ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1616

Title and Content Revised June 2019
(Replaced May 2019 version)

Education and Geographical Mobility: The Role of the Job Surplus

Michael Amior

Abstract
Better-educated workers form many more long-distance job matches, and they move more quickly following local employment shocks. I argue this is a consequence of larger dispersion in wage offers, independent of geography. In a frictional market, this generates larger surpluses for workers in new matches, which can better justify the cost of moving - should the offer originate from far away. The market is then "thinner" but better integrated spatially. I motivate my hypothesis with new evidence on mobility patterns and subjective moving costs; and I test it using wage returns to local and long-distance matches over the jobs ladder.

Key words: geographical mobility, job search, education
JEL Codes: J61; J64; R23

This paper was produced as part of the Centre's Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

I am grateful to Alan Manning for his support at LSE. This paper was part of my PhD thesis at UCL, and I thank my PhD supervisors Steve Machin and Jeremy Lise and examiners Marco Manacorda and Ian Preston for their support and advice. I also thank David Albouy, Dan Black, Mitch Downey, Mike Elsby, Eric French, Georg Graetz, David Green, Caroline Hoxby, Kevin Hutchinson, John Kennan, Felix Koenig, Hamish Low, Barbara Petrongolo, Steve Pischke, Jan Stuhler and Coen Teulings for helpful comments, as well as participants at the EEA Annual Congress, RES Annual Conference, Annual Search and Matching Conference, and seminars at Uppsala, LSE, Tel Aviv, Hebrew University and Cambridge. I gratefully acknowledge financial support from the Economic and Social Research Council and the Royal Economic Society. This paper supersedes an earlier (2015) version entitled "Why are Higher Skilled Workers More Mobile Geographically? The Role of the Job Surplus" (CEP Discussion Paper 1338).
Michael Amior, Hebrew University of Jerusalem and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Geographical mobility is known to be crucial to the adjustment of local labor markets (Blanchard and Katz, 1992; Amior and Manning, 2018). But there is severe inequity in the incidence of mobility. Better educated Americans make many more long-distance moves, especially near the beginning of their careers: one year after leaving school, 9 percent of postgraduate degree holders move state annually, compared to just 2 percent of high school dropouts (Figure 1).[1] At the same time, the low educated move more sluggishly in response to local employment shocks (Topel, 1986; Bound and Holzer, 2000; Wozniak, 2010; Notowidigdo, 2011), contributing to substantial persistence in local jobless rates (Amior and Manning, 2018). This is of particular concern today, given the severe contraction of local manufacturing industries: see Autor, Dorn and Hanson (2013), Acemoglu et al. (2016) and Charles, Hurst and Notowidigdo (2016).

The evidence suggests the effect of education is causal (Malamud and Wozniak, 2012; Machin, Salvanes and Pelkonen, 2012), but what explains it? One view is that better educated workers face larger local differentials in expected utility, a consequence perhaps of local variation in the returns to human capital (e.g. Costa and Kahn, 2000; Wheeler, 2001; Lkhagvasuren, 2014; Davis and Dingel, 2019).[2] But one might reasonably argue the reverse, that it is the *low educated* who face larger differentials: after all, they suffer larger local shocks to wages and employment (Hoynes, 2000; Gregg, Machin and Manning, 2004), and the impact of these shocks persists much longer (Amior and Manning, 2018). For these reasons and others, many have concluded that the low educated face prohibitive migration costs, whether due to financial constraints, lack of information or home attachment[3] (Greenwood, 1973; Topel, 1986; Bound and Holzer, 2000; Wozniak, 2010; Moretti, 2011; Kennan, 2015; Caldwell and Danieli, 2018). However, it has proven difficult to identify exactly which costs might be responsible. Indeed, migration costs are typically estimated as a residual, conditional on the assumptions of the particular model (e.g. Kennan and Walker, 2011).

[1]Figure 1 is based on Current Population Survey (CPS) waves since 1999. Certainly, cross-state mobility has declined since the 1980s (Molloy, Smith and Wozniak, 2011), but this decline was fairly uniform across education groups. Appendix A.2 shows the large education differentials have persisted since at least the 1960s.

[2]Better educated individuals may also face larger geographical differentials in expected utility for other reasons. Diamond (2016) finds they have stronger preferences over local utilities. And Notowidigdo (2011) argues the low educated are better insured against local shocks by compensating transfer payments and fluctuations in housing costs.

[3]Gregg, Machin and Manning (2004) and Balgova (2018) suggest that long-distance job search may be less costly for better educated workers; as I claim in Section 3.6 below, this can be rationalized as an *outcome* of larger job surpluses. Gregg, Machin and Manning (2004) argue further that college graduates have weaker home attachment, having already left home to study - though Malamud and Wozniak (2012) dispute the latter claim. Bound and Holzer (2000) suggest a lack of assets may constrain the set of location choices.

In this paper, I offer an alternative explanation: high-educated mobility is a consequence of larger dispersion in wage offers, independent of geography. In a frictional market, this generates larger job surpluses for workers in new matches (in excess of their reservations), particularly at the beginning of workers' careers. Crucially, these surpluses will more frequently justify the otherwise prohibitive cost of moving, even if the offer distribution is identical everywhere. They will also amplify the migratory response to local job creation, as they place more workers on the margin of moving. In this way, offer dispersion makes the labor market "thinner" (match quality varies considerably) but better integrated spatially. Theoretically, larger offer dispersion may be motivated intuitively by a notion of specialized skills - or supermodularity between workers' abilities and job attributes (such as task complexity or firm quality), in the spirit of Sattinger (1975). Gottfries and Teulings (2016) find that offer dispersion is indeed increasing in education; and Lise, Meghir and Robin (2016), Bagger and Lentz (2018) and Lopes de Melo (2018) offer empirical support for positive sorting.

To take a practical example, a good job may plausibly motivate a young engineer to move between two similar cities (at sizeable cost), but not his older counterpart (who already has a good match) or somebody who cuts hair for a living (for whom a good match yields little reward). A "thin" market for specialized jobs is crucial to this hypothesis: otherwise, the young engineer would wait for an equivalent local offer. This reflects evidence on job transitions from Nimczik (2018), who shows that better educated workers participate in labor markets which are more expansive spatially but more specialized in particular industries.

My hypothesis builds on early insights from Schwartz (1976) and Wildasin (2000), who discuss how specialized markets can motivate mobility - though to my knowledge, it is absent from the recent literature. My contribution is to explore the theoretical implications using a simple jobs ladder model - and to evaluate the job surplus hypothesis empirically against competing explanations. The idea is also closely related to Manning (2003*b*) and Van Ommeren, Rietveld and Nijkamp (1997), who argue that search frictions and the associated job surplus can help explain why workers accept offers with long commutes. I apply this insight to the long-standing debate over education differentials in geographical mobility.

In Section 2, I motivate my hypothesis with three new descriptive facts. First, the education differentials in Figure 1 are entirely driven by individuals who report moving for the sake of a *specific* job, rather than to "look for work" or for non-work reasons: this affirms the central role of long-distance job matching. Second, new evidence on self-reported willingness to move suggests that moving costs are remarkably similar across education groups. And third, as is well known, net migratory flows between states are small (relative to gross flows); but I also show they are not increasing in education, even within detailed occupation-defined labor markets. This casts doubt on the importance of differential local returns in driving the mobility gap. In light of this evidence, the purpose of this paper is to offer an alternative (and non-geographical) explanation for the mobility gap, based on market frictions and the job surplus.

In Section 3, I set out a model of migration embedded in a jobs ladder, building on Schmutz and Sidibé (forthcoming). Workers draw random offers at a finite rate from an exogenous wage offer distribution, independent of geography. They search both on and off the job, which gives rise to a jobs ladder - following the logic of Burdett (1978). Job offers may arise locally or elsewhere, as in Jackman and Savouri (1992), Molho (2001) and Lutgen and Van der Linden (2015). In the latter case, a random moving cost is drawn; and the worker accepts the offer (and moves) if the associated surplus can justify the cost.

My treatment of migration as long-distance job matching deviates from Kennan and Walker (2011), the seminal study on migration choices. There, workers only draw their wage *after* moving (so locations are "experience goods"); but here, the wage offer is known *ex ante* - so workers move with a job lined up, conditional on a sufficiently attractive offer. This is crucial to my claim that the rate of migration depends on wage offer dispersion *within* geographical areas.[4] In an abstract sense, the model describes a jobs ladder with both wage and non-wage dimensions. Sorkin (2018) uses such a framework to estimate the value of jobs' non-wage characteristics more broadly.

The model predicts a number of effects of offer dispersion. First, for workers with given initial match quality, larger dispersion (independent of geography) increases the rate of (costly) long-distance matching, but has no effect on (costless) local matching. Second, the impact on long-distance matching is greater for workers with initially lower quality matches, because they have more rungs of the jobs ladder to climb. And third, larger offer dispersion amplifies the migratory response to local offer rate shocks. All these results depend crucially on search frictions: in their absence, there is no job surplus and no geographical mobility.

In Section 4, I show that patterns in local and long-distance job matching are consistent with my hypothesis. Conditional on initial job tenure (which I use to proxy match quality), better educated workers form many more cross-state matches, but a similar number of within-state matches. The education gap in cross-state mobility is also much larger among workers with lower initial tenure, and this can help account for similar effects over the support of experience (Figure 1). However, as I show below, these patterns are not only consistent with education differences in job surplus, but also in moving costs.

The empirical challenge then is to disentangle the contribution of job surplus and moving costs. To this end, I turn to evidence on wage transitions. First, in Section 5, I identify variation in within-area offer dispersion and job surplus using the wage returns to *within-state* job matches. These returns are strongly increasing in education, consistent with earlier work by Bartel and Borjas (1981) and Mincer (1986). As the model predicts, these differentials are

[4]Given that Kennan and Walker's model cannot account for this effect, any impact of offer dispersion on mobility would be identified as part of the residual migration cost. This is analogous to Schmutz and Sidibé's argument regarding the impact of long-distance job search frictions.

larger among workers with initially low match quality - as proxied by either tenure or experience. And on imposing a uniform distribution on moving costs (which yields a convenient functional form), I show that mean education differentials in *local* job surplus can quantitatively account for the bulk of the observed gap in the rate of *long-distance* matching.

Second, in Section 6, I show the wage returns to cross-state matches are disproportionately large (relative to within-state) for better educated workers. Under my model's assumptions, this implies that workers' realized moving costs (conditional on moving) are steeply *increasing* in education - by a compensating differentials argument. This is a natural consequence of larger offer dispersion: better educated workers are more likely to select into migration *because* of large job surplus and *despite* steep moving costs. Crucially, the large wage returns to cross-state matches (among college graduates) cannot be explained by differential local prices of observable human capital - or even of detailed occupation-defined tasks. My analysis here builds on earlier studies which estimate the wage returns to geographical mobility (such as Lkhagvasuren, 2014, or Huttunen, Moen and Salvanes, 2018), but I implement an empirical specification and offer an interpretation based on my particular theoretical model.

To summarize, I claim there is strong evidence that within-area offer dispersion and job surplus drive the mobility gap - as well as good theoretical reasons to expect it. Though I focus on education differentials, this paper offers new insights for understanding geographical immobility more generally. In the following section, I present my three new descriptive facts. Section 3 sets out the jobs ladder model and derives the key results, Section 4 tests its predictions for local and cross-state job matching rates, and Sections 5 and 6 estimate the wage returns to local and cross-state matches respectively.

Before moving on, it is worth briefly discussing the role of firms. Given my focus is the behavior of workers, I take the offer distribution as given. But in a previous iteration of this paper (Amior, 2015), I show that accounting for firms' decisions can offer valuable insights. In particular, dispersion in wage offers can be modeled as an outcome of dispersion in idiosyncratic match productivity. To the extent that the returns to matching are shared, both workers and firms will expect larger surplus as dispersion increases. Indeed, in the same paper, I offer evidence that firms invest more heavily in the recruitment of better educated workers (to secure this larger surplus), in terms of applications received, applicants interviewed and human resource hours. Interestingly, firms in this model also invest in geographically broader recruitment (e.g. fly-outs of potential hires), and this may *endogenously* reduce the cost of long-distance job search for better educated workers. This mechanism can amplify any direct effect of match quality dispersion on geographical mobility. Any such endogenous long-distance search costs (driven by match quality dispersion) should be distinguished from the exogenous "moving costs" of my model. The latter represent the cost of *consummating* a long-distance offer - as opposed to the cost of encountering one. I return to these questions in Section 3.6 below.

2 Motivating facts

2.1 The mobility gap is driven by workers moving for a specific job

Long-distance job matching is integral to my hypothesis, and this approach is supported by evidence on subjective reasons for moving. Figure 2 shows the mobility gap (between education groups) is entirely driven by individuals who report moving for reasons related to a *specific* job - whether due to a new job, job transfer, or shorter commute. The effect is strongest for the young, though I show in Appendix B that it is not driven by former students returning home.[5]

Individuals very rarely move speculatively to "look for work": only 5 percent of cross-state migrants report this motivation, compared to the 47 percent who move for a specific job. This is perhaps unsurprising given the associated risks (Molho, 1986). Interestingly, these speculative movers are disproportionately low-educated. Amior (2015) attributes this to different levels of investment in long-distance search and recruitment (see also Balgova, 2018), itself a consequence of offer dispersion and job surplus; and I return to this point below. There is also a mild negative education gradient in the residual "non-job" migration, which consists primarily of family and housing-related motivations.

Appendix A.3 offers a more detailed breakdown of reported reasons for both cross-state and cross-county mobility, together with associated education gradients. Interestingly, the (positive) education slope is proportionally steeper for cross-state than cross-county moves, and my model offers a rationale (job surplus should matter more for costlier longer-distance moves: see Section 3.6). There may be concern that household dependents simply report the migration reasons of the household breadwinners, but Appendix A.4 shows that restricting the sample to top earners in each household makes no difference to the results.

2.2 Subjective moving costs are unrelated to education

The cost of moving is typically estimated as a residual, conditional on the assumptions of the particular model. But here, I offer more direct estimates of moving costs, based on a unique set of hypothetical questions on willingness to move in the Panel Study of Income Dynamics (PSID) in the 1970s. These estimates suggest that moving costs are remarkably similar across education groups. Though this is not contemporary data, mobility differentials between education groups in the 1970s are similar to today (see Appendix A.2).

[5] Kennan and Walker (2011) emphasize the role of return migration in migratory flows, and Kennan (2015) shows it is an important factor in the migration decisions of recent graduates.

In 1969-72 and 1979-80, employed household heads[6] were asked: "Would you be willing to move to another community if you could earn more money there?" And in 1969-72, those who answered affirmatively were also asked: "How much would a job have to pay for you to be willing to move?" These questions speak to the cost of moving conditional on receiving a job offer, so they exclude the costs of long-distance job search. The first question identifies the set of "marginal residents" who would be willing to accept a good long-distance match (should one materialize). And the second allows me to impute the distribution of moving costs among these marginal residents.

In the first panel of Figure 3, I plot the share of respondents who report being "willing" to move. 55 percent answer yes. Though this share is declining somewhat in experience, there is remarkably little variation between education groups.[7] Of course, these subjective responses are only useful if the low-educated do not disproportionately overstate their willingness to move. And it turns out they are indeed realistic about their meager migration prospects. The PSID asks: "Do you think you might move in the next couple of years?" and "Why might you move?" Based on this data, the second panel of Figure 3 plots the share of respondents who claim they "might" move for work. The results here reflect the familiar education-experience mobility patterns from Figures 1 and 2 above.[8] The contrast with the first panel is striking: the fact that low-educated workers expect low mobility is apparently unrelated to their hypothetical "willingness" to move.

I now turn to the distribution of costs among marginal residents. For each individual, I compute the difference between the log reservation wage (for accepting a long distance offer) and the log current wage. This difference may be interpreted as the annuitized value of a hypothetical fixed moving cost (see the discussion in Section 5.3 below) or the disutility of living away from home (see the extension in Section 3.6). One might alternatively study a dollar (rather than log) difference. The appropriate formulation would depend on the model for utility, but the log difference is more conservative - in that it would (if anything) bias me against finding that better educated workers face lower costs.

[6]Household heads are always male, unless there is no husband (or cohabiting partner) present or the husband is too ill to respond to the survey.

[7]I have grouped undergraduate and postgraduate degree holders together because of small samples. Between 1970 and 1980, the PSID also asked unemployed individuals: "Would you be willing to move to another community if you could get a good job there?" 77 percent agree; and as with the employed, the fraction answering yes varies little with education.

[8]24.3 percent of college graduates who claimed they "might" move residence for job reasons actually did so (citing those same reasons) in the subsequent year, and the number is very similar (25.2) for non-graduates. This suggests there is little systematic difference by education in the accuracy of these subjective expectations.

The distribution of these imputed costs is implicitly truncated: only those who are "willing to move" (i.e. with sufficiently low costs) report a long-distance reservation. Conveniently though, the fraction who do so varies little with education (Figure 3), so selection should not be a problem for group comparisons: for each education group, I observe the (approximately) bottom 50 percent of costs.

In Figure 4, I report kernel densities of imputed costs separately for college graduates and non-graduates with 1-30 years of experience. The heterogeneity in costs is substantial, as stressed by Kennan and Walker (2011). Reassuringly, Appendix C shows these imputed costs do have significant predictive power for future migration decisions. The graduate and non-graduate plots look remarkably similar, with mean imputed costs of 0.35 and 0.38 respectively - relative to workers' current wages. In Section 6.3 below, I compare these numbers with existing estimates of migration costs in the literature.

2.3 The mobility gap is not driven by net migratory flows

It is commonly argued that the mobility gap is driven by local variation in the returns to human capital. If this is the case, net migratory flows of better educated workers between areas should be disproportionately large (relative to gross flows). But I find no evidence of this - both on aggregate (as has previously been documented) and even for net flows within detailed occupation-defined labor markets.

I estimate the cross-state net migration rate as $\frac{1}{2n}\Sigma_s|n_s^{in} - n_s^{out}|$, where n is the total sample of individuals, n_s^{in} is the number of in-migrants to state s, and n_s^{out} is the out-migrants from state s. Dividing the expression by 2 ensures that migrants are not double-counted. Notice the gross migration rate is equal to $\frac{1}{n}\sum_s n_s^{in}$ or equivalently $\frac{1}{n}\sum_s n_s^{out}$. I base my estimates on American Community Survey (ACS) samples between 2000 and 2009.[9] Migrants are defined as people who lived in a different state 12 months previously.

Table 1 reports gross and net migration rates separately by education. As is well known[10], net flows (column 2) are dwarfed by gross flows (column 1). It is less well known that this

[9]The 2000-9 ACS samples offer a consistent occupation classification based on the census 2000 scheme, and they offer much larger samples than the CPS - important for a detailed occupation decomposition. There are 221,000 cross-state migrants in my ACS sample, compared to 18,000 in the CPS in the same years. I take ACS data from the IPUMS database (Ruggles et al., 2017).

[10]See Shryock (1959); Schwartz (1971); Jackman and Savouri (1992); Coen-Pirani (2010); Monras (2015).

is especially true of better educated workers. Though gross migration is steeply increasing in education, net flows are remarkably flat; so the ratio of net to gross migration is decreasing in education: from 0.14 for dropouts to 0.08 for postgraduates (column 3). Thus, high-educated migration has a much weaker directional component: see Folger and Nam (1967), Schwartz (1971) and Lkhagvasuren (2014).

This does not entirely rule out the possibility that local returns to human capital are driving the mobility gap - if these returns are tied to particular task specializations. This can be tested by studying net and gross flows within detailed occupation-defined labor markets: for each education category, I estimate the within-occupation net migration rate as $\sum_o \gamma_o \frac{1}{2n_o} \Sigma_s |n_{os}^{in} - n_{os}^{out}|$, where γ_o is the fraction of individuals employed in occupation group o, and n_{os}^{in} and n_{os}^{out} are the numbers of in/out-migrants to/from state s employed in occupation o. I restrict the sample to individuals employed at the time of survey, when occupations are recorded.[11] I report estimates separately using 98 2-digit occupations (columns 4-6) and 466 3-digit occupations (7-9). The education gradient in net migration (columns 5 and 8) remains remarkably flat, even using the 3-digit definition. And again, the net-gross ratios are strongly decreasing in education. Appendix D reproduces these migration rates separately for different experience groups, and the key results are preserved.

This suggests the mobility gap cannot be explained by better educated workers converging on particular states - even within detailed occupation-defined markets. For example, there may be many software engineers flocking to California (relative to hairdressers), but there are also many moving in the opposite direction. This is not to say California does not offer particular productive advantages to software engineers. But in the context of spatial arbitrage (in the spirit of Roback, 1982), any such skill-location complementarities may be manifested more in the geographical distribution of population stocks than population flows - in an environment with inelastic housing or other congestion externalities.

3 Jobs ladder model of migration

3.1 Overview

I set the model in continuous time. The model and its parameters are defined for an individual worker i; but to ease the notation, I suppress the subscript i until I set out the empirical specification. Workers are distributed across J areas denoted by subscript j. In the main exposition, I assume these areas are identical - at least in the steady-state. One can certainly model local invariance in worker values as an *outcome* of spatial arbitrage (as in Roback, 1982); but since

[11] Note this is immediately *after* the twelve month period in which migration occurs. This is the appropriate time to measure occupation for this exercise: an individual's ex post occupation is a good indicator of the job market in which they were originally searching.

my focus is on worker behavior, I have chosen to take local wage offers and prices as given. Still, I do briefly consider the implications of endogenous wage-setting in Section 3.6 below - together with a range of other theoretical extensions.

Unemployed workers receive income b, and employed workers are paid a wage equal to:

$$w(\varepsilon, X) = \gamma' X + \sigma^{\varepsilon}(X)\varepsilon \tag{1}$$

where X is a vector of individual characteristics defining human capital, and ε is an idiosyncratic term representing job match quality. One might suppose that returns to human capital, γ, vary locally. But, the evidence on net flows (Section 2.3) suggests this variation does not contribute to the mobility gap between education groups. I have chosen to rule out differential local returns in the main exposition, but I study its implications in an extension below.

The importance of match quality depends on the offer dispersion σ^{ε}, which is permitted to vary with an individual's human capital X. My central claim is that σ^{ε} is increasing in education. This may be motivated intuitively by a notion of specialized skills. Or, in the spirit of Sattinger (1975), it may be characterized as supermodularity between human capital X and job attributes (such as task complexity or firm quality) represented by ε. To ease notation, I suppress the X argument of σ^{ε} in the analysis below.

Both employed and unemployed workers in area j draw local job offers at a finite exogenous rate λ_j, and they draw offers from other areas $k \neq j$ at rate $\pi\lambda_k$. One might suppose $\pi < 1$ if long-distance job search is more costly. For example, many job offers arrive through personal networks (Granovetter, 1995), and networks are weaker at longer distances. I take π as given; but as I point out below, there are good reasons to believe it may vary with education. I study a steady-state with $\lambda_j = \lambda$ for all j, though I do also consider local deviations from this state.

Job offers are characterized by match quality draws ε from an exogenous distribution, F^{ε}. Both F^{ε} and its density f^{ε} are continuous and differentiable over their support, and the hazard rate $\frac{f^{\varepsilon}(\varepsilon)}{1-F^{\varepsilon}(\varepsilon)}$ is monotonically increasing in ε. If an offer happens to arrive from outside area j, workers also draw a fixed moving cost equal to $m = \sigma^{\mu}(X)\mu$, payable on acceptance of the offer. The parameter $\mu \sim F^{\mu}$ is stochastic and bounded below by zero, where both F^{μ} and its density f^{μ} are continuous and differentiable. I also assume the density of f^{μ}, i.e. $\mu\frac{f^{\mu'}(\mu)}{f^{\mu}(\mu)}$, is monotonically decreasing in μ. This is a stronger condition than log concavity (or a monotone hazard), but it is satisfied by standard distributions with finite lower supports[12]. σ^{μ} determines the size of moving costs and is permitted to vary with human capital X; though as before, I suppress the X argument below.

Workers condition their acceptance of long-distance offers on their ε and μ draws. Implicitly, the zero lower bound on μ rules out "non-job" motivations for moving (Figure 2 suggests

[12]E.g. the two-parameter Weibull, two-parameter gamma, exponential, chi-squared, log normal and uniform. See Andersen (1996) and Nocke and Yeaple (2008)

these do not contribute to the mobility gap). But in an extension below, I do briefly consider the implications of negative cost draws; and I also show how one can generate equivalent results in a model with heterogeneous local preferences (instead of fixed moving costs).

Workers can exit jobs in two ways: either through a quit (if they receives a more attractive offer) or involuntary separation (to unemployment). The latter arrive randomly at rate δ. On separation, workers remain (at least initially) in their current area of residence. I rule out the possibility of moving without a job in hand: i.e. from unemployment in one area to unemployment in another. In practice, the evidence in Figure 2 suggests that such speculative migration is rare; but I briefly consider this possibility in an extension below.

3.2 Workers' value and job matching rates

Given the offer rate is independent of employment status, the reservation wage for unemployed workers is simply the out-of-work income b (Burdett and Mortensen, 1998). Conditional on human capital X, the reservation match quality which generates a wage offer equal to b is:

$$\varepsilon_R = \frac{b - \gamma' X}{\sigma^\varepsilon} \tag{2}$$

The unemployment value in area j is therefore equal to $V_j(\varepsilon_R)$, where $V_j(\varepsilon)$ is the value of a local job with match quality ε:

$$\begin{aligned} rV_j(\varepsilon) &= \gamma' X + \sigma^\varepsilon \varepsilon + \delta\left[V_j(\varepsilon_R) - V_j(\varepsilon)\right] + \lambda_j \int_\varepsilon^\infty \left[V_j(\varepsilon') - V_j(\varepsilon)\right] dF^\varepsilon(\varepsilon') \\ &\quad + \pi \sum_{k \neq j} \lambda_k \int_0^\infty \left[\int_{-\infty}^\infty \max\left\{V_k(\varepsilon') - V_j(\varepsilon) - \sigma^\mu \mu, 0\right\} dF^\varepsilon(\varepsilon')\right] dF^\mu(\mu) \end{aligned} \tag{3}$$

To ease notation, I have suppressed X as an argument in V_j. Workers discount utility at rate r. The first term, $\gamma' X + \sigma^\varepsilon \varepsilon$, is the flow utility. The second term accounts for the asset loss associated with job separations, which arrive at rate δ. The final two terms describe the value of local and long-distance search respectively. Workers accept any local offer yielding ε' (distributed F^ε) exceeding ε, where ε is the initial match quality. And conditional on initial quality ε and a cost draw μ, workers accept an offer ε' outside j if the worker's job surplus $V_k(\varepsilon') - V_j(\varepsilon)$ exceeds the moving cost $\sigma^\mu \mu$.

Imposing that offer rates are the same everywhere, i.e. $\lambda_j = \lambda$, $V_j(\varepsilon)$ collapses to:

$$\begin{aligned} rV(\varepsilon) &= \gamma' X + \sigma^\varepsilon \varepsilon + \delta\left[V(\varepsilon_R) - V(\varepsilon)\right] + \lambda \int_\varepsilon^\infty \left[V_j(\varepsilon') - V_j(\varepsilon)\right] dF^\varepsilon(\varepsilon') \\ &\quad + \tilde{\pi} \lambda \int_0^\infty \left[\int_\varepsilon^\infty \max\left\{V(\varepsilon') - V(\varepsilon) - \sigma^\mu \mu, 0\right\} dF^\varepsilon(\varepsilon')\right] dF^\mu(\mu) \end{aligned} \tag{4}$$

where $\tilde{\pi} = \pi(J-1)$. Conditional on initial match quality ε, workers form local matches at rate:

$$\rho_L(\varepsilon) = \lambda\left[1 - F^{\varepsilon}(\varepsilon)\right] \tag{5}$$

and cross-area matches at rate:

$$\rho_C(\varepsilon) = \tilde{\pi}\lambda \int_{\varepsilon}^{\infty} \left[F^{\mu}\left(\frac{V(\varepsilon') - V(\varepsilon)}{\sigma^{\mu}}\right)\right] dF^{\varepsilon}(\varepsilon') \tag{6}$$

Notice that $\rho_C(\varepsilon)$ is increasing in the job surplus, $V(\varepsilon') - V(\varepsilon)$. Intuitively, in the absence of spatial disparities, a surplus is necessary to justify the cost of moving.

3.3 The crucial role of search frictions

In an environment with identical areas j, a finite offer rate λ is crucial to motivating geographical mobility. To see precisely why, it is useful to rewrite the job surplus in (6) as:

$$V(\varepsilon + z) - V(\varepsilon) = \int_{\varepsilon}^{\varepsilon+z} V'(x)\,dx = \sigma^{\varepsilon}\Omega(z, \varepsilon) \tag{7}$$

where

$$\Omega(z|\varepsilon) = \int_{\varepsilon}^{\varepsilon+z} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx \tag{8}$$

is the surplus *in ε units* accruing to a match which raises match quality by z, for a worker with initial match quality ε; and $\sigma^{\varepsilon}\Omega(z|\varepsilon)$ is the surplus in wage units. Using (7), the cross-area matching rate in (6) can be expressed as:

$$\rho_C(\varepsilon) = \tilde{\pi}\lambda \int_{\varepsilon}^{\infty} \left[F^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}} \int_{\varepsilon}^{\varepsilon'} \frac{1}{r + \delta + \rho_L(x) + \rho_C(x)} dx\right)\right] dF^{\varepsilon}(\varepsilon') \tag{9}$$

Search frictions matters for two reasons. First, for given initial match quality ε, a worker will only accept a long-distance offer (at large expense) if he does not expect an equivalent local offer to arrive soon. In terms of (9), a small offer rate λ ensures ρ_L (in the denominator) is small. This expands the (discounted) value of long-distance matches and increases their frequency - relative to local matches.

Second, as I show in Appendix G, the offer rate λ affects the equilibrium distribution of match quality ε across workers. For an infinite offer rate λ (or zero separation rate δ), all workers will benefit from the maximum match quality in equilibrium; so there will be no job surplus and no cross-area matching. But as λ declines relative to δ (i.e. the market becomes "thinner"), workers increasingly find themselves at lower "rungs" of the jobs ladder (i.e. lower ε); so job matches will yield larger surpluses - which facilitates greater geographical mobility.

3.4 Impact of σ^ε and σ^μ on job matching rates

I now consider the impact of offer dispersion σ^ε and the size of moving costs σ^μ on local and cross-area matching rates, across the support of match quality ε.

Proposition 1. *Given a worker's initial match quality ε, the local matching rate $\rho_L(\varepsilon)$ is independent of the offer dispersion σ^ε. But the cross-area matching rate $\rho_C(\varepsilon)$ is increasing in σ^ε and decreasing in the size of moving costs σ^μ.*

The independence of local job matching follows from (5). Intuitively, local job transitions are costless, so a strictly positive job surplus is not necessary for the acceptance of an offer. So, larger offer dispersion σ^ε (and the larger associated surplus) will have no effect on $\rho_L(\varepsilon)$.

In contrast, cross-area matching is costly, so job surplus does matter. Surpluses are increasing in offer dispersion σ^ε, but they justify fewer moves if σ^μ is larger: this follows directly from (9). Alternatively, a larger offer dispersion σ^ε can be interpreted as "thinning" the market; and an equivalent argument to Section 3.3 then applies.

Proposition 2. *For sufficiently large moving costs (i.e. sufficiently large σ^μ), the positive effect of offer dispersion σ^ε on the cross-area matching rate $\rho_C(\varepsilon)$ is decreasing in a worker's initial match quality ε. Similarly, the negative effect of σ^μ on $\rho_C(\varepsilon)$ is decreasing in ε.*

Intuitively, workers with larger initial match quality ε have fewer rungs of the jobs ladder left to climb. Consequently, any increase in offer dispersion σ^ε will have less effect on their surpluses in future matches and, therefore, on their cross-area matching rate. Similarly, workers with larger ε will be less sensitive to changes in moving costs σ^μ: they are unlikely to move either way.

To see this more formally, consider the derivative of (6) with respect to $\log\frac{\sigma^\varepsilon}{\sigma^\mu}$:

$$\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^\varepsilon}{\sigma^\mu}} = \tilde{\pi}\lambda\int_\varepsilon^\infty \left\{ \left[\frac{d\log\Omega(\varepsilon'-\varepsilon|\varepsilon)}{d\log\frac{\sigma^\varepsilon}{\sigma^\mu}} + 1 \right] \frac{\sigma^\varepsilon}{\sigma^\mu}\Omega(\varepsilon'-\varepsilon|\varepsilon) f^\mu\left(\frac{\sigma^\varepsilon}{\sigma^\mu}\Omega(\varepsilon'-\varepsilon|\varepsilon) \right) \right\} dF^\varepsilon(\varepsilon') \tag{10}$$

This expression is positive, consistent with Proposition 1. And in Appendix H.1, I show further that it is decreasing in initial match quality ε - for sufficiently large moving costs σ^μ. This effect is driven by the $\Omega(\varepsilon'-\varepsilon|\varepsilon)$ term (representing the job surplus), which is decreasing in ε.

In principle, the vagaries of the moving cost distribution F^μ may upset the result, especially if there are many workers on the margin of moving. A sufficient condition for Proposition 2 is that $\mu f^\mu(\mu)$ is unambiguously increasing in μ; or equivalently, the elasticity of the density f^μ globally exceeds -1. Alternatively, as I show in Appendix H.1, Proposition 2 must also hold for sufficiently large σ^μ - given my assumption (above) that the elasticity of the density is decreasing. Intuitively, both the increasing $\mu f^{\mu}(\mu)$ and large σ^{μ} ensure a substantial fraction of moving cost draws are "large". Finally, I show in the appendix that the term in square brackets[13] in (10) does not undermine the result.

3.5 Response to local job creation

Until now, I have focused on the implications of within-area offer dispersion σ^{ε} for steady-state mobility. But as I now show, it can also help account for differences in the migratory responses to local employment shocks. The model in this paper is not entirely suitable for assessing local labor market adjustment, as I take the firm-side of the economy (i.e. the offer distribution and arrival rate) as given. Nevertheless, the model does offer a simple intuition for the migratory response at the moment of the shock; and in a more complete model (with endogenous labor market conditions), this should speak to the speed of local adjustment.

Specifically, I consider the migratory response (both outflows and inflows) to a shock to the local offer rate λ_j. This can be interpreted in terms of creation or destruction of local jobs. As outlined above, the evidence has consistently shown that local employment shocks elicit larger migratory responses from better educated workers (Bound and Holzer, 2000; Wozniak, 2010; Notowidigdo, 2011; Amior and Manning, 2018). These studies have relied on initial local industrial composition to identify these shocks (as in Bartik, 1991): some industries have consistently shed employment, while others have expanded rapidly. My focus here on shocks to λ_j is consistent with this approach. For simplicity, I study the response around a steady-state where offer rates $\lambda_j = \lambda$ and worker values $V_j = V$ are spatially invariant.

Proposition 3. *For sufficiently large moving costs (i.e. sufficiently large σ^{μ}), the migratory response to the local offer rate λ_j is increasing in offer dispersion σ^{ε}.*

I begin with the response of outflows. The outflow rate from area j is:

$$\rho_{Cj}^{Out}(\varepsilon) = \tilde{\pi}\lambda \int_{\varepsilon}^{\infty} \left[F^{\mu}\left(\frac{V(\varepsilon') - V_j(\varepsilon)}{\sigma^{\mu}} \right) \right] dF^{\varepsilon}(\varepsilon') \tag{11}$$

and the derivative with respect to the local offer rate λ_j is:

$$\frac{d\rho_{Cj}^{Out}(\varepsilon)}{d\lambda_j} = -\frac{\tilde{\pi}\lambda}{\sigma^{\mu}} \frac{dV_j(\varepsilon)}{d\lambda_j} \int_{\varepsilon}^{\infty} \left[f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}} \Omega(\varepsilon' - \varepsilon|\varepsilon) \right) \right] dF^{\varepsilon}(\varepsilon') \tag{12}$$

I show in Appendix H.2 that, for sufficiently large moving costs σ^{μ}, the response of local worker value $\frac{dV_j(\varepsilon)}{d\lambda_j}$ is positive and increasing (proportionally) with σ^{ε}. Intuitively, workers

[13] The impact of σ^{ε} and σ^{μ} on $\Omega(\varepsilon' - \varepsilon|\varepsilon)$, for given ε, comes entirely through their effect on cross-area matching $\rho_C(\varepsilon)$. Looking at (8), changes in $\rho_C(\varepsilon)$ affect the rate at which the job surplus is discounted.

who expect larger match surplus will benefit more from local job creation. And so, the outflow response $\frac{d\rho_{Cj}^{Out}(\varepsilon)}{d\lambda_j}$ must be negative and increasing (in magnitude) in σ^{ε}. Similarly to Proposition 2, the vagaries of the moving cost distribution F^{μ} may in principle upset the result. But for the same reasons as before, the proposition will hold for sufficiently large σ^{μ}.

Next, I turn to the inflow response. The total inflow from all areas outside j is:

$$\rho_{Cj}^{In}(\varepsilon) = \tilde{\pi}\lambda_j \int_{\varepsilon}^{\infty} \left[F^{\mu} \left(\frac{V_j(\varepsilon') - V(\varepsilon)}{\sigma^{\mu}} \right) \right] dF^{\varepsilon}(\varepsilon') \tag{13}$$

and the response to the local offer rate is:

$$\frac{d\rho_{Cj}^{In}(\varepsilon)}{d\lambda_j} = \frac{\tilde{\pi}\lambda}{\sigma^{\mu}} \int_{\varepsilon}^{\infty} \left[\frac{dV_j(\varepsilon')}{d\lambda_j} f^{\mu} \left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}} \Omega(\varepsilon' - \varepsilon | \varepsilon) \right) \right] dF^{\varepsilon}(\varepsilon') + \frac{\rho_C(\varepsilon)}{\lambda} \tag{14}$$

Based on the same reasoning as above, the term in square brackets must be increasing in offer dispersion σ^{ε}. And given Proposition 1, the $\frac{\rho_C(\varepsilon)}{\lambda}$ term is also increasing in σ^{ε}. Intuitively, a larger σ^{ε} places more workers (outside area j) on the margin of moving, so local job creation will elicit larger migratory inflows. This direct effect of λ_j does not exist for outflows; and this may help explain why, in practice, local population adjustment materializes largely through changes in inflows rather than outflows: see e.g. Monras (2015), Dustmann, Schoenberg and Stuhler (2017), Amior (2018) and Amior and Manning (2018).

3.6 Other considerations

Before moving to the evidence, I briefly consider a number of other pertinent issues: (i) the distribution of match quality ε among workers, (ii) the determinants of the offer distribution, (iii) the determinants of long-distance search intensity π, (iv) speculative migration, (v) negative moving costs, (vi) migration distance, (vii) heterogeneous local preferences, and (viii) the implications of differential local returns to human capital.

Distribution of match quality. In the analysis above, I have focused on the determination of matching rates $\rho_L(\varepsilon)$ and $\rho_C(\varepsilon)$ for given ε. But the equilibrium distribution of ε across workers may itself vary with education, and this will have implications for mean matching rates. In Appendix G, I show the mean matching rate is increasing in the separation rate δ and decreasing in the reservation match quality ε_R. Intuitively, larger δ and smaller ε_R push workers down the jobs ladder, so a larger fraction of job offers become viable.

Determinants of offer distribution. I have chosen to take the offer distribution as given. But in a more complete model, offer dispersion may originate from dispersion in idiosyncratic match productivity, driven in turn by complementarities between human capital and job attributes. This would generate larger total surpluses in job matches, which are shared by both firms and workers - according to the wages posted by firms or negotiated by the parties.[14] To the extent that the surplus is not fully captured by firms, dispersion in match productivity will pass through to wage offers. In an earlier version of this paper (Amior, 2015), I offer such a model with optimizing firms; and I also bring evidence that firms recruiting better educated workers invest more resources in the recruitment process (in terms of applications received, applicants interviewed and human resource hours). As I argue there, this is consistent with these firms securing larger match surplus themselves.

Determinants of search intensity. I have taken π, the intensity of long-distance job search, as given. But there is good reason to believe that π is increasing in education. Suppose workers (and firms in a more complete model) choose how much effort they invest in long-distance job search (and recruitment, e.g. through fly-outs of potential hires). In the model I have described here, the amount of effort should be increasing in the expected surplus (net of moving costs) in cross-area matches. And therefore, π should be increasing in offer dispersion σ^{ε} (or in match productivity dispersion, in a model with firms) and decreasing in moving costs σ^{μ}. In this way, endogenous search effort would serve to *amplify* any effects of σ^{ε} and σ^{μ} - which would reinforce my hypothesis. Still, I show below that observable differentials in job surplus by education may be sufficient to account for the mobility gap without resorting to endogenous search costs - at least under a particular parameterization of the costs distribution.

Speculative migration. Until now, I have ruled out the possibility of individuals moving speculatively to look for work: i.e. from unemployment in one area to unemployment in another. Of course, this is only relevant in an environment where local utilities differ; and few workers actually make speculative moves in practice (Figure 2). Suppose that, at rate ξ, unemployed workers are given the option of making a speculative move, conditional on a μ draw from the moving cost distribution F^{μ}. Notice there is no incentive to make a costly speculative move if $\pi = 1$, i.e. if workers already have full access to all jobs. As π decreases though, workers are more likely to make speculative moves and less likely to engage in cross-area matching (i.e. $\rho_C(\varepsilon)$ falls). In this way, speculative migration and cross-area matching are substitutes. And if π is indeed larger for better educated workers (for the reasons discussed above), this may help explain why they make fewer speculative moves (Figure 2). See Amior (2015), an earlier version of this paper, for further discussion on this point.

Negative moving costs. The results above are derived for a moving cost distribution whose support is bounded below by zero. However, as Figure 2 shows, many long-distance moves are motivated by broadly defined "non-job" (primarily family and housing) reasons; and low-

[14] The particular multi-location environment of this model offers interesting implications for wage setting. To the extent that match complementarities encourage geographical mobility (by expanding job surpluses), outside options in distant locations become more viable. Firms would then be compelled to offer higher wages to compete for workers over longer distances (as in Caldwell and Danieli, 2018), and this would amplify any initial effect of these complementarities on the job surplus and mobility.

educated movers disproportionately report such reasons. This result can be rationalized by differences in offer dispersion σ^{ε}, if these non-job reasons are characterized as negative cost draws (as in Kennan and Walker, 2011). Given a meager job surplus (due to low σ^{ε}), low-educated movers are disproportionately selected from workers who draw low (and especially negative) moving costs. And similarly, given the meager surplus, the low-educated are more likely to break their existing job match and move elsewhere for the sake of a given non-job benefit. I explore the selection idea further in Section 6, when I study the wage returns to long-distance matching.

Migration distance. I have not explicitly modeled distance above, but its effect can be interpreted through the lens of moving cost draws. Conditional on a higher cost draw, the cross-area matching rate will be more elastic to offer dispersion: a consequence of the monotone hazard rate of the offer distribution F^{ε}.[15] So, job surplus should matter more for longer-distance migration decisions - to the extent that these are more costly. This is consistent with evidence in Appendix A.3 that the (positive) education slope is proportionally steeper for cross-state than cross-county moves. See also Schwartz (1973) and Davis and Dingel (2019) for similar evidence on the effect of distance.

Heterogeneous local preferences. In the main exposition, I have modeled the migration friction as a fixed one-off cost m. But one can generate equivalent results by replacing this moving cost with heterogeneous preferences over locations (as in e.g. Moretti, 2011, Gyourko, Mayer and Sinai, 2013, and Hilber and Robert-Nicoud, 2013). Suppose a worker i residing in area j receives a flow utility equal to $w(\varepsilon, X) + \sigma^{\alpha}(X)\alpha_{ij}$, where the α_{ij} matches are i.i.d. random draws from a distribution F^{α} with a maximum of 0, and where σ^{α} determines the strength of preferences. Given human capital X, the value of a match can be summarized by the pair (ε, α). A worker with current match quality (ε, α) will accept a job offering (ε', α') if $\sigma^{\varepsilon}(\varepsilon' - \varepsilon) + \sigma^{\alpha}(\alpha' - \alpha) \geq 0$. So, the cross-area matching rate is:

$$\rho_C(\varepsilon, \alpha) = \tilde{\pi}\lambda \int_{-\infty}^{\infty} \left[1 - F^{\alpha}\left(\alpha + \frac{\sigma^{\varepsilon}}{\sigma^{\alpha}}(\varepsilon - \varepsilon')\right)\right] dF^{\varepsilon}(\varepsilon') \tag{15}$$

As σ^{α} becomes large, the equilibrium distribution of workers' amenity matches α will converge to a mass point at zero (its maximum value). And so, the expected cross-area matching rate $\mathbb{E}[\rho_C(\varepsilon, \alpha)|\varepsilon]$ for given ε will converge to $\tilde{\pi}\lambda \int_{\varepsilon}^{\infty} \left[1 - F^{\alpha}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\alpha}}(\varepsilon - \varepsilon')\right)\right] dF^{\varepsilon}(\varepsilon')$. This is analogous to (9), and Propositions 1 and 2 follow immediately, though with σ^{μ} (the size of moving costs) replaced by σ^{α} (the strength of local preferences). One advantage of this

[15] Consider a worker who draws a moving cost with annuitized value $\tilde{m}$. I define the "annuitized cost" more formally in equation (20) below: it summarizes the minimum wage improvement required to justify a given long-distance move. The probability of accepting such a cross-area match (with ex ante unspecified wage offer) is $\left[1 - F^{\varepsilon}\left(\varepsilon + \frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right]$, and the elasticity of this probability to offer dispersion σ^{ε} is $\frac{f^{\varepsilon}\left(\varepsilon + \frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{1 - F^{\varepsilon}\left(\varepsilon + \frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}$. Given the monotone hazard assumption, the probability is more elastic for larger cost draws $\tilde{m}$.

specification is that it can motivate return migration: workers who leave home for a productive match (i.e. a high ε) may later return to restore their α. Kennan and Walker (2011) emphasize that return migration accounts for many long-distance moves, though I show in Appendix B that it does not drive education differentials in mobility.

Differential local returns. I have not studied the implications of differential local returns to human capital in the analysis above. But, the model is capable of accounting for this - by placing an area j subscript on the γ parameter in (1):

$$w_j(\varepsilon, X) = \gamma_j' X + \sigma^{\varepsilon}\varepsilon \tag{16}$$

where the vector X accounts for a multi-dimensional index of human capital. For example, software engineers may expect larger wage offers in California than New York, and the reverse for bankers. Interestingly, local dispersion in γ_j will actually *reduce* the rate of cross-area matching in steady-state. This effect is analogous to the effect of larger σ^{α} in the previous paragraph. Intuitively, individuals will be reluctant to leave locations which offer larger returns to their particular human capital. However, suppose workers retire and are replaced by new entrants with different Xs. Differential local returns may then boost geographical mobility, as new entrants sort to the areas which offer them the largest returns.

Still, to the extent that differential returns account for the mobility gap between education groups, we should expect to see large net flows of high educated individuals - especially within detailed occupation groups. But this is not what the evidence shows (Section 2.3), even when I restrict the sample to new labor market entrants (Appendix D). In line with (16), we should also expect that the wage return to long-distance matching is largely driven by area-specific returns to human capital. But in Section 6 below, I find no evidence that the return is driven by differential local prices of observable human capital - or even of occupation-defined tasks.

4 Evidence on job matching rates

In this section, I show that empirical patterns in within and cross-state matching rates are consistent with better educated workers facing larger offer dispersion σ^{ε} (and enjoying larger job surpluses). However, using matching rates alone, it is not possible to identify the effect of σ^{ε} separately from differences in moving costs σ^{μ}: see Propositions 1 and 2 above. I attempt to disentangle these effects in Sections 5 and 6, using evidence on the wage returns to matching.

4.1 Data

I base my analysis on the Survey of Income and Program Participation (SIPP). The SIPP offers substantial samples and high-frequency waves, just four months apart. I study transitions between outcomes recorded at the end of each wave[16] for individuals with 1 to 30 years of labor market experience and no business income, in SIPP panels beginning 1996, 2001, 2004 and 2008 (which cover the period 1996-2013).

A "job match" occurs when an individual works for an employer at the end of wave t for whom they did not work in $t-1$. For individuals with multiple jobs at the end of t, I restrict attention to cases where the new job is the "primary" job: that is, the job which occupies the most weekly hours.[17] A cross-state job match is one which is accompanied with a change in state of residence. In principle, it would be preferable to use the state of the individual's workplace (to account for workers who choose to commute), but this information is not available.

4.2 Mean job matching rates

Table 2 summarizes mean within-state and cross-state matching rates by education. Identifying the "areas" in Section 3 with states, these represent the mean $\rho_L(\varepsilon)$ and $\rho_C(\varepsilon)$ across the distribution of workers' ε, i.e. $\mathbb{E}[\rho_L(\varepsilon)]$ and $\mathbb{E}[\rho_C(\varepsilon)]$. Column 1 shows the within-state rate is decreasing in education: about 10 percent of high school dropouts begin a new local job in each wave, compared to 6 percent of postgraduate degree holders. The patterns are similar for the initially employed (column 2), though with smaller numbers. This may be indicative of better educated workers facing a smaller offer rate λ, but notice they find jobs a little quicker from unemployment (column 3). Looking at equation (5), a more plausible explanation for their lower $\mathbb{E}[\rho_L(\varepsilon)]$ is that, on average, they enjoy larger match quality ε. This a natural consequence of their much lower separation rates to non-employment (column 9): lower separation rates allow them to maintain higher positions on the jobs ladder (i.e. larger ε), so fewer job offers can tempt them. See the discussion in Section 3.6.[18]

The mean cross-state matching rates $\mathbb{E}[\rho_C(\varepsilon)]$ in columns 5-8 are of course much smaller. But the education slopes are consistently positive and proportionally much steeper. If better educated workers do indeed enjoy higher match quality ε, they must therefore also face substantially larger cross-area matching rates $\rho_C(\varepsilon)$ for given ε. And in light of Proposition 1, this is consistent with larger offer dispersion σ^{ε} or smaller moving costs σ^{μ}.

[16]SIPP respondents report their earnings at the end of each month, but I do not exploit these monthly frequencies. The SIPP is known to suffer from severe seam bias (see e.g. Marquis and Moore, 2010), presumably due to poor recall: monthly changes in individuals' outcomes tend to be larger between months at the seam of two waves than within the same wave.

[17]Individuals report up to two jobs in each wave. In cases where there are two jobs with equal hours, I define the primary job as the first one reported by the individual.

[18]Among the non-employed, better educated workers do find work somewhat more quickly (columns 3-4). In the model above, this may be interpreted in terms of a lower reservation match quality ε_R. However, the impact of this on the mean matching rate (in column 1) is apparently more than offset by the much starker differences in separation rates in column 9. See also Mincer (1991) on education differences in employment transitions.

As one would expect, unemployed workers (who lie at the bottom of the jobs ladder) form more cross-state matches: 0.39 percent over four-month waves on average, compared to 0.18 percent for employed workers. And interestingly, the education gradient is also much steeper for the unemployed. Based on Proposition 2, this is again consistent with better educated workers facing larger σ^{ε} or smaller σ^{μ}: cross-state mobility should respond more heavily at the bottom of the jobs ladder.

4.3 Job matching rates within employment cycles

Table 2 sheds some light on Proposition 2 by focusing on one dimension of the jobs ladder: initial employment status. I now dig further into Proposition 2 by studying heterogeneity in match quality ε among the initially employed. To this end, I restrict attention to individuals within so-called "employment cycles" (in the language of Wolpin, 1992, or Barlevy, 2008). That is, I study the rate at which workers transition from one job at the end of wave $t-1$ to another at the end of t, without reporting an unemployment or layoff spell in between.

Match quality ε is not observable: though I do see wages, these are conflated with differences in offer dispersion σ^{ε}. Instead, I proxy initial ε with initial (i.e. $t-1$) job tenure. Intuitively, conditional on the offer rate λ, a longer tenure reflects a higher valuation of a particular match, relative to the outside options which may arise: see e.g. Mincer and Jovanovic (1979) or Burdett (1978). This is a natural consequence of a frictional environment where jobs are modeled as "pure search goods" (as in Section 3 above) or "experience goods" (as in e.g. Jovanovic, 1979), or if tenure drives the accumulation of firm-specific human capital (e.g. Topel, 1991). Importantly, as I show in Appendix E, tenure is only informative about match quality at larger values of tenure: for workers who have just begun their job, tenure tells us little about match quality.

In Figure 5, I study how within-state and cross-state matching rates vary with initial match quality, within employment cycles. Panel A plots the within-state rate on the log of initial job tenure (measured in months), separately by education. I have divided the support of tenure into decile bins, within each education group. Each data point identifies the matching rate and mean log tenure in a given decile bin. Trivially, within-state matching rates are decreasing in initial tenure. Looking at the decile markers, tenure is typically larger for better educated workers. This can account for part of the negative education differentials in mean matching rates in Table 2: at least conditional on high initial tenure, education has little effect in Panel A. The gaps are larger however for lower tenure: intuitively, tenure provides little information on match quality for new job matches.

The cross-state matching rates in Panel B look very different. As before, these are decreasing in initial tenure. But the education differentials are now clearly positive and strongly decreasing in tenure - reaching zero at the top of the support. Based on Propositions 1 and 2, these patterns are consistent with better educated workers facing larger σ^{ε} or smaller σ^{μ}.

As an aside, these effects may help account for patterns in matching rates by labor market experience. First, Panel C shows that within-state matching rates are decreasing in experience. Intuitively, older workers have spent more years searching, so they should typically enjoy larger match quality ε and make fewer job transitions (Topel and Ward, 1992; Manning, 2003*a*). Gottfries and Teulings (2016) find this accumulated "search capital" can explain a third of the return to labor market experience. One might make an alternative argument based on firm-specific human capital, with equivalent implications for match quality. Notice the education gaps are larger than in Panel A: this reflects the higher match quality enjoyed by better educated workers.

Second, Panel D shows the mobility gap between education groups is much larger at the beginning of workers' careers - which reflects the pattern of Panel B. By inspection, the support of tenure (in Panel B) generates more than half the variation (in cross-state matching rates) compared to the support of experience (in Panel D): see the axis ranges. I show in Appendix F that, statistically, initial tenure can explain away about one third of the experience effect. To the extent that tenure is an imprecise proxy for match quality, this presumably understates the true contribution of match quality ε to the experience effect. But of course, experience may itself contain independent information on moving costs.[19] Though the effect of experience on mobility is an interesting question, I do not pursue it further: my focus here is the effect of education.

To summarize, the evidence in this section offers empirical support for a jobs ladder model of migration. Based on Propositions 1 and 2, the patterns of within-state and cross-state job matching are consistent with better educated workers facing larger offer dispersion σ^{ε} and/or lower moving costs σ^{μ}. In what follows, I attempt to disentangle these two effects.

5 Wage returns to within-state job matching

I now attempt to identify variation in expected job surplus (and offer dispersion σ^{ε}) across education groups, by estimating the mean wage return to *within-state* job matching. I focus on within-state returns to exclude a contribution from moving costs. And I then consider whether this variation is sufficiently large to account for differentials in *cross-state* job matching.

[19]In Figure 3 above, older workers report less willingness to move for work. A human capital explanation is that older workers have fewer years to benefit from the sunk cost of moving (see e.g. Kennan and Walker, 2011).

5.1 Empirical specification

Using equation (1), the expected wage change between $t-1$ and t for some individual i can be written as:

$$\mathbb{E}(w_{it} - w_{it-1}) = \sigma^{\varepsilon}\mathbb{E}(\varepsilon_{it} - \varepsilon_{it-1}) \cdot I[NewJob_{it} = 1] + \gamma'(X_{it} - X_{it-1}) \tag{17}$$

where $\mathbb{E}(\varepsilon_{it} - \varepsilon_{it-1})$ is the expected change in match quality for individuals who take a new job between $t-1$ and t (for whom the indicator function I takes 1), and $\gamma'(X_{it} - X_{it-1})$ is the contribution from changes in human capital. The focus on *mean* returns is motivated by empirical convenience. But as I show below, the mean is the relevant moment for predicting long-distance matching rates under certain (specifically uniform) distributional assumptions on moving costs.

Until now, I have interpreted w as a dollar wage. In principle though, w may represent a log wage, yielding a model in log utility.[20] This may not be theoretically innocuous since workers cannot borrow or save in this model. However, it will yield more conservative empirical results: any education gradient in *proportional* returns will be flatter than in *dollar* returns, because better educated workers earn substantially more. (17) then suggests the following empirical specification:

$$\Delta w_{it} = \beta_0 + \beta_1 NewJob_{it} + \beta_2' X_{it} + \beta_t + u_{it} \tag{18}$$

where Δw_{it} is the change in worker i's log wage between $t-1$ and t, controlling for a vector X_{it} of demographic characteristics[21] and time effects β_t. I restrict the sample to transitions which lie within "employment cycles" (as defined in Section 4.3): that is, I exclude transitions which include an unemployment or layoff spell. Without this restriction, the initial wage cannot be interpreted as a reservation - in which case I cannot identify the job surplus. I also exclude transitions which involve cross-state residential moves. So, β_1 identifies the expected wage return to a within-state match against the counterfactual of remaining in the same job. This analysis necessarily excludes individuals out of employment in $t-1$. But, it is worth emphasizing that education differentials in cross-state matching among the initially employed closely approximate those of the full population: compare columns 5 and 6 of Table 2.

[20]Grogger and Hanson (2011) show that a Roy model with linear utility and skill-invariant moving costs can better explain the observed selection of high and low-educated migrants across countries than an alternative specification with log utility and moving costs which are proportional to income. However, it is not clear whether this result is generalizable to internal migration in the US, where wage gains are much smaller.

[21]See notes under Table 3 for details.

I have no interest here in identifying a *causal* effect of an "exogenous" job change. Rather, the model makes predictions on the conditional mean wage change - and this is the moment that equation (18) identifies. Of course, this conditional mean is driven by selection on job offers, but it is precisely this selection which interests me. In particular, better educated workers should expect larger job surpluses if they face larger offer dispersion σ^{ε}. And based on Proposition 2, the effect of education should be especially large for workers with lower initial match quality ε, i.e. lower down the jobs ladder. To test these claims, I interact the $NewJob_{it}$ dummy with a set of education effects, and I estimate the model separately for different experience and tenure groups.

There is already a literature which estimates wage returns to job mobility, using similar specifications to (18). These wage returns are known to be increasing in education and decreasing in age: see Bartel and Borjas (1981), Mincer (1986), Topel and Ward (1992), Manning (2003*a*) and Gottfries and Teulings (2016). To my knowledge, this study is the first to link these differential returns to geographical mobility.

5.2 Empirical estimates

I present estimates of (18) in Table 3, based on the same SIPP sample described in Section 4.1. I identify w_{it} with log hourly wages at the end of each four-month wave t[22], and I restrict the sample to jobs which command at least 15 hours per week. Column 1 shows that the mean wage return to a new local job is 0.034. In the next column, I interact the $NewJob_{it}$ dummy with a set of education effects (all of which are included in the demographic controls). There is a steep education gradient, stretching from 0.01 (and statistically insignificant) for high school dropouts (the omitted category) to 0.05-0.06 for those with college degrees. The interaction effects are precisely estimated, with standard errors of about 0.01.

As the model predicts, these wage returns appear to be larger for workers with lower initial match quality - which I proxy with experience and initial job tenure. Columns 3 to 5 show the education differentials are largely driven by the young: for those with under 10 years of experience, β_1 ranges from 0.01 for dropouts to 0.08 for postgraduates. There is no significant return to job mobility for those with 21-30 years of experience. Similarly, there is no education gradient for individuals with more than 7 years of initial tenure.

Until now, I have assumed that match quality ε is fixed over the duration of a job. But to the extent that the wage return for better educated workers is delayed (due to seniority returns or the accumulation of firm-specific human capital), the β_1 coefficient will understate their true returns - relative to the low-educated.

[22] See notes under Table 3 for further details on wage data.

5.3 Quantifying the effect of wage returns on geographical mobility

I now consider whether these estimates of high-educated wage returns (in local matches) are sufficiently large to explain the education differentials in cross-state job matching. Such an exercise necessitates courageous assumptions on the distribution of moving costs. But it can nevertheless offer some useful guidance.

An important concept in this exercise is the "annuitized cost" of a long-distance move in terms of current wages, and I begin by defining this more precisely. Consider a worker with initial match quality ε who receives a long-distance offer with a moving cost draw equal to $m = \sigma^{\mu}\mu$. The (equivalent variation) annuitized cost, which I denote $\tilde{m}$, is the wage increase (over the worker's current wage) which would make him indifferent to accepting the offer. This is the $\tilde{m}$ which satisfies:

$$V\left(\varepsilon + \frac{\tilde{m}}{\sigma^{\varepsilon}}\right) - V(\varepsilon) = \sigma^{\varepsilon}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}|\varepsilon\right) = \sigma^{\mu}\mu \tag{19}$$

where the first equality follows from (7), and the second describes the indifference relationship. Ω is the job surplus in ε units (associated with an increase of match quality of $\frac{\tilde{m}}{\sigma^{\varepsilon}}$), as defined by (8). I have normalized the annuitized cost $\tilde{m}$ by the offer dispersion σ^{ε} to express it in ε units. Rearranging this expression for $\tilde{m}$ yields:

$$\tilde{m}(\mu|\varepsilon) = \sigma^{\varepsilon}\Omega^{-1}\left(\frac{\sigma^{\mu}}{\sigma^{\varepsilon}}\mu|\varepsilon\right) \tag{20}$$

where the annuitized cost $\tilde{m}$ is increasing in the cost draw μ and decreasing in initial match quality ε. This is equivalent to the concept of "mobility-compatible indifference wages" in Schmutz and Sidibé (forthcoming), and it offers a theoretical basis for the subjective annuitized costs of Section 2.2 above. Note that Ω can be inverted because it is monotonically increasing: see equation (8). Ω^{-1} is an implicit function; but to aid intuition, consider a first order approximation of Ω in (8) around the initial match quality ε. Substituting this approximation into (20) gives:

$$\tilde{m}(\mu|\varepsilon) \approx \sigma^{\mu}\mu\left[r + \delta + \rho_L(\varepsilon) + \rho_C(\varepsilon)\right] \tag{21}$$

where the fixed moving cost $\sigma^{\mu}\mu$ is discounted by the interest rate r and the rate of separation from a job with match quality ε. This expression is decreasing in ε: matches with lower ε do not last as long, and workers are unlikely to move for the sake of short-term jobs.

I now impose distributional assumptions on the moving cost draws. Specifically, suppose the draws of $\mu \sim F^{\mu}$ are distributed uniformly between 0 and a maximum normalized to 1: so moving costs m range from 0 to σ^{μ}. And suppose also there are no wage offers which can justify moving at the maximum cost draw: that is, for every initial match quality ε, $1 - F^{\varepsilon}\left(\varepsilon + \frac{\tilde{m}(1|\varepsilon)}{\sigma^{\varepsilon}}\right) = 0$. The odds ratio of cross-area to local matching can then be expressed as:

$$\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} = \tilde{\pi} \int_0^1 \left[\frac{1 - F^{\varepsilon}\left(\varepsilon + \frac{\tilde{m}(\mu|\varepsilon)}{\sigma^{\varepsilon}}\right)}{1 - F^{\varepsilon}(\varepsilon)} \right] d\mu \tag{22}$$

And as I show in Appendix H.3, applying the linear approximation in (21) then yields:

$$\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} \approx \frac{\tilde{\pi}\sigma^{\varepsilon}\mathbb{E}_L[\varepsilon' - \varepsilon|\varepsilon' \geq \varepsilon]}{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]} \tag{23}$$

where $\mathbb{E}[\tilde{m}(\mu|\varepsilon)]$ is the mean annuitized cost, and $\mathbb{E}_L[\varepsilon' - \varepsilon|\varepsilon' \geq \varepsilon]$ is the expected improvement in match quality ε arising from a local (subscript L) job match; so $\sigma^{\varepsilon}\mathbb{E}_L[\varepsilon' - \varepsilon|\varepsilon' \geq \varepsilon]$ is the expected wage return. Now, taking the expectation of (23) over the distribution of match quality ε for initially employed workers, and abusing Jensen's inequality:

$$\frac{\mathbb{E}[\rho_C(\varepsilon)]}{\mathbb{E}[\rho_L(\varepsilon)]} \approx \frac{\tilde{\pi}}{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}\beta_1 \tag{24}$$

where β_1 is identified by the estimating equation (18) and is equal to the expectation of local wage returns $\sigma^{\varepsilon}\mathbb{E}_L[\varepsilon' - \varepsilon|\varepsilon' \geq \varepsilon]$ for the initially employed. The $\frac{\tilde{\pi}}{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}$ term can be interpreted as a composite (inverse) measure of the constraints on cross-area job matching, accounting for both long-distance search intensity $\tilde{\pi}$ and the annuitized cost of moving.

Given my estimates of the odds ratio $\frac{\mathbb{E}[\rho_C(\varepsilon)]}{\mathbb{E}[\rho_L(\varepsilon)]}$ and β_1, I now back out values for the composite constraints $\frac{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}{\tilde{\pi}}$ for different education groups. The first three rows of Table 4 report the mean cross-state and within-state matching rates (together with the odds ratio), for the initially employed. These simply replicate the numbers from columns 2 and 6 of Table 2, but I also report matching rates for two aggregate groups (college graduates and non-graduates). The fourth row reports the mean wage return to within-state matching β_1, based on the estimates from columns 2 of Table 2. I also report β_1 estimates for college graduates and non-graduates, based on an equivalent regression with an interaction between $NewJob_{it}$ and a graduate dummy (instead of the full set of education effects).

In the final row, I impute values for the composite constraints on long-distance matching (i.e. $\frac{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}{\tilde{\pi}}$), by dividing β_1 by the odds ratio (in line with the approximation in (24)).

Given the size of the SIPP sample, the standard errors on the matching rates are negligible. So I compute standard errors for the composite constraints by simply dividing the standard errors of the β_1 estimates by the odds ratio. There is no systematic trend in the imputed values across the five basic education groups: it is lowest for high school dropouts and postgraduate degree-holders, though the standard error is very large for the former. However, the values are very similar for the aggregated college graduate (1.25) and non-graduate (1.33) groups. The difference between them is statistically insignificant, though they are precisely estimated (with standard errors of just 0.13 and 0.15).

In this sense, conditional on my courageous distributional assumptions, the differences in expected local job surplus (encapsulated by β_1) between college graduates and non-graduates can account for the gap in geographical mobility - without resorting to cost differentials. Of course, these results are entirely contingent on my distributional assumptions on moving costs[23] - though at least they can offer a useful guide. Also, it is worth stressing that the "composite constraints" measure accounts not only for moving costs, but also for long-distance search (or recruitment) intensity π. As I have argued in Section 3.6, education differentials in π are themselves a plausible (endogenous) outcome of differences in offer dispersion σ^{ε}. So even if there were systematic education differences in this imputed measure, one need not attribute this to moving costs.

6 Wage returns to cross-state job matching

I now offer additional evidence for the job surplus hypothesis which does not impose such strong assumptions, based on the returns to long-distance job matching. As I show below, these returns depend on the relative importance of selection on wage and moving cost draws. To the extent that workers select into migration because of large job surplus (and despite large costs), this will be manifested in larger long-distance returns. This insight offers an alternative means to identify the source of the mobility gap between education groups. In particular, I am able to identify bounds on the realized costs of moving - which are not conflated by long-distance search intensity π.

[23] For example, suppose the uniform distribution understates the true probability of high cost draws. Then, I would be understating the impact of the β_1 differentials on mobility: as I note in the discussion on migration distance in Section 3.6, mobility is more elastic to job surplus if costs are higher. On the other hand, if the support of the cost distribution extends to negative territory, the analysis here would be overstating the impact of the β_1 differentials.

6.1 Selection on wage and cost draws

To interpret these selection effects, I first derive the theoretical distribution of *realized* moving costs - conditional on accepting a cross-area job offer. For the purposes of this exercise, I focus on the annuitized costs $\tilde{m}$ (as defined by (20)) rather than the basic fixed costs m. This approach can aid the interpretation of the empirical estimates which follow. Conditional on accepting a cross-area match, the probability of having drawn an annuitized moving cost exceeding $\tilde{m}$ is:

$$1 - Z(\tilde{m}|\varepsilon) = \frac{\int_{\tilde{m}}^{\infty} \left[1 - F^{\varepsilon}\left(\varepsilon + \frac{x}{\sigma^{\varepsilon}}\right)\right] f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}|\varepsilon\right)\right) dx}{\int_{0}^{\infty} \left[1 - F^{\varepsilon}\left(\varepsilon + \frac{x}{\sigma^{\varepsilon}}\right)\right] f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}|\varepsilon\right)\right) dx} \tag{25}$$

The numerator describes the probability of drawing *and* accepting a cross-area offer with annuitized cost exceeding $\tilde{m}$, and the denominator describes the probability of accepting *any* cross-area offer. I make the following claim:

Proposition 4. *For sufficiently large moving costs (i.e. sufficiently large σ^{μ}), and given a worker's initial match quality ε, the expected realized annuitized cost is increasing in both σ^{μ} and σ^{ε}.*

This follows from my assumptions on the offer distribution (monotone hazard) and cost distribution (decreasing elasticity of density), and I leave the proof to Appendix H.4. Intuitively, a larger σ^{μ} implies larger unconditional cost draws, so realized costs will also be larger. And a larger σ^{ε} expands job surpluses, so workers are more likely to accept cross-area offers with high cost draws. The condition of a sufficiently large σ^{μ} ensures stability in the relationship between moving costs and their annuitized values in (20).[24]

If these realized costs could be observed, this would offer a useful test on the origins of the mobility gap. If high-educated mobility is driven by low costs (i.e. low σ^{μ}), realized costs would be *decreasing* in education. But if it is driven by large offer dispersion (large σ^{ε}) and job surplus, realized costs would be *increasing* in education. Intuitively, in the latter case, better educated workers would be moving *because* of large surplus - and *despite* the associated costs.

Clearly, these costs are unobserved. But the model does offer a way to identify upper and lower bounds on the expectation of realized annuitized costs:

Proposition 5. *(i) The expected wage return to cross-area job matching identifies an upper bound on the expectation of realized annuitized costs. (ii) For non-negative cost draws, the differential between the expected return to cross-area and local matches identifies a lower bound on the expected realized costs.*

[24] Intuitively, σ^{ε} and σ^{μ} affect the cross-area matching rate $\rho_C(\varepsilon)$, and this rate matters for the discounting of job surplus Ω in (8). However, for a sufficiently large σ^{μ}, cross-area mobility contributes relatively little to job matching overall, so changes in σ^{ε} and σ^{μ} (which only affect cross-area matching) have little effect on the Ω relationship. See Appendix H.4.

The intuition for the upper bound is simple. Let $\mathbb{E}_C[\varepsilon' - \varepsilon|\varepsilon]$ denote the expected gain in match quality on accepting a cross-area (subscript C) offer, for given initial match quality ε. The associated wage return can then be expressed as:

$$\sigma^{\varepsilon}\mathbb{E}_C\left[\varepsilon' - \varepsilon|\varepsilon\right] = \int_0^{\infty} \mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon' - \varepsilon\right)|\sigma^{\varepsilon}\left(\varepsilon' - \varepsilon\right) \geq \tilde{m}\right] dZ\left(\tilde{m}|\varepsilon\right) \tag{26}$$

where the $\mathbb{E}_L$ term is the expected return to local matches, conditional on the return exceeding the annuitized moving cost $\tilde{m}$. To derive the expected return to cross-area matches, I integrate this expression over the distribution of realized annuitized costs Z. Since the $\mathbb{E}_L$ term must exceed $\tilde{m}$, the expected return must exceed the expected annuitized costs $\int_0^{\infty} \tilde{m} dZ(\tilde{m}|\varepsilon)$ associated with those matches. Intuitively, workers will only accept cross-area offers if the associated job surplus exceeds the cost of moving.

The lower bound is identified by the differential between the expected wage return to cross-area and local job matching:

$$\begin{aligned} & \sigma^{\varepsilon}\mathbb{E}_C\left[\varepsilon' - \varepsilon|\varepsilon\right] - \sigma^{\varepsilon}\mathbb{E}_L\left[\varepsilon' - \varepsilon|\varepsilon\right] \\ = & \int_0^{\infty} \left\{\mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon' - \varepsilon\right)|\sigma^{\varepsilon}\left(\varepsilon' - \varepsilon\right) \geq \tilde{m}\right] - \mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon' - \varepsilon\right)|\varepsilon' - \varepsilon \geq 0\right]\right\} dZ\left(\tilde{m}|\varepsilon\right) \end{aligned} \tag{27}$$

Given my assumption that the offer distribution F^{ε} has a monotonically increasing hazard rate, the term in curly brackets in (27) must be less or equal to $\tilde{m}$ for all annuitized cost draws $\tilde{m} \geq 0$. See Appendix H.5 for a proof. And consequently, the cross-area/local differential in expected wage returns will identify a lower bound on the expected realized annuitized costs, $\int_0^{\infty} \tilde{m} dZ(\tilde{m}|\varepsilon)$. Intuitively, the moving cost is not always binding: there are some local offers which would be sufficient to justify a cross-area match. And so, the differential in job surplus should understate the magnitude of realized costs. This is effectively a compensating differentials argument, but accounting for search frictions: workers cannot choose from the universe of jobs.

Importantly, the lower bound result of Proposition 5 is contingent on non-negative moving costs. To see why, consider the extreme case where workers only move with negative cost draws. The differential between the expected wage returns to cross-area and local matches will then be negative. Applying the same logic as above, this differential will identify a lower bound on the expected costs in absolute terms: that is, the true expected costs will be *even more negative* than the differential. This is of course a more important consideration for low educated movers, among whom "non-job" motivations feature more heavily: see Figure 2 and the discussion in Section 3.6. And it necessitates some caution in interpreting the estimates which follow.

6.2 Empirical estimates

Motivated by Propositions 4 and 5, I now offer estimates of the wage returns to cross-state matching. Based on the sample of wage changes which involve job transitions (i.e. conditional on $NewJob_{it} = 1$), I estimate the following empirical specification:

$$\Delta \log w_{it} = \theta_0 + \theta_1 Move_{it} + \theta_X' X_{it} + \theta_t + v_{it} \tag{28}$$

where $Move_{it}$ is a dummy variable taking 1 if the individual moved state between $t-1$ and t. Based on Proposition 5, a lower bound on the expected annuitized cost of movers can be identified by the coefficient θ_1 (the premium to cross-state relative to local matching). The upper bound can be identified by $\theta_1 + \beta_1$, where β_1 is the return to local matching in (18): this is the total wage return to a long-distance match. Given the specification is in log wages, these bounds will approximate the expected costs as a fraction of initial wages; and they will also be comparable to the subjective cost estimates in Figure 4. To study how the coefficients vary with education, I interact $Move_{it}$ with education effects. Given the small sample of high school dropout movers, I aggregate high school graduates and dropouts into a single category.

I present estimates of (28) in Table 5. My basic estimate of θ_1 in column 1 is 0.06, with a standard error of 0.02. This implies the expected annuitized costs of movers are bounded below by 0.06 and above by $\beta_1 + \theta_1 = 0.09$, as a fraction of a worker's initial wage (where β_1 is taken from column 1 of Table 3).

The effect is largely driven by college graduates. Column 2 reports a slightly negative (though statistically insignificant) θ_1 for non-college workers (the omitted category). After accounting for the relevant β_1 estimates in Table 3 (column 2), this implies that both the upper and lower bounds on their expected realized costs are close to zero. In contrast, θ_1 is 0.08 ($= 0.130 - 0.054$) for undergraduate degree holders and 0.20 ($= 0.252 - 0.054$) for postgraduates (which identify the lower bounds). And adding these to the β_1 estimates in Table 3, I retrieve upper bounds of 0.15 and 0.26 respectively.[25] Columns 5-8 suggest these large graduate returns are mostly driven by the young, though there is no clear effect of initial tenure.

Of course, these returns to cross-state matching may simply reflect local differentials in the returns to human capital (see Lkhagvasuren, 2014, and the discussion in Section 3.6), rather than the job surpluses accruing to progression up a within-state jobs ladder. To study this further, I purge log wages of state-specific returns to human capital. Specifically, I regress log

[25] Based on column 2 of Table 3, I use β_1 estimates of 0.065 ($= 0.008 + 0.057$) for undergraduate degree holders and 0.056 ($= 0.008 + 0.048$) for postgraduates.

wage levels on a full set of state effects (using the full wage sample), each interacted with a full set of education indicators (five categories), and all preceding variables interacted with a quadratic in experience. I then use individual-level changes in the regression residuals as the dependent variable for (28). As it happens though, this has a negligible effect on the θ_1 estimates: see column 3.

Now, there may also be selection on local returns to particular abilities *within* education-experience groups: for example, high educated software engineers may be better rewarded in California, and high educated bankers in New York. To address this, I purge individual wages (again, using the full sample) additionally of triple interactions between state fixed effects, education fixed effects (five categories) and 3-digit occupation fixed effects (between 415 and 496 categories, depending on SIPP panel): this amounts to 68,000 fixed effects in total. But remarkably, this too makes almost no difference to the θ_1 estimates: see column 4. Certainly, it is not possible to rule out selection on returns to *unobserved* components of human capital. But the strong evidence on observables casts doubt on their importance.

To summarize, the estimates here point to much larger realized costs among better educated movers. According to Proposition 4, these larger costs must be driven by larger offer dispersion σ^{ε} or larger moving costs σ^{μ}. In other words, better educated workers typically move *because* of large job surplus and *despite* large costs; whereas low educated workers move *because* of low cost draws and *despite* meager surplus. This is consistent with the subjective reasons for moving reported in Figure 2. The results here also suggest that that the mobility differentials cannot be explained by selection on local returns to human capital.

6.3 Comparison with existing estimates of moving costs

Reassuringly, the realized annuitized costs implied by this exercise are similar in magnitude to the subjective costs from Section 2. Conditional on being "willing to move" for work, the mean annuitized cost in my PSID sample is 0.37 - relative to workers' wages. This is somewhat larger than the 0.26 upper bound implied by the SIPP wage returns, but this should be expected: the PSID estimates condition on "willingness" to move, whereas the SIPP estimates condition on *actually* moving. So the latter should be selected from lower down the costs distribution.

There are of course several other studies which have estimated moving costs. In order to compare mine with theirs, I first convert my annuitized cost estimates $\tilde{m}$ into one-off cost equivalents $m = \sigma^{\mu}\mu$. Taking expectations of (21) over the distribution of match quality ε, and again abusing Jensen's inequality:

$$\mathbb{E}[m] \approx \frac{\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}{r+\delta+\mathbb{E}[\rho_L(\varepsilon)+\rho_C(\varepsilon)]} \tag{29}$$

In line with Shimer (2005), I take a value of 0.03 for the monthly separation rate δ, and 0.03 for the mean job-to-job transition rate $\mathbb{E}\left[\rho_L(\varepsilon)+\rho_C(\varepsilon)\right]$. Interest rates r are typically negligible in comparison, so I use an overall discount rate of 0.06. Based on my estimates above, the low-educated typically move with negligible realized costs. Consider instead the case of post-graduate degree holders, who face the highest realized costs. The upper and lower bounds on expected costs are 0.26 and 0.20 respectively (see above), so take a mid-point of 0.23. Average monthly earnings for postgraduates in my SIPP sample are \$5,100 (2015 prices). Taking 23 percent yields a monthly annuitized cost of \$1,173. And dividing \$1,173 by the discount rate 0.06 gives an expected one-off moving cost (conditional on moving) of about \$20,000.

How does this compare with existing estimates? These do vary substantially, partly because they identify different objects. Most studies do not allow for individual heterogeneity in costs, which rules out the selection effects described above. Bayer and Juessen (2012) estimate a cross-state moving cost of \$34,000, using a dynamic structural model. Lkhagvasuren (2014) calibrates a Roy model and estimates a moving cost of \$28,000 to \$54,000 for moving between census divisions. Davies, Greenwood and Li (2001) estimate cross-state moving costs of around \$200,000 in a conditional logit framework. And using a similar model to mine (though estimating it structurally and imposing fixed costs), Schmutz and Sidibé (forthcoming) estimate a cost of about €15,000 between French metro areas.

In contrast, Kennan and Walker (2011) do allow for individual heterogeneity in costs. They estimate a large (unconditional) average cost of \$312,000[26], though the cost for actual cross-state movers is typically negative: this is because most moves are motivated by idiosyncratic amenity payoffs, which they factor into the cost. Importantly, their sample is restricted to high school graduates, who are more likely to move for amenity reasons: see Figure 2. Indeed, my estimates in Table 5 point to negligible realized costs for the low educated.

7 Conclusion

In this paper, I have argued that better educated workers are more mobile because they face larger dispersion in wage offers. In a frictional labor market, this generates larger job surpluses as workers climb the jobs ladder, irrespective of geography. This makes the labor market better integrated spatially, particularly for younger workers who are just beginning their careers. While these surpluses are unimportant for local job matching, they play a crucial role in driving long-distance matching - given the associated moving costs. Larger surpluses not only drive more long-distance mobility in the steady-state, but they also make migratory flows more responsive to local job creation. This can help explain why the low educated face much more persistence in local jobless rates. Though I focus on education differentials, this paper offers new insights for understanding geographical immobility more generally.

[26] Schmutz and Sidibé put most of this down to search frictions.

My hypothesis is attractive firstly because it is theoretically intuitive. Larger offer dispersion may be motivated by a notion of specialized skills or supermodularity between workers' abilities and job attributes (such as task complexity or firm quality). And second, it has strong empirical foundations. Using initial job tenure as a proxy for match quality, I show that patterns in within and cross-state matching are consistent with better educated workers facing larger offer dispersion. And crucially, I estimate large education differentials in the wage returns to within-state matching. Under certain assumptions on the distribution of moving costs, these differentials can quantitatively account for the bulk of the mobility gap - without resorting to competing explanations.

Nevertheless, I do consider the competing explanations more directly - and two in particular. The first is that better educated workers face lower moving costs. However, moving costs imputed from subjective willingness to move (which predict future mobility) are remarkably similar across education groups. I also show that wage returns are much larger for better educated workers in cross-state job matches. Using a compensating differentials argument, this suggests that better educated workers typically select into migration *because* of large job surplus and *despite* steep moving costs. In contrast, among the low-educated, returns are similarly small for both local and cross-state matches.

The second possibility is that better educated workers face larger local differentials in expected utility, driven perhaps by differential local returns to human capital. But this view is difficult to reconcile with evidence on the direction of migratory flows: high educated mobility is not driven by large net flows to particular states, even within detailed occupation-defined labor markets. Moreover, I show that the large wage returns to cross-state matches (among college graduates) cannot be explained by differential local prices of observable human capital - or even of detailed occupation-defined tasks.

Bibliography

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price. 2016. "Import Competition and the Great US Employment Sag of the 2000s." *Journal of Labor Economics*, 34(S1): S141–S198.

Amior, Michael. 2015. "Why are Higher Skilled Workers More Mobile Geographically? The Role of the Job Surplus." Centre for Economic Performance Discussion Paper No. 1338.

Amior, Michael. 2018. "The Contribution of Foreign Migrants to Local Labor Market Adjustment." Centre for Economic Performance Discussion Paper No. 1582.

Amior, Michael, and Alan Manning. 2018. "The Persistence of Local Joblessness." *American Economic Review*, 108(7): 1942–70.

Andersen, Torben M. 1996. "Rationing of Sales and Price Setting." *European Economic Review*, 40(7): 1441–1451.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Bagger, Jesper, and Rasmus Lentz. 2018. "An Empirical Model of Wage Dispersion with Sorting." *Review of Economic Studies*, 86(1): 153–190.

Balgova, Maria. 2018. "Why Don't Less Educated Workers Move? The Role of Job Search in Migration Decisions." https://mariabalgova.github.io/.

Barlevy, Gadi. 2008. "Identification of Search Models using Record Statistics." *The Review of Economic Studies*, 75(1): 29–64.

Bartel, Ann P., and George J. Borjas. 1981. "Wage Growth and Job Turnover: An Empirical Analysis." In *Studies in Labor Markets*. , ed. Sherwin Rosen, 65–90. Chicago: University of Chicago Press.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Bayer, Christian, and Falko Juessen. 2012. "On the Dynamics of Interstate Migration: Migration Costs and Self-Selection." *Review of Economic Dynamics*, 15(3): 377–401.

Blanchard, Olivier J., and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, 23(1): 1–76.

Bound, John, and Harry J. Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

Burdett, Kenneth. 1978. "A Theory of Employee Job Search and Quit Rates." *American Economic Review*, 68(1): 212–220.

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

Caldwell, S., and O. Danieli. 2018. "Outside Options in the Labor Market." https://scholar.harvard.edu/danieli.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo. 2016. "The Masking of the Decline in Manufacturing Employment by the Housing Bubble." *Journal of Economic Perspectives*, 30(2): 179–200.

Coen-Pirani, Daniele. 2010. "Understanding Gross Worker Flows across US States." *Journal of Monetary Economics*, 57(7): 769–784.

Costa, Dora L., and Matthew E. Kahn. 2000. "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990." *Quarterly Journal of Economics*, 115(4): 1287–1315.

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

Davis, Donald R., and Jonathan I. Dingel. 2019. "A Spatial Knowledge Economy." *American Economic Review*, 109(1): 153–70.

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

Dustmann, Christian, Uta Schoenberg, and Jan Stuhler. 2017. "Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment." *Quarterly Journal of Economics*, 123(1): 435–483.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2018. "Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]." Minneapolis: University of Minnesota.

Folger, John K., and Charles B. Nam. 1967. *Education of the American Population.* Washington, DC: U.S. Census Bureau.

Gottfries, Axel, and Coen N Teulings. 2016. "Returns to On-the-Job Search and the Dispersion of Wages." Centre for Macroeconomics Discussion Paper No. 1629, Tinbergen Institute Discussion Paper 16-080/VI.

Granovetter, Mark. 1995. *Getting a Job: A Study of Contacts and Careers.* Chicago: University of Chicago Press.

Greenwood, Michael J. 1973. "The Geographic Mobility of College Graduates." *The Journal of Human Resources*, 8(4): 506–515.

Gregg, Paul, Stephen Machin, and Alan Manning. 2004. "Mobility and Joblessness." In *Seeking a Premier League Economy.* , ed. David Card, Richard Blundell and Richard B. Freeman. Chicago: University of Chicago Press.

Grogger, Jeffrey, and Gordon H. Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics*, 95(1): 42–57.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar Cities." *American Economic Journal: Economic Policy*, 5(4): 167–199.

Hilber, Christian A.L., and Frédéric Robert-Nicoud. 2013. "On the Origins of Land Use Regulations: Theory and Evidence from US Metro Areas." *Journal of Urban Economics*, 75: 29–43.

Hoynes, Hilary W. 2000. "The Employment and Earnings of Less Skilled Workers over the Business Cycle." In *Finding Jobs: Work and Welfare Reform.* , ed. Rebecca Blank and David Card. New York: Russell Sage Foundation.

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

Jackman, Richard, and Savvas Savouri. 1992. "Regional Migration in Britain: An Analysis of Gross Flows Using NHS Central Register Data." *The Economic Journal*, 102(415): 1433–1450.

Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(5): 972–990.

Kaplan, Greg, and Sam Schulhofer-Wohl. 2012. "Interstate Migration has Fallen Less than you Think: Consequences of Hot Deck Imputation in the Current Population Survey." *Demography*, 49(3): 1061–1074.

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

Kennan, John. 2015. "Spatial Variation in Higher Education Financing and the Supply of College Graduates." NBER Working Paper No. 21065.

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

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin. 2016. "Matching, Sorting and Wages." *Review of Economic Dynamics*, 19: 63–87.

Lkhagvasuren, Damba. 2014. "Education, Mobility and the College Wage Premium." *European Economic Review*, 67: 159–173.

Lopes de Melo, Rafael. 2018. "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Journal of Political Economy*, 126(1): 313–346.

Lutgen, Vanessa, and Bruno Van der Linden. 2015. "Regional Equilibrium Unemployment Theory at the Age of the Internet." *Regional Science and Urban Economics*, 53: 50–67.

Machin, Stephen, Kjell G. Salvanes, and Panu Pelkonen. 2012. "Education and Mobility." *Journal of the European Economic Association*, 10(2): 417–450.

Malamud, Ofer, and Abigail Wozniak. 2012. "The Impact of College on Migration Evidence from the Vietnam Generation." *Journal of Human Resources*, 47(4): 913–950.

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

Manning, Alan. 2003*b*. "The Real Thin Theory: Monopsony in Modern Labour Markets." *Labour Economics*, 10(2): 105–131.

Marquis, Kent H., and Jeffrey C. Moore. 2010. "Measurement Errors in SIPP Program Reports." SIPP Working Paper Series No. 9008, Washington, DC: U.S. Census Bureau.

Mincer, Jacob. 1986. "Wage Changes in Job Changes." In *Research in Labor Economics*. Vol. 8, , ed. Ronald G. Ehrenberg, 171–97. Greenwich, CT: JAI Press.

Mincer, Jacob. 1991. "Education and Unemployment." NBER Working Paper No. 3838.

Mincer, Jacob, and Boyan Jovanovic. 1979. "Labor Mobility and Wages." NBER Working Paper No. 357.

Molho, Ian. 1986. "Theories of Migration: A Review." *Scottish Journal of Political Economy*, 33(4): 396–419.

Molho, Ian. 2001. “Spatial Search, Migration and Regional Unemployment.” *Economica*, 68(270): 269–283.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. “Internal Migration in the United States.” *Journal of Economic Perspectives*, 25(3): 173–96.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2017. “Job Changing and the Decline in Long-Distance Migration in the United States.” *Demography*, 54(2): 631–653.

Monras, Joan. 2015. “Economic Shocks and Internal Migration.” IZA Discussion Paper No. 8840.

Moretti, Enrico. 2011. “Local Labor Markets.” In *Handbook of Labor Economics*. Vol. 4B, , ed. David Card and Orley Ashenfelter, 1237–1313. New York: Elsevier.

Nimczik, Jan S. 2018. “Job Mobility Networks and Endogenous Labor Markets.” https://sites.google.com/site/jannimczik.

Nocke, Volker, and Stephen Yeaple. 2008. “An Assignment Theory of Foreign Direct Investment.” *The Review of Economic Studies*, 75(2): 529–557.

Notowidigdo, Matthew J. 2011. “The Incidence of Local Labor Demand Shocks.” NBER Working Paper No. 17167.

Roback, Jennifer. 1982. “Wages, Rents, and the Quality of Life.” *Journal of Political Economy*, 90(6): 1257–1278.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. “Integrated Public Use Microdata Series: Version 7.0 [dataset].” Minneapolis: University of Minnesota. Minneapolis: University of Minnesota.

Sattinger, Michael. 1975. “Comparative Advantage and the Distributions of Earnings and Abilities.” *Econometrica*, 43(3): 455–468.

Schmutz, Benoît, and Modibo Sidibé. forthcoming. “Frictional Labor Mobility.” *Review of Economic Studies*.

Schwartz, Aba. 1971. “On Efficiency of Migration.” *Journal of Human Resources*, 193–205.

Schwartz, Aba. 1973. “Interpreting the Effect of Distance on Migration.” *Journal of Political Economy*, 81(5): 1153–1169.

Schwartz, Aba. 1976. “Migration, Age, and Education.” *Journal of Political Economy*, 84(4): 701–719.

Shimer, Robert. 2005. “The Cyclicality of Hires, Separations, and Job-to-Job Transitions.” Federal Reserve Bank of St Louis Review.

Shryock, H.S. Jr. 1959. “The Efficiency of Internal Migration in the United States.” Proceedings, International Population Conference, Vienna.

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

Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1): 145–176.

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

Topel, Robert H., and Michael P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 439–479.

Van Ommeren, Jos, Piet Rietveld, and Peter Nijkamp. 1997. "Commuting: In Search of Jobs and Residences." *Journal of Urban Economics*, 42(3): 402–421.

Wheeler, Christopher H. 2001. "Search, Sorting, and Urban Agglomeration." *Journal of Labor Economics*, 19(4): 879–899.

Wildasin, David E. 2000. "Labor-Market Integration, Investment in Risky Human Capital, and Fiscal Competition." *American Economic Review*, 90(1): 73–95.

Wolpin, Kenneth I. 1992. "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth." *Journal of Political Economy*, 100(3): 535–560.

Wozniak, Abigail. 2010. "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*, 45(4): 944–970.

Appendix Contents

A CPS sample description and supplementary estimates

A.1 Sample description

My Current Population Survey (CPS) estimates in the main text are based on March waves between 1999 and 2018, taken from the IPUMS database (Flood et al., 2018). The March waves include the Annual Social and Economic Supplement, which reports whether respondents lived in a different state 12 months previously. Since 1999, individuals have also given their primary reason for moving.

I consider five education groups: high school dropouts (less than 12 years of schooling), high school graduates (12 years), some college (less than an undergraduate degree), and undergraduate and postgraduate degree holders. Potential labor market experience is defined as age minus years of education minus 6 (or age minus 16, whichever is smaller). I set years of schooling to 13 for individuals with some college but no degree, 14 for associate degrees, 16 for undergraduate, 18 for Master's, 19 for professional and 21 for doctorate degrees. I restrict the sample to individuals with 2-30 years of experience at the survey date: this excludes people with less than one year of experience at the time of moving. I also restrict attention to individuals living in the US one year previously.

Kaplan and Schulhofer-Wohl (2012) show there are inconsistencies in the CPS's procedure for imputing migration status in non-response cases: the imputed data artificially inflate the cross-state migration rate between 1999 and 2005. As it happens, the non-response rate for migration status varies little with education: 13 percent for college graduates and 14 percent for non-graduates. I choose to drop all these observations.

A.2 Historical changes in mobility differentials

The CPS analysis in the main text is restricted to the period 1999-2018, for which I have information on reasons for moving. But the mobility gap between education groups goes back many decades. In Figure A1, I plot annual cross-state migration rates using CPS March waves from 1964 to 2018.[27]

As is well known, migration rates have declined over this period: see e.g. Molloy, Smith and Wozniak (2011). Kaplan and Schulhofer-Wohl (2017) argue this was driven by the declining geographical specificity of occupational returns, coupled with improvements in communications technology. Molloy, Smith and Wozniak (2017) attribute it to a declining rate of labor market transitions. Either way, the decline was fairly uniform across education groups. The ratio of graduate to non-graduate mobility has mostly hovered around 1.4, and there is no clear upward or downward trend over the period as a whole.

A.3 Breakdown of migration by reported reasons for moving

In Table A1, I present detailed disaggregations of cross-state and cross-county migration in the CPS by reported reason for moving. The first column gives the percentage of the full sample who changed state (in the previous 12 months) for each recorded reason, and the second column expresses these numbers as a percentage of cross-state migrants. The final two columns repeat this exercise for cross-county moves: these consist of both moves across states and across counties within states.

[27] I omit 1995 because the relevant migration question was not asked that year.

The bottom row shows that, each year, 2.4 percent of the sample move across states and 5.4 percent across counties. About half of cross-state moves are motivated by a specific job, compared with a third of cross-county moves. These are mostly due to a job change or transfer, but some workers also report commuting reasons. The commuting motivation can be interpreted in the context of a long-distance match: after accepting a distant job offer (with a long associated commute), the worker eventually changes residence. In contrast, it is rare to move to *look* for work without a job lined up. This sort of speculative job search accounts for just 5 percent of cross-state and 4 percent of within-state moves. This is unsurprising: moving without a job in hand is a costly and risky strategy. In terms of non-job migration, family and housing motivations account for most moves.

In Table A2, I report the cross-state and cross-county migration rates (in columns 1 and 3 of Table A1) separately by education group: high school dropouts (HSD), high school graduates (HSG), some college (SC), undergraduate degree (UG) and postgraduate (PG). As before, the first row reports the rate of job-motivated migration. Notice the (positive) education slope is steeper in proportional terms for cross-state than cross-county moves. I offer a rationale for this result in Section 3.6: to the extent that cross-state migration is more costly, education differences in offer dispersion and job surplus should matter more. Also, consistent with Figure 2 in the main text, better educated individuals make fewer speculative moves to "look for work".

On aggregate, there is also a mild negative education gradient in non-job migration, which is stronger for cross-county moves. This effect is driven by a broad range of motivations: mostly to "establish own household", "other family reasons", "cheaper housing", "other housing reasons" and "better neighborhood". There are just two non-job motivations with (largely) positive education slopes: the desire to purchase a home and attending or leaving college.

The final row reports total migration rates - for all motivations combined. Notice this is much flatter for cross-county migration. Mechanically, this reflects the flatter (positive) gradient of job-motivated migration, the steeper (negative) gradient of non-job migration, and the greater dominance of non-job motivations for cross-county movers.

A.4 Robustness to top earner restriction

Importantly, the CPS question on reasons for moving is addressed to *individuals* within households. But of course, migration decisions are made in the context of the household. This ambiguity may yield some problems for interpretation: for example, household dependents may choose to simply report the reasons of the breadwinners. This is most clearly illustrated for children (though they are excluded from my sample): in households with at least one adult moving for a specific job, 80 percent of under-16s also report moving for the same reason.

To address this concern, I recompute migration rates by reason for moving (and by education), but this time restricting the sample to those individuals with the greatest annual earnings in each household. In households with joint top-earners, I divide the person weights by the number of top-earners. This restriction excludes 44 percent of the original sample. But as Table A3 shows, it makes little difference to the education slopes of job-specific, speculative or non-job migration. The first, third and fifth columns replicate the cross-state migration rates from Table A2, and these look very similar to the remaining columns which impose the top-earner restriction.

B Contribution of returning students

In this section, I check whether returning students may be contributing to education differentials in mobility. Table A2 shows that workers who report moving primarily to leave or attend college account for a negligible part of these differentials. But even if this is not the primary stated motivation, it may be an underlying factor for those who report job-related reasons - at least for the young. Indeed, Kennan and Walker (2011) emphasize that a large fraction of long-distance movers in the US are returning to former places of residence; and Kennan (2015) studies the tendency of individuals to return home after studying in another state.

The contribution of this return migration can be assessed in the Panel Study of Income Dynamics (PSID). Similarly to the CPS analysis, I define a migrant as somebody living in a different state 12 months previously. I restrict attention to individuals with 2-10 years of potential labor market experience at the end of each 12-month interval, in annual PSID waves between 1990 and 1997. I exclude waves after 1997 because these are biennial: it is not possible to track migration at annual frequencies. The first row of Table A4 reports the fraction of individuals in each education group who were recently students (either in the current or previous annual wave). This is largest for high school dropouts (18 percent) and smallest for individuals with undergraduate and postgraduate degrees (5 and 2 percent respectively).

The remaining rows report annual cross-state migration rates by education. The second row computes these for the full sample, illustrating the familiar positive education gradient. I exclude recent students in the third row; and unsurprisingly, this makes little difference - given they comprise such a small fraction of the college graduate sample. Interestingly, the graduate migration rates are actually a little larger once recent students are excluded.

However, excluding recent students does not address the concerns entirely, because ex-students may yet return to their home state several years after completing their education. In the final two rows, I disaggregate the cross-state migration rate into "return" and "non-return" moves. Return moves consist of moves to any state where the individual has previously resided since the panel began (i.e. since 1968). The education gradient is clearly positive for both return and non-return rates, and the gradient is about twice as steep for the latter. This can be rationalized by the model: if non-return moves are more costly, the rate of non-return migration should be more elastic to job surplus (and hence to education). This is an equivalent idea to that of migration distance in Section 3.6. To summarize then, the evidence shows that returning students (and return migration in general) cannot account for the mobility gap.

C Predictive power of imputed costs

Given that the imputed costs in Section 2.2 are based on the subjective judgments of respondents, there may be doubts over their accuracy. But reassuringly, the cost measures do have significant predictive power for future migration decisions. Suppose the instantaneous cross-area matching rate for some individual i is constant within the time interval $t-1$ to t, and denote this matching rate as ρ_{Cit}. The probability of moving within this interval is then:

$$\Pr\left(Move_{it}=1\right)=1-\exp\left(-\rho_{Cit}\right) \tag{A1}$$

This motivates a complementary log-log model:

$$\Pr\left(Move_{it}=1\right)=1-\exp\left(-\exp\left(\beta_m\tilde{m}_{it-1}+\beta_X'X_{it}+\beta_t\right)\right) \tag{A2}$$

where I express ρ_{Cit} as a function of the initial annuitized moving cost $\tilde{m}_{it-1}$ (as defined in (20)), human capital indicators[28] X_{it}, and a full set of year effects β_t. The advantage of this specification is that β_m can intuitively be interpreted as the elasticity of the instantaneous migration rate with respect to $\tilde{m}_{it-1}$. And assuming a constant hazard, this interpretation is independent of the time horizon associated with the migration variable.

[28]Specifically: experience and experience squared; four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in experience; black and Hispanic dummies; and a gender dummy interacted with all previously mentioned variables.

I report my estimates in Table A5. I only have imputed costs for employed household heads between 1969 and 1973, so I restrict attention to the mobility decisions of these individuals in the annual intervals between 1969 and 1974. The first two columns report the elasticity of cross-state migration (within each 12-month interval) to a binary indicator for "willingness to move" (at the beginning of each interval). Willingness to move adds about 100 log points to the cross-state migration rate, and an interaction with a college graduate dummy reveals no significant difference in the response by education.

In the final three columns, I restrict the sample to those who are "willing to move" and estimate elasticities to the imputed costs (again, at the beginning of each interval). The imputed cost is the difference between (i) the log of a worker's stated reservation for accepting a long-distance offer and (ii) the log of the worker's actual wage: see Section 2.2 in the main text for further details. Column 3 reports an elasticity of -0.2, but it is statistically insignificant. This estimate will presumably be attenuated by classical measurement error, but there is also a more systematic problem. If the long-distance reservation is noisier than a worker's wage, the imputed cost (the difference between the two) will be artificially negatively correlated with wages. But as the model shows, to the extent that wages reflect match quality, workers with higher wages will be less likely to move. This should bias the estimated effect of imputed costs towards zero. To address this problem, in column 4, I control additionally for the initial log wage. Both the imputed cost and the wage now take strong negative effects, with elasticities of -1.3 and standard errors of 0.4 in each case. In column 5, I allow for education heterogeneity in these effects, but the interactions are not statistically significant.

The key message here is that the subjective costs *do* have predictive power for future mobility - which suggests they are informative about the true costs of moving. This reinforces the validity of the claim in Section 2.2 that moving costs vary little with education.

D Net migratory flows by potential experience

In Table A6, I reproduce the results in Table 1 in the main text - but this time separately for individuals with 2-10 and 11-30 years of potential labor market experience. At least for young college-educated individuals, there is a discernible positive effect of education on net migration rates. But this effect is relatively small: for all experience groups and occupation schemes, the ratio of net to gross migration is still decreasing in education. I conclude from this that the mobility differentials between education groups are not driven by large net flows to particular states, even within detailed occupation-defined markets and within distinct experience categories. This reinforces the general message in Section 2.3 in the main text.

E Tenure as a proxy for job match quality

In this paper, I have used initial job tenure as a proxy for job match quality ε. The quality of this proxy is increasing in tenure: intuitively, for workers who have just begun their job, tenure offers little information about match quality. In this appendix, I show this more formally.

For large moving costs, the overall matching rate $\rho(\varepsilon)$ converges to the local matching rate $\rho_L(\varepsilon) = \lambda\left[1 - F^{\varepsilon}(\varepsilon)\right]$, as defined in (5). Conditional on the offer arrival rate λ (which I take as given throughout), match quality ε is fully identified by the corresponding job matching rate ρ. So in this analysis, it is sufficient to study the informativeness of tenure as a proxy for the ε-specific matching rate ρ.

The overall separation rate from a job is equal to $\delta + \rho$, where δ is the exogenous transition rate to unemployment, and ρ is the quit rate to better jobs. Given this is a constant hazard, the probability that a worker is still in a given job match at tenure τ (i.e. the survivor function) is equal to $\exp(-\tau(\delta+\rho))$. Now, let $G^{\rho}(\rho)$ be the unconditional distribution of matching rates ρ across workers (which I take as given), where ρ is bounded below by zero. And let $G^{\rho}(\rho|\tau)$ be the distribution of matching rates ρ for workers with tenure τ. By Bayes' theorem, the density of the latter is:

$$g^{\rho}(\rho|\tau) = \frac{\exp(-\tau(\delta+\rho))\, g^{\rho}(\rho)}{\int \exp(-\tau(\delta+x))\, g^{\rho}(x)\, dx} = \frac{\exp(-\tau\rho)\, g^{\rho}(\rho)}{\int \exp(-\tau x)\, g^{\rho}(x)\, dx} \tag{A3}$$

As $\tau \to 0$, the conditional density $g^{\rho}(\rho|\tau)$ collapses to the unconditional density $g^{\rho}(\rho)$; so tenure τ offers no information on ρ (or equivalently, on match quality). But as τ increases, the conditional distribution of ρ becomes less dispersed. And in the limit, as $\tau \to \infty$, the conditional distribution collapses to a unit mass at $\rho = 0$, which corresponds to the maximum match quality.

To summarize, tenure can indeed serve as a proxy for match quality (conditional on the offer rate λ), and the precision of the proxy is increasing in the level of tenure.

F Effect of tenure and experience on cross-state matching

Geographical mobility is strongly decreasing in experience. One intuition for this effect, arising from my model, is that older workers enjoy larger match quality - and hence are tempted by fewer long-distance job offers. In Section 4.3, I propose using job tenure as a proxy for match quality, ε. And indeed, I show there that the rate of cross-state matching is decreasing in initial job tenure.

In this appendix, I ask the following question: statistically, to what extent can initial tenure (as an imperfect proxy for match quality) account for the negative effect of experience on cross-state matching? Applying the complementary log-log model in (A2), the probability of forming a cross-state match within the time interval $t-1$ to t can be written as:

$$\Pr\left(Move_{it}=1\right)=1-\exp\left(-\exp\left(\beta' X_{it}\right)\right) \tag{A4}$$

where X_{it} contain a set of education effects, as well as experience or initial log tenure. The β coefficients can be interpreted as elasticities of the cross-state matching rate.

I present the complementary log-log estimates in Table A7. Column 1 reports the basic education effects, which are increasing monotonically. A postgraduate degree adds 170 log points to the cross-state matching rate, relative to high school dropouts (the omitted category). In column 2, I control additionally for a quadratic in labor market experience. One extra year of experience reduces the cross-state matching rate by 11 log points; the squared term is statistically insignificant. In column 3, I replace the experience controls with initial log tenure: the elasticity of the matching rate is -0.4. Notice also that the education effects become somewhat larger in column 3, which is consistent with better educated workers enjoying higher match quality ε: see the discussion in Section 4.3.

In column 4, I control for experience and initial log tenure simultaneously. Comparing this with column 2, initial tenure accounts statistically for about one third of the effect of experience on cross-state matching. To the extent that initial tenure is an imprecise proxy for match quality (see Appendix E), this presumably understates the true contribution of match quality ε to the experience effect. But of course, experience itself may contain independent information on moving costs. For example, in Figure 3 above, older workers report less willingness to move for work. A human capital explanation is that older workers have fewer years to benefit from the sunk cost of moving (see e.g. Kennan and Walker, 2011).

G Distribution of match quality and mean matching rates

In this section, following Burdett and Mortensen (1998), I derive the equilibrium distribution of match quality ε. And I then consider the implications for the mean local and cross-area matching rates. For any match quality ε, let:

$$\rho\left(\varepsilon\right)=\rho_L\left(\varepsilon\right)+\rho_C\left(\varepsilon\right) \tag{A5}$$

be the total job matching rate, i.e. the sum of the local and cross-area rates. This is equal to $\rho(\varepsilon_R)$ for the unemployed, so the steady-state unemployment rate is:

$$u = \frac{\delta}{\delta + \rho(\varepsilon_R)} \tag{A6}$$

Now, consider the set of employed workers with match quality below ε. The inflow of workers to this set must equal the outflow in equilibrium:

$$u\left[\rho(\varepsilon_R) - \rho(\varepsilon)\right] = (1-u)\,G(\varepsilon)\left[\delta + \rho(\varepsilon)\right] \tag{A7}$$

where $G(\varepsilon)$ is the distribution of ε among employed workers. The inflow is composed entirely of the unemployed, who enter jobs with match quality below ε at rate $[\rho(\varepsilon_R) - \rho(\varepsilon)]$. The outflow is composed of employed workers with match quality below ε who (i) are separated to unemployment (at rate δ) or (ii) find jobs yielding utility exceeding ε. Substituting (A6) for u gives:

$$G(\varepsilon) = \frac{\delta}{\delta + \rho(\varepsilon)} \cdot \frac{\rho(\varepsilon_R) - \rho(\varepsilon)}{\rho(\varepsilon_R)} \tag{A8}$$

This equation demonstrates the importance of market frictions for my hypothesis. For all ε, $G(\varepsilon)$ converges to zero as the offer rate λ (and therefore the matching rate $\rho(\varepsilon)$) becomes large relative to the separation rate δ. So, in a frictionless world, all workers will benefit from the maximum match quality - and there will be no job surplus to justify geographical mobility.

The distribution $G(\varepsilon)$ accounts for employed workers only. To extend this to the unemployed, notice that they behave identically to workers with match quality ε_R. In this vein, I can define a distribution function $\hat{G}(\varepsilon)$: the fraction of all workers (irrespective of employment status) who receive *effective* match quality below ε. Specifically:

$$\hat{G}(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < \varepsilon_R \\ u + (1-u)\,G(\varepsilon) = \frac{\delta}{\delta + \rho(\varepsilon)} & \text{if } \varepsilon \geq \varepsilon_R \end{cases} \tag{A9}$$

with probability density:

$$\hat{g}(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < \varepsilon_R \\ \frac{\delta}{\delta + \rho(\varepsilon_R)} & \text{if } \varepsilon = \varepsilon_R \\ -\frac{\delta \rho'(\varepsilon)}{[\delta + \rho(\varepsilon)]^2} & \text{if } \varepsilon > \varepsilon_R \end{cases} \tag{A10}$$

where the unemployed are treated as receiving ε_R. This is effectively a left-censored distribution, with a discrete probability mass (i.e. the unemployed) at the censored value of ε_R.

In an elaboration of the discussion in Section 3.6, I now consider the determinants of the mean job matching rate across all workers, both employed and unemployed. The mean local and cross-area matching rates can be expressed as:

$$\mathbb{E}\left[\rho_X(\varepsilon)\right] = \int_{\varepsilon_R}^{\infty} \rho_X(\varepsilon)\, d\hat{G}(\varepsilon) \tag{A11}$$

for $X = \{L, C\}$. Consider first the mean local rate $\mathbb{E}\left[\rho_L(\varepsilon)\right]$. Based on (5), (A9) and (A11), it can be fully summarized by the offer rate λ, the separation rate δ, and the the reservation match quality ε_R. Evidently, $\mathbb{E}\left[\rho_L(\varepsilon)\right]$ is increasing in λ. It is also increasing in δ: this is because a larger δ raises $\hat{G}(\varepsilon)$ for all ε, and $\rho_L(\varepsilon)$ is decreasing in ε. Intuitively, workers have less time to rise up the ladder before they fall to the bottom (through a separation), so more offers will be acceptable to them on average. Finally, $\mathbb{E}\left[\rho_L(\varepsilon)\right]$ is decreasing in ε_R. As I note above, ε_R can be interpreted as the censoring value of a left-censored distribution. Therefore, a larger reservation ε_R causes $\hat{G}(\varepsilon)$ to decline for given ε in the neighborhood of ε_R. Intuitively, if workers are more demanding, they will be located at higher ε in equilibrium; and given $\rho_L(\varepsilon)$ is decreasing in ε, fewer offers will be acceptable to them on average. Based on Proposition 1, the mean cross-area matching rate $\mathbb{E}\left[\rho_C(\varepsilon)\right]$ additionally depends (positively) on the offer dispersion σ^{ε} and (negatively) on the size of moving costs σ^{μ}.

H Theoretical proofs and derivations

H.1 Proof of Proposition 2 in Section 3.4

Proposition 2 states that, for sufficiently large σ^{μ}, the (positive) response of cross-area job matching to $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}$ is decreasing in initial match quality ε. In Section 3.4, I offer the following expression for $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$:

$$\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}} = \tilde{\pi}\lambda \int_{\varepsilon}^{\infty} \left\{ \left[\frac{d\log\Omega(\varepsilon'-\varepsilon|\varepsilon)}{d\log\sigma^{\varepsilon}} + 1\right] \frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\, f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\right)\right\} dF^{\varepsilon}(\varepsilon') \tag{A12}$$

I begin by considering the $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\, f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\right)$ term, and I return to the term in square brackets later. $\Omega(\varepsilon'-\varepsilon|\varepsilon)$ is unambiguously decreasing in ε. So, to ensure that $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\, f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\right)$ is decreasing in ε, a sufficient condition is that $\mu f^{\mu}(\mu)$ is increasing in μ; or equivalently, that the elasticity of the density $\mu\frac{f^{\mu\prime}(\mu)}{f^{\mu}(\mu)} > -1$. In Proposition 2 though, I rely on an alternative sufficient condition: that σ^{μ} is sufficiently large. Note I have assumed throughout that the elasticity of the density f^{μ} is monotonically decreasing. Therefore, as $\mu \to 0$ from above, the elasticity of f^{μ} grows and will eventually exceed -1. It follows that, for sufficiently large σ^{μ}, $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\, f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega(\varepsilon'-\varepsilon|\varepsilon)\right)$ must be decreasing in ε.

It remains to consider the implications of the $\frac{d\log\Omega(\varepsilon'-\varepsilon|\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ term. Notice that:

$$\frac{d\log\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}=-\int_{\varepsilon}^{\varepsilon'}\frac{\frac{d\rho_C(x)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}}{\left[r+\delta+\rho_L\left(x\right)+\rho_C\left(x\right)\right]^2}dx\left[\int_{\varepsilon}^{\varepsilon'}\frac{1}{r+\delta+\rho_L\left(x\right)+\rho_C\left(x\right)}dx\right]^{-1} \tag{A13}$$

This is a weighted average of $\frac{d\rho_C(x)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ across different values of initial match quality x. With this in mind, the proposition can be demonstrated by contradiction. I begin by considering the response at the top of the support of match quality ε, and I then move down the distribution. Suppose the response $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ is non-decreasing in ε at the top of the support of ε. Based on (A12) and assuming that σ^{μ} is "sufficiently large", such that $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)\right)$ is decreasing in ε, it then follows that $\frac{d\log\Omega(\varepsilon'-\varepsilon|\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ must be increasing in ε at the top of the support. And using (A13), this in turn implies that $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ must be decreasing in ε (notice the negative sign in (A13)) at the top of the support of ε. But this is a contradiction. One can then make a similar argument for every other ε, moving sequentially down the distribution of F^{ε} - which implies that $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ is decreasing in ε over the full support of ε.

H.2 Proof of Proposition 3 in Section 3.4

Proposition 3 states that, for sufficiently large σ^{μ}, the responses of migratory outflows and inflows to changes in the local offer rate λ_j are increasing (in magnitude) in offer dispersion σ^{ε} - around a steady-state where local areas are identical, i.e. $\lambda_j=\lambda$ and $V_j=V$ for all j.

Consider first the response of outflows. Using (12), this can be expressed as:

$$\frac{d\rho_{Cj}^{Outflow}\left(\varepsilon\right)}{d\lambda_j}=-\tilde{\pi}\lambda\left(\frac{1}{\sigma^{\varepsilon}}\frac{dV_j\left(\varepsilon\right)}{d\lambda_j}\right)\int_{\varepsilon}^{\infty}\left[\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)\right)\right]dF^{\varepsilon}\left(\varepsilon'\right) \tag{A14}$$

Based on the same argument from the previous section (and due to the decreasing elasticity of the density f^{μ}), the impact of σ^{ε} on $\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)\right)$ must be positive for sufficiently large σ^{μ}. To demonstrate that $\frac{d\rho_{Cj}^{Outflow}(\varepsilon)}{d\lambda_j}$ is increasing (in magnitude) in σ^{ε}, it is therefore sufficient to show that $\frac{dV_j(\varepsilon)}{d\lambda_j}$ is increasing (at least) proportionally with σ^{ε}.

To see this, notice that local worker value (3) can be written as:

$$\begin{aligned} rV_j\left(\varepsilon\right) &= \gamma' X+\sigma^{\varepsilon}\varepsilon+\delta\left[V_j\left(\varepsilon_R\right)-V_j\left(\varepsilon\right)\right]+\lambda_j\int_{\varepsilon}^{\infty}\left[V_j\left(\varepsilon'\right)-V_j\left(\varepsilon\right)\right]dF^{\varepsilon}\left(\varepsilon'\right) \\ &\quad +\tilde{\pi}\lambda\int_0^{\infty}\left[\int_{-\infty}^{\infty}\max\left\{V\left(\varepsilon'\right)-V_j\left(\varepsilon\right)-\sigma^{\mu}\mu,0\right\}dF^{\varepsilon}\left(\varepsilon'\right)\right]dF^{\mu}\left(\mu\right) \end{aligned} \tag{A15}$$

This expression equates values and offer rates in all areas $k \neq j$ in equation (3). Differentiating with respect to the local offer rate λ_j, around a steady-state with $\lambda_j = \lambda$:

$$\begin{aligned} r\frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda} &= \delta\left[\frac{dV_j(\varepsilon_R)}{d\lambda_j}|_{\lambda_j=\lambda} - \frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda}\right] + \sigma^{\varepsilon}\int_{\varepsilon}^{\infty}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)dF^{\varepsilon}\left(\varepsilon'\right) \\ &+\rho_L(\varepsilon)\int_{\varepsilon}^{\infty}\frac{dV_j(\varepsilon')}{d\lambda_j}|_{\lambda_j=\lambda}dF^{\varepsilon}\left(\varepsilon'\right) - \rho_L(\varepsilon)\frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda} - \rho_C(\varepsilon)\frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda} \end{aligned} \tag{A16}$$

where the function Ω is defined by (8). Rearranging this expression:

$$\begin{aligned} \frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda} &= \frac{\delta}{r+\delta+\rho(\varepsilon)}\frac{dV_j(\varepsilon_R)}{d\lambda_j}|_{\lambda_j=\lambda} + \frac{\rho_L(\varepsilon)}{r+\delta+\rho(\varepsilon)}\int_{\varepsilon}^{\infty}\frac{dV_j(\varepsilon')}{d\lambda_j}|_{\lambda_j=\lambda}dF^{\varepsilon}\left(\varepsilon'\right) \\ &+\frac{\sigma^{\varepsilon}}{r+\delta+\rho(\varepsilon)}\int_{\varepsilon}^{\infty}\Omega\left(\varepsilon'-\varepsilon|\varepsilon\right)dF^{\varepsilon}\left(\varepsilon'\right) \end{aligned} \tag{A17}$$

where $\rho(\varepsilon) \equiv \rho_L(\varepsilon) + \rho_C(\varepsilon)$. Both $\Omega(\varepsilon' - \varepsilon|\varepsilon)$ and $\rho(\varepsilon)$ are functions of the cross-state matching rate $\rho_C(\varepsilon)$, so they are both in principle sensitive to offer dispersion σ^{ε}. But as σ^{μ} becomes large, this sensitivity goes to zero.[29] By inspection of (A17), $\frac{dV_j(\varepsilon)}{d\lambda_j}|_{\lambda_j=\lambda}$ must then be increasing proportionally with σ^{ε}.

Given this result for $\frac{dV_j(\varepsilon)}{d\lambda_j}$, a parallel argument can be made for the response of migratory inflows in equation (14). And as a result, the inflow response must also be increasing in σ^{ε}: see the discussion in Section 3.5.

H.3 Derivation of equation (23) in Section 5.1

The aim here is to derive the approximation for the odds ratio of cross-area to local job matching, $\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)}$, in equation (23). Beginning with (22):

$$\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} = \tilde{\pi}\int_0^1\left[\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}(\mu|\varepsilon)}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}(\varepsilon)}\right]d\mu = \tilde{\pi}\int_0^{\infty}\left[\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}(\mu|\varepsilon)}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}(\varepsilon)}\right]d\mu \tag{A18}$$

where the second equality replaces the integral's upper bound with ∞, based on the assumption that no wage offers are sufficient for a worker to accept a cross-area offer with the top moving

[29] Looking at (8), the job surplus $\Omega(\varepsilon' - \varepsilon|\varepsilon)$ is discounted by the overall rate of job matching $\rho(\varepsilon)$. But intuitively, as σ^{μ} becomes large, the contribution of cross-area matching $\rho_C(\varepsilon)$ to the overall matching rate $\rho(\varepsilon)$ becomes small. More formally, this effect can be appreciated from the argument in the previous section. According to Proposition 2, $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ must be decreasing in ε for sufficiently large σ^{μ}. Looking at (A12), it must therefore also be true that $\frac{d\rho_C(\varepsilon)}{d\log\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}}$ decreases to zero as σ^{μ} becomes large.

cost draw $\mu = 1$. This expression can then be approximated as:

$$\begin{aligned}
\frac{\rho_C(\varepsilon)}{\rho_L(\varepsilon)} &\approx \tilde{\pi}\int_0^{\infty}\left[\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{\sigma^{\mu}\mu}{\sigma^{\varepsilon}}\left[r+\delta+\rho(\varepsilon)\right]\right)}{1-F^{\varepsilon}(\varepsilon)}\right]d\mu \\
&= \frac{\tilde{\pi}\sigma^{\varepsilon}}{\sigma^{\mu}\left[r+\delta+\rho(\varepsilon)\right]}\int_0^{\infty}\left[\frac{1-F^{\varepsilon}(\varepsilon+x)}{1-F^{\varepsilon}(\varepsilon)}\right]dx \\
&= \frac{\tilde{\pi}\sigma^{\varepsilon}}{\sigma^{\mu}\left[r+\delta+\rho(\varepsilon)\right]}\int_0^{\infty}\left[\frac{xf^{\varepsilon}(\varepsilon+x)}{1-F^{\varepsilon}(\varepsilon)}\right]dx \\
&\approx \frac{\tilde{\pi}\sigma^{\varepsilon}\mathbb{E}_L\left[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon\right]}{2\mathbb{E}\left[\tilde{m}(\mu|\varepsilon)\right]}
\end{aligned} \tag{A19}$$

In the first line of (A19), I have replaced the annuitized moving cost $\tilde{m}(\mu|\varepsilon)$ with the linear approximation in (21), where $\rho(\varepsilon) \equiv \rho_L(\varepsilon)+\rho_C(\varepsilon)$. I next define $x \equiv \frac{\sigma^{\mu}\mu}{\sigma^{\varepsilon}}\left[r+\delta+\rho(\varepsilon)\right]$, and I use it to replace μ in the second line. The third line is derived through integration by parts. Notice the expression $\sigma^{\varepsilon}\int_0^{\infty}\left[\frac{xf^{\varepsilon}(\varepsilon+x)}{1-F^{\varepsilon}(\varepsilon)}\right]dx$ is equal to the expectation of the local wage returns, $\sigma^{\varepsilon}\mathbb{E}_L\left[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon\right]$. Finally, using (21), the expectation of annuitized moving costs can be approximated as:

$$\mathbb{E}\left[\tilde{m}(\mu|\varepsilon)\right] \approx \int_0^1 \sigma^{\mu}\mu\left[r+\delta+\rho(\varepsilon)\right]d\mu = \frac{1}{2}\sigma^{\mu}\left[r+\delta+\rho(\varepsilon)\right] \tag{A20}$$

assuming, as before, that μ is uniformly distributed over a $[0,1]$ support. Substituting (A20) for $\sigma^{\mu}\left[r+\delta+\rho(\varepsilon)\right]$ in the third line of (A19) then yields the final line.

H.4 Proof of Proposition 4 in Section 6.1

Proposition 4 states that the expected annuitized cost is increasing in both σ^{ε} and σ^{μ}, for given initial match quality ε and for sufficiently large σ^{μ}. It is sufficient to show that increases in σ^{ε} and σ^{μ} cause hazard rate dominating transformations of the realized annuitized cost distribution, $Z(\tilde{m}|\varepsilon)$. In this section, the distribution function $Z(\cdot|\varepsilon)$ and job surplus $\Omega(\cdot|\varepsilon)$ are both defined for given match quality ε, but I omit the "$|\varepsilon$" from here on to ease notation.

Importantly, the Ω relationship depends on the level of σ^{ε} and σ^{μ}. This is because σ^{ε} and σ^{μ} affect the cross-area matching rate $\rho_C(\varepsilon)$, and this rate matters for the discounting of job surplus in (8). However, for a sufficiently large σ^{μ}, cross-area matching contributes relatively little to job matching overall, so changes in σ^{μ} and σ^{ε} have little effect on the Ω relationship. See footnote 29 in Appendix H.2. From here on, I therefore neglect the impact of σ^{ε} and σ^{μ} on the Ω relationship.

Using (25), the hazard rate of Z can be written as:

$$\frac{z(\tilde{m})}{1-Z(\tilde{m})}=\left[\int_{m}^{\infty}\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}\cdot\frac{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}dx\right]^{-1} \tag{A21}$$

The $\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}$ term is increasing in σ^{ε}, conditional on ε, for all $x>\tilde{m}$. To see this, consider the derivative of its log with respect to σ^{ε}

$$\frac{d}{d\sigma^{\varepsilon}}\left\{\log\frac{1-F^{\varepsilon}\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}\right\}=\frac{1}{(\sigma^{\varepsilon})^{2}}\left(x\frac{f^{\varepsilon}\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}-\tilde{m}\frac{f^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{1-F^{\varepsilon}\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}\right)>0 \tag{A22}$$

which exceeds zero due to my assumption that F^{ε} has a monotonically increasing hazard rate.

I now show the second term in (A21), $\frac{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}$, is increasing in both σ^{μ} and σ^{ε}. Consider first the impact of σ^{μ}:

$$\frac{d}{d\sigma^{\mu}}\left\{\log\frac{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}\right\}=\frac{1}{\sigma^{\mu}}\left[\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\frac{f^{\mu\prime}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}-\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\frac{f^{\mu\prime}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}\right]>0 \tag{A23}$$

which exceeds zero for $x>\tilde{m}$, due to my assumption that the density's elasticity, i.e. $\mu\frac{f^{\mu\prime}(\mu)}{f^{\mu}(\mu)}$, is monotonically decreasing in μ. Now, consider the impact of σ^{ε}:

$$\begin{aligned}\frac{d}{d\sigma^{\varepsilon}}\left\{\log\frac{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}\right\} &= \frac{\frac{d}{d\sigma^{\varepsilon}}\left\{-\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right\}}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\frac{f^{\mu\prime}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)} \\ &\quad -\frac{\frac{d}{d\sigma^{\varepsilon}}\left\{-\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right\}}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)}\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\frac{f^{\mu\prime}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}\end{aligned} \tag{A24}$$

Given my assumption on the elasticity's density, a sufficient condition for this expression to exceed zero is:

$$\frac{\frac{d}{d\sigma^{\varepsilon}}\left\{-\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right\}}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}>\frac{\frac{d}{d\sigma^{\varepsilon}}\left\{-\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right\}}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)} \tag{A25}$$

which, using the definition of Ω in (8), implies:

$$\begin{aligned}
\frac{\frac{\tilde{m}}{(\sigma^{\varepsilon})^2}\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega'\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)-\frac{1}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)} &> \frac{\frac{x}{(\sigma^{\varepsilon})^2}\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega'\left(\frac{x}{\sigma^{\varepsilon}}\right)-\frac{1}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)}{\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)} \qquad \text{(A26)}\\
\tilde{m}\frac{\Omega'\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)} &> x\frac{\Omega'\left(\frac{x}{\sigma^{\varepsilon}}\right)}{\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)}\\
\tilde{m}\frac{\left[r+\delta+\rho\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right]^{-1}}{\int_0^{\frac{\tilde{m}}{\sigma^{\varepsilon}}}\left[r+\delta+\rho\left(\varepsilon+s\right)\right]^{-1}ds} &> x\frac{\left[r+\delta+\rho\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)\right]^{-1}}{\int_0^{\frac{x}{\sigma^{\varepsilon}}}\left[r+\delta+\rho\left(\varepsilon+s\right)\right]^{-1}ds}\\
\left[\frac{\sigma^{\varepsilon}}{\tilde{m}}\int_0^{\frac{\tilde{m}}{\sigma^{\varepsilon}}}\frac{r+\delta+\rho\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{r+\delta+\rho\left(\varepsilon+s\right)}dq\right]^{-1} &> \left[\frac{\sigma^{\varepsilon}}{x}\int_0^{\frac{x}{\sigma^{\varepsilon}}}\frac{r+\delta+\rho\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{r+\delta+\rho\left(\varepsilon+s\right)}ds\right]^{-1}\\
\frac{\sigma^{\varepsilon}}{x}\int_0^{\frac{x}{\sigma^{\varepsilon}}}\frac{r+\delta+\rho\left(\varepsilon+\frac{x}{\sigma^{\varepsilon}}\right)}{r+\delta+\rho\left(\varepsilon+s\right)}ds &> \frac{\sigma^{\varepsilon}}{\tilde{m}}\int_0^{\frac{\tilde{m}}{\sigma^{\varepsilon}}}\frac{r+\delta+\rho\left(\varepsilon+\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)}{r+\delta+\rho\left(\varepsilon+s\right)}ds
\end{aligned}$$

where $\rho(\varepsilon) \equiv \rho_L(\varepsilon) + \rho_C(\varepsilon)$ is the total job matching rate. The final line must be true because of the monotonicity of $\rho(\varepsilon)$. This confirms that $\frac{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{x}{\sigma^{\varepsilon}}\right)\right)}{f^{\mu}\left(\frac{\sigma^{\varepsilon}}{\sigma^{\mu}}\Omega\left(\frac{\tilde{m}}{\sigma^{\varepsilon}}\right)\right)}$ in (A21) is indeed increasing in both σ^{μ} and σ^{ε}.

As a result, the hazard rate in (A21) must be decreasing in both σ^{μ} and σ^{ε} for given $\tilde{m}$ and ε. This means that increases in σ^{μ} and σ^{ε} cause hazard rate dominating transformations of the realized annuitized cost distribution, $Z(\tilde{m})$. And so, the expected annuitized costs must be increasing in both σ^{μ} and σ^{ε}.

H.5 Proof of Proposition 5 in Section 6.1

Finally, I prove that (27) can serve as a lower bound on the expected realized annuitized costs. Following the argument given in Section 6.1, it suffices to show that the expression in the curly brackets in (27) is less or equal to $\tilde{m}$, i.e.:

$$\mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)|\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)\geq\tilde{m}\right]-\mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)|\varepsilon'-\varepsilon\geq 0\right]\leq\tilde{m} \qquad \text{(A27)}$$

conditional on the initial match quality ε, where the operator $\mathbb{E}_L$ denotes the expected improvement in match quality ε arising from a local (subscript L) match. This can be rewritten as:

$$\mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)-\tilde{m}|\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)-\tilde{m}\geq 0\right]\leq\mathbb{E}_L\left[\sigma^{\varepsilon}\left(\varepsilon'-\varepsilon\right)|\varepsilon'-\varepsilon\geq 0\right] \qquad \text{(A28)}$$

Since I have assumed the annuitized cost $\tilde{m}$ always exceeds zero, it is sufficient to show that:

$$\frac{d}{dx}\log\mathbb{E}_L\left[\varepsilon'-x|\varepsilon'-x\geq 0\right]\leq 0 \qquad \text{(A29)}$$

for all $x \equiv \varepsilon + \frac{\tilde{m}}{\sigma^{\varepsilon}}$. Writing this in terms of the offer distribution F^{ε}:

$$\begin{aligned}
\frac{d}{dx}\log\mathbb{E}_L\left[\varepsilon' - x|\varepsilon' - x \geq 0\right] &= \frac{d}{dx}\log\frac{\int_x^{\infty}\varepsilon f^{\varepsilon}(\varepsilon)\,d\varepsilon}{1 - F^{\varepsilon}(x)} \\
&= \frac{d}{dx}\log\frac{\int_x^{\infty}\left[1 - F^{\varepsilon}(\varepsilon)\right]d\varepsilon}{1 - F^{\varepsilon}(x)} \\
&= -\frac{1 - F^{\varepsilon}(x)}{\int_x^{\infty}\left[\frac{1 - F^{\varepsilon}(\varepsilon)}{f^{\varepsilon}(\varepsilon)}\right]f^{\varepsilon}(\varepsilon)\,d\varepsilon} + \frac{f^{\varepsilon}(x)}{1 - F^{\varepsilon}(x)}
\end{aligned} \tag{A30}$$

where the second line follows from integration by parts. Now, I have assumed that F^{ε} has a monotonically increasing hazard rate; that is, $\frac{f^{\varepsilon}(\varepsilon)}{1-F^{\varepsilon}(\varepsilon)} \geq \frac{f^{\varepsilon}(x)}{1-F^{\varepsilon}(x)}$ for all $\varepsilon \geq x$. Therefore:

$$\frac{d}{dx}\log\mathbb{E}_L\left[\varepsilon' - x|\varepsilon' - x \geq 0\right] \leq -\frac{1 - F^{\varepsilon}(x)}{\int_x^{\infty}\left[\frac{1 - F^{\varepsilon}(x)}{f^{\varepsilon}(x)}\right]f^{\varepsilon}(\varepsilon)\,d\varepsilon} + \frac{f^{\varepsilon}(x)}{1 - F^{\varepsilon}(x)} = 0 \tag{A31}$$

so equation (A29) is satisfied.

Tables and figures

Table 1: Net cross-state migration rates by education

	Basic			Flows within 2-digit occ markets			Flows within 3-digit occ markets		
	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HS dropout	2.51	0.34	0.14	2.22	0.83	0.38	2.22	1.12	0.51
HS graduate	2.71	0.30	0.11	2.27	0.55	0.24	2.27	0.81	0.36
Some college	3.04	0.29	0.10	2.57	0.62	0.24	2.57	0.97	0.38
Undergraduate	3.71	0.29	0.08	3.26	0.68	0.21	3.26	1.05	0.32
Postgraduate	3.91	0.32	0.08	3.56	0.76	0.21	3.56	1.12	0.31

This table reports annual gross and net cross-state migration rates within education groups. The cross-state net migration rate is estimated as $\frac{1}{2n}\Sigma_j |n_j^{in} - n_j^{out}|$, where n is the total sample of individuals, n_j^{in} is the number of in-migrants to state j, and n_j^{out} is the number of out-migrants from state j. The first three columns report the overall migratory flows by education group, and the final six report flows within 2-digit and 3-digit occupation-defined labor markets. For each education group, these are constructed by weighting occupation-specific migration rates (gross and net) by occupational employment shares. Migrants are defined as individuals who lived in a different state 12 months previously. The sample consists of individuals with 2 to 30 years of potential experience (at the end of the 12 month window) in the ACS between 2000 and 2009, and this is further restricted to the employed in columns 4-9. Employment status and occupation are recorded at time of survey. Occupational codes are based on the census 2000 scheme.

Table 2: Job matching and separation rates: 4-month intervals

	Within-state matching rate (%)				Cross-state matching rate (%)				Separation rate to non-emp (%)
	All indiv	By emp status in $t-1$			All indiv	By emp status in $t-1$			
		Emp	Unemp	Inactive		Emp	Unemp	Inactive	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HS dropout	10.29	8.06	26.23	8.63	0.10	0.11	0.20	0.07	10.46
HS graduate	10.14	6.82	30.17	11.69	0.13	0.12	0.21	0.12	6.94
Some college	10.09	6.76	32.68	13.87	0.19	0.15	0.41	0.18	6.25
Undergraduate	7.96	5.33	36.16	13.35	0.31	0.25	0.94	0.28	4.04
Postgraduate	6.09	3.98	36.28	14.84	0.38	0.33	1.29	0.55	2.85

This table reports within-state and cross-state matching rates across four-month waves, together with separation rates to non-employment, for individuals with 1-30 years of potential experience and with no business income, based on the SIPP panels of 1996, 2001, 2004 and 2008. The full sample consists of 1.2m individual wave transitions. A "job find" occurs when an individual works for an employer at the end of wave t for whom they did not work in $t-1$. For individuals with multiple jobs at the end of t, I restrict attention to cases where the new job is the "primary" job: that is, the job which occupies the most weekly hours. A cross-state job find is one which is accompanied with a change in state of residence.

Table 3: Wage returns to within-state job matching

	All individuals		Experience groups			Initial tenure (months)		
			1-10	11-20	21-30	1-24	25-84	85+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New job (NJ)	0.034***	0.008	0.008	0.007	0.009	0.007	0.011	-0.005
	(0.003)	(0.007)	(0.010)	(0.012)	(0.013)	(0.009)	(0.013)	(0.024)
NJ * HS grad		0.014*	0.028**	0.004	0.003	0.026**	0.023	-0.040
		(0.008)	(0.012)	(0.015)	(0.017)	(0.011)	(0.017)	(0.028)
NJ * Some coll		0.022***	0.037***	0.017	0.000	0.039***	0.005	-0.007
		(0.008)	(0.012)	(0.015)	(0.016)	(0.011)	(0.016)	(0.028)
NJ * Undergrad		0.057***	0.075***	0.066***	0.009	0.057***	0.079***	0.028
		(0.010)	(0.015)	(0.018)	(0.022)	(0.013)	(0.020)	(0.030)
NJ * Postgrad		0.048***	0.079***	0.046*	0.008	0.069***	0.044*	0.031
		(0.015)	(0.023)	(0.027)	(0.028)	(0.023)	(0.027)	(0.034)
Demog controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	647,228	647,228	185,300	223,147	238,781	210,483	219,624	217,121
% New job	5.785	5.785	8.444	4.896	4.031	8.635	4.814	2.774

This table reports estimates of (18), based on four-month transitions in SIPP panels beginning 1996, 2001, 2004 and 2008, for observations with no cross-state migration. Throughout, I control for a full set of wave effects and a detailed set of demographic characteristics, specifically: experience and experience squared; four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in experience; black and Hispanic race dummies; foreign-born and native-born dummies (the omitted category contains the 7 percent of respondents who do not answer the relevant survey module); and a gender indicator which is also interacted with all previously mentioned variables. I use hourly wage data for workers paid by the hour, and I impute hourly wages for salaried workers using monthly earnings and hours. I exclude workers with multiple jobs or business income at the end of a wave, and I exclude wage observations below the 1st or above the 99th percentiles (within SIPP panels and education groups). The sample is restricted to individuals with 1-30 years of experience and working at least 15 hours per week, and it excludes wage transitions with intervening unemployment or layoff spells. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Quantifying the effect of wage returns

	Five edu groups					Two edu groups	
	HS dropout	HS grad	Some coll	Undergrad	Postgrad	Non-grad	Coll grad
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cross-state rate (%): $\mathbb{E}[\rho_C(\varepsilon)]$	0.099	0.118	0.157	0.268	0.331	0.134	0.289
Within-state rate (%): $\mathbb{E}[\rho_L(\varepsilon)]$	8.377	7.151	6.97	5.513	4.025	7.235	5.025
Odds ratio: $\frac{\mathbb{E}[\rho_C(\varepsilon)]}{\mathbb{E}[\rho_L(\varepsilon)]}$	0.012	0.016	0.023	0.049	0.082	0.019	0.057
Mean local wage returns: β_1	0.008	0.026	0.029	0.071	0.074	0.025	0.072
	(0.006)	(0.005)	(0.004)	(0.008)	(0.015)	(0.003)	(0.007)
Composite constraints: $\frac{2\mathbb{E}[\tilde{m}(\mu\|\varepsilon)]}{\tilde{\pi}}$	0.650	1.563	1.293	1.463	0.897	1.334	1.250
	(0.501)	(0.277)	(0.196)	(0.165)	(0.187)	(0.154)	(0.125)

This table calibrates a composite measure of constraints on long-distance job matching, $\frac{\tilde{\pi}}{2\mathbb{E}[\tilde{m}(\mu|\varepsilon)]}$, education. This is imputed from (i) $\mathbb{E}[\rho_C(\varepsilon)]$, the flow of cross-state job matches over four-month waves (within employment cycles), (ii) $\mathbb{E}[\rho_L(\varepsilon)]$, the flow of within-state matches, and (iii) β_1, the expected wage returns to within-state matches. The identification of these parameters is described in Section 5.3.

Table 5: Wage returns to cross-state job matching

	All individuals				Experience groups		Initial tenure (months)	
					1-10	11-30	0-24	25+
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Move	0.059***	-0.054*	-0.047	-0.062**	-0.055	-0.057	-0.030	-0.105*
	(0.021)	(0.033)	(0.034)	(0.025)	(0.038)	(0.056)	(0.039)	(0.057)
Move * Some coll		0.096**	0.085*	0.075*	0.120**	0.070	0.036	0.201**
		(0.049)	(0.049)	(0.046)	(0.060)	(0.080)	(0.060)	(0.083)
Move * Undergrad		0.130**	0.129**	0.120***	0.188***	0.029	0.156**	0.124
		(0.053)	(0.054)	(0.045)	(0.064)	(0.088)	(0.068)	(0.086)
Move * Postgrad		0.252***	0.252***	0.207***	0.315***	0.155*	0.241**	0.292***
		(0.068)	(0.068)	(0.068)	(0.085)	(0.093)	(0.114)	(0.084)
Demog controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Purged of state*HC	No	No	Yes	Yes	No	No	No	No
Purged of state*HC*occ	No	No	No	Yes	No	No	No	No
Sample	33,161	33,161	33,161	33,161	14,661	18,500	17,999	15,162
% Move	2.397	2.397	2.397	2.397	3.143	1.681	2.476	2.323

This table reports estimates of (28). The sample is identical to Table 3, except I now restrict it to individuals who change job (NewJob = 1). See notes under Table 3 for details on the demographic controls. In column 3, I use individual wages which are purged of triple interactions between state effects, education effects and a quadratic in experience. For column 4, I purge individual wages additionally of triple interactions between state effects, education effects and 3-digit occupation effects. See Section 6.2 for further details. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A1: Breakdown of primary reasons for moving

	State moves		County moves	
Primary reason	% full sample	% state migrants	% full sample	% county migrants
DUE TO SPECIFIC JOB	1.11	47.02	1.90	35.53
New job or job transfer	0.94	39.82	1.38	25.82
Easier commute	0.05	2.30	0.33	6.23
Other job reasons	0.12	4.91	0.19	3.48
LOOK FOR WORK	0.13	5.48	0.22	4.03
NON-JOB REASONS	1.12	47.49	3.23	60.44
Family	0.55	23.48	1.40	26.12
Change in marital status	0.10	4.42	0.33	6.08
Establish own household	0.09	3.79	0.36	6.76
Other family reasons	0.36	15.28	0.71	13.28
Housing	0.26	10.86	1.16	21.75
Want to own home	0.04	1.69	0.27	5.04
New or better housing	0.06	2.56	0.35	6.47
Cheaper housing	0.06	2.56	0.22	4.19
Other housing reasons	0.10	4.05	0.32	6.04
Environment	0.13	5.01	0.27	4.91
Better neighborhood	0.04	1.50	0.15	2.71
Climate, health, retirement	0.08	3.51	0.12	2.20
Attend/leave college	0.10	4.39	0.21	3.92
Other reasons	0.09	3.75	0.20	3.74
ALL REASONS	2.35	100	5.35	100

This table presents migration rates by primary reason in CPS March waves between 1999 and 2018. The first column reports the percentage of the *full sample* who changed state, for each given reason, over the previous twelve months. The second column expresses these numbers as a percentage of *state-movers*. The final two columns repeat the exercise for cross-county moves. I include individuals moving because of foreclosure or eviction in the CPS's "other housing reasons" category; and I include individuals moving because of natural disasters in the "other reasons" category. See Appendix A.1 for sample details.

Table A2: Education gradients by primary reasons for moving

	State moves					County moves				
Primary reason	HSD	HSG	SC	UG	PG	HSD	HSG	SC	UG	PG
DUE TO SPECIFIC JOB	0.43	0.74	0.92	1.63	2.20	0.93	1.37	1.75	2.72	3.03
New job or job transfer	0.33	0.59	0.77	1.42	1.95	0.58	0.91	1.21	2.09	2.46
Easier commute	0.04	0.05	0.05	0.06	0.08	0.22	0.29	0.36	0.42	0.33
Other job reasons	0.06	0.10	0.11	0.15	0.17	0.13	0.17	0.18	0.22	0.23
LOOK FOR WORK	0.21	0.15	0.11	0.12	0.07	0.39	0.25	0.19	0.17	0.10
NON-JOB REASONS	1.15	1.18	1.16	1.11	0.82	3.42	3.45	3.48	3.07	2.16
Family	0.63	0.64	0.58	0.46	0.36	1.57	1.62	1.49	1.19	0.82
Change in marital status	0.08	0.12	0.12	0.09	0.09	0.22	0.35	0.37	0.32	0.25
Establish own household	0.11	0.10	0.10	0.07	0.05	0.37	0.42	0.40	0.33	0.19
Other family reasons	0.44	0.42	0.37	0.29	0.22	0.97	0.85	0.72	0.54	0.38
Housing	0.29	0.27	0.23	0.27	0.23	1.27	1.22	1.16	1.17	0.92
Want to own home	0.03	0.04	0.04	0.05	0.05	0.20	0.25	0.27	0.34	0.27
New or better housing	0.08	0.07	0.05	0.06	0.04	0.36	0.36	0.34	0.38	0.28
Cheaper housing	0.08	0.08	0.05	0.05	0.04	0.31	0.27	0.22	0.17	0.13
Other housing reasons	0.10	0.09	0.09	0.10	0.10	0.40	0.35	0.33	0.28	0.24
Environment	0.16	0.15	0.14	0.11	0.08	0.35	0.30	0.29	0.22	0.15
Better neighborhood	0.06	0.04	0.04	0.02	0.02	0.21	0.16	0.15	0.12	0.08
Climate, health, retirement	0.09	0.09	0.09	0.08	0.05	0.14	0.14	0.12	0.09	0.07
Attend/leave college	0.00	0.05	0.13	0.19	0.10	0.02	0.10	0.34	0.30	0.16
Other reasons	0.08	0.09	0.09	0.10	0.07	0.21	0.22	0.21	0.19	0.12
ALL REASONS	1.79	2.07	2.2	2.86	3.09	4.73	5.07	5.42	5.96	5.29

This table reports cross-state and cross-county migration rates by primary reason for moving (as in Table A1), but now disaggregated by education. I consider five education groups: high school dropouts (HSD), high school graduates (HSG), some college (SC), undergraduate degree (UG) and postgraduate (PG).

Table A3: Migration rates (%) for all individuals and household top earners

	Specific job		Look for work		Non-job	
	All indiv	Top earners	All indiv	Top earners	All indiv	Top earners
HS dropout	0.43	0.46	0.21	0.21	1.15	1.11
HS graduate	0.74	0.76	0.15	0.15	1.18	1.13
Some college	0.92	0.95	0.11	0.09	1.16	1.18
Undergrad	1.63	1.67	0.12	0.11	1.11	1.12
Postgrad	2.20	2.36	0.07	0.06	0.82	0.79

This table reports annual cross-state job migration rates by reported reason for moving, separately for all individuals (identical to Figure 1) and for household top earners, and based on CPS March waves between 1999 and 2018. See Appendix A.1 for further sample details.

Table A4: Students and return movers: Individuals with 2-10 years of experience

	HS dropout (1)	HS grad (2)	Some coll (3)	Undergrad (4)	Postgrad (5)
% recent students by completed education	0.18	0.08	0.16	0.05	0.02
Migration rate (%): all cross-state moves					
Full sample	3.10	4.31	5.23	8.78	9.58
Excl. recent students	3.44	4.52	5.79	8.97	9.82
Migration rate (%): return moves	1.66	1.94	1.98	3.35	3.19
Migration rate (%): non-return moves	1.43	2.37	3.25	5.42	6.39
Observations	4,396	9,977	6,125	3,172	1,244

This table reports annual cross-state migration rates by education group, based on all (annual) PSID waves between 1990 and 1997. Migration rates are constructed using reported state of residence 12 months previously. The first row gives the fraction of the sample who were recently students (in the current or previous wave). The second row reports cross-state migration rates for the full sample, and the third row reports these rates excluding recent students. The fourth and fifth rows disaggregate the migration rate (for the full sample) into return and non-return moves. Return moves include all moves to (i) states where the individual has resided previously in the panel or (ii) the state where the individual reports having grown up. The sample consists of all individuals with 2-10 years of potential labor market experience at the end of each 12-month interval.

Table A5: Log point responses of cross-state migration to imputed costs

	Unconditional sample		Conditional sample (willing to move)		
	(1)	(2)	(3)	(4)	(5)
Initial willingness to move	0.987***	0.868***			
	(0.170)	(0.191)			
Initial willingness to move * Grad		0.404			
		(0.406)			
Initial imputed cost			-0.206	-1.346***	-1.345***
			(0.349)	(0.423)	(0.462)
Initial imputed cost * Grad					0.043
					(0.992)
Initial log wage				-1.321***	-1.631***
				(0.373)	(0.419)
Initial log wage * Grad					1.353*
					(0.796)
Demographic controls, year effects	Yes	Yes	Yes	Yes	Yes
Observations	8,836	8,836	2,904	2,904	2,904
Cross-state mig rate	0.035	0.035	0.059	0.059	0.059

This table reports responses of cross-state migration (over 12-month intervals) to subjective costs and wages (at the beginning of each interval), based on complementary log-log regressions. I study the response to both a binary indicator of "willingness" to move for work; and conditional on being willing to move, the response to the imputed moving cost. These measures are described in greater detail in the notes under Figure 4. In columns 2 and 5, I also allow for interactions between the cost measures and a college graduate dummy. Coefficients should be interpreted as the log point effect of each measure on the instantaneous cross-state migration rate, conditional on the empirical model described by equation (A2). I only have imputed costs for employed household heads between 1969 and 1973, so I restrict attention to the mobility decisions of these individuals in the annual intervals between 1969 and 1974. I exclude individuals with less than 2 or more than 30 years of potential experience at the end of each interval. Household heads in the PSID are always male, unless there is no husband (or cohabiting partner) present or the husband is too ill to respond to the survey. All specifications control for a full set of year effects and demographic controls, specifically experience and experience squared, four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in experience, black and Hispanic dummies, and a gender dummy interacted with all previously mentioned controls. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Net cross-state migration rates by education and experience

	Basic			Within 2-digit occs			Within 3-digit occs		
	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio	Gross mig rate (%)	Net mig rate (%)	Net-gross ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Individuals with 2-10 years of experience									
HS dropout	3.37	0.45	0.13	3.24	1.55	0.48	3.24	1.96	0.61
HS graduate	4.29	0.37	0.09	3.78	1.13	0.30	3.78	1.64	0.43
Some college	4.42	0.36	0.08	3.91	1.22	0.31	3.91	1.83	0.47
Undergraduate	6.52	0.51	0.08	5.82	1.51	0.26	5.82	2.27	0.39
Postgraduate	7.50	0.66	0.09	7.04	1.84	0.26	7.04	2.65	0.38
Individuals with 11-30 years of experience									
HS dropout	2.11	0.32	0.15	1.83	0.79	0.43	1.83	1.06	0.58
HS graduate	2.01	0.27	0.14	1.65	0.51	0.31	1.65	0.73	0.44
Some college	2.32	0.28	0.12	1.92	0.60	0.31	1.92	0.90	0.47
Undergraduate	2.38	0.24	0.10	2.03	0.58	0.28	2.03	0.87	0.43
Postgraduate	2.64	0.24	0.09	2.34	0.66	0.28	2.34	0.94	0.40

This table reports annual gross and net cross-state migration rates within education groups, separately for individuals with 2-10 and 11-30 years of potential experience. See notes under Table 1 in main text for sample details and construction of variables.

Table A7: Elasticity of cross-state job matching to experience and initial job tenure

	(1)	(2)	(3)	(4)
HS grad	0.416**	0.425**	0.542***	0.502***
	(0.191)	(0.191)	(0.191)	(0.191)
Some coll	0.601***	0.557***	0.737***	0.656***
	(0.182)	(0.182)	(0.182)	(0.183)
Undergrad	1.192***	1.151***	1.416***	1.305***
	(0.179)	(0.180)	(0.180)	(0.180)
Postgrad	1.656***	1.766***	1.989***	1.963***
	(0.181)	(0.183)	(0.182)	(0.183)
Experience		-0.106***		-0.071***
		(0.018)		(0.018)
Experience ^ 2		0.001		0.000
		(0.001)		(0.001)
Initial log tenure			-0.398***	-0.266***
			(0.021)	(0.027)
Sample	874,582	874,582	874,582	874,582
Matching rate (%)	0.129	0.129	0.129	0.129

This table reports complementary log-log regressions, of the form of (A4), for the incidence of cross-state job matching, based on four-month transitions in SIPP panels beginning 1996, 2001, 2004 and 2008. I restrict attention to job finds within employment cycles, i.e. without intervening unemployment or layoff spells. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

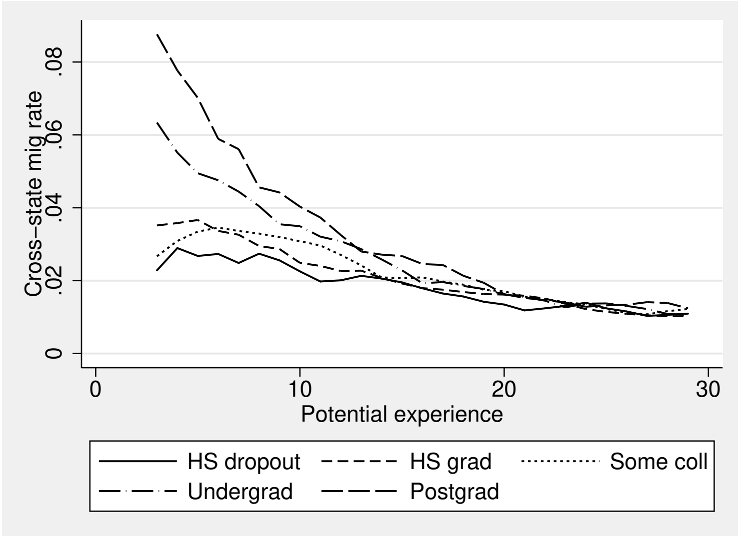

Figure 1: Annual cross-state migration rates

This figure reports the fraction of individuals living in a different state 12 months previously in Current Population Survey (CPS) March waves between 1999 and 2018, taken from the IPUMS database (Flood et al., 2018). Data points represent three-year moving averages (separately by education) over the support of potential labor market experience. I exclude individuals with less than one year of experience at the beginning of the 12 month window, those living abroad one year previously, and those with imputed migration status: Kaplan and Schulhofer-Wohl (2012) show there are inconsistencies in the imputation procedure in non-response cases. The non-response rate for migration status is 14 percent in my sample, and this varies little with education. See Appendix A.1 for definitions of variables.

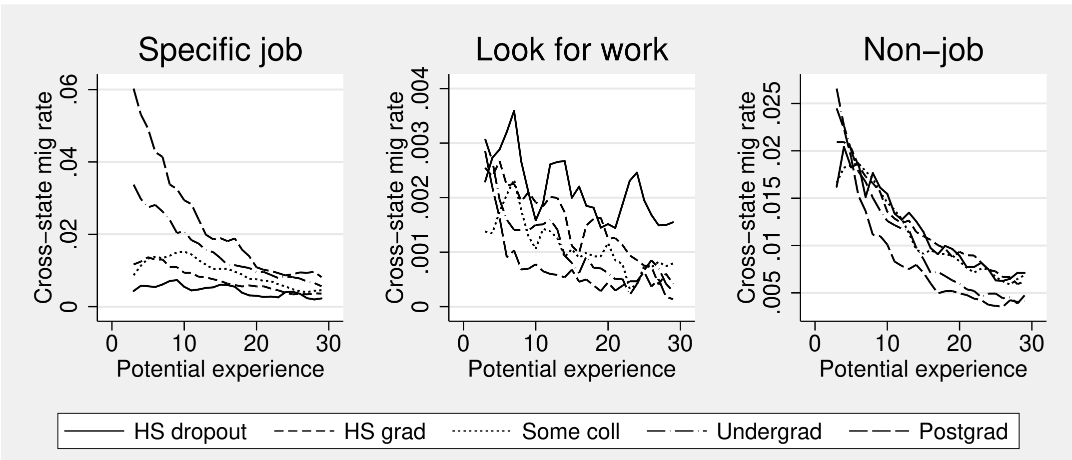

Figure 2: Annual cross-state migration rates by reported reason for moving

The first panel reports the fraction of individuals who moved state primarily for reasons related to a specific job (new job, job transfer, shorten a commute, other job reason) in the previous 12 months. The second panel does the same for individuals who report moving to "look for work", and the third panel for reasons unrelated to employment. Data is based on the March CPS between 1999 and 2018. See notes below Figure 1 for further details.

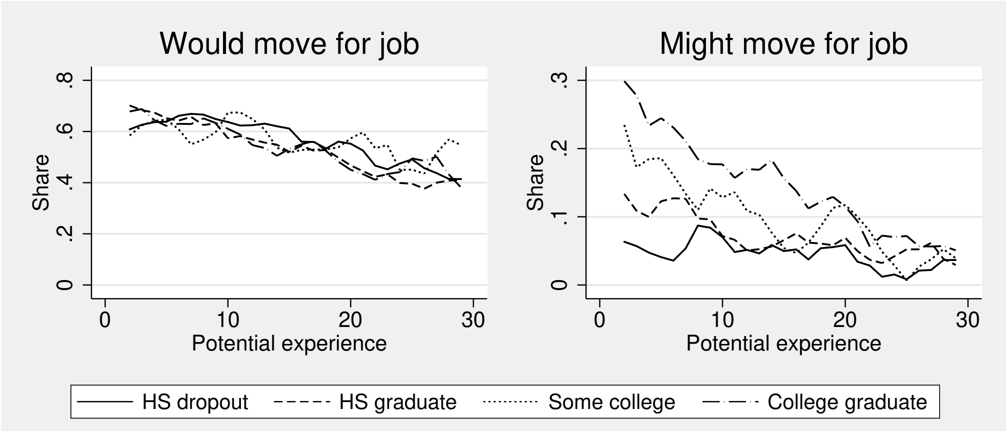

Figure 3: Share who "would move" and "might move" for better job

The first panel reports the share of employed household heads who report being willing to move for work. This is based on responses to "Would you be willing to move to another community if you could earn more money there?" The second panel reports the share of employed heads who both (i) answer affirmatively to the question "Do you think you might move in the next couple of years?" and (ii) report job-related reasons in answer to the question "Why might you move?" Household heads in the PSID are always male, unless there is no husband (or cohabiting partner) present or the husband is too ill to respond to the survey. The sample is restricted to employed heads with 1-30 years of experience in the years 1969-72 and 1979-80, when both questions were asked. The full sample consists of 16,947 observations.

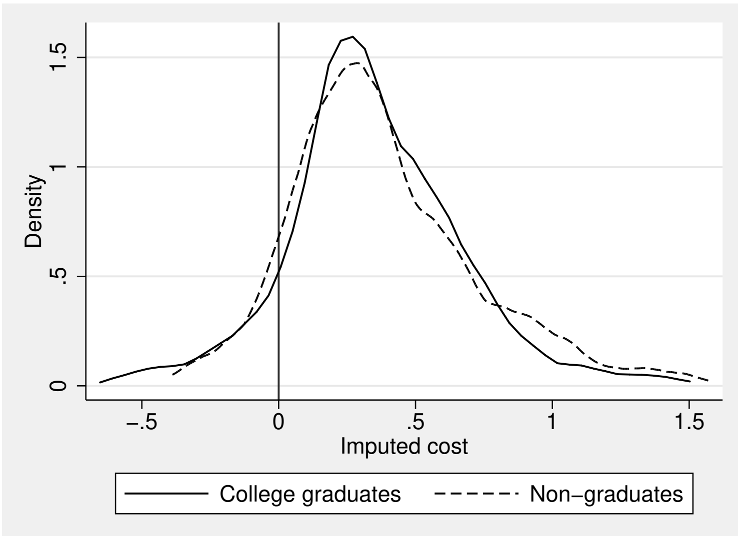

Figure 4: Conditional distribution of imputed annuitized costs

This figure plots kernel distributions of the imputed annuitized costs of migration. Conditional on expressing willingness to move, employed household heads (in the PSID, 1969-72) answer the question: "How much would a job have to pay for you to be willing to move?" Imputed costs are computed as the difference between the log of this reservation wage and the worker's current wage (or more specifically, the worker's average wage over the previous 12 months), where wages are measured in hourly terms. I drop all observations with top-coded reservation wages. Importantly, these are unbalanced: 31 percent of graduate observations are top-coded, compared to just 6 percent for non-graduates. To address concerns about selection, I additionally exclude the top 25 percent of the (remaining) non-graduate reservation wage observations. Beyond this, I also drop the top and bottom 2 percent of the imputed cost distribution within each education group. I restrict the sample to individuals with 1-30 years of experience. The sample consists of 398 college graduates and 2,744 non-graduates. Plots are based on Epanechnikov kernel functions, with a bandwidth of 0.07.

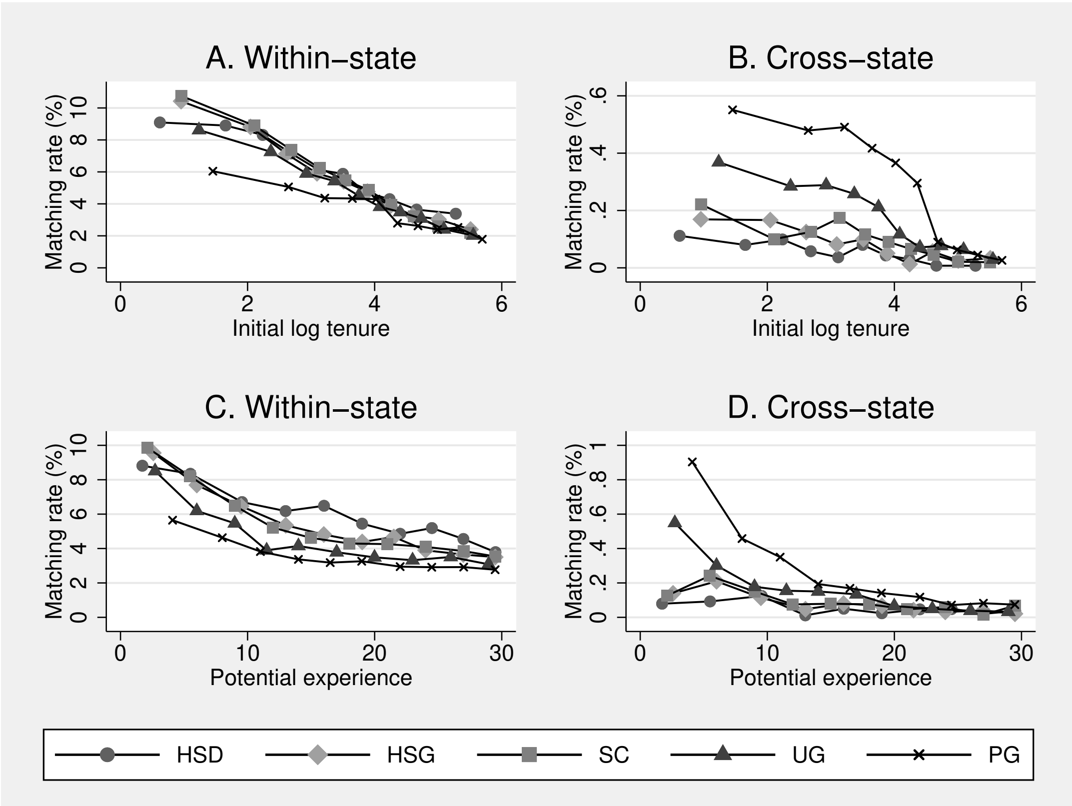

Figure 5: Job matching rates within employment cycles, by tenure and experience

This figure plots the rate of within-state and cross-state job matching on both the log of initial (i.e. $t-1$) job tenure (measured in months) and experience, separately by education. I restrict attention to job finds within employment cycles, i.e. without intervening unemployment or layoff spells. I have divided the support of both experience and tenure into decile bins, within each education group. Each data point identifies the mean experience (or tenure) and the matching rate in a given decile bin.

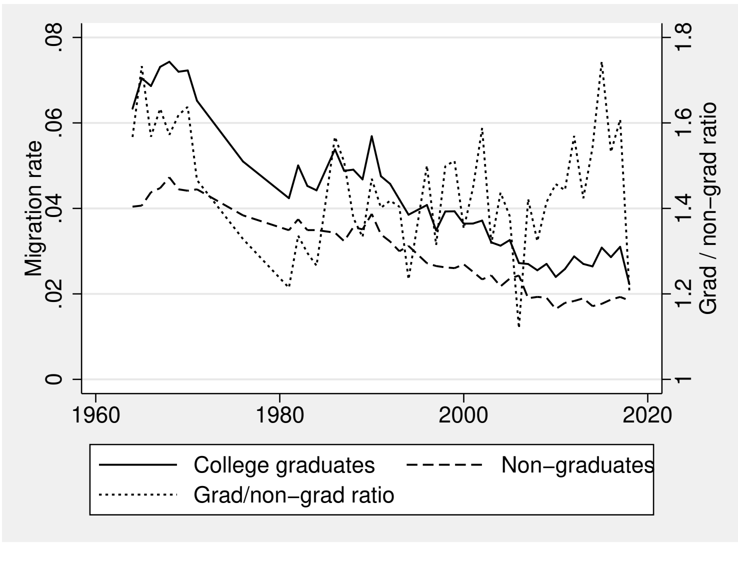

Figure A1: Annual cross-state migration rates by education (1964-2018)

This figure reports annual rates of cross-state migration over time among individuals with 2-30 years of potential experience in the CPS, separately for college graduates and non-graduates. The right-hand scale gives the ratio of the two. I exclude all individuals living abroad one year previously, and I also exclude observations for which the CPS has imputed migration status: see Appendix A.1 for further details.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1615	Ria Ivandic Tom Kirchmaier Stephen Machin	Jihadi Attacks, Media and Local Hate Crime
1614	Johannes Boehm Swati Dhingra John Morrow	The Comparative Advantage of Firms
1613	Tito Boeri Pietro Garibaldi	A Tale of Comprehensive Labor Market Reforms: Evidence from the Italian Jobs Act
1612	Esteban Aucejo Teresa Romano Eric S. Taylor	Does Evaluation Distort Teacher Effort and Decisions? Quasi-experimental Evidence from a Policy of Retesting Students
1611	Jane K. Dokko Benjamin J. Keys Lindsay E. Relihan	Affordability, Financial Innovation and the Start of the Housing Boom
1610	Antonella Nocco Gianmarco I.P. Ottaviano Matteo Salto	Geography, Competition and Optimal Multilateral Trade Policy
1609	Andrew E. Clark Conchita D'Ambrosio Marta Barazzetta	Childhood Circumstances and Young Adult Outcomes: The Role of Mothers' Financial Problems
1608	Boris Hirsch Elke J. Jahn Alan Manning Michael Oberfichtner	The Urban Wage Premium in Imperfect Labour Markets
1607	Max Nathan Anna Rosso	Innovative Events
1606	Christopher T. Stanton Catherine Thomas	Missing Trade in Tasks: Employer Outsourcing in the Gig Economy

1605	Jan-Emmanuel De Neve Christian Krekel George Ward	Employee Wellbeing, Productivity and Firm Performance
1604	Stephen Gibbons Cong Peng Cheng Keat Tang	Valuing the Environmental Benefits of Canals Using House Prices
1603	Mary Amiti Stephen J. Redding David Weinstein	The Impact of the 2018 Trade War on U.S. Prices and Welfare
1602	Greer Gosnell Ralf Martin Mirabelle Muûls Quentin Coutellier Goran Strbac Mingyang Sun Simon Tindermans	Making Smart Meters Smarter the Smart Way
1601	Antoine Dechezleprêtre Caterina Gennaioli Ralf Martin Mirabelle Muûls Thomas Stoerk	Searching for Carbon Leaks in Multinational Companies
1600	Jeremiah Dittmar Skipper Seabold	New Media and Competition: Printing and Europe's Transformation after Gutenberg
1599	Kilian Huber Volker Lindenthal Fabian Waldinger	Discrimination, Managers, and Firm Performance: Evidence from “Aryanizations” in Nazi Germany
1598	Julia Cajal-Grossi Rocco Macchiavello Guillermo Noguera	International Buyers’ Sourcing and Suppliers’ Markups in Bangladeshi Garments

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE