Online Appendix to Geographic Mobility over the Life Cycle

Antonia Díaz[a] Álvaro Jáñez[b] Felix Wellschmied[c]

[a]ICAE, Universidad Complutense de Madrid, and CEPR
[b]Stockholm School of Economics
[c]Universidad Carlos III de Madrid

January 8, 2026

A Data details

Data aggregation and definitions in the Census: The Census reports the residence at the municipality level whenever a municipality has more than 20,000 inhabitants. We aggregate the data to *Urban Areas*, whose definition is similar to that of a commuting zone in the US and it is meant to represent the local economy where people work and live. Therefore, an urban area can consist of multiple municipalities that are close by. Urban areas represent 69% of Spain's total population, and 75% of the non-covered people live in rural municipalities with fewer than 20,000 inhabitants. The data provides 3,888,692 individual observations for the 1991 Census, 2,039,274 for the 2001 Census, and 4,107,465 for the 2011 Census.

We classify a person as employed in her current urban area when she reports holding a job.[1] The unemployed are those reporting to search for a job. Finally, those non-employed who report being retired, disabled, or have other reasons not to search for a job are classified as out of the labor force. Given this individual information, we compute the unemployment rate of an urban area as the total number of unemployed individuals relative to those in the labor force. The aggregate unemployment rate has large cyclical fluctuations in Spain. As we are interested in long-run decisions, we compute the time-averaged unemployment rate across the three Censuses at the urban area level.[2]

The 2001 and 2011 Censuses included a question on the location of residence during the previous

[1]We assume that all people are working in the urban areas where they live. According to the INE, less than 3% of workers were working from home in 2011. Moreover, according to the Ministry of Transport, Mobility, and Digital Agenda, the number of people whose commuting time was longer than 60 minutes comprised 3.7% of the workforce. 90.5% of the workforce needed less than 45 minutes to commute to work.

[2]The ranking of urban areas according to their unemployment rate is very stationary across censuses.

Census, i.e., 10 years ago. This allows us to compute decennial flows of people who flow into a specific urban area and have lived in a different urban area before, IN_{it}, as well as those who flow out from a specific urban area, OUT_{it}.[3] To compute rates, we use as convention the size of the urban area in the previous Census, i.e., the inflow rate of an urban area is the sum of all people who have arrived at that urban area over the period of 10 years relative to the size of the urban area at the beginning of that period: $IR_{it} = \frac{IN_{it}}{N_{it-1}}$, and $OR_{it} = \frac{OUT_{it}}{N_{it-1}}$.

Definitions in the MCVL: We identify the workplace of the individual using the contribution account codes of the firm, which allows us to identify municipalities with a population of more than 40,000 inhabitants.[4] We group municipalities in urban areas as we did with the Census samples. We exclude job spells of the Basque Country and Navarre residents as well as the self-employed, as the MCVL does not collect data on earnings for these individuals.[5] We also omit job spells in agriculture, fishing, forestry, mining, and extractive industries because their fiscal regime allows them to self-report earnings and the number of working days. Finally, we discard foreign workers because we do not have information about their employment history before migrating to Spain. Similarly, we omit workers born before 1962 as we do not have information on job spells before 1980. This selection results in 329,418 workers and 7,366,678 observations.

The MCVL provides two sources of income information for the reference year of each panel (2006-2008). First, annual uncoded earnings from tax administration records. Second, monthly top-coded earnings from Social Security records[6]. We allocate uncoded yearly earnings across months according to the fraction of top-coded earnings that the worker earns each month. In the monthly data, we regard a worker as employed whenever she has positive social security contributions. In the yearly data, we count a worker as employed when she contributes for at least six months in a year to Social Security. Finally, we define a worker's current employer using the ID of the job with the highest earnings. The employer identifier also allows us to identify job-to-job transitions.

[3]To this end, we include persons who move from and to municipalities who are not part of an urban area. Yet, our data still does not cover all people joining and leaving an urban area as it excludes deaths, those individuals who were younger than 16 years old in the previous Census, and those migrating from and to Spain.

[4]Since the data does not identify municipalities with fewer than 40,000 inhabitants, we have information on 78 out of the 86 existing Urban Areas. In particular, we do not identify the urban areas of Eivissa, La Orotava, Melilla, Ceuta, Blanes, Sant Feliu de Guíxols, Soria, and Teruel.

[5]However, we include Basque Country and Navarre residents when studying labor market transitions.

[6]The data contains top-coded monthly earnings used to calculate social security contributions since 1980. Because of the heavy censoring, we do not use that information in the baseline regressions.

B Appendix to Empirical Results

B.1 Urban area characteristics

Table B.1: Labor market characteristics of urban areas

	Gap relative to T3 (pp.)		**Baseline** (%)
	T1	**T2**	**T3**
Unemp. to emp. probability (U2E)	4.4	2.1	32.6
Emp. to unemp. probability (E2U)	-2.2	-1.6	9.8
Job-to-job probability (JTJ)	1.8	1.0	14.7
Share job-to-job down (JTJ down)	-4.3	-2.3	41.5

Sources: (a) Unemployment: Time-averaged values from the Census; (b) Flow rates: Time-averaged values from the MCVL 2006-2008. The "Employment to unemployment" rate is the share of non-employed workers who find a job in the next year. The "Unemployment to employment" rate is the share of employed workers who are non-employed in the next year. The "job-to-job" rate (J2J) is the share of employed workers who change employer, and "J2J down" reports the share of J2J that experience earnings losses.

Table B.2: Reduced-form evidence for labor market characteristics

	Probit AME (pp.)		**FE AME** (pp.)		**Baseline** (%)
	T1	**T2**	**T1**	**T2**	**T3**
Unemp. to emp. (U2E)	2.8^{***} (0.04)	0.6^{***} (0.04)	3.6^{***} (0.31)	2.0^{***} (0.30)	8.1
Emp. to unemp. (E2U)	-0.5^{***} (0.01)	-0.3^{***} (0.01)	-2.2^{***} (0.08)	-0.8^{***} (0.08)	2.5
Share JTJ down (JTJ down)	0.3^{***} (0.01)	0.02^{**} (0.01)	0.3^{***} (0.07)	0.1^{*} (0.07)	1.2
Share JTJ down (JTJ down)	-0.1 (0.28)	-0.9^{**} (0.33)	-1.7 (3.9)	-2.5 (4.3)	71.3

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Sources: (a) Unemployment: Time-averaged values from the Census; (b) Flow rates: Time-averaged values from the MCVL 2006-2008. The "Employment to unemployment" rate is the share of non-employed workers who find a job in the next year. The "Unemployment to employment" rate is the share of employed workers who are non-employed in the next year. The "job-to-job" rate (J2J) is the share of employed workers who change employer, and "J2J down" reports the share of J2J that experience earnings losses.

Labor market flows. Section 3.1 documents differences in labor market opportunities across urban areas in Spain. In our baseline specification, we exploit cross-sectional differences in job-finding, job-destruction, and job-to-job mobility rates across individuals with similar *observable* sociodemographic characteristics living in different urban areas. However, these cross-sectional differences in labor market outcomes might still reflect the innate abilities of people. To address potential selection bias from high-ability individuals sorting into low-unemployment urban areas, consider the following Linear Probability Model (LPM) with individual fixed-effects:

$$P(y_{i\ell t} = 1) = \tau_t + \varphi_i + \alpha_\ell + \beta' \mathbf{X_{ijt}} + \varepsilon_{i\ell t} \tag{B.1}$$

where $y_{i\ell t}$ refers to a binary labor market outcome of worker i in an urban area of tercile ℓ at time t, τ_t is a time-fixed effect, α_ℓ is an urban area (of unemployment tercile ℓ) fixed effect, and $\mathbf{X_{ijt}}$ is a vector of time-varying characteristics. Our baseline specification of Section 3.1 uses annual data because the quantification of the structural model uses the year as the model period. Here, we turn to monthly data to have more observations for each individual and, therefore, exploit within-worker, across-time variation in labor market outcomes. For this reason, we also provide the results of regressing the baseline Probit regression using monthly data to facilitate comparability of the results.

Table B.2 reports the results from the Probit and Within Fixed Effect regressions. Table B.2 shows that, even after controlling for workers' ability, low-unemployment urban areas still offer higher job-finding rates for both employed and unemployed workers, as well as lower job-destruction rates and lower downward job mobility. Moreover, if anything, results from the fixed effects regression are stronger.

We relegate the FE estimates to the Appendix because, unlike the Probit regression, the FE regression relies on a linear specification to remove the worker fixed effects. The linear specification presents two relevant challenges. Firstly, a linear model is less effective in modeling extreme probabilities that are very close to zero, such as labor market flows. Secondly, related to the extrapolation of the effect to the entire covariate support, a linear model imposes a stricter restriction compared to the non-linear Probit model.

Earnings. We are interested in the properties of locations that rise earnings and returns to experience for workers with identical characteristics. Our baseline regression of Equation (3.2) is based on the sample period from 2006 and 2008, just before the onset of the Great Recession where the construction sector was experiencing a great economic boom. Thus, as locations differ in their sectoral composition, one potential problem is that we might attribute unusual sectoral-specific time trends to the properties of locations. Indeed, we find that the construction sector only represents 7.8 percent of employment in low-unemployment urban areas, while it represents 13 percent in high-unemployment urban areas. Therefore, we might underestimate the earnings premium of low-unemployment urban areas because workers in high-unemployment urban areas are more likely to work in construction, and this sector was experiencing an unusual economic boom.

Table B.3: Reduced-form earnings equation controlling for sectoral shocks

	T1	**T2**
Urban area fixed effect, α_ℓ (%)	5.7^{***}	1.6^{***}
	(0.21)	(0.22)
Urban area returns to experience, δ_ℓ (%)	1.0^{***}	0.25^{***}
	(0.03)	(0.04)
Overall returns to experience, γ_1 (%)	7.5^{***}	
	(0.07)	
Overall returns to experience^2, γ_2 (%)	-0.20^{***}	
	(0.01)	
Sector × time FE	Yes	
City, worker, time FE	Yes	
N	6,354,675	
R^2	0.81	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Excessive reallocation and unemployment

	Excessive reallocation rate	
Unemployment rate	-0.41	-0.59
Further observables	No	Yes
N	170	170
R^2	0.07	0.40

Sources: Census 1991, 2001, and 2011. The further observables are dummies for the urban area shares of 4 age groups, the shares of three education groups, and the share of workers with jobs with high socio-economic status.

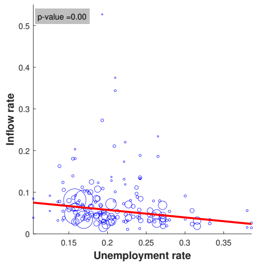
To address this issue, we additionally include an interaction term of sector and time dummies that capture sector-specific trends in earnings. Table B.3 reports the results from this regression. We find that controlling for sector-specific trends does not significantly change our results. Therefore, our main conclusions from the baseline regression are robust to controlling for sector-specific trends. That is, low-unemployment urban areas pay higher average earnings and higher returns to experience conditional on worker characteristics.

B.2 The role of sociodemographic characteristics in explaining excess reallocation

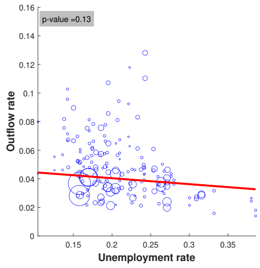
The different observed mobility patterns across urban areas may be the result of individual attributes rather than those of urban areas. To control for individual attributes in explaining mobility, we regress the gross reallocation rate of each urban area on the unemployment rate, while adding urban area controls for age, education, and socio-economic status. Table B.4 shows that the negative relationship between the excess reallocation rate and the unemployment rate becomes yet more negative when controlling for individual observable characteristics.

B.3 Mobility flows by age

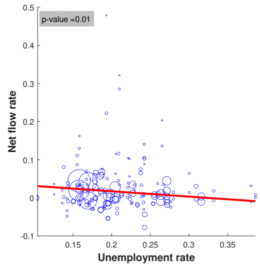
FIGURE B.1: Mobility flows, unemployment, and age.



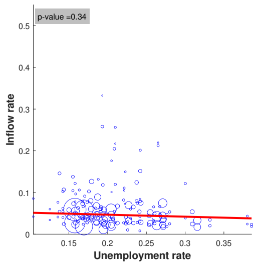
(a) IR young



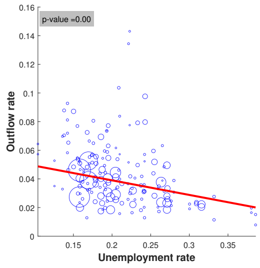
(b) OR young



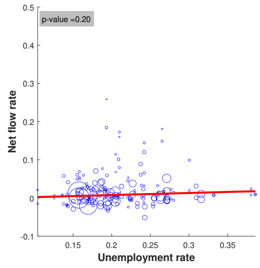
(c) NR young



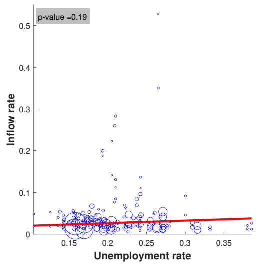
(d) IR prime age



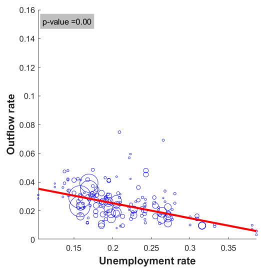
(e) OR prime age



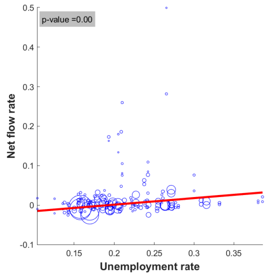
(f) NR prime age



(g) IR old



(h) OR old



(i) NR old

Notes: The lines show size-weighted OLS regression slopes. Young: age 25-35; Prime-age: ages 36-49; Old: ages 50-80. Source: 1991, 2001, and 2011 Censuses.

Section 3.2 shows that gross mobility is higher in low-unemployment urban areas compared to high-unemployment urban areas and that people sort into urban areas with different unemployment rates over age. The top row of Figure B.1 displays the inflow rates at the individual urban area level underlying these patterns. The second row does the same for the outflow rates. To highlight the sorting pattern over age, we divide the population into three age groups.[7] The figure shows that the inflow rates of young people (ages 25–35) fall rapidly with the urban area unemployment rate with very few young people joining urban areas with unemployment rates of 35% or higher.[8] In contrast, outflow rates show only a weak relationship with the unemployment rate. As a result,

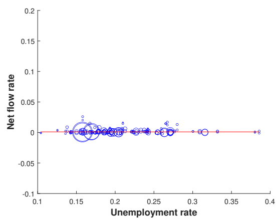
[7]We define rates using the age-specific flow of people in the numerator and the total urban area size in the denominator. This way, the total flow rate can be decomposed additively into the flow rates displayed in Figure 1.

[8]We discard people younger than age 25 as, given the decennial measure, their mobility may have resulted from the mobility decisions of their parents. Including those people leaves the results unchanged.

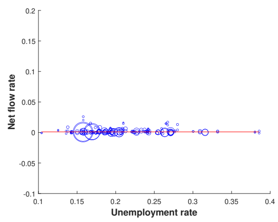
as the last row shows, the net flow is decreasing in the unemployment rate, i.e., young people move on net to low-unemployment urban areas. Turning to prime-aged workers, the outflow rate displays a stronger negative relationship with the unemployment rate, and the inflow rate shows only a weak negative relationship with the unemployment rate. As a result, the net flow rate is weakly increasing in the unemployment rate. Finally, the outflow rates of the elderly (ages 50+) also display a strong negative relationship with the unemployment rate. and the inflow rate shows a weak positive relationship with the unemployment rate. As a result, the net outflow of old people displays a strong positive relationship with the unemployment rate, i.e., the elderly sort into high-unemployment urban areas.

B.4 The role of education in explaining migration

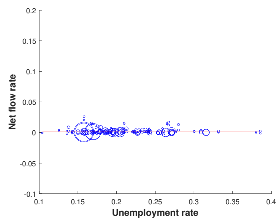
FIGURE B.2: Mobility by education



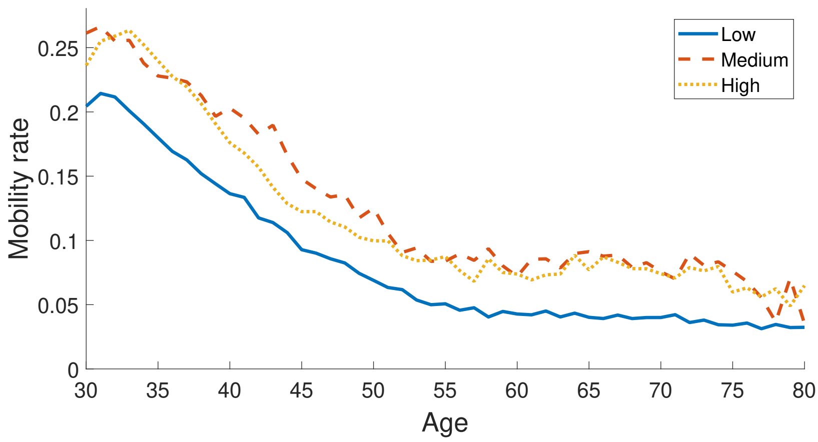
(a) Low



(b) Medium



(c) High



(d) Mobility rate by age and education

The lines show size-weighted OLS regression slopes. Low: less than secondary education; Medium: secondary education; High: More than secondary education. The bottom panel shows the mean decennial mobility rate of individuals over age. Source: 1991, 2001, and 2011 Censuses.

Our analysis abstracts from education differences among workers. This appendix shows that the Spanish data does, indeed, suggest that education differences,different from age differences, are of second-order importance to understanding mobility patterns in comparison to life cycle heterogeneity. The top panel of Figure B.2 shows that there is no systematic sorting into urban areas with different unemployment rates based on people's education. This becomes particularly apparent when comparing the gradients in the unemployment rate with those in Figure B.1 for different age groups. Moreover, the bottom panel shows that also people's migration hazards over age look very similar for different education groups. The only difference is a somewhat lower average mobility rate of the lowest educated group throughout the life cycle. This lower mobility rate may present a confounding factor for our analysis if low-educated people were strongly sorted into high-unemployment urban areas. Though they are indeed over-represented in those areas, the differences are small: Their population share is 60% in urban areas in the first decile of the urban area unemployment distribution, 64% in the second, and 67% in the third. Given that we find the largest differences in mobility rates between urban areas in the first and second tercile, these relatively minor differences in education shares explain only a small fraction of the differences in observed mobility rates.

C Appendix to Model

C.1 Microfounding the spatial search friction

This section provides a micro-foundation for the spatial search friction (μ) used in the main text. We adapt the arguments from the rational inattention literature to our framework (e.g., see, Ellis, 2018; Dean, 2022). The key insight from this simple model of mobility is that an equilibrium may arise in which only a fraction of individuals behaves as if acquiring information is a necessary preliminary step before making mobility decisions ($\mu < 1$). In this framework, individuals may refrain from acquiring information either because their information costs are too high or because they are sufficiently convinced that they already live in the best location. Moreover, we argue that, unlike non-frictional models, this framework is consistent with the micro-evidence on the effects of providing information on migration decisions (Bergman et al., 2024; Wilson, 2021).

Environment. For analytical transparency, we consider here static model. Individuals, indexed by $i \in [0,1]$,currently live in one of a finite set of locations $\mathcal{F}$. There is a unit measure of individuals who are ex-ante heterogeneous in two dimensions. First, they have idiosyncratic prior beliefs about the state of the world $\Pi_i : \Omega \to [0,1]$ such that $\sum_{\omega \in \Omega} \Pi_i(\omega) = 1$, where the state of the world is a finite set $\Omega = \{\omega_1, \ldots, \omega_N\}$ and ω_n may be an infinite-dimensional object. Thus, the prior belief function can be summarized as a vector of probabilities $(\Pi(\omega_1), \ldots, \Pi(\omega_N))$. In our model, the realization of the state of the world w_n reflects differences across locations, such as the joint distribution of job offers, unemployment risk, housing costs, etc. Second, individuals have idiosyncratic costs of acquiring information $K_i \in \mathbb{R}_+$, which reflect non-pecuniary costs of the time and effort to scan locations and get information about the state of the world. It is reasonable to expect substantial heterogeneity in these costs in a given period: they may range from very small (e.g., individuals hear by chance about a job offer in some location) to very large costs (e.g., aversion to scan locations due to the large amount of available information).

Regarding the ex-ante distribution of individuals, we define $P : [0,1]^N \times \mathbb{R}_+ \to [0,1]$ as the joint probability distribution of prior beliefs and information acquisition costs. We assume the initial distribution of agents across locations is independent of P. This assumption is not essential; rather, it allows us to illustrate beliefs and information costs as the key determinants of spatial search frictions.

Preferences and actions. Individuals' preferences depend on the state of the world (ω) and the location (ℓ_i) where they live before making the location decision. The utility from living in location

$\ell_i \in \mathcal{F}$ in the state of the world $\omega \in \Omega$ is given by $u(\ell_i, \omega)$. Individuals make two decisions. First, they choose whether to acquire information. Second, they decide in which location to live given an information structure.

Information structure. Information comes in terms of a signal $s_m \in S = \{s_1, \ldots, s_M\}$, which provides information about the state of the world. Specifically, individuals know that each signal is associated with a probability conditional on the state of the world $\gamma(s_m \mid \omega_n)$, $\forall m \in \{1, \ldots, M\}$. For example, a signal may consist of a sample of job offers in Madrid and Alicante, reflecting that the labor market in Madrid provides higher lifetime earnings to identical individuals than that in Alicante. Individuals thus understand that such a signal s_m is more likely to arise when the underlying distribution of job offers w_m indeed yields higher lifetime earnings in Madrid than in Alicante, and they use this information to update their beliefs about the state of the world. In particular, we assume that individuals update their beliefs using the Bayesian rule. Thus, after receiving a signal s_m, the updated probability $\lambda(\tilde{\omega})$ of being in the state $\tilde{\omega} \in \Omega$ is given by:

$$\lambda_i(\tilde{\omega}) = \frac{Pr(\tilde{\omega} \cap s_m)}{Pr(s_m)} = \frac{\Pi_i(\tilde{\omega}) \cdot \gamma(s_m \mid \tilde{\omega})}{\sum_{\omega \in \Omega} \Pi_i(\omega) \cdot \gamma(s_m \mid \omega)}.$$

We denote the information structure that individuals may acquire as $\Gamma_i = (S, \lambda_i)$, which involves the set of signals and the updated beliefs. Note that the information structure is idiosyncratic because we allow for idiosyncratic prior beliefs.

Value of information. Consider the expected utility $V_i(\Gamma_i)$ of individual i when choosing a location ℓ'_i under the information structure Γ_i, which is given by:

$$V_i(\Gamma_i) = \sum_{\omega \in \Omega} \Pi_i(\omega) \sum_{s \in S} \gamma(s \mid \omega) \cdot v(\lambda_i, \mathcal{F}) \; - K_i, \tag{C.1}$$

$$v_i(\lambda_i, \mathcal{F}) = \max_{\ell'_i \in \mathcal{F}} \sum_{\omega \in \Omega} \lambda(\omega) u(\ell'_i, \omega). \tag{C.2}$$

This expectation makes it explicit the three key factors that determine the value of information within the model. First, the conditional distribution of available signals (γ). Second, the expected utility, $v(\lambda_i, \mathcal{F})$, of choosing from a location set ($\mathcal{F}$) under the updated beliefs (λ_i). Third, the idiosyncratic cost, K_i, of acquiring the information structure.

In addition, individuals may choose a location without acquiring any information about the state of the world. In this case, the expected utility of an individual i depends on the prior beliefs, $V_i(\Pi_i)$, and is given by:

$$V_i(\Pi_i) = \max_{\ell'_i \in \mathcal{F}} \sum_{\omega \in \Omega} \Pi_i(\omega) u(\ell'_i, \omega). \tag{C.3}$$

In our quantitative model, only a subset of individuals make mobility decisions. We interpret this framework as a simplification of a model in which individuals behave *as if* acquiring information is a preliminary step before choosing a location. Hence, we restrict our analysis to a world, where individuals never move without acquiring information, which is a natural assumption. That is, in the absence of information acquisition, individuals always find it optimal to remain in their current location. Equation (C.3) shows that this is equivalent to restricting prior beliefs such that all movers acquire information leading to the following assumption:

A1. In the absence of information acquisition, the agent's prior belief vector Π_i belongs to a subset of the probability simplex, $\Pi_i \in \tilde{\Delta}^N$, such that

$$\ell_i = \arg\max_{\ell'_i \in \mathcal{F}} \sum_{\omega \in \Omega} \Pi_i(\omega) u(\ell'_i, \omega), \quad \text{so } V_i(\Pi_i) = \sum_{\omega \in \Omega} \Pi_i(\omega) u(\ell_i, \omega).$$

The spatial search friction as an equilibrium outcome. To close the model, we need to solve for the fraction of individuals who acquire information. Agents find it optimal to acquire information whenever $V_i(\Gamma_i) \geq V_i(\Pi_i)$. Hence, the mass of individuals who acquire information and possibly migrate, which we define as μ following the notation of the spatial search friction in the quantitative model, is given by:

$$\mu \;=\; P(I), \quad \text{where } I = \Big\{(\Pi_i, K_i) \in [0,1]^N \times \mathbb{R}_+ : \sum_{\omega \in \Omega} \Pi(\omega) = 1,\ V_i(\Gamma_i) \geq V_i(\Pi_i)\Big\}. \tag{C.4}$$

Hence, a possible equilibrium outcome is that only a fraction of individuals behave as if acquiring information is a preliminary step to make mobility choices ($\mu < 1$). Equation (C.4) highlights that this equilibrium outcome may arise because only a subset of people gain from acquiring information due to the presence of ex-ante heterogeneity in prior beliefs about the state of the world (Π_i) or information costs (K_i). Moreover, this equilibrium outcome may arise even when large spatial economic differences exist net of fixed mobility costs (i.e., for a given realization of $\omega \in \Omega$). Lastly, consistent with evidence on the response of migration decisions to information access (Bergman et al., 2024; Wilson, 2021), this model may endogenously generate changes in mobility rates due

Table C.1: Payoffs

	Choices	
State	M	S
H	$-\bar{u}$	0
C	0	0
L	0	$-\bar{u}$

to changes in information. We note that our quantitative model is the limiting case of the present model where either people's prior beliefs are very dispersed or the idiosyncratic costs to information are very dispersed such that search is random.

Example. To illustrate the intuition of the model and show that the above-mentioned equilibrium is possible, consider the following toy economy. There are two locations, so the location choice reduces to moving (M) or staying (S). There are three possible states of the world, $\Omega \in \{L, C, H\}$. In state H, the agent lives in the desirable location, so moving entails a loss relative to staying. In state L, the opposite holds: moving yields a gain compared to staying. State C is a central situation in which both actions deliver the same flow utility. Table C.1 displays a payoff matrix consistent with this environment, where the loss from the strictly worst action is $\bar{u} \in \mathbb{R}_{++}$. In this simple economy, we just need to compute the probability of living in the desirable/undesirable location to choose the best action.

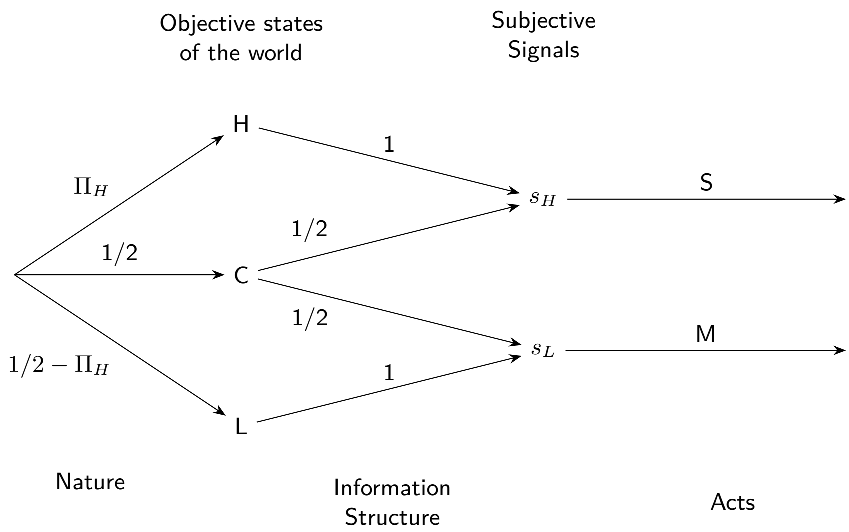
Figure C.1 illustrates the prior beliefs and the available information structure. The agent assigns a prior probability $\Pi_H \in [0, 1/2]$ of currently living in the desirable location H, and a prior probability $1/2 - \Pi_H$ to the state L in which she currently lives in the undesirable location. The remaining probability corresponds to the state in which the individual views both locations as identical. Conditional on the prior beliefs, the agent prefers to stay in the current location whenever $\Pi_H \in [1/4, 1/2]$, which is the condition for **A1** to hold. Intuitively, the agent prefers to stay whenever she assigns a relatively higher probability to living in the desirable in comparison to the undesirable location. Then, the expected utility of not acquiring information is $V_i(\Pi_i) = -\bar{u} \cdot (1/2 - \Pi_H)$.

In contrast, the informational structure may induce the person to move even under **A1**. In particular, there is an information structure with two possible signals. The individual understands that receiving signal s_H implies the following updated Bayesian beliefs about the state of the world:

$$\gamma(H) + \gamma(C) = 1, \quad \text{and} \quad \gamma(L) = 0.$$

Hence, the signal s_H implies that the agent believes she is not living in the undesirable location.

Figure C.1: An example of an information structure and acts



Note: This example is an adaptation from Dean (2022).

As a result, conditional on receiving signal s_H, the individual stays. Conversely, conditional on receiving signal s_L, the individual thinks the opposite and moves. The expected value of this information structure is given by $V(\Gamma) = -K$.

Thus, agents acquire information whenever:

$$-K \;\geq\; -\bar{u} \cdot \left(\frac{1}{2} - \Pi_H\right) \;\Longleftrightarrow\; K + \Pi_H \;\leq\; \frac{1}{2}\bar{u}. \tag{C.5}$$

Given a loss $\bar{u}$, Equation (C.5) summarizes the key insights from our model. A mass of individuals may choose not to acquire information for two reasons. First, their idiosyncratic cost of acquiring information is too high (K is high). Second, they are sufficiently convinced that they already live in the best location (Π_H is high).

C.2 Characterizing the stationary equilibrium

To define the equilibrium we need to keep track of the population size of each location type. Formally, we define the population at the beginning of the period as a measure of people of different characteristics. Let L denote the set of all possible location types and let S denote the set of amenity values. Let $X^R \equiv L \times S$ be the set of state variables for the retirees. Let $N_t^R : \mathcal{X}^R \rightarrow [0, 1]$ denote the density of retirees of age t where $\mathcal{X}^R$ is the Borel σ-algebra on X^R. Likewise, $\mathbb{E}$ is the set of all possible values of experience and Z is the set of labor productivities. Let us define $X^U \equiv \mathbb{E} \times S$ as the set of state variables for the unemployed. Likewise, $X^E \equiv Z \times X^U$.

Likewise, we can define $\mathcal{X}^U$, $\mathcal{X}^E$, N_t^U and N_t^E. Hence, the population at a location of type ℓ is $\mathrm{N}(\ell) = \sum_{t=R+1}^{T} \int_S N_t^R(\ell, s)\ f_S(s)\, ds + \sum_{t=1}^{R} \sum_{\mathbb{E}} \int_S N_t^U(\ell, s, e)\ f_S(s)\, ds$
$+ \sum_{t=1}^{R} \sum_{\mathbb{E}} \int_S \int Z N_t^E(\ell, s, e, z)\ f_S(s)\, f_Z(z)\, ds\, dz$. Before we define the stationary equilibrium, we need to define flows in the economy across urban areas. We denote inflows by $IF(\ell)$ and outflows by $OF(\ell)$. In Section 3.2 we described their data counterparts, IR, and OR, relative to population size, to characterize mobility patterns across urban areas. These flows are computed using the individual's migration policy function and aggregating across individuals. For instance, let $OF_i^E(\ell, s, e, z)$ denote the amount of i years old worker with state (ℓ, s, e, z) who leave a location of type ℓ. It satisfies:

$$OF_i^E(\ell, s, e, z) = N_i^E(\ell, s, e, z)\ \Xi_i^E(\ell, s, e, z)\,, \tag{C.6}$$

where $\Xi_i^E(\ell, s, e, z)$ denotes the overall probability of migration, which depends on all possible migration opportunities and the individual's migration decision:

$$\begin{aligned}\Xi_i^E(\ell, s, e, z) = \mu_\ell^E \sum_{\ell'} \frac{1}{3} \int_{s'} (1-\phi_{\ell'}) g_i^{EU}(\ell, s, e, z, \ell', s')\ f_S(s')\, ds' + \\ \mu_\ell^E \sum_{\ell'} \frac{1}{3} \int_{s'} \phi_{\ell'}\ f_S(s') \int_{z'} g_i^{EE}(\ell, s, e, z, \ell', s', z')\, f_Z(z')\, ds'\, dz'.\end{aligned} \tag{C.7}$$

The evolution of the population is given by the law of motion

$$N(\ell)' = N(\ell) + IF(\ell) - OF(\ell) + N_1(\ell)' - N_T(\ell), \tag{C.8}$$

where $N_1(\ell)'$ is the overall measure of newborns at a location of type ℓ and $N_T(\ell)$ is the measure of T years old who died at the end of the previous period. Now we are ready to define the stationary equilibrium.

Definition 1. *A recursive stationary equilibrium, given subsidies $\{b_U, b_R\}$, is a vector of rental prices, $\{r_\ell\}_1^L$, a set of value functions and optimal decision rules for retirees, $\left\{V_t^R, W_t^R, \Omega_t^R, g_t^{R,\mu}, g_t^{R,h}\right\}_{t=R+1}^T$, for unemployed individuals, $\left\{V_t^U, W_t^U, \Omega_t^{UU}, \Omega_t^{UE}, \Omega_R^{UR}, \Psi^{EU}, g_t^{UU,\mu}, g_t^{UE,\mu}, g_t^{U,z}, g_t^{U,h}\right\}_{t=1}^R$, for workers, $\left\{V_t^E, W_t^E, \Omega_t^{EU}, \Omega_t^{EE}, \Omega_R^{ER}, \Psi_t, \Psi_t^{EE}, \Psi_t^{ER} g_t^{EU,\mu}, g_t^{EE,\mu}, g_t^{EE,z}\right\}_{t=1}^{R-1}$ and population measures $\left\{N_t^R\right\}_{t=R+1}^T$, and $\left\{N_t^U, N_t^E\right\}_{t=1}^R$ such that:*

1. *Value functions and policy functions solve individual problems shown in Equations (4.7) to (4.20),*

2. *the housing markets clear, $H_\ell^D = \overline{H}_\ell$, for all ℓ where the demand function is given by* **??**,

3. *all population measures, $\left\{N_t^R\right\}_{t=R+1}^{T}$, and $\left\{N_t^U, N_t^E\right\}_{t=1}^{R}$, given by, Appendix C.2, are constant over time and their laws of motion satisfy Equation (C.8).*

Proposition 1 In the main text, we use the fact that all urban areas of the same type have the same equilibrium rental price. The intuition for the proposition is as follows: Suppose that there are two locations, 1 and 2, of productivity ℓ, and that location 1 is cheaper than 2. If it its rental price is cheaper, Equation (4.21) implies that some population group is smaller in location 1: either retirees, unemployed of a particular age and experience, or employed individuals. However, this cannot be, as the inflows to location 1 must be greater than those to location 2 and its outflows must be lower. Let us focus our attention on retirees. Take two retirees identical in all respects (age and current residence) but the first one has the opportunity to migrate to 1 and the second one has the opportunity to migrate to 2. Since migration opportunities across locations of the same productivity type are drawn from a uniform distribution, the law of large numbers ensures that there is always a positive measure of people from any location ℓ who have a migration opportunity to either 1 or 2. The gain of moving to 1 is larger than the gain of moving to 2,

$$\Omega_t^R(\ell, s, 1) > \Omega_t^R(\ell, s, 2), \tag{C.9}$$

since 1 is cheaper. Hence, agents need to draw a higher amenity value to migrate to location 2 than to migrate to location 1. Since the distribution of amenity draws is the same across locations, inflows of retirees to location 1 are larger than inflows to location 2. Conversely, outflows from 1 to a given location ℓ are lower than the similar outflow from 2. The reason, again, is that retirees located in 1 have to draw a higher amenity value to move to ℓ than the similar retiree in location 2. The same reasoning applies to unemployed people of a given age and experience, location of residence, and current amenity value. The key is that, in any given location, there is always a positive measure of people that are offered to move to 1 and another measure of who are offered to move to location 2 under the same labor conditions. Since location 1 is cheaper, people moving to location 2 have to be compensated for the rental differential with a higher amenity value. Thus, the inflows to 2 is lower than the inflows to 1. Similarly happens to employed individuals. Hence, it follows that the population must be strictly larger in location 1, arriving at a contradiction.

Table C.2: Moments in model and data

Moment and parameter	Model T1	Model T2	Model T3	Data T1	Data T2	Data T3
Labor markets						
EU rate (%); λ_ℓ	7.17	7.72	9.77	7.30	7.70	9.85
U rate (%); ϕ_ℓ	16.67	20.61	27.47	16.20	20.10	27.10
Job-to-Job rate (%); Λ		11.79			11.81	
Job-to-Job share losses (%); λ_d		40.29			41.97	
Std of job switchers; σ_Z		0.32			0.31	
Mobility						
Relative people turnover; ω_ℓ	1	0.91	0.83	1	0.89	0.82
Mobility rate (%); p^U		9.52			9.60	
Ratio of E to U movers; p^E		2.69			2.71	
Mobility age 45; κ		9.52			19.2	
Mobility ages 76–80; κ		3.34			3.64	
Share T1 to T1 prime-age; σ_S		0.39			0.39	

The table displays the endogenously model calibrated moments and the corresponding data moments. Those moments that are urban area (UA) specific, are reported for each, otherwise, only one common number is reported.

C.3 Calibration details

C.4 A Model without search frictions

Section 6.1.3 compares our model to a model without search frictions to show the importance of those frictions for mobility patterns across urban areas. This section describes the model without mobility frictions across urban areas. Local labor markets are modeled identically to the baseline model and so are preferences. Hence, Equations (4.14) to (4.20) take the same form, and we restrict us here to describing the migration stage.[9] For comparability, we will use the same notation as in the main text.

We follow much of the migration literature and assume that people optimally decide each period in which urban area to search given some realization of i.i.d. shocks for each urban area type, $\ell = 1, 2, 3$. The literature usually assumes that these shocks follow extreme value distributions as this simplifies migration decisions. For a better comparison to the baseline model, we keep here the assumption that these shocks are log-normally distributed. Hence, at the migration stage, the value of a retiree of age $t = R+1, \ldots, T-1$, who lives in a location of type ℓ and amenity value s solves:

[9] For parsimony, we also omit the value functions in the last period of working life and the last period of life which have different continuation values.

$$V_t^R(\ell,s) = \int_1\int_2\int_3 \max\left\{\beta\, W_{t+1}^R(\ell,s)\,,\beta\Omega^R(s',s'',s''')-\kappa\right\} f_S(s')f_S(s'')f_S(s''') \tag{C.10}$$

$$\Omega^R(s',s'',s''') = \max\left\{W_{t+1}^R(1,s')\,,W_{t+1}^R(2,s'')\,,W_{t+1}^R(3,s''')\right\}. \tag{C.11}$$

$W_{t+1}^R(\ell,s)$ is the value of staying in the current location, and $\Omega^R(s',s'',s''')$ is the value of moving to the best alternative location.

Similarly, the unemployed also choose the optimal place to search. In doing so, they take into account that different locations provide different probabilities to be offered a job, $\phi_{\ell'}$, and that they have the choice to move after having observed the type of job offer:

$$\begin{aligned} V_t^U(\ell,s,e') =& \\ &\int_1\int_2\int_3 \max\left\{\beta\, W_{t+1}^U(\ell,s,e')\,,\beta\,\Omega^U(\ell,s,e',s',s'',s''')-\kappa\right\} f_S(s')f_S(s'')f_S(s'''), \end{aligned} \tag{C.12}$$

and the value of moving also includes possible job offers:

$$\Omega^U(\ell,s,e',s',s'',s''') = \max\left\{\bar{\Omega}^U(\ell,s,e',1,s')\,,\bar{\Omega}^U(\ell,s,e',2,s'')\,,\bar{\Omega}^U(\ell,s,e',3,s''')\right\}, \tag{C.13}$$

$$\begin{aligned} \bar{\Omega}^U(\ell,s,e',\ell',s') =\ & \phi_{\ell'}\int \max\{W_{t+1}^U(\ell,s,e')\,,W_{t+1}^E(\ell',s',e',z')\}f_Z(z') \\ & +(1-\phi_{\ell'})\max\{W_{t+1}^U(\ell,s,e')\,,W_{t+1}^U(\ell',s',e')\}. \end{aligned} \tag{C.14}$$

Finally, the employed face a similar trade-off as the unemployed with the only difference that they can stay at their current place as employed:

$$\begin{aligned} V_t^E(\ell,s,e',z) =& \\ &\int_1\int_2\int_3 \max\left\{\beta\, W_{t+1}^E(\ell,s,e',z)\,,\beta\Omega^E(\ell,s,e',z,s',s'',s''')-\kappa\right\} f_S(s')f_S(s'')f_S(s'''), \end{aligned} \tag{C.15}$$

and, likewise, the value of moving include possible job offers:

$$\Omega^E(\ell, s, e', z, s', s'', s''') = \max\left\{\bar{\Omega}^E(\ell, s, e', z, 1, s'), \bar{\Omega}^U(\ell, s, e', z, 2, s''), \bar{\Omega}^U(\ell, s, e', z, 3, s''')\right\} \tag{C.16}$$

$$\begin{aligned}\bar{\Omega}^E(\ell, s, e', z, \ell', s') = &\ \phi_{\ell'}\int \max\{W^E_{t+1}(\ell, s, e', z), W^E_{t+1}(\ell', s', e', z')\} f_Z(z') \\ &+ (1-\phi_{\ell'})\max\{W^E_{t+1}(\ell, s, e', z), W^U_{t+1}(\ell', s', e')\}.\end{aligned} \tag{C.17}$$

C.5 Welfare analysis

Let us define as ξ_ℓ the compensation in lifetime consumption (both housing and non-housing) needed for an individual to be indifferent between being born in location types $\ell \in \{2, 3\}$ relative to location type 1. Note that the indirect utility function is $u(c, h, s) = \theta^\theta (1-\theta)^{1-\theta} y / \left(r_\ell^{(1-\theta)}\right) + s$, where y is income. s is the amenity value that the current location yields to the individual. In the case in which the agent has received an opportunity offer and it was advantageous, the agent pays the cost κ_t. Next, define the expected welfare, given the compensation, of living in ℓ:

$$EW_\ell(\xi_\ell) \equiv E_{0,\ell}\left\{\sum_{t=0}^{T}\beta^t\left((1+\xi_\ell)\,\theta^\theta(1-\theta)^{1-\theta}\,\frac{y_t}{r_t^{1-\theta}} + s_t - I_t\,\kappa_t\right)\right\}. \tag{C.18}$$

This expectation comprises the fact that labor markets are different across locations and, therefore, there are static differences (so that the initial distribution of employment across newborns is different) but also the expected horizon is different as each location provides different jobs, migration opportunities, and return to experience. I_t is the mobility choice. The value ξ_ℓ is obtained so that

$$EW_\ell(\xi_\ell) = EW_1 \equiv E_{0,1}\left\{\sum_{t=0}^{T}\beta^t\left(\theta^\theta(1-\theta)^{1-\theta}\,\frac{y_t}{r_t^{1-\theta}} + s_t - I_t\,\kappa_t\right)\right\}.$$

Defining the amenities net of moving costs, $\widetilde{s}_t = s_t - I_t\,\kappa_t$, we then rewrite Equation (C.18) as:

$$\begin{aligned}EW_\ell(\xi_\ell) &= (1+\xi_\ell)\ E_{0,\ell}\left\{\sum_{t=0}^{T}\beta^t\left(\theta^\theta(1-\theta)^{1-\theta}\,\frac{y_t}{r_t^{1-\theta}} + \frac{\widetilde{s}_t}{(1+\xi_\ell)}\right)\right\}, \\ &= (1+\xi_\ell)\ E_{0,\ell}\left\{\sum_{t=0}^{T}\beta^t\left(\theta^\theta(1-\theta)^{1-\theta}\,\frac{y_t}{r_t^{1-\theta}} + \widetilde{s}_t - \frac{\xi_\ell\,\widetilde{s}_t}{(1+\xi_\ell)}\right)\right\}.\end{aligned}$$

Note that the expected value of being born in location ℓ in period 0 is given by

$$EW_\ell = E_{0,\ell}\left\{\sum_{t=0}^{T}\beta^t\left(\theta^\theta(1-\theta)^{1-\theta}\,\frac{y_t}{r_t^{1-\theta}} + \widetilde{s}_t\right)\right\}.$$

Therefore, the expected welfare given the compensations is given by:

$$EW_\ell(\xi_\ell) = (1 + \xi_\ell)\ EW_\ell - \xi_\ell\, E_{0,\ell} \sum_{t=0}^{T} \beta^t \widetilde{s}_t.$$

Hence, using the definition, we then show that ξ_ℓ satisfies

$$\xi_\ell = \frac{EW_1 - EW_\ell}{EW_\ell - E_{0,\ell} \sum\limits_{t=0}^{T} \beta^t \widetilde{s}_t}.$$

Notice that EW_ℓ comprises expectations about labor maret realizations right when agents are born. The term $E_{0,\ell} \sum_{t=0}^{T} \widetilde{s}_t$ varies across locations because of the interaction of migration decisions, the amenities realizations and the migration cost. Thus, ξ_ℓ is the extra lifetime consumption needed to compensate for the difference in the present yield of income in location 1 plus the difference in the present value of expected amenities relative to the present yield of consumption in location ℓ.

References

Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., and Palmer, C. (2024). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. *American Economic Review*, 114(5):1281–1337.

Dean, M. (2022). GR6943: Behavioral Economics, Rational Inattention – Lecture 1. https://www.columbia.edu/~md3405/Behave_Col_Bounded_3_19.pdf.

Ellis, A. (2018). Foundations for optimal inattention. *Journal of Economic Theory*, 173:56–94.

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