quad (\text{A.31})$$

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 .⁵²

The integral in (A.31) 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:

$$\int \ln p_{0\ell't+1}(\xi_{t+1}, \zeta_{t+1}) dF(\xi, \zeta) \approx \frac{1}{D} \sum_{d=1}^D \ln p_{0\ell't+1}(\xi_d, \zeta_d) \quad (\text{A.32})$$

where (ξ_d, ζ_d) is the d th 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

$$\Psi + \Psi \odot RR' \quad (\text{A.33})$$

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.33) comes about because the forecasting shocks are assumed to be normally distributed and independent over

⁵² Ψ is estimated by computing the covariance of the AR(1) residuals for all equations in both AR(1) systems.

time.

A.3.3 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. The 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.⁵³

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

$$L = \sum_i \ln \left(\sum_{r=1}^R \bar{\pi}_r \mathcal{L}_{d,i|r} \mathcal{L}_{w,i|r} \mathcal{L}_{\pi,i|r} \right) \quad (\text{A.34})$$

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

$$L = \sum_i \ln \left(\sum_{r=1}^R \bar{\pi}_r \prod_j \prod_\ell \prod_t \left\{ [P_{ij\ell rt} \Lambda_{i\ell rt} h_{i\ell rt}]^{d_{ij\ell t}} \right\}^{1[j=1]} [P_{ij\ell rt}^{d_{ij\ell t}}]^{1[j=0]} \right) \quad (\text{A.35})$$

where the following hold:

- P is a multinomial logit function:

$$P_{ij\ell rt} = \frac{\exp(\cdot)}{\sum_{k=1}^{J \times L} \exp(\cdot)} \quad (\text{A.36})$$

- 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

⁵³For 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.

labor market outcome. Λ is a product of two binary logit likelihoods.

$$\Lambda_{i\ell rt} = \left[\pi_{i\ell rt}^{1[y_{i\ell t}=e]} (1 - \pi_{i\ell rt})^{1[y_{i\ell t} \neq e]} \right]^{1[y_{i\ell t-1}=e]} \left[\pi_{i\ell rt}^{1[y_{i\ell t}=e]} (1 - \pi_{i\ell rt})^{1[y_{i\ell t} \neq e]} \right]^{1[y_{i\ell t-1} \neq e]} \quad (\text{A.37})$$

where $\pi_{i\ell rt}$ is as defined in (A.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 A.1), the likelihood is given by

$$h_{i\ell rt} = \frac{1}{\sigma_\eta} \phi^h \left(\frac{\ln w_{i\ell t} - \psi_{0\ell t} - \psi_1 G(x_{it}) - \psi_2 \tau_{ir}}{\sigma_\eta} \right) \quad (\text{A.38})$$

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:

$$q_{ir} = \frac{\bar{\pi}_r \mathcal{L}_{d,i|r} \mathcal{L}_{w,i|r} \mathcal{L}_{\pi,i|r}}{\sum_{r'=1}^R \bar{\pi}_{r'} \mathcal{L}_{d,i|r'} \mathcal{L}_{w,i|r'} \mathcal{L}_{\pi,i|r'}} \quad (\text{A.39})$$

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

$$\bar{\pi}_r = \frac{1}{N} \sum_i q_{ir} \quad (\text{A.40})$$

The key insight of (A.39) is that (A.34) can be rewritten as

$$\tilde{L} = \sum_i \sum_{r=1}^R \{ q_{ir} \ln \mathcal{L}_{d,i|r} + q_{ir} \ln \mathcal{L}_{w,i|r} + q_{ir} \ln \mathcal{L}_{\pi,i|r} \} \quad (\text{A.41})$$

where the first order conditions of \tilde{L} are equal to the first order conditions of L in (A.34). 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.34) and (A.41). Most importantly, (A.41) 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.39) and $\bar{\pi}_r$'s according to (A.40).
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

This section details the calculation of the moving cost and amenity values. For simplicity, consider only the location dimension of choice and only the amenity and moving cost components of the the flow utility. That is, suppose that the flow utility of i for choosing location j in period t is

$$u_{ijt} = \alpha_j + \gamma_0 \ln w_{ijt} + \theta_0 \cdot 1[d_{it-1} \neq j] + \varepsilon_{ijt} \quad (\text{A.42})$$

where $1[d_{it-1} \neq j]$ indicates that i chose a location different from j in the previous period, i.e. θ_0 is the parameter representing the fixed cost of moving.

A.4.1 Moving costs

Moving costs are paid once, rather than on a per-period basis. Because the model is dynamic, a natural question to ask is, “what is the present value of the increase in (flow) earnings required to make the person indifferent between moving and not moving, holding fixed all other aspects of utility?”

Assume that the person has T periods left in his working life, and that a move is made in period τ . Denote w as the status-quo earnings and w' as the earnings received after moving.

If the person moves, the present value stream of utility is

$$PV_{\text{move}} = \sum_{t=0}^{\tau-1} \beta^t u(w) + \beta^\tau u(w', \text{move}) + \sum_{t=\tau+1}^T \beta^t u(w'). \quad (\text{A.43})$$

And if the person does not move, the present value stream of utility is

$$PV_{\text{stay}} = \sum_{t=0}^{\tau-1} \beta^t u(w) + \beta^\tau u(w) + \sum_{t=\tau+1}^T \beta^t u(w). \quad (\text{A.44})$$

Setting the two equal and canceling out the first $\tau - 1$ periods gives

$$\begin{aligned} \beta^\tau u(w', \text{move}) + \sum_{t=\tau+1}^T \beta^t u(w') &= \beta^\tau u(w) + \sum_{t=\tau+1}^T \beta^t u(w) \\ \beta^\tau u(w', \text{move}) &= \beta^\tau u(w) + \sum_{t=\tau+1}^T \beta^t [u(w) - u(w')] \\ \beta^\tau [\gamma_0 \ln w' + \theta_0] &= \beta^\tau \gamma_0 \ln w + \sum_{t=\tau+1}^T \beta^t [\gamma_0 \ln w - \gamma_0 \ln w'] \\ \beta^\tau \left[\ln w' + \frac{\theta_0}{\gamma_0} \right] &= \beta^\tau \ln w + \sum_{t=\tau+1}^T \beta^t [\ln w - \ln w'] \\ -\frac{\theta_0}{\gamma_0} &= \sum_{t=\tau}^T \beta^{t-\tau} \left[\ln \left(\frac{w'}{w} \right) \right] \end{aligned} \quad (\text{A.45})$$

One can use a variety of numerical methods to solve the above equation for w'/w (and then subtract 1 to obtain a percentage change). To obtain the values reported in Table 7, I plug in the parameter estimates from Table 6, set T to be 65 minus the age, and set $\tau = 1$. Rather than use θ_0 in each cell of Table 7, I plug in the linear index implied by the parameter estimates, e.g. $\theta_0 + \theta_5$

for the fixed cost for an employed person of unobserved type 2, or $\theta_0 + \theta_5 + \theta_7$ for an employed person of unobserved type 1.

A.4.2 Amenity values

Because the amenity values are flow values, I can easily obtain their valuation by using a straightforward willingness to pay (WTP) formula (Koşar, Ransom, and van der Klaauw, Forthcoming).

As an example, I explain the calculation for living in one's state of birth.

Calculation of the WTP implies that the following indifference condition is satisfied:

$$\begin{aligned}
 u(w - WTP, \text{ in birth state}) &= u(w, \text{ not in birth state}) \\
 \gamma_0 \ln(w - WTP) + \gamma_3 &= \gamma_0 \ln(w) \\
 \gamma_0 [\ln(w - WTP) - \ln w] &= -\gamma_3 \\
 \ln\left(\frac{w - WTP}{w}\right) &= -\frac{\gamma_3}{\gamma_0} \\
 \frac{w - WTP}{w} &= \exp\left(-\frac{\gamma_3}{\gamma_0}\right) \\
 \frac{WTP}{w} &= 1 - \exp\left(-\frac{\gamma_3}{\gamma_0}\right) \\
 WTP &= \underbrace{\left[1 - \exp\left(-\frac{\gamma_3}{\gamma_0}\right)\right]}_{\text{fraction of flow earnings}} w
 \end{aligned} \tag{A.46}$$

A.4.3 Converting flow-equivalent values to net present values

The moving costs and amenity values are expressed in terms of a percentage of flow earnings. They can equivalently be expressed in terms of net present value. The formula to compute this is given below:

$$NPV = \sum_{t=\tau}^T \beta^{t-\tau} \frac{pct}{100} 12 \exp(\overline{\ln w}) \tag{A.47}$$

where pct is the percentage difference in flow earnings computed in (A.45) and (A.46), and $12 \exp(\overline{\ln w})$ is 12 times the exponential of the average monthly log earnings. For the numbers

reported in Panel A of Table 7, I use the same values of τ and T that are used in the calculation of the flow earnings equivalents above. I evaluate $12 \exp(\overline{\ln w})$ at different ages: \$34,530 (the unconditional sample average) for the fixed cost of moving, \$41,365 (the average among 39 year olds) for the average mover, and \$27,142 (the average among 25 year olds) for a young mover.

A.5 Further discussion of counterfactual simulations: where moving subsidy recipients relocate

I can 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).⁵⁴ 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 A9 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 A9 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.⁵⁵ 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

⁵⁴These 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.

⁵⁵These 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.

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 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.

A.6 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 Online Appendix Table A1.

A.7 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.48) 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):

$$\text{INDEX}_j = \prod_g \left(\frac{p_g^j}{p_g^0} \right)^{s_g} \quad (\text{A.48})$$

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.48) 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):

$$\bar{v} = \max_z U(z) + \lambda \left[w_j - \sum_g p_g^j z_g \right]. \quad (\text{A.49})$$

Log-linearizing (A.49) 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 :

$$\ln(w_0) = \ln(w_j) - \sum_g s_g [\ln(p_g^j) - \ln(p_g^0)] \quad (\text{A.50})$$

Taking the exponential of both sides and rearranging terms yields equation (A.48).

A.8 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.

A.9 Model Extension: Firm Switching Costs

In this appendix section, I present a dynamic model of firm choice that illustrates how mobility frictions might give rise to employer market power. This model cannot be estimated on data, since I do not have access to the identities of the firms at which the workers in my sample are employed.

However, I can use parameter estimates from my empirical model and combine those estimates with results in the literature to calibrate plausible parameter values.

The setting I use fuses together the “new classical monopsony” literature discussed in [Card et al. \(2018\)](#), [Lamadon, Mogstad, and Setzler \(2019\)](#), [Azar, Berry, and Marinescu \(2019\)](#) and [Manning \(Forthcoming\)](#) with the “modern monopsony” literature discussed in [Manning \(2003\)](#), [Hirsch et al. \(2019\)](#) and [Manning \(Forthcoming\)](#). In “new classical” models, workers have idiosyncratic tastes for wage and non-wage amenities offered by firms. Only a small number of firms offer similar bundles of wages and amenities. This then generates monopsony power, since only a small number of firms are “comparable” from the worker’s perspective. In “modern” models, workers and firms cannot immediately match, and the time it takes for the match to resolve gives employers some monopsony power.

The following model is meant to show how the presence and extent of switching costs can convey some monopsony power to firms. It is not based on actual data and is not meant to be a model of the US labor market. Rather, it is a model that has some basis in previous studies, and that has some relationship to the empirical model presented in this paper. Where this model departs from my empirical model is in modeling worker-firm matching. My empirical model abstracts from workers’ choice over firms, so this model attempts to bridge that gap.

In the model, workers have idiosyncratic tastes over firms’ wage and non-wage amenities. Workers also face a cost to switching firms, which acts as a market friction. Together, these two features both generate monopsony power to the firm. I focus on both features because my empirical model presented earlier in this paper incorporates both.

A.9.1 Firm productivity and worker preferences

Suppose that there are F firms in the economy and L identical workers. Because workers are identical, each worker i at firm f is paid an identical wage. Firms may differ either in the wages or non-wage amenities they offer. The degree of dispersion across firms in wages and amenities creates variation in firm market shares, and hence, monopsony power. Workers’ preferences for non-wage amenities also confers market power.

I now detail the model, which resembles [Card et al. \(2018\)](#) and [Lamadon, Mogstad, and Setzler \(2019\)](#), but adds switching costs (and, hence, dynamics) to the worker’s decision.

Firms are endowed with time-varying productivity and permanent amenities. Firms post wages based on their productivity. In keeping with the theory of compensating differentials, amenities and wages are negatively correlated in the population of firms. Workers are assumed to be able to observe each firm's wages and non-wage amenities before making a decision about where to work, and firms are assumed to know workers' preferences over each firm's wage and amenity bundle. Firms are also assumed to know the value of workers' switching costs. Finally, firms are assumed to not be able to strategically interact when setting their wages.

Workers have preferences over wages and amenities. They also have idiosyncratic preferences so that i 's preference for working at firm f in period t is given by

$$U_{ift} = \alpha_f + \gamma w_{ft} + \theta 1[d_{it-1} \neq f] + \eta_{ift} \quad (\text{A.51})$$

$$= u_{ift} + \eta_{ift} \quad (\text{A.52})$$

where w_{ft} is the natural logarithm of the wage posted (and paid) by firm f , $1[\cdot]$ is the indicator function, d_{it-1} is i 's decision in the previous period, and η_{ift} is an idiosyncratic taste shock. The parameter γ measures how responsive workers are to higher wages offered, α measures preferences for non-wage amenities, and θ represents the utility cost of switching firms. This switching cost could arise from geographical or industrial distance, ties to the current firm or its environs, or from psychologically having to adjust to a new firm or geographical location. Switching costs may also increase with distance. In what follows, I let θ be a 2-dimensional vector, where the first element, labeled θ_{move} , corresponds to the utility cost of switching geographical markets, while the second element, θ_{switch} , gives the utility cost of switching firms within a market.

Suppose that workers discount the future with discount factor β and that firm wages evolve according to an AR(1) process where $w_{ft} = \rho w_{ft-1} + \xi_{ft}$. Workers' conditional value function (i.e. present discounted value of utility) for choosing to work at f in period t is then

$$v_{ift} = u_{ift} + \beta \ln \sum_h \exp(v_{iht+1}) \quad (\text{A.53})$$

assuming that the η_{ift} 's are distributed Type I extreme value.

The probability that i chooses to work at f is given by

$$P_{ift} = \frac{\exp(v_{ift})}{\sum_h \exp(v_{iht})} \quad (\text{A.54})$$

Employment at firm f in period t is then given by

$$\begin{aligned} N_{ft} &= \sum_i P_{ift} L \\ &= s_{ft} L \end{aligned} \quad (\text{A.55})$$

where s_{ft} is the market share of firm f in time t and I assume for ease of exposition that L is fixed over time and with respect to wages and non-wage amenities.

A.9.2 Employment transitions and wage elasticities

Following [Hirsch et al. \(2019\)](#), consider employment transitions into and out of firm f . The change in employment will be equal to the number of new recruits minus the quit rate of existing employees.

$$N_{ft+1}(\cdot) - N_{ft}(\cdot) = R_{ft+1}(\alpha, \gamma, w_t, \theta) - q_{ft+1}(\alpha, \gamma, w_t, \theta) N_{ft}(\cdot) \quad (\text{A.56})$$

where R denotes the number of recruits and q denotes the quit rate. w_t here denotes the *entire* set of posted wages at all firms and α denotes the entire set of firm amenities.

Equation (A.56) shows that, in a steady state,

$$\begin{aligned} 0 &= R_{ft+1}(\alpha, \gamma, w_t, \theta) - q_{ft+1}(\alpha, \gamma, w_t, \theta) N_{ft}(\cdot) \\ N_{ft}(\alpha, \gamma, w_t, \theta) &= \frac{R_{ft+1}(\alpha, \gamma, w_t, \theta)}{q_{ft+1}(\alpha, \gamma, w_t, \theta)} \end{aligned} \quad (\text{A.57})$$

This equation then gives the labor supply elasticity to the firm as

$$\varepsilon_{NW} = \varepsilon_{RW} - \varepsilon_{qW} \quad (\text{A.58})$$

A.9.3 Computing ε_{NW} from employment transitions

Let Ψ_t be the $F \times F$ Markov transition matrix, where the (f, f') element reports the probability of transition to firm f' from firm f at time t . Ψ_t is a function of $(\alpha, \gamma, w_t, \theta)$ but I suppress this for ease of exposition.

The quit rate. The quit rate from firm f , which appears in (A.58), can be written in terms of the Markov transition matrix as follows:

$$q_{ft}(\alpha, \gamma, w_t, \theta) = \sum_{f' \neq f} \Psi_{t(f, f')} \quad (\text{A.59})$$

where $\Psi_{t(f, f')}$ denotes the (f, f') element of Ψ_t .

I can compute ε_{qW} by taking the derivative of $q_{ft}(\cdot)$ with respect to w_{ft} . That is, consider the situation where firm f raises its wage by a very small amount (e.g. .01 log points), but all other firms keep their wage the same, and then see how much lower the quit rate from f is as a result. I perform this calculation as

$$\begin{aligned} \frac{\partial q_{ft}(\alpha, \gamma, w_t, \theta)}{\partial w_{ft}} &\approx \frac{\sum_{f' \neq f} \Psi_{t(f, f')}(w_{ft} + .01, \cdot) - \sum_{f' \neq f} \Psi_{t(f, f')}(w_{ft}, \cdot)}{.01} \\ &\equiv \varepsilon_{qW, ft} \end{aligned} \quad (\text{A.60})$$

where $\Psi_{t(f, f')}(w_{ft} + .01, \cdot)$ and $\Psi_{t(f, f')}(w_{ft}, \cdot)$ indicate that $\Psi_{t(f, f')}$ is evaluated at $w_{ft} + .01$ or w_{ft} , holding all other arguments constant.

Note that $\varepsilon_{qW, ft}$ is specific to f and t . Each firm's $\varepsilon_{qW, ft}$ depends on its wage and non-wage endowment. Henceforth, I report the mean of the distribution of $\varepsilon_{qW, ft}$.

Number of recruits. The number of recruits coming to firm f , which appears in (A.58), can be written in terms of the Markov transition matrix as follows:

$$R_{ft}(\alpha, \gamma, w_t, \theta) = \sum_{f' \neq f} N_{f't} \Psi_{t(f', f)} \quad (\text{A.61})$$

where all terms are as defined previously.

I can then compute ε_{RW} in the same way as ε_{qW} .

A.9.4 How does ε_{NW} change with moving costs?

To assess how ε_{NW} changes with moving costs, I simulate my model under a number of different parameter values for $(\alpha, \gamma, w_t, \theta)$. I set the number of markets at 35, in accordance with the empirical model estimated earlier in the paper. I then examine (i) what is the value of ε_{NW} that roughly corresponds with the level of switching costs that match the rate of moving and job turnover observed in the SIPP? and (ii) how would ε_{NW} change if these switching costs were to change?

It is important to point out that this simulation is underidentified and thus requires some parameters to be calibrated. There are six parameters to choose, but only three empirical moments to match. The six parameters are:

1. degree of cross-firm wage dispersion, $\text{Var}(w)$
2. degree of cross-firm amenity dispersion, $\text{Var}(\alpha)$
3. degree of correlation between wages and amenities, $\text{Corr}(\alpha, w)$
4. preference intensity of wages, γ
5. magnitude of market moving costs, θ_{move}
6. magnitude of firm switching costs, θ_{switch}

The three SIPP empirical moments to match are the annualized migration rate (3.4%) and annualized job switching rate (21.1%), as well as the overall wage variance (0.27). Estimates from [Lamadon, Mogstad, and Setzler \(2019\)](#) indicate that roughly one-third of the variation in wages in the US is due to across-firm wage dispersion. Thus, a reasonable value for $\text{Var}(w)$ would be 0.1. Similarly,

While I can set γ to be the $\hat{\gamma}_0$ estimated in the empirical model, it is unclear how this estimate would change if the empirical model had included firm choice (as opposed to strictly locational choice). Thus, I present results corresponding to several values of γ , including the value estimated in my empirical model.

The results are reported in Table [A13](#) for an economy with $F = 700$ firms and 35 different geographical markets. As mentioned previously, the switching cost θ consists of two components: θ_{switch} which is a cost of switching firms, and θ_{move} which is a cost of moving to a different

geographical market. I choose values of γ and the two θ 's that roughly match the overall migration and job switching rates in my subsample the SIPP (quoted above).

The rows of Table A13 are divided into three groups, corresponding to the imposed value of γ . Within each group, I present implied values of ε_{NW} corresponding to five different scenarios: (i) if switching costs were infinite; (ii) if switching costs were such that the firm switching probability resembles what is in the SIPP; (iii) if switching costs were smaller than the status quo, such that the firm switching probability were roughly double what it is in the SIPP; (iv) if within-market switching costs were zero, and (v) if all switching costs were zero.

Regardless of the value of γ , $\varepsilon_{NW} = 0$ when switching costs are infinite. When switching costs are set to the level that resembles the amount of switching in the SIPP, this implies an elasticity of approximately 0.39γ . For example, when $\gamma = 1$, the model implies a very low average labor supply elasticity of 0.39, which corresponds to a $\frac{1}{1+0.39} = 72\%$ wage markdown. This number is a similar magnitude to that reported in Dube et al. (2020). Even if switching costs were lower such that the average switching rate is double what is observed in the SIPP, the elasticity increases to approximately 0.6γ . If there continue to be large inter-market switching costs, but no intra-market switching costs, the elasticity approaches γ . When there are no switching costs, the elasticity is equal to γ .

One remaining question regarding Table A13 is what the best value of γ is. In the estimates of my empirical model, $\hat{\gamma}_0$ is very close to unity. However, this estimate comes from a model with no aspect of within-location firm choice. It is reasonable to assume that workers are more responsive to wage differences across firms within than across locations. Thus, a more reasonable value for γ may be two or three.

Online Appendix References

- 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.
- 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.
- . 2019. “Nonstationary Dynamic Models with Finite Dependence.” *Quantitative Economics* 10 (3):853–890.
- Azar, José, Steven Berry, and Ioana Marinescu. 2019. “Estimating Labor Market Power.” Working paper, IESE Business School.
- Balgova, Maria. 2018. “Why Don’t Less Educated Workers Move? The Role of Job Search in Migration Decisions.” Working paper, Oxford University.
- Baum-Snow, Nathaniel and Ronni Pavan. 2012. “Understanding the City Size Wage Gap.” *Review of Economic Studies* 79 (1):88–127.
- Bishop, Kelly. 2012. “A Dynamic Model of Location Choice and Hedonic Valuation.” Working paper, Washington University in St. Louis.
- Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. “Firms and Labor Market Inequality: Evidence and Some Theory.” *Journal of Labor Economics* 36 (S1):S13–S70.
- Coate, Patrick. 2013. “Parental Influence on Labor Market Outcomes and Location Decisions of Young Workers.” Working paper, Duke University.
- 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).
- Dube, Arindrajit, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. 2020. “Monopsony in Online Labor Markets.” *American Economic Review: Insights* 2 (1):33–46.

- 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.
- 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.
- 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.
- Hirsch, Boris, Elke J. Jahn, Alan Manning, and Michael Oberfichtner. 2019. Discussion Paper 1608, Center for Economic Performance.
- 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.
- James, Jonathan. 2012. “Learning and Occupational Sorting.” *Federal Reserve Bank of Cleveland Working Paper 12-25* .
- 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.
- Koşar, Gizem, Tyler Ransom, and Wilbert van der Klaauw. Forthcoming. “Understanding Migration Aversion Using Elicited Counterfactual Choice Probabilities.” *Journal of Econometrics* .
- Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2019. “Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market.” Working Paper 25954, National Bureau of Economic Research.
- Magnac, Thierry and David Thesmar. 2002. “Identifying Dynamic Discrete Decision Processes.” *Econometrica* 70 (2):801–816.
- Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.
- . Forthcoming. “Monopsony in Labor Markets: A Review.” *ILR Review* .
- McFadden, Daniel. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontiers in Econometrics*, edited by Paul Zarembka. Academic Press, New York, 105–142.

- 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.
- 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.
- Schluter, Christian and Guillaume Willems. 2018. "A Dynamic Empirical Model of Frictional Spatial Job Search." Working paper, Aix-Marseille University.
- Schmutz, Benoît and Modibo Sidibé. 2019. "Frictional Labour Mobility." *Review of Economic Studies* 86 (4):1779–1826.
- Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference." *Quarterly Journal of Economics* 133 (3):1331–1393.
- Sullivan, Paul and Ted To. 2014. "Search and Nonwage Job Characteristics." *Journal of Human Resources* 49 (2):472–507.
- Taber, Christopher and Rune Vejlin. 2020. "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica* 88 (3):1031–1069.
- 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.

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

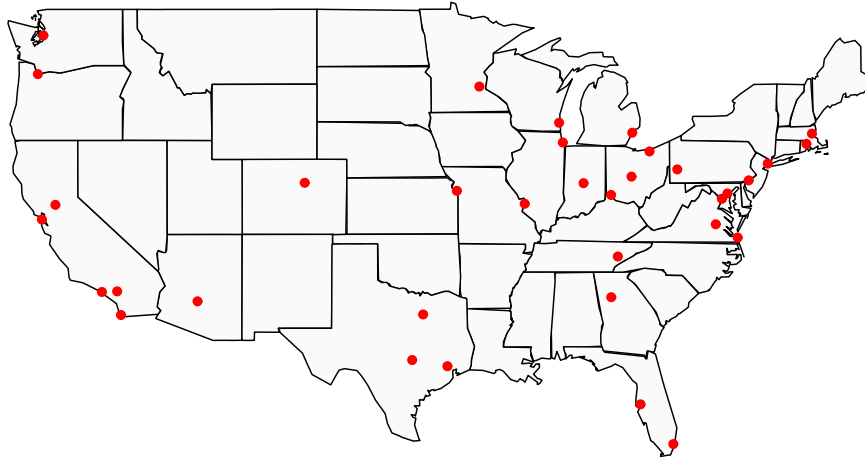
Data	Source	Years
Earnings and location & employment transitions	SIPP, 2004 and 2008 Panels	2004–2013
CBSA population	Census Bureau	2000
County unemployment rate	Bureau of Labor Statistics (BLS)	1990–2013
Local price level	American Chamber of Commerce Researchers Assoc. (ACCRA)	1990–2008

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

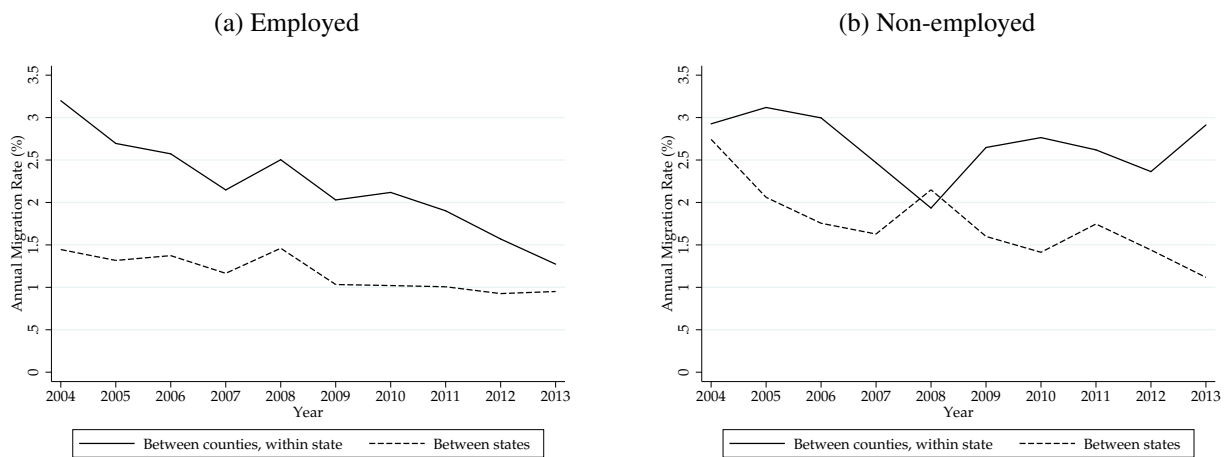


Note: Dots correspond to CBSA centroids of cities that are included in the model.

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: Annual migration rates by lagged employment status and migration distance for conventional definitions of employment and labor force participation



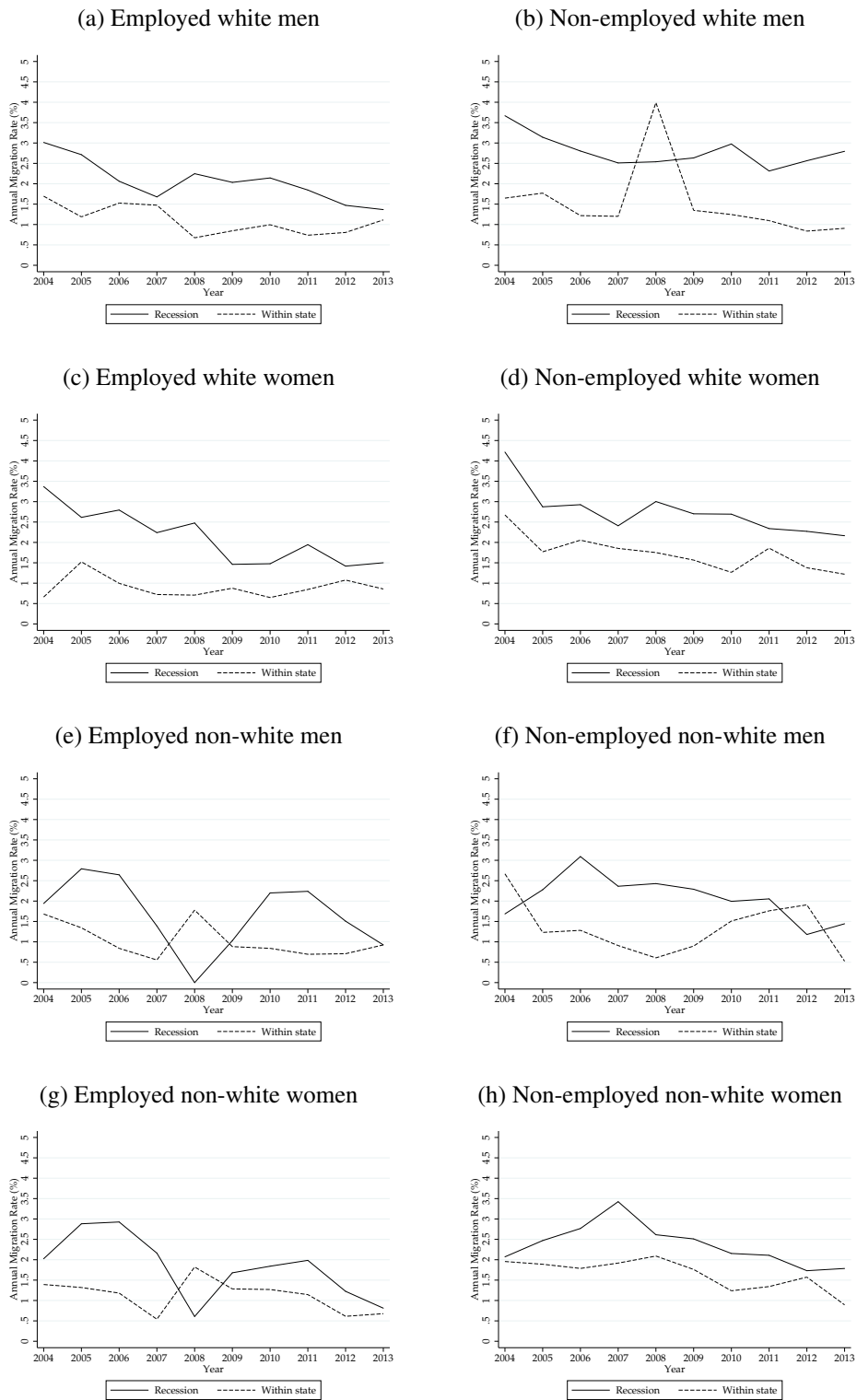
Source: 2004 and 2008 Panels of the Survey of Income and Program Participation. Figures include all non-college graduates aged 18-55 who have completed their schooling. Employment is defined as any amount of employment. Compare with Figure 1.

Table A6: Robustness of stylized facts to a more conventional definition of employment and labor force participation

Variable	Prev. employed		Prev. non-employed	
	Coeff	Std Err	Coeff	Std Err
Constant	0.8570***	0.0041	0.2337***	0.0084
Experience	0.0096***	0.0003	0.0060***	0.0006
Experience ² /100	-0.0168***	0.0007	-0.0113***	0.0016
Lagged state unempl. rate	-0.0044***	0.0004	-0.0049***	0.0009
Mover dummy	-0.1073***	0.0047	0.0948***	0.0083
Race × gender dummies	✓		✓	
Observations	120,748		40,633	

Notes: Compare with Table 3, which uses a less conventional definition of labor force participation and employment. Results are estimates from a pair of linear probability models where the dependent variable is an indicator for being employed at all in the current period. Sample includes all non-college graduates aged 18-55 in the 2004 and 2008 panels of the public-use SIPP who have completed their schooling. *** p<0.01; ** p<0.05; * p<0.10.

Figure A3: Annual migration rates by lagged employment status and migration distance for various demographic sub-groups



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. Compare with Figure 1.

Table A7: Robustness of stylized facts to non-college-educated demographic sub-groups

Variable	Prev. employed		Prev. non-employed	
	Coeff	Std Err	Coeff	Std Err
<i>Panel A: Non-Hispanic White Men</i>				
Constant	0.7516***	0.0104	0.2331***	0.0115
Experience	0.0115***	0.0008	0.0081***	0.0009
Experience ² /100	-0.0193***	0.0017	-0.0190***	0.0022
Lagged state unempl. rate	-0.0057***	0.0009	-0.0075***	0.0013
Mover dummy	-0.0936***	0.0113	0.0799***	0.0157
Observations	37,237		23,112	
<i>Panel B: Non-Hispanic White Women</i>				
Constant	0.6820***	0.0131	0.1078***	0.0076
Experience	0.0117***	0.0010	0.0044***	0.0006
Experience ² /100	-0.0168***	0.0022	-0.0019	0.0017
Lagged state unempl. rate	-0.0026**	0.0011	-0.0047***	0.0009
Mover dummy	-0.1627***	0.0148	0.0356***	0.0103
Observations	26,249		33,239	
<i>Panel C: Non-White Men</i>				
Constant	0.6648***	0.0156	0.1592***	0.0127
Experience	0.0147***	0.0012	0.0119***	0.0010
Experience ² /100	-0.0264***	0.0029	-0.0278***	0.0030
Lagged state unempl. rate	0.0002	0.0015	-0.0056***	0.0016
Mover dummy	-0.088***	0.0213	0.0435**	0.0211
Observations	13,827		12,965	
<i>Panel D: Non-White Women</i>				
Constant	0.6845***	0.0159	0.1337***	0.0111
Experience	0.0132***	0.0013	0.0106***	0.0009
Experience ² /100	-0.0231***	0.0030	-0.0218***	0.0027
Lagged state unempl. rate	-0.0019	0.0015	-0.0046***	0.0014
Mover dummy	-0.1558***	0.0208	0.0281	0.0175
Observations	14,155		15,502	

Notes: Compare with Table 3, which pools all demographic groups in the SIPP. Results are estimates from pairs of linear probability models where the dependent variable is an indicator for being employed at all in the current period. Sample includes non-college graduates aged 18-55 in the 2004 and 2008 panels of the public-use SIPP who have completed their schooling. *** p<0.01; ** p<0.05; * p<0.10.

Table A8: Employment probability equation estimates, with and without controlling for marital status

Variable	No Control for Marital Status				Control for Marital Status			
	Prev. employed		Prev. non-employed		Prev. employed		Prev. non-employed	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
Constant	1.3056***	0.2220	0.2566	0.2237	1.2220***	0.2224	0.2361	0.2257
Experience	0.0858***	0.0091	0.0359***	0.0086	0.0715***	0.0093	0.0141	0.0088
Experience ² /100	-0.1228***	0.0208	-0.0285	0.0219	-0.0991***	0.0211	0.0070	0.0222
Lagged local unempl. rate	-0.0314***	0.0104	-0.0922***	0.0110	-0.0303***	0.0104	-0.0929***	0.0111
Mover dummy	-0.9257***	0.1280	0.1929	0.1557	-0.9325***	0.1282	0.2331	0.1569
Location fixed effects	✓		✓		✓		✓	
Marital status dummy					✓		✓	
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. The first four columns coincide with results presented in Table 4. *** p<0.01; ** p<0.05; * p<0.10.

Table A9: Characteristics of destination location given moving cost subsidy to unemployed workers in various origin cities, year 2007

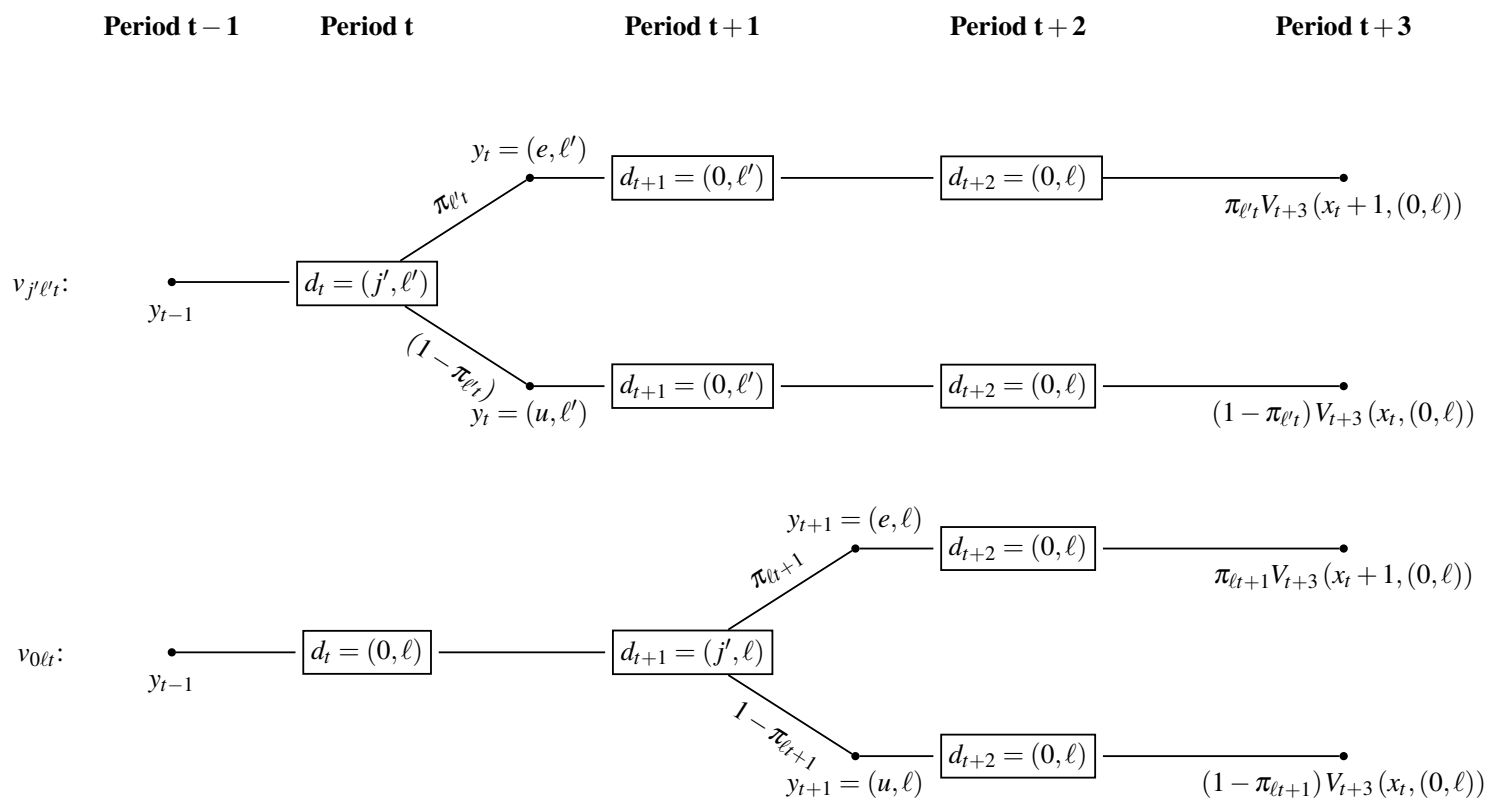
(a) Amenities								
	Type 1 workers				Type 2 workers			
	high amenity city		low amenity city		high amenity city		low amenity city	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
constant	0.3405	0.1790	0.4587	0.2750	0.3000	0.1643	0.4205	0.2643
amenities	0.0284*	0.0086	0.0436*	0.0132	0.0261*	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

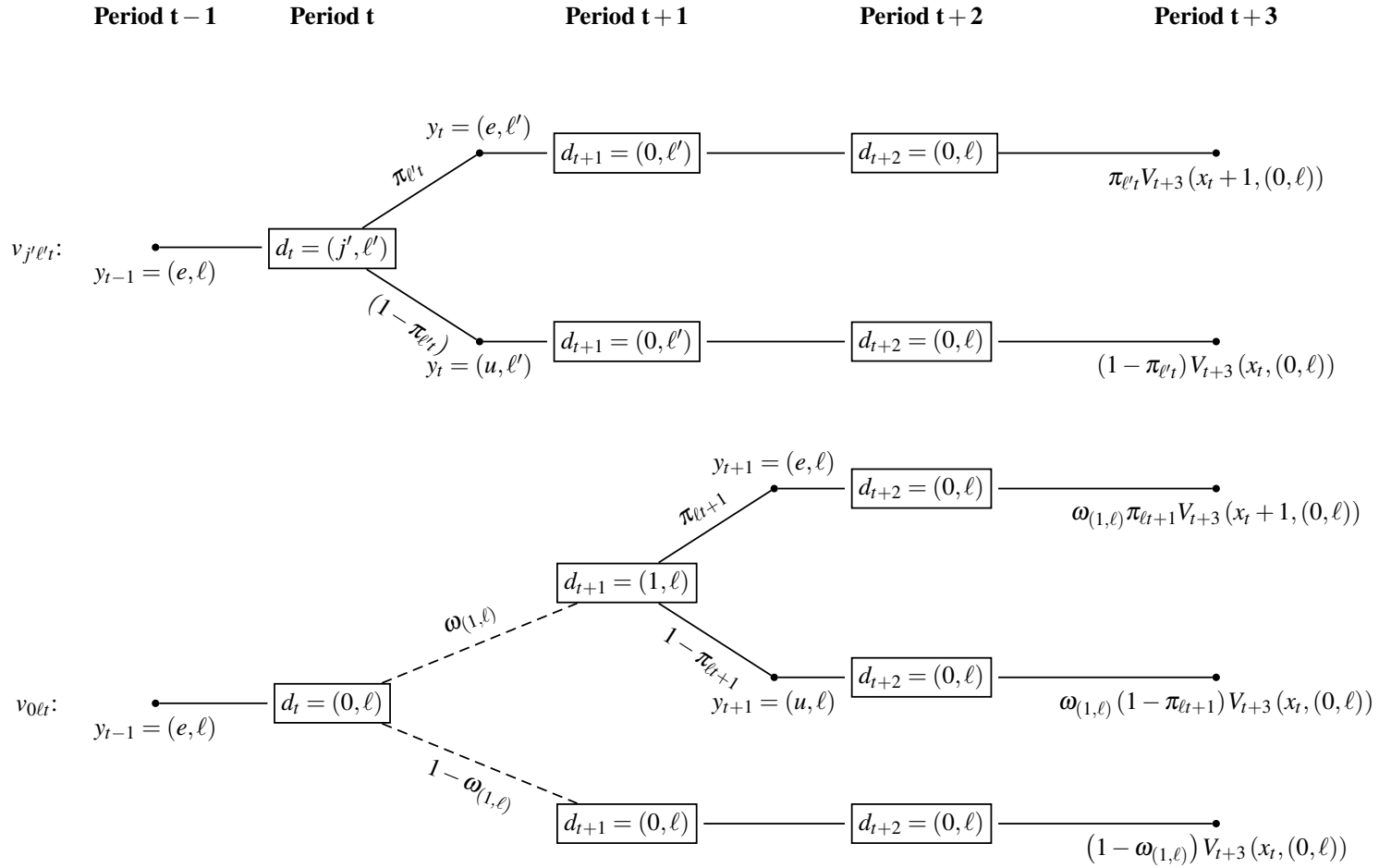
Figure A4: Finite dependence paths conditional on y_{t-1}



A45

Note: This figure depicts the evolution of the state space given the finite dependence paths described in Section A.3.2. 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 A5: Expanded finite dependence paths conditional on y_{t-1}



A46

Note: This figure depicts the evolution of the state space given the finite dependence paths described in Section A.3.2. Cancellation of the future value terms occurs when $\omega_{(1,\ell)} = \frac{\pi_{\ell t}}{\pi_{\ell t+1}}$, as described in Equations (A.27) and (A.28).

Table A10: 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 A11: Earnings forecasting estimates

Parameter	Symbol	Mean	Std Dev
<i>Earnings</i>			
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 A12: Determinants of local labor market attributes

(a) Amenities, earnings, and employment levels

Variable	Amenities (α_ℓ)		Earnings (w_ℓ)		Job destruction (δ_ℓ)		Job offer (λ_ℓ)	
	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	

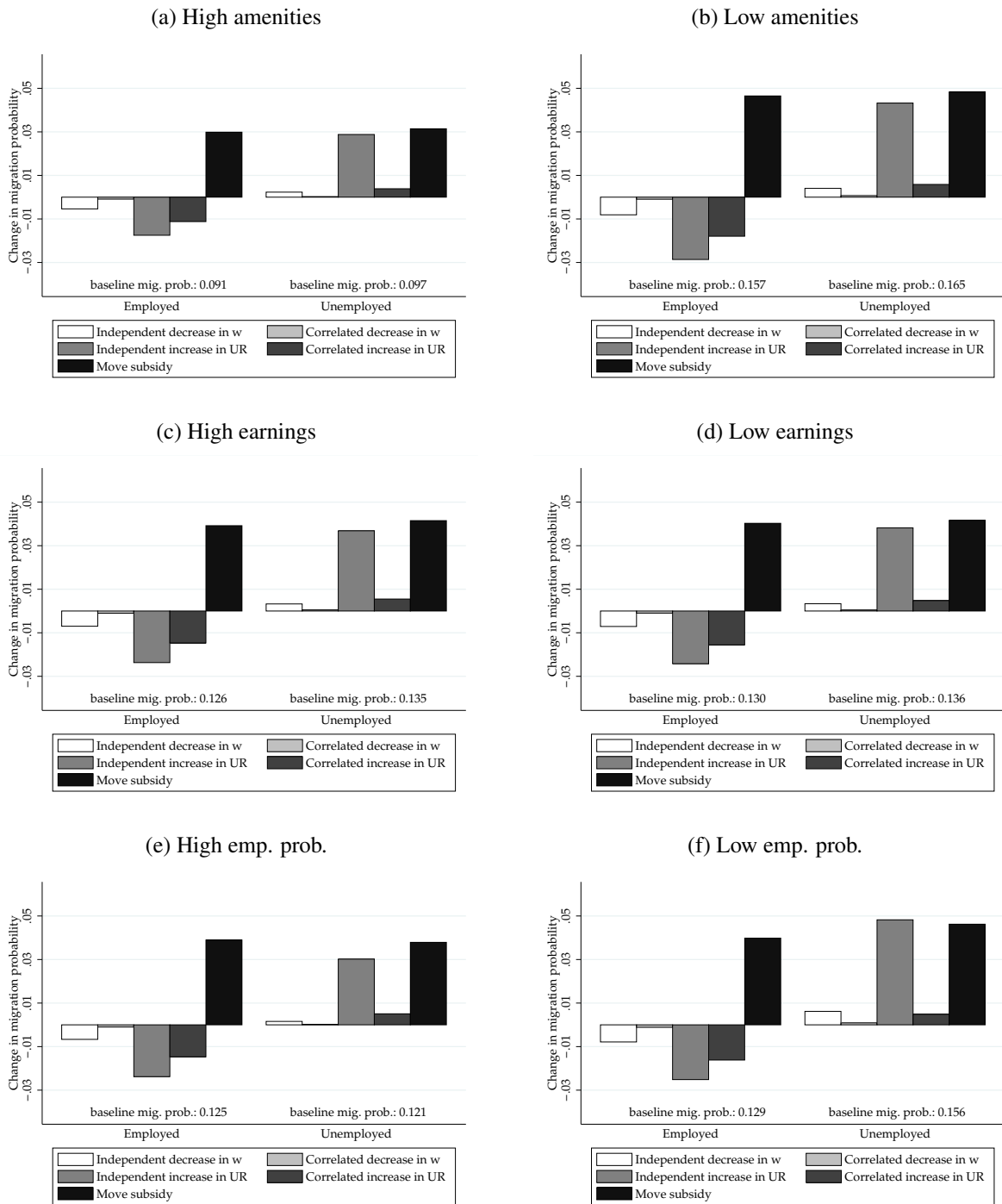
Notes: Each column is a separate regression with 350 observations (35 cities, 10 time periods) of the corresponding model parameter on the log population of the location and Census division dummies. Amenities and AR(1) shock standard deviations do not vary over time, so these regressions have 35 observations.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

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

Variable	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})	
	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	

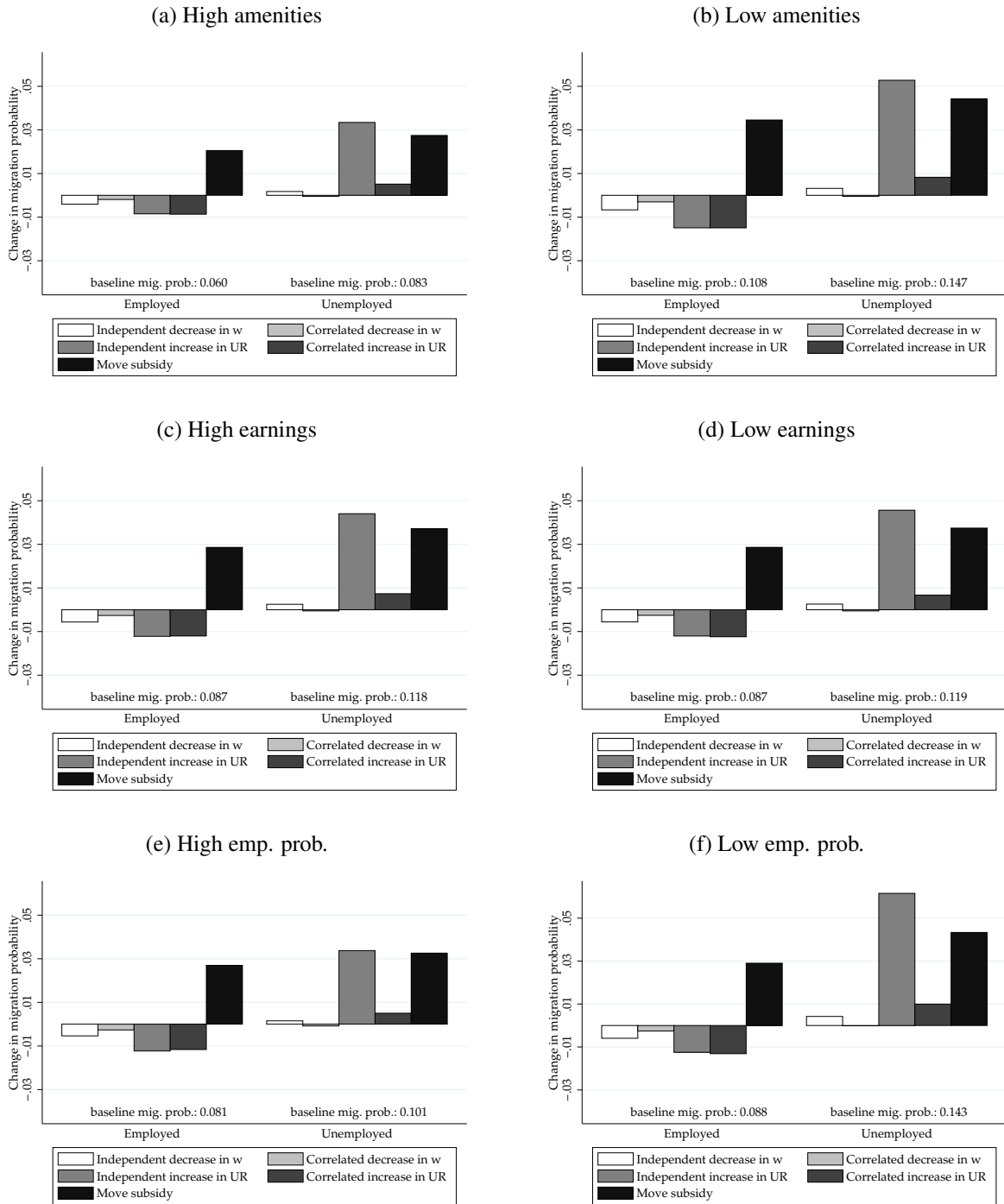
Notes: “UR” denotes unemployment rate. Each column is a separate regression with 350 observations (35 cities, 10 time periods) of the corresponding model parameter on the log population of the location and Census division dummies. Amenities and AR(1) shock standard deviations do not vary over time, so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

Figure A6: Counterfactual change in migration for Type 1 individuals, by origin city and prior employment status



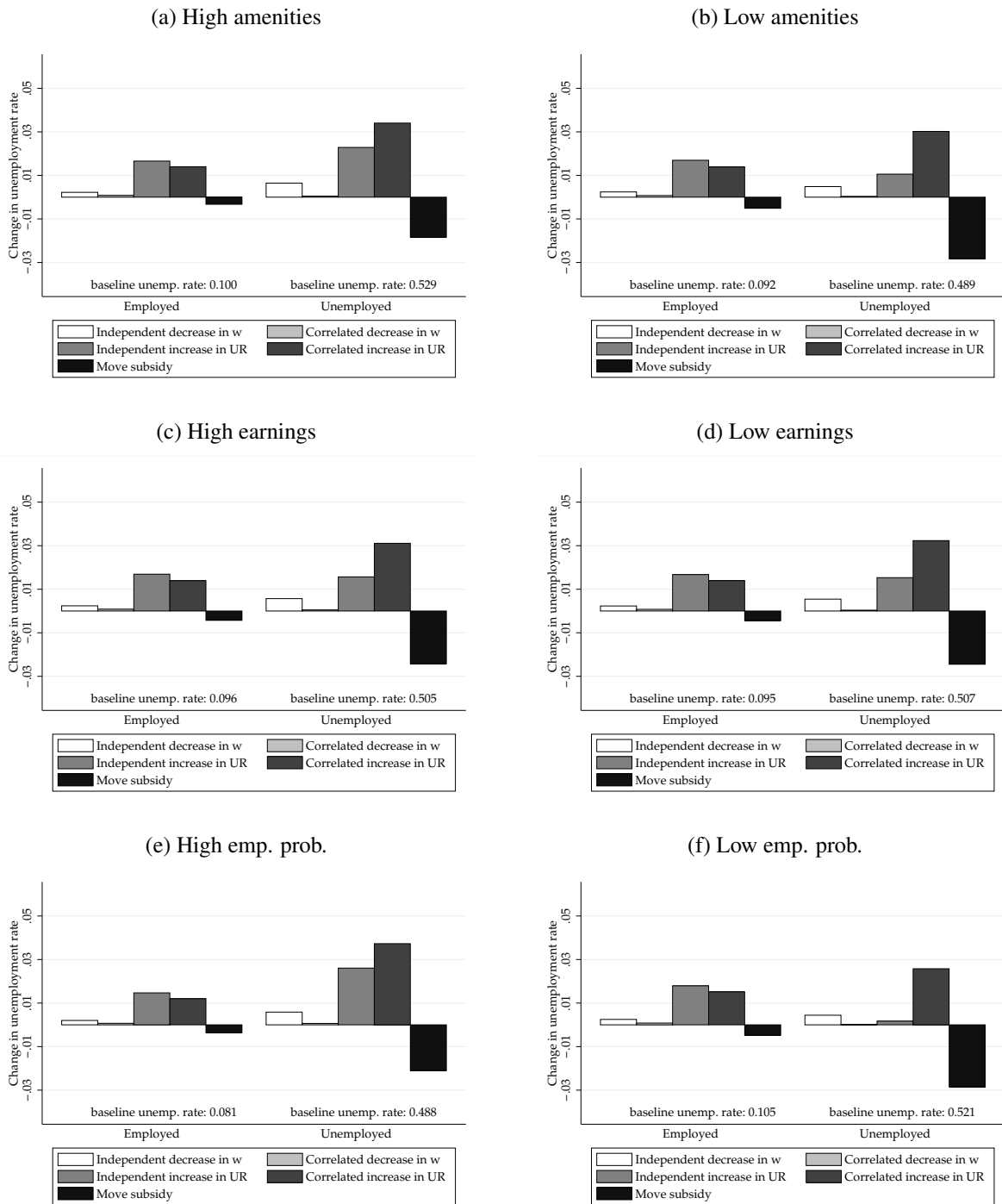
Notes: Each panel corresponds to a different origin city. Bar heights refer to the change in the out-migration rate from the specified location in response to the listed counterfactual. All figures are for 25-year-olds who were not born in the origin location. “high” refers to a location in the 75th 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 A7: Counterfactual change in migration for Type 2 individuals, by origin city and prior employment status



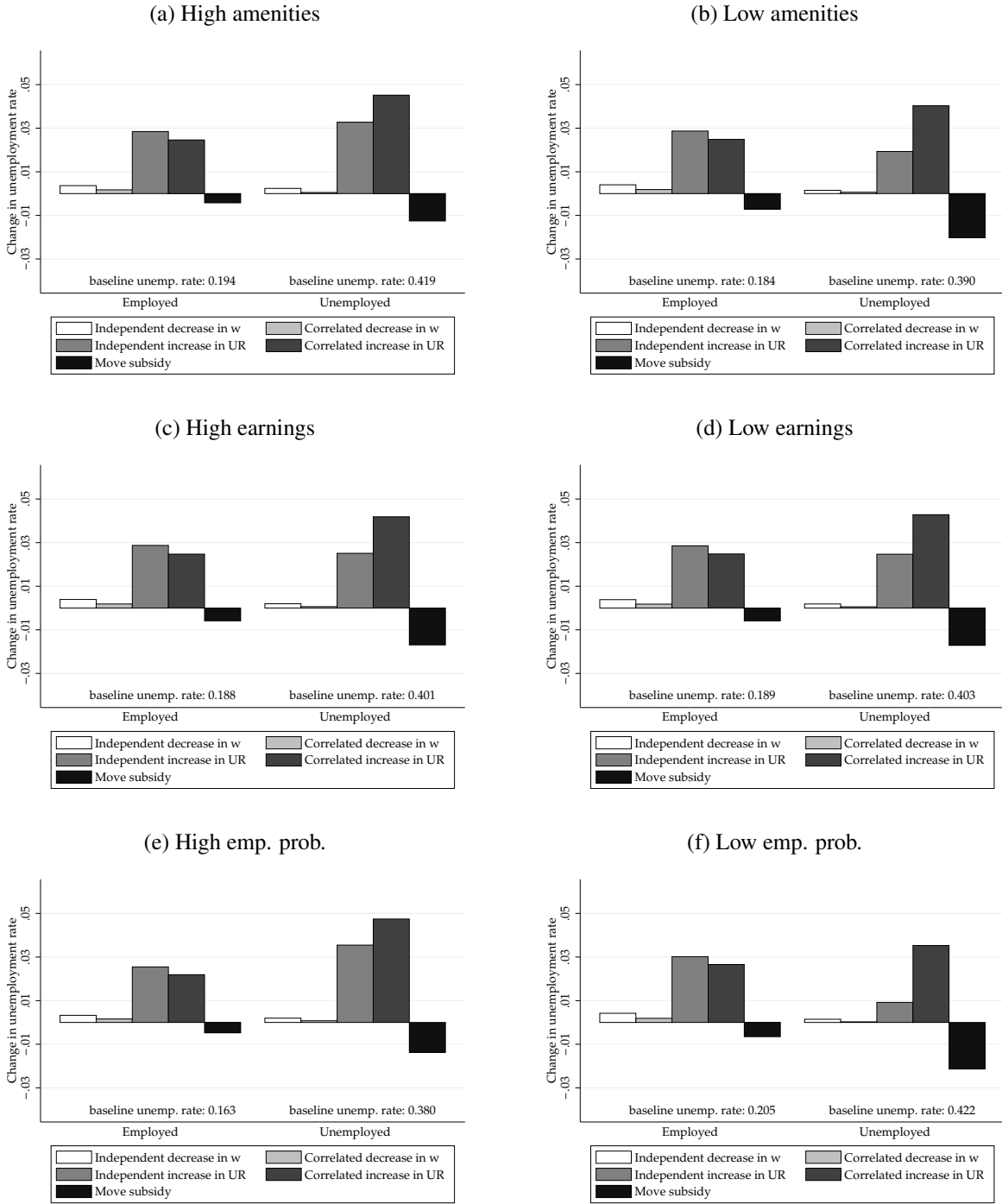
Notes: See notes to Figure 2.

Figure A8: Counterfactual change in unemployment rate for Type 1 individuals, by origin city and prior employment status



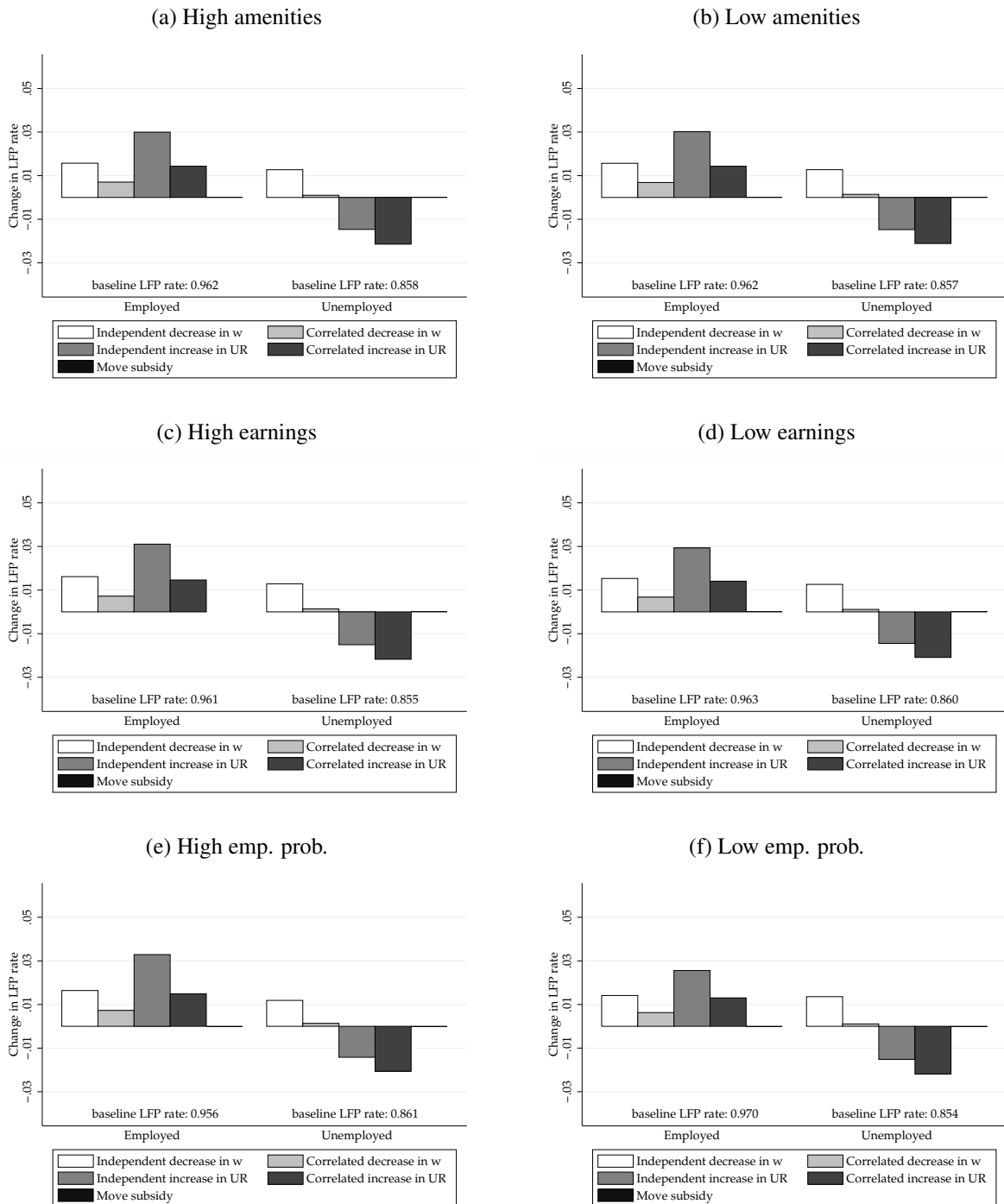
Notes: Each panel corresponds to a different origin city. Bar heights refer to the change in the out-migration rate from the specified location in response to the listed counterfactual. All figures are for 25-year-olds who were not born in the origin location. “high” refers to a location in the 75th 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 A9: Counterfactual change in unemployment rate for Type 2 individuals, by origin city and prior employment status



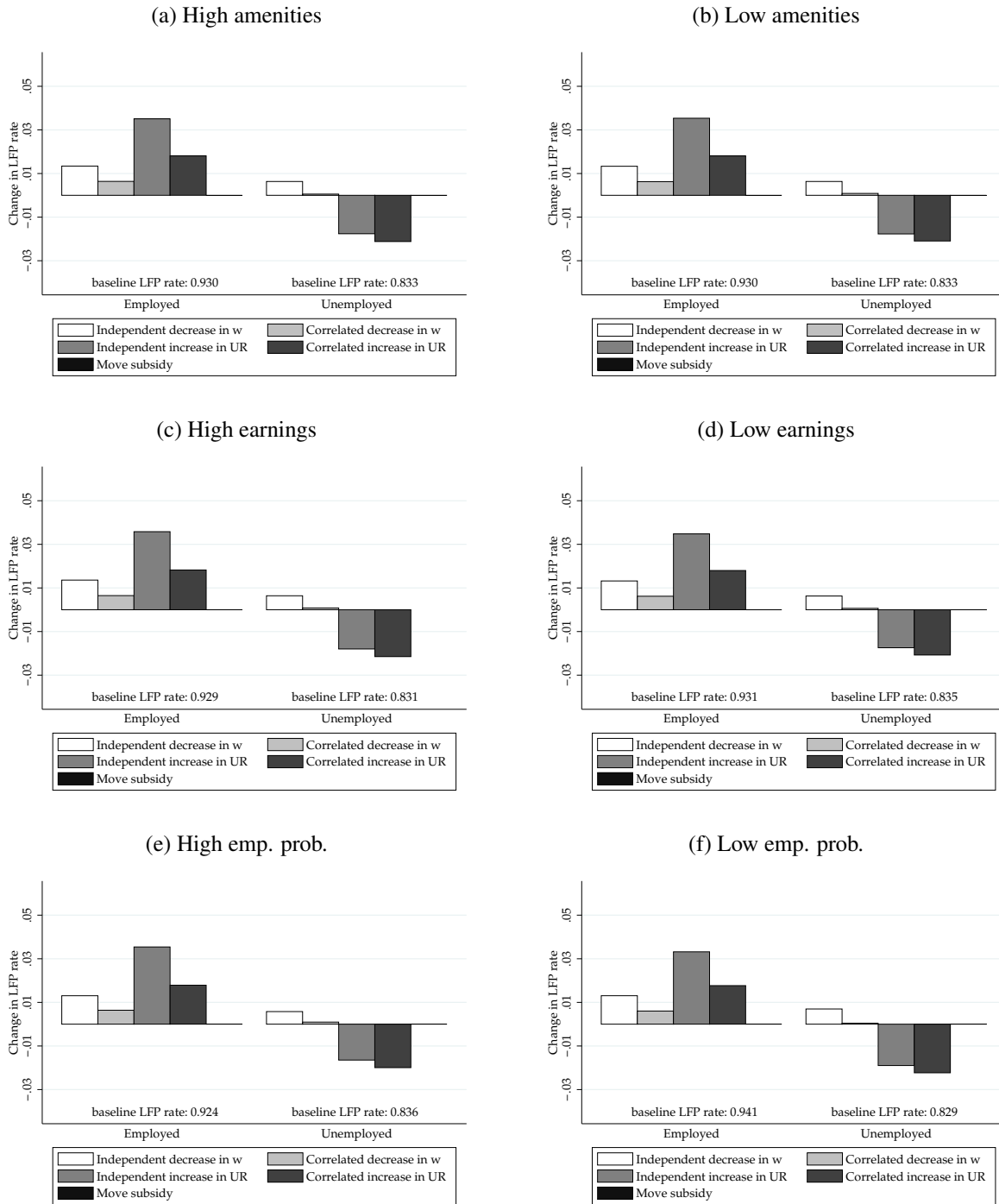
Notes: See notes to Figure 2.

Figure A10: 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 A11: Counterfactual change in labor force participation rate for Type 2 individuals, by origin city and prior employment status



Notes: See notes to Figure 2.

Table A13: How elasticity of labor supply (ϵ_{NW}) changes with switching costs

No. firms	No. markets	σ_w^2	σ_α^2	$\text{Cov}(w, \alpha)$	γ	θ_{move}	θ_{switch}	Mean(ϵ_{NW})	Std(ϵ_{NW})	Migration rate (%)	Job Switch rate (%)
700	35	0.10	0.14	-0.07	1	$-\infty$	$-\infty$	0.00	0.00	0.00	0.00
700	35	0.10	0.14	-0.07	1	-5	-4.5	0.39	0.04	4.28	22.1
700	35	0.10	0.14	-0.07	1	-5	-3.5	0.66	0.12	8.35	42.9
700	35	0.10	0.14	-0.07	1	-6	0	0.95	0.92	7.80	95.4
700	35	0.10	0.14	-0.07	1	0	0	1.00	0.32	97.1	99.9
700	35	0.10	0.14	-0.07	2	$-\infty$	$-\infty$	0.00	0.00	0.00	0.00
700	35	0.10	0.14	-0.07	2	-5.5	-4.5	0.78	0.15	2.90	22.7
700	35	0.10	0.14	-0.07	2	-5	-3.5	1.34	0.42	8.75	44.9
700	35	0.10	0.14	-0.07	2	-6	0	1.89	1.74	7.85	95.4
700	35	0.10	0.14	-0.07	2	0	0	2.00	1.03	97.1	99.9
700	35	0.10	0.14	-0.07	3	$-\infty$	$-\infty$	0.00	0.00	0.00	0.00
700	35	0.10	0.14	-0.07	3	-5	-5	1.04	0.38	3.95	20.3
700	35	0.10	0.14	-0.07	3	-5	-4	1.73	0.82	7.46	38.2
700	35	0.10	0.14	-0.07	3	-6	0	2.79	2.29	7.97	95.4
700	35	0.10	0.14	-0.07	3	0	0	3.00	2.57	97.1	99.9

Notes: σ_w^2 is the variance in wages across firms, σ_α^2 the variance of amenities across firms, and $\text{Cov}(w, \alpha)$ the covariance between the two.

γ measures wage responsiveness of workers

θ_{move} is a cost of switching markets (i.e. a moving cost) while θ_{switch} is a cost of switching firms (regardless of whether the switch is within or across markets).

In the SIPP, the migration rate is 3.4% and the job switching rate is 21.1%.

Table A14: Greek symbol notation glossary

Greek symbol	Equation of first reference	Description
α	(A.9)	Local amenities
β	(3.1)	Discount factor
γ	(A.9)	Flow utility parameters
$\bar{\gamma}$	(A.22)	Euler's constant
δ	(A.3)	Job destruction probability
ε	(A.6)	Preference shocks
ζ	(A.2)	Shocks to evolution of earnings parameters
η	(A.1)	Earnings measurement error
θ	(A.13)	Moving and switching cost parameters
λ	(A.3)	Job offer probability
μ	(A.15)	Parameters in estimation of employment probabilities
ξ	(A.4)	Shocks to evolution of unemployment rate
π	(A.7)	Employment probabilities
$\bar{\pi}$	(A.34)	Population unobserved type probabilities
ρ	(A.2)	Parameters governing evolution of earnings parameters
σ_{ζ}	(A.2)	Std deviation of shocks to evolution of earnings parameters
σ_{η}	(A.1)	Std deviation of earnings measurement error
σ_{ξ}	(A.4)	Std deviation of shocks to evolution of unemployment rate
ϕ	(A.4)	Parameters governing evolution of employment probabilities
ϕ^h	(A.38)	Density of wage equation errors
ψ	(A.1)	Earnings parameters
ω	(A.26)	Value function weights
τ	(A.1)	Unobserved type (discrete)
Δ	(A.9)	Moving cost
Λ	(A.36)	Likelihood contribution of employment probabilities
Θ	(A.15)	Employment probability determinants
Ξ	(A.9)	Switching cost
Ψ	(A.31)	Covariance of local labor market shocks