# Why are Higher Skilled Workers More Mobile Geographically? The Role of Job Rents

Michael Amior*

November 2016

**Abstract**

Skill differences in geographical mobility are entirely driven by (typically young) workers who report moving for a new job. I argue this is a natural consequence of the specialized nature of skills, independent of geography. In a "thin" labor market, given the importance of job match quality, skilled workers will accrue substantial rents as they climb the jobs ladder. This is particularly so for younger workers, who are just beginning their careers. Though the origin of these rents is unrelated to geography, I claim these rents are crucial in explaining geographical mobility given that moving is typically costly. Using data from the US, I show that skill differences in wage rents are large enough to plausibly explain the mobility gap. I find little support for the view that skilled mobility is driven principally by low migration costs. In fact, based on estimates of the wage returns to cross-state matches, I show that workers' realized migration costs are steeply *increasing* in skill - conditional on moving. This is a natural consequence of selection on large wage offers. I also present new evidence on subjective migration costs which supports my claims.

## 1 Introduction

It is well known that better educated individuals are more mobile geographically over long distances. The rate of cross-state migration is almost twice as high for college graduates, and this is largely due to younger workers (Figure 1).[1] Furthermore, the low skilled population

*St. Catharine's College, University of Cambridge, and Centre for Economic Performance, London School of Economics; E-mail: maa88 at cam.ac.uk. I am grateful to Alan Manning for his support at LSE. This study was part of my PhD thesis at University College London, 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, Hamish Low, Barbara Petrongolo, Steve Pischke, Jan Stuhler and Coen Teulings for helpful comments, as well as seminar participants. I gratefully acknowledge financial support from the Economic and Social Research Council and the Royal Economic Society while at UCL.

[1] Figure 1 is based on data since 1999 (since these waves of the Current Population Survey report reasons for moving, which I exploit below), but I show in Appendix A.4 that these patterns are much older than this. There has been a decline in overall migration rates for all education groups since the 1980s (Molloy, Smith and Wozniak, 2011), perhaps due to declining geographical specificity of occupational returns (Kaplan and Schulhofer-Wohl, 2012*b*) or a decline in labor market transitions (Molloy, Smith and Wozniak, 2014). But nonetheless, large skill differentials have persisted. Appendix A.3 also presents a breakdown of these results by single-year age categories.

adjusts more sluggishly to local business cycle fluctuations (e.g. Bound and Holzer, 2000; Wozniak, 2010; Notowidigdo, 2011; Amior and Manning, 2015). This is concerning, given they suffer disproportionately from local volatility (Hoynes, 2002).

[Figure 1 here]

Recent evidence suggests the effect of education is causal (Malamud and Wozniak, 2012; Machin, Salvanes and Pelkonen, 2012), but the specific mechanism is still debated. Figure 2 shows the mobility gap is entirely driven by individuals who report moving for job-related reasons rather than "non-job" (primarily family and housing) reasons.[2] It is worth emphasizing that the job-motivated movers almost always have a job lined up at their destination. Only 5 percent of cross-state migrants report moving speculatively to "look for work"[3]; and these speculative movers are in fact disproportionately low skilled (see Appendix B.3).

[Figure 2 here]

Building on early work by Schwartz (1976), later echoed by Wildasin (2000), I argue these patterns are a natural consequence of the specialized nature of skills, independent of geography. Skill specialization generates large dispersion in the productivity of job matches. In a "thin" labor market, skilled workers therefore accrue substantial wage rents as they climb the jobs ladder (and improve their match), irrespective of location. This is particularly so for the young, who are just beginning their careers. What are the implications for migration? Workers only move if the wage rents associated with a distant job offer exceed the cost of moving. And given that costs are typically large, these wage rents play a crucial role. For example, a better job may motivate a computer scientist to move from New York to Houston, but not somebody who cuts hair for a living. However, while wage rents are potentially decisive for long-distance matching, they matter little for *local* job matching (where transition costs are low); and indeed, Figure 3 shows that education has little effect on the *overall* flow of new jobs.

[Figure 3 here]

[2] There is a slight positive gradient for the under-35s in "non-job" migration, but I show in Appendix B.2 that this is entirely driven by individuals who report moving to attend or leave college. I also show these results are robust to controlling for demographic characteristics; and Appendix B.3 presents skill gradients for a more detailed breakdown of reasons for moving. In the same Appendix, I also estimate skill gradients for cross-county moves - within states. There is still a large positive effect of education on job-motivated migration, though the slope is not as steep. This result is consistent with the model below, to the extent that cross-state migration is more costly (see Proposition 1).

[3] This is unsurprising: moving without a job in hand is a costly and risky strategy (Molho, 1986).

Critically, this hypothesis makes no claims on the geographical structure of these job rents. If (and only if) labor markets are thin, an aggregate-level difference in the wage offer distribution is sufficient to generate differentials in geographical mobility. Following the earlier example, the computer scientist moves from New York to Houston because that is where the job offer happens to materialize - and not because Houston (ex ante) offers higher returns to his particular skills. In other words, my hypothesis stresses the role of the worker-*firm* match (independent of geography), rather than the worker-*location* match. This is important because the evidence suggests geographical variation plays little role in driving these patterns: I show below that net migratory flows between states are not increasing in education, even within detailed occupation groups. In contrast, the importance of (aggregate-level) match quality in skilled markets is both theoretically intuitive and supported by recent empirical work: see Gottfries and Teulings (2016) and Lise, Meghir and Robin (2016).

I make two contributions to the literature. First, I offer a new model of migration embedded in a simple jobs ladder framework - in the spirit of Burdett and Mortensen (1998). This framework was not available to Schwartz (1976) when he was writing, and it offers a simple explanation for age differences in the mobility gap. The model yields a series of predictions on wage changes conditional on changing job and conditional on changing location. My second contribution is to test these predictions in the data. The results offer strong support for the wage rents hypothesis and reject the view that skilled mobility is principally explained by low migration costs. As a supplementary exercise, I also offer new subjective measures of these migration costs.

In the model, both employed and unemployed workers draw random wage offers from an exogenous distribution. Following the logic of Burdett and Mortensen (1998), on-the-job search gives rise to a jobs ladder - with the ladder's rungs corresponding to job match quality. Job offers may arise in a worker's home location or elsewhere. If the offer happens to come from elsewhere, the worker draws a random migration cost - which I express in terms of the disamenity of living away from home.[4] And the worker accepts the offer (and moves) if the associated wage gains exceed this amenity cost. At its most abstract level, this model describes a jobs ladder in two dimensions: productivity and amenities. While I focus here on *residential* amenities, its applications are certainly broader.

My approach to modeling multi-location job search diverges from the seminal study in the field, Kennan and Walker (2011). There, workers only draw their wage *after* arriving in a new location; but here, the wage offer is known *ex ante* - so workers move with a job lined up, conditional on a sufficiently attractive offer.[5] This assumption is instrumental to my claim that

[4]The migration cost can alternatively be modelled as a one-off moving cost (I offer an extension below), but this does not affect the intuition.

[5]See also Jackman and Savouri (1992), Molho (2001) and Lutgen and Van der Linden (2015), who interpret internal migration as long-distance job matching.

skilled migration is driven by job match quality.

Workers are more likely to migrate from their home area if the dispersion of wage offers is large relative to preferences over amenities: that is, workers care more about jobs than places. Also, the impact of wage dispersion is greater for workers with lower quality matches (at lower rungs of the jobs ladder), because they still have many rungs to climb. To the extent that younger workers are concentrated at these lower rungs, this offers a ready explanation for why they are largely responsible for the skill mobility gap.[6]

To test these claims, I study the wage returns to new job matches in the Survey of Income and Program Participation (SIPP). I am not interested here in deriving causal estimates of the returns to "exogenous" job transitions. Instead, the model makes predictions on selection into new jobs - and this is precisely what I want to study. First, I show the wage returns to job finding are steeply increasing in education (confirming evidence from Bartel and Borjas, 1981, and Mincer, 1986) and especially for the young (consistent with Gottfries and Teulings, 2016). Remarkably, the larger impact of education on younger workers' rents is fully explained (statistically) by differences in workers' initial wage. This is consistent with the hypothesis that the disproportionate effect on the young is entirely driven by the existence of a jobs ladder.

Importantly, the magnitude of these skilled rents is sufficiently large to theoretically explain the observed skill differences in geographical mobility - conditional on some distributional assumptions on preferences over locations.

In contrast, I find little support for the view that skilled mobility is driven principally by low migration costs. In particular, I show the wage returns to *cross-state* matches are disproportionately large for skilled workers. This suggests that workers' realized migration costs are steeply *increasing* in skill - conditional on moving. Intuitively, given large wage dispersion, skilled workers are more likely to select into migration *because* of large wage rents and *despite* steep migration costs. This can explain why they disproportionately report moving for job-related rather than amenity reasons (Figure 2). Among the low skilled though, the returns to cross-state and local matches are remarkably similar. This suggests they typically move *because* of a low cost draw and *despite* meager rents.

I also present more direct evidence on migration costs, imputing them from subjective data in the Panel Study of Income Dynamics (PSID). In the 1970s[7], respondents were asked whether they would relocate for higher pay - and what wage would tempt them to move. I impute migration costs from the answers, but these vary little with education. Reassuringly, these cost estimates are consistent with estimates of migrants' wage rents in the SIPP. Of course, this exercise is only valuable if these subjective responses are informative about true costs - and I

[6]Indeed, the existing evidence shows that a significant part of lifecycle earnings progression is driven by the jobs ladder (Manning, 2003) - and especially for skilled workers (Gottfries and Teulings, 2016).

[7]Of course, this is old data. But, as I show in Appendix F.1, age and education differentials in cross-state mobility in my 1970s PSID sample look very similar to those in Figure 1 above.

show they do indeed have substantial explanatory power for future migration decisions.

In short, I claim there is strong empirical support for the view that aggregate-level job rents drive the mobility gap - as well as good theoretical reasons to expect it. In the next section, I contrast my explanation for the mobility gap against the existing literature. In Section 3, I offer evidence (based on net migratory flows) that the mobility gap is not driven by geographical variation. Section 4 sets out the jobs ladder model and derives the key results on wage rents and geographical mobility; and I test these predictions in Section 5. In Section 6, I study the implications of selection into migration on amenity cost draws - both theoretically and in the data. Section 7 sets out the evidence on subjective migration costs, and I conclude in Section 8.

# 2 Related literature

For the most part, the migration literature has relied on a location choice model, where workers trade off differential values of locations (which depend on expected local wages, housing costs and amenities) with the cost of moving. This confines us to two possible explanations for the skill mobility gap. Either (i) skilled workers face larger geographical differentials in expected utility, whether due to local skill agglomeration (e.g. Costa and Kahn, 2000; Wheeler, 2001; Davis and Dingel, 2012), the worker-location productivity match (Lkhagvasuren, 2014), compensating transfer payments and housing costs (Notowidigdo, 2011), or skill-varying preferences over local amenities (Diamond, 2016). Or alternatively, (ii) the low skilled face higher migration costs, whether due to financial constraints, lack of information or home attachment (Greenwood, 1973; Topel, 1986; Bound and Holzer, 2000; Wozniak, 2010; Moretti, 2011; Kennan, 2015).[8]

But neither of these explanations are entirely satisfying. First, it has proven difficult to identify exactly which costs might be responsible for the mobility gap. For this reason, Davis and Dingel (2012) have criticized the costs explanation for assuming low skilled "immobility" without an adequate theoretical basis. And regarding the utility differentials view, it can plausibly be argued that local disparities are in fact larger for the *low skilled*: after all, they suffer from greater local fluctuations in wages and employment rates (Hoynes, 2002; Gregg, Machin and Manning, 2004). And indeed, the evidence below shows that net flows between states contribute little to the skill mobility gap - even *within* detailed occupation groups. Instead, by allowing for thin labor markets, I move away from the confines of the location choice model. In this environment, the worker-firm match can drive geographical mobility - even if locations

[8] Gregg, Machin and Manning (2004) suggest that college graduates have weaker home attachment, having already left home to study, though Malamud and Wozniak (2012) dispute this hypothesis. Both these studies propose that long-distance job search is more costly for lower skilled workers, whether due to a lack of information or fewer social contacts in other regions. In addition to these ideas, Bound and Holzer (2000) suggest a lack of assets may constrain the set of location choices, especially if house prices are higher in desirable cities.

are (ex ante) identical.

Finally, it is important to discuss the role of return migration. Recent work by Kennan and Walker (2011) has shown this is an important feature of migratory flows. And there is a legitimate concern that the mobility gap is largely driven by students returning home from college - given that the skill gradient is steepest for the young (Figure 1). Indeed, Kennan (2015) has shown this is an important factor in the migration decisions of recent graduates. But, I show in Appendix D that returning students contribute little to the mobility gap for the youngest individuals in my sample (aged 25-34).

# 3 Net and gross migratory flows

If the mobility gap is indeed driven by geographical differentials in expected utility, this yields a simple testable implication: the mobility gap should be empirically explained by large net flows of skilled workers to particular locations. But I show here that this is not supported by the evidence - both on aggregate and even within detailed occupation categories.

It is well known that gross migration flows vastly exceed net flows across locations (Shryock, 1959; Schwartz, 1971; Jackman and Savouri, 1992; Wildasin, 2000; Coen-Pirani, 2010): that is, the bulk of internal migration is due to people moving in opposite directions. It turns out this is especially true for better educated workers: skilled migration has a much weaker directional component. This has previously been shown by Folger and Nam (1967), Schwartz (1971) and Lkhagvasuren (2014).

I begin by replicating this result. 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 number of out-migrants from state $s$. Dividing the expression by 2 ensures that migrants are not double-counted. Notice the gross migration rate is simply equal to $\frac{1}{n}\sum_s n_s^{in}$ or $\frac{1}{n}\sum_s n_s^{out}$. My sample includes individuals aged 25 to 64 in the American Community Survey (ACS) between 2000 and 2009.[9] Migration rates are estimated by comparing individuals' current state of residence with their residence 12 months previously.

[Table 1 here]

Table 1 reports gross and net migration rates separately for each education group. Gross migration is steeply increasing in education (column 1); but remarkably, there is no systematic relationship between education and net migration (column 2). As a result, the ratio of net to

[9] Occupational classification is consistent in these years, based on the census 2000 scheme. The occupational decomposition is very demanding empirically, and the ACS offers much larger samples than the Current Population Survey (CPS), which I use for Figures 1 and 2. There are 390,000 cross-state migrants in my ACS sample, compared to just 35,000 in the CPS (between 1999 and 2015).

gross migration is decreasing in education: from 0.15 for dropouts to 0.09 for postgraduates (column 3).

The evidence in Table 1 shows that net flows within education groups explain very little of the skill mobility gap. But, this does not entirely rule out the possibility that local utility differentials are driving the mobility gap - if these differentials are tied to particular task specializations. A natural way to test this is to study net and gross flows within detailed occupation groups. Specifically, for each education group, I estimate the within-occupation net migration rate as $\sum_o \gamma_o \frac{1}{2n_o} \Sigma_s |n^{in}_{os} - n^{out}_{os}|$, where $\gamma_o$ is the fraction of individuals employed in occupation group $o$, and $n^{in}_{os}$ and $n^{out}_{os}$ are the number of in/out-migrants to/from state $s$ employed in occupation $o$.

In constructing these within-occupation rates, I restrict the sample to individuals employed at the time of survey; and occupations are also recorded at the time of survey.[10] I report estimates separately based on 98 2-digit occupations (columns 4-6) and 466 3-digit occupations (7-9); these codes are based on the 2000 census scheme.

Gross migration rates in the occupation sample are slightly lower than in the basic sample (because of the restriction to employed individuals), though the steep education gradient is unaffected. The key point is that the skill gradient in net migration remains remarkably flat - even within 3-digit occupations. Consequently, columns 6 and 9 show a strong negative effect of education on the net-gross ratio: ranging from 0.4 for dropouts to 0.2 for postgraduates in the 2-digit classification, and from 0.5 to 0.3 for 3 digits. The net-gross ratios are larger for more detailed categories, which is a natural consequence of smaller cells.

These results suggest the skill mobility gap cannot be explained by skilled workers converging on particular states - even within detailed occupation categories. For example, computer scientists are much more mobile than hairdressers overall; but net flows of computer scientists across states are not much greater than net flows of hairdressers. There may be many computer scientists flocking to Northern California, but there are also many moving in the opposite direction. In Appendix C, I re-estimate these migration rates separately within different age groups - and I show the key results are preserved. I also estimate gross and net migration rates for each occupation group individually; and I show the gross rates are steeply increasing in a measure of occupational skill, while there is little relationship with the net rates.

[10]Note this is immediately *after* the twelve month period in which migration occurs. There may be a concern here that occupation and the migration decision are then simultaneously determined. But arguably, this is a useful time to measure occupation for this particular exercise - since an individual's ex post occupation is a good indicator of the job market in which they were searching. In any case, I draw comfort from the strength of the patterns in Table 1.

# 4 Jobs ladder model of migration

## 4.1 Overview

I set the model in continuous time. Workers are either employed or unemployed. A worker $i$ residing in area $j$ receives a flow utility:

$$v_{ij} = w_i - c_{ij} \tag{1}$$

where $w_i$ denotes the wage or (for the unemployed) out-of-work income $b_i$. This may represent a dollar or log value, and I follow the latter interpretation in the empirical analysis below. $c_{ij}$ is a local amenity cost. While areas are ex ante identical, I allow for heterogeneous preferences over locations. Each worker $i$ is assigned a "home area" $j$ with $c_{ij}$ normalized to zero. And for the same worker $i$, the remaining $c_{ij}$ matches are strictly positive draws of i.i.d random variables:

$$c_i(\varepsilon^c) = \sigma_i^c \varepsilon^c \tag{2}$$

where $\varepsilon^c \sim F^c$. The $\sigma_i^c$ parameter determines the strength of preferences over local amenities and may vary by individual $i$.

In this model, the $c_{ij}$ draws are the sole origin of costs associated with migration. This approach has precedent in other work; see e.g. Hilber and Robert-Nicoud (2010); Moretti (2011); Gyourko, Mayer and Sinai (2013). Kennan and Walker (2011) offer evidence that heterogeneous amenity valuations play a fundamental role in determining migration decisions. An alternative approach is to impose one-off moving costs, and I offer an extension in this direction below.

Both employed and unemployed workers receive job offers at rate $\mu$. It is possible to allow this parameter to vary by employment status, though this does not affect the results of interest. A fraction $\pi$ of offers originate from outside a worker's home area, irrespective of where they are currently living. For example, workers may be conducting a more intensive search in their most preferred location. In the exposition that follows, I take the parameter $\pi$ as given - but as I point out below, there are good reasons to believe it may vary with skill.

An individual $i$ draws wage offers equal to:

$$w_i(\varepsilon^w) = \gamma' X_i + \sigma_i^w \varepsilon^w \tag{3}$$

where $X_i$ is a vector of characteristics defining individual $i$'s human capital, and $\varepsilon^w$ is an idiosyncratic term representing the quality of the worker-firm match. The importance of match quality depends on $\sigma_i^w$, which varies across individuals $i$. To the extent that skills are specialized, $\sigma_i^w$ will be larger for skilled workers. Alternatively, in the log wage specification, $\sigma_i^w$ may define the importance of productive complementarities between a unidimensional index of firm quality - as represented by the $\varepsilon^w$ draw - and individual human capital, $\gamma' X_i$.

The $\varepsilon^w$ parameter is drawn from an exogenous distribution $F^w$. Both $F^w$ and its density function $f^w$ are continuous and differentiable over its support, and I assume its hazard rate $\frac{f^w(\varepsilon^w)}{1-F^w(\varepsilon^w)}$ is monotonically increasing in $\varepsilon^w$.

Critically, the match distribution $F^w$ is invariant across geography. This assumption is motivated by the evidence in Section 3, which suggests local differentials in (ex ante) utility are unimportant for explaining the skill mobility gap. Of course, there is substantial local variation in wages in practice, but it is generally thought that this is offset by corresponding variation in housing costs in spatial equilibrium (see e.g. Glaeser and Gottlieb, 2009; Moretti, 2011). Instead, I restrict attention to skill differences in the distribution of wage offers (specifically, in the parameter $\sigma_i^w$) which are entirely independent of geography.

A worker can exit a job in two ways: either if he receives a better offer, or through a random job separation (to unemployment). These separations arrive at rate $\delta$. On separation, workers optimally choose to return to their home area (and receive $c_{ij} = 0$) if they happen to have been employed elsewhere.

Migration in this model is manifested in two ways: first, through "non-home" job finding, where workers accept job offers outside their home area, despite the amenity cost; or second, if workers choose to return to their home area at a later time, motivated (either partially or entirely) by the associated amenity gains.[11] Figure 2 suggests it is the former component which drives the skill mobility gap. In what follows, I study how the rate of "non-home" job finding (i.e. matching with jobs outside a worker's home area) responds to the dispersion of match productivity $\sigma_i^w$ and amenity preferences $\sigma_i^c$.

## 4.2 Worker's value

It is useful to define $\varepsilon$ as the aggregate match component of utility, covering both the productivity $\varepsilon^w$ and amenity $\varepsilon^c$ dimensions. I normalize $\varepsilon$ by wage dispersion $\sigma^w$ to express it in units of $\varepsilon^w$:

$$\varepsilon \equiv \varepsilon^w + \frac{\sigma_i^c}{\sigma_i^w}\varepsilon^c \tag{4}$$

so the flow utility $v_i$ is simply equal to $\gamma' X_i + \sigma_i^w \varepsilon$. The $\varepsilon$ parameter summarizes a worker's position on the jobs ladder, and all individual choices depend on this state variable alone.

Conditional on human capital $X_i$, employed workers accept any job offer yielding a larger $\varepsilon$

[11]Kennan and Walker (2011) show empirically that return migration accounts for a substantial portion of migratory decisions.

than their current level. For an unemployed worker $i$, I denote the reservation match quality as $\varepsilon_{Ri}$. Unemployed workers optimally choose to live in their home area, so they face no amenity penalty. Therefore, they accept any job offer whose match quality satisfies: $\gamma' X_i + \sigma_i^w \varepsilon \geq b_i$, where $b_i$ is out-of-work income. So:

$$\varepsilon_{Ri} = \frac{b_i - \gamma' X_i}{\sigma_i^w} \tag{5}$$

Since $\varepsilon_{Ri}$ is the lowest viable match quality, it defines the bottom of the jobs ladder. Given a match quality of $\varepsilon$ and human capital $X_i$, a worker $i$'s value can then be expressed as:

$$\begin{aligned} rV_i(\varepsilon) \quad = \quad & \gamma' X_i + \sigma_i^w \varepsilon + \delta\left[V_i(\varepsilon_{Ri}) - V(\varepsilon)\right] + (1-\pi)\mu \int_{\varepsilon}^{\infty} \left[V_i(\varepsilon^w) - V_i(\varepsilon)\right] dF^w \qquad (6) \\ & + \pi\mu \int_0^{\infty} \left\{ \int_{\varepsilon + \frac{\sigma_i^c}{\sigma_i^w}\varepsilon^c}^{\infty} \left[ V_i\left(\varepsilon_i^w - \frac{\sigma_i^c}{\sigma_i^w}\varepsilon^c\right) - V_i(\varepsilon) \right] dF^w \right\} dF^c \end{aligned}$$

where $r$ is the interest rate. The first term, $\gamma' X_i + \sigma_i^w \varepsilon$, is the flow utility. The second term accounts for the possibility of job separation, which randomly occurs at rate $\delta$. Notice the unemployment value is equal to $V_i(\varepsilon_{Ri})$, so $V_i(\varepsilon_{Ri}) - V_i(\varepsilon)$ is the associated capital loss. The final two terms describe the value of home area search (beginning $1-\pi$) and non-home search (beginning $\pi$) respectively. Workers accept any home area offer yielding $\varepsilon^w$ (distributed $F^w$) exceeding $\varepsilon$, where $\varepsilon$ is the initial match quality. For non-home offers, the reservation draw of $\varepsilon^w$ is equal to $\varepsilon + \frac{\sigma_i^c}{\sigma_i^w}\varepsilon^c$, where $\varepsilon$ is the worker's initial match quality and $\varepsilon^c \sim F^c$ is the amenity draw.

To ease the notation, I suppress the subscript $i$ from here on - until I define the empirical specification. It suffices to keep in mind that the model is defined for an individual of given human capital $X_i$ and a given set of parameters $\sigma_i^w$, $\sigma_i^c$ and $b_i$.

## 4.3 Job flows

Let $\rho(\varepsilon)$ be the job finding rate for employed workers initially on match quality $\varepsilon$:

$$\rho(\varepsilon) = \rho_H(\varepsilon) + \rho_N(\varepsilon) \tag{7}$$

where

$$\rho_H(\varepsilon) = (1-\pi)\mu\left[1 - F^w(\varepsilon)\right] \tag{8}$$

is the home area finding rate, and

$$\rho_N(\varepsilon) = \pi\mu \int_0^{\infty} \left[1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)\right] dF^c \tag{9}$$

is the non-home rate. Notice the job finding rate is $\rho(\varepsilon_R)$ for the unemployed; so the steady-state unemployment rate is:

$$u = \frac{\delta}{\delta + \rho(\varepsilon_R)} \tag{10}$$

Next, following the method of Burdett and Mortensen (1998), I derive the equilibrium distribution of match quality $\varepsilon$ among employed workers, which I denote by $G$. Consider the set of employed workers receiving match quality below $\varepsilon$. The inflow of workers to this set must equal the outflow in equilibrium:

$$u[\rho(\varepsilon_R) - \rho(\varepsilon)] = (1-u)G(\varepsilon)[\delta + \rho(\varepsilon)] \tag{11}$$

The inflow is composed entirely of the unemployed, who enter jobs yielding match quality below $\varepsilon$ at rate $[\rho(\varepsilon_R) - \rho(\varepsilon)]$. The outflow is composed of employed workers on match quality below $\varepsilon$ who (i) are separated to unemployment (at rate $\delta$) and (ii) find jobs yielding utility exceeding $\varepsilon$. Substituting (10) for $u$ gives:

$$G(\varepsilon) = \frac{\delta}{\delta + \rho(\varepsilon)} \cdot \frac{\rho(\varepsilon_R) - \rho(\varepsilon)}{\rho(\varepsilon_R)} \tag{12}$$

This equation demonstrates the importance of "thin" markets to my hypothesis. Given the specification of $\rho(\varepsilon)$ above, $G(\varepsilon)$ converges (at all $\varepsilon$) to zero as the offer rate $\mu$ becomes large relative to the separation rate $\delta$. Thus, in a world without search frictions, all workers will benefit from the maximum match quality - and there will be no wage rents to justify geographical mobility.

Of course, the distribution $G(\varepsilon)$ accounts for employed workers only. Notice that the unemployed behave identically to workers with match quality $\varepsilon_R$. And in this vein, it is useful to define a distribution function $\hat{G}(\varepsilon)$: the fraction of all workers (irrespective of employment status) who receive *effective* match quality below $\varepsilon$. Specifically:

$$\hat{G}(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < \varepsilon_R \\ u + (1-u)G(\varepsilon) = \frac{\delta}{\delta+\rho(\varepsilon)} & \text{if } \varepsilon \geq \varepsilon_R \end{cases} \tag{13}$$

with probability density given by:

$$\hat{g}(\varepsilon) = \begin{cases} 0 & \text{if } \varepsilon < \varepsilon_R \\ \frac{\delta}{\delta+\rho(\varepsilon_R)} & \text{if } \varepsilon = \varepsilon_R \\ -\frac{\delta\rho'(\varepsilon)}{[\delta+\rho(\varepsilon)]^2} & \text{if } \varepsilon \geq \varepsilon_R \end{cases} \tag{14}$$

where the unemployed are treated as receiving $\varepsilon_R$. This is effectively a left-censored distribution, with a discrete probability mass (corresponding to the unemployed) at the censored value of $\varepsilon_R$.

To simplify the analysis, I study outcomes as $\sigma^c$ becomes large relative to $\sigma^w$: that is, workers place an increasing value on the amenity match. Intuitively, this reflects the low rate of migration in the data. So the overall rate of job finding can be approximated as the home area finding rate:

$$\lim_{\frac{\sigma^c}{\sigma^w}\to\infty} \rho(\varepsilon) = \rho_H(\varepsilon) = (1-\pi)\,\mu\,[1-F^w(\varepsilon)] \tag{15}$$

In the limit, all workers will live in their home area in equilibrium - so the distribution of match quality $\varepsilon$ (among the employed) collapses onto the distribution of workers' productivity matches $\varepsilon^w$.

## 4.4 Extension with one-off migration costs

In this model, I have characterized the cost of migration in terms of amenity penalties. An alternative approach is to use one-off migration costs, and I show here how these might be modelled. This does not affect the basic interpretation of the model, though it does offer some additional insights.

Suppose that on drawing a non-home job offer, workers also draw a cost of moving $m \geq 0$ (rather than an amenity match draw $\varepsilon^c$) from some distribution $M$. This moving cost is payable on acceptance of a job offer. I assume there are no amenity costs, so match quality $\varepsilon$ is simply equal to the productivity match $\varepsilon^w$; and offers are distributed according to: $\varepsilon \sim F^w$. Conditional on human capital $X$, the value of being employed at match quality $\varepsilon$ is now:

$$\begin{aligned} rV(\varepsilon) \;=\;& \gamma X + \sigma^w\varepsilon + \delta\,[V(\varepsilon_R) - V(\varepsilon)] + (1-\pi)\,\mu\int_{\varepsilon}^{\infty}[V(\varepsilon^w) - V(\varepsilon)]\,dF^w \\ & + \pi\mu\int_0^{\infty}\left[\int_{\varepsilon}^{\infty}\max\{V(\varepsilon^w) - V(\varepsilon) - m, 0\}\,dF^w\right]dM \end{aligned} \tag{16}$$

Workers currently on match quality $\varepsilon$ accept a non-home offer $\varepsilon^w$ if $V(\varepsilon^w) - V(\varepsilon) \geq m$. Or equivalently, using (16), the offer is accepted if:

$$\int_{\varepsilon}^{\varepsilon^w} V'(x)\,dx = \int_{\varepsilon}^{\varepsilon^w} \frac{\sigma^w}{r+\delta+\rho(x)}\,dx \geq m \tag{17}$$

And taking a first order approximation around the initial match quality $\varepsilon$, this can be simplified to:

$$\sigma^w(\varepsilon^w - \varepsilon) \gtrapprox [r+\delta+\rho(\varepsilon)]\,m \tag{18}$$

So workers accept a non-home offer if the utility gain exceeds the "flow-equivalent migration cost", where the one-off cost $m$ is scaled by the interest rate $r$, separation rate $\delta$ and job finding rate $\rho(\varepsilon)$. Intuitively, workers are less likely to accept a non-home offer (and pay the one-off cost $m$) if the new job is unlikely to last long (large $\delta$) or they are likely to find a local job soon (large $\rho(\varepsilon)$).

## 4.5 Comparative statics

### 4.5.1 Impact of wage dispersion $\sigma^w$ on non-home job finding

To the extent that skills are specialized, it is intuitive that better educated workers should face larger $\sigma^w$. In this section, I study the response of non-home job finding to changes in $\sigma^w$ relative to amenity preferences $\sigma^c$. Let $\rho_N$ be the average rate of non-home job finding across all workers. This can be summarized by integrating over the distribution $\hat{G}$ of (effective) match quality:

$$\rho_N = \int_{\varepsilon_R}^{\infty} \rho_N(\varepsilon)\,\hat{g}(\varepsilon)\,d\varepsilon \tag{19}$$

As $\frac{\sigma^c}{\sigma^w}$ becomes large (following the logic of (15)), the non-home finding rate at match quality $\varepsilon$ can usefully be written as:

$$\rho_N(\varepsilon) = \frac{\pi}{1-\pi}\rho(\varepsilon)\,\Omega(\varepsilon) \tag{20}$$

which depends on (i) the relative non-home contact rate $\frac{\pi}{1-\pi}$, (ii) the overall job finding rate $\rho(\varepsilon)$, and (iii) a term characterizing the contribution of job rents:

$$\Omega(\varepsilon) = \int_0^{\infty} \left[\frac{1-F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1-F^w(\varepsilon)}\right] dF^c(\varepsilon^c) \tag{21}$$

which is the probability that a match quality draw $\varepsilon^w$ exceeds $\varepsilon+\frac{\sigma^c}{\sigma^w}\varepsilon^c$, conditional on it exceeding $\varepsilon$, averaged over the distribution of amenity draws $\varepsilon^c$. Alternatively, for a worker with initial match quality $\varepsilon$, this is the fraction of acceptable home area job offers which yield improvements in match quality of at least $\frac{\sigma^c}{\sigma^w}\varepsilon^c$. Based on (3), this is equivalent to wage rents exceeding $\sigma^c\varepsilon^c$.

Equation (20) can be interpreted using a jobs ladder analogy: $\rho(\varepsilon)$ depends on the number of remaining "rungs" above match quality $\varepsilon$, and $\Omega(\varepsilon)$ describes the "slope" of the ladder at these rungs. A steeper slope means that given wage rents of $\sigma^c\varepsilon^c$ can be extracted by climbing a smaller number of rungs. Critically, these considerations are independent of geography: all that matters is the characteristics of the jobs ladder.

What is the impact of an increase in wage dispersion, $\sigma^w$, on $\rho(\varepsilon)$ and $\Omega(\varepsilon)$? Conditional on initial match quality $\varepsilon$, there is no effect on $\rho(\varepsilon)$: see equation (8). Intuitively, this is because home area job finding is costless, so strictly positive rents are not necessary for acceptance of a job offer. Using the ladder analogy, only the number of remaining rungs matters - and not the steepness of the ladder.

In contrast, the job rents term $\Omega(\varepsilon)$ is increasing in the ratio of $\sigma^w$ to $\sigma^c$ - so match dispersion does matter for non-home job finding. Intuitively, workers are more likely to move if they care more about their productivity match than their amenity match. Notice the effect of $\sigma^w$ on $\Omega(\varepsilon)$ is increasing in the magnitude of amenity costs, $c = \sigma^c \varepsilon^c$. That is, wage rents only matter to the extent that moving is costly. To summarise:

**Proposition 1.** *Given a worker's initial match quality $\varepsilon$, $\Omega(\varepsilon)$ and therefore $\rho_N(\varepsilon)$ are increasing in the ratio of $\sigma^w$ to $\sigma^c$. The effect of $\sigma^w$ is larger if workers have stronger preferences over amenities ($\sigma^c$ larger).*

Before moving on, it is worth briefly considering the significance of $\pi$, the non-home share of offers. There is reason to believe this may be larger in skilled markets. To the extent that workers - and also firms (in a more complete model) - expect larger rents in employment relationships, they may invest harder in non-home search and recruitment because a greater fraction of non-home offers are viable. This would amplify any impact of wage rents on geographical mobility. However, I leave the analysis of these effects to future research - and take $\pi$ as exogenous. It is worth emphasizing that, conditional on certain distributional assumptions, observed wage rents alone can quantitatively account for the skill differences in geographical mobility (without relying on $\pi$): see Section 5 below.

The model also has important implications for job rents and geographical mobility over the lifecycle, and I consider these next. And I conclude this section with a discussion of the effect of $\sigma^w$ on $\hat{G}$, the distribution of $\varepsilon$ across workers: as (20) shows, this distribution matters for the overall rate of non-home job finding, $\rho_N$. This is because access to job rents depends on workers' positions $\varepsilon$ on the jobs ladder - as I show below.

### 4.5.2 Lifecycle effects

As Figure 1 makes clear, the lifecycle plays an important role in the skill mobility gap. Though I have not explicitly incorporated the lifecycle in the model above, the model has clear implications for these questions. For example, suppose workers live for a fixed time period $T$. And suppose also that the labor market is "thin"; that is, the ratio of the separation rate $\delta$ to the offer rate $\mu$ exceeds zero. One would then expect that (among those in work) older workers should typically benefit from larger match utility $\varepsilon$ - merely because they have had more time to find the ideal match. That is, they have accumulated more "search capital". Indeed, Manning (2003) and Gottfries and Teulings (2016) offer evidence that this accumulated search capital can explain a substantial portion of the returns to labor market experience. In this way, an individual's age can be proxied by his match quality $\varepsilon$. And I show below that:

**Proposition 2.** *(i) $\Omega(\varepsilon)$ is decreasing in initial match quality $\varepsilon$, and (ii) the (positive) effect of $\sigma^w$ on $\Omega(\varepsilon)$ is decreasing in $\varepsilon$. The latter result is contingent on the distribution $F^c$ of amenity draws $\varepsilon^c$. A sufficient criterion is that the elasticity of the density $\varepsilon^c \frac{f^{c'}(\varepsilon^c)}{f^c(\varepsilon^c)}$ exceeds -1 for all $\varepsilon^c$.*

The first result is simple to show. Taking the derivative of the job rents term $\Omega(\varepsilon)$ with respect to $\varepsilon$:

$$\Omega'(\varepsilon) = \int_0^\infty \left[ \frac{f^w(\varepsilon)}{1 - F^w(\varepsilon)} - \frac{f^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)} \right] \frac{1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1 - F^w(\varepsilon)} dF^c(\varepsilon^c) \tag{22}$$

And given my assumption that $F^w$ has a monotonically increasing hazard rate, $\Omega'(\varepsilon)$ must be negative. Intuitively, the monotone hazard rate ensures the upper tail of the productivity distribution is not too thick, so workers will expect lower job rents in subsequent matches as they move higher up the ladder.

To show the second result, I differentiate $\Omega(\varepsilon)$ with respect to $\sigma^w$. Conditional on the worker's initial match quality $\varepsilon$:

$$\frac{d\Omega(\varepsilon)}{d\sigma^w} = \frac{\sigma^c}{(\sigma^w)^2} \int_0^\infty \varepsilon^c f^c(\varepsilon^c) \left[ \frac{f^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1 - F^w(\varepsilon)} \right] d\varepsilon^c \tag{23}$$

Notice the term in square brackets is the density of the match productivity draw at $\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c$, conditional on the draw exceeding $\varepsilon$. The hazard rate of this distribution at $\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c$ is $\frac{f^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}$. Given my assumption that $F^w$ has a monotonically increasing hazard rate, a smaller $\varepsilon$ causes the hazard rate to decrease at every $\frac{\sigma^c}{\sigma^w}\varepsilon^c$. That is, there is a dominating transformation of the conditional distribution - by the hazard rate criterion. So in the integral in (23), relatively more density is concentrated at larger values of $\varepsilon^c f^c(\varepsilon^c)$.

The effect of $\varepsilon$ on equation (23) then depends on the shape of $F^c$. A sufficient condition for the response to $\sigma^w$ to be decreasing in $\varepsilon$ is that $\varepsilon^c f^c(\varepsilon^c)$ is unambiguously increasing in $\varepsilon^c$ - or equivalently, the elasticity of the density $\varepsilon^c \frac{f^{c'}(\varepsilon^c)}{f^c(\varepsilon^c)}$ exceeds -1 for all $\varepsilon^c$. This ensures that amenity cost draws are not heavily concentrated at the bottom of the distribution. For example, it is sufficient that $F^c$ is uniform. A uniform assumption on $F^c$ is in fact a useful one for deriving a tractable empirical specification - and I return to this point below.

This result is simple to interpret: the job rents on offer are larger at lower rungs of the jobs ladder; and workers initially at lower match quality $\varepsilon$ benefit more from an expansion of rents.

This offers a simple intuition for why the skill mobility gap is largely driven by the young: as the beginning of their career, workers will have greater scope for accumulating wage rents; and these rents will justify more long-distance moves.

Of course, this argument depends on the claim that match quality $\varepsilon$ can proxy for a worker's age - conditional on human capital. But this claim is empirically testable: in particular, once I control for an individual's initial wage (to absorb match quality), any dependence of job rents on age (in any skill group) should be washed away. I take this prediction to the data in Section 5 below.

#### 4.5.3 Impact of wage dispersion $\sigma^w$ on $\hat{G}(\varepsilon)$, distribution of (effective) match quality across workers

Until now, I have studied the effect of $\sigma^w$ conditional on a worker's initial match quality $\varepsilon$. But of course, the distribution $\hat{G}(\varepsilon)$ across workers is liable to shift following a change in $\sigma^w$. All else equal, if workers are located at higher rungs of the jobs ladder (larger $\varepsilon$), job finding rates $\rho(\varepsilon)$ will be smaller *on average* - across the distribution of $\varepsilon$. And the same will be true of wage rents and non-home job finding $\rho_N(\varepsilon)$: see Proposition 2.

In practice, average job finding rates (across the distribution of $\varepsilon$) vary little by education: see Figure 3. This indicates that these distributional effects are unlikely to be important. But in any case, it is worth briefly considering the theoretical mechanisms at play.

Based on (13), conditional on the exogenous parameters $\delta$ and $\mu$, the distribution $\hat{G}(\varepsilon)$ of (effective) match quality is fully determined by the reservation match quality $\varepsilon_R$. Notice that $\varepsilon_R$ can be interpreted as the censoring value of a left-censored distribution. Therefore, a larger reservation $\varepsilon_R$ causes $\hat{G}(\varepsilon)$ to decline for given $\varepsilon$ in the neighborhood of $\varepsilon_R$. Intuitively, if workers are more demanding, they will be located at higher $\varepsilon$ in equilibrium. So it suffices to consider how the reservation match quality $\varepsilon_R$ responds to $\sigma^w$. Based on (5), this is theoretically ambiguous. [12] - so the same must be true of the effect on the match quality distribution $\hat{G}$. To summarise:

**Proposition 3.** *The effect of $\sigma^w$ on $\hat{G}(\varepsilon)$, the distribution of (effective) match quality across workers, is theoretically ambiguous.*

In practice (in the data), the finding rate $\rho(\varepsilon_R)$ among jobless workers is increasing in education - which suggests a smaller reservation $\varepsilon_R$. The effect is small among the unemployed (see, for example, Mincer, 1991) but more substantial if the economically inactive[13] are included. See Appendix E for empirical estimates of $\rho(\varepsilon_R)$. All else equal, this should imply

[12] It depends on whether $\varepsilon_R$ is positive or negative - or intuitively, on whether larger match dispersion $\sigma^w$ concentrates more of the density above or below the reservation threshold. Of course, this will partly depend on how out-of-work income $b$ varies with human capital $X$.

[13] These workers may be modelled as having larger $b$, and so, larger reservation $\varepsilon_R$.

skilled workers typically have lower match quality (with positive consequences for geographical mobility) in equilibrium.

However, all else is not equal. In particular, skilled workers also face a lower separation rate $\delta$ (see, again, Mincer, 1991, and Appendix E for estimates) which, based on (13), makes $\hat{G}(\varepsilon)$ smaller for all $\varepsilon$. Intuitively, workers have more time to rise up the ladder before they fall to the bottom (through a separation). Ultimately then, the effect of skill on the distribution of match quality is an empirical question. The fact that the overall flow of new jobs varies little by education (Figure 3) suggests these various effects offset each other in practice.

# 5 Evidence on wage rents

## 5.1 Empirical specification

In this section, I offer estimates of the wage returns to new job matches. In particular, I study how these wage rents - as characterized by $\Omega(\varepsilon)$ in (21) - vary with education, conditional on accepting a job offer. It turns out that $\Omega(\varepsilon)$ can be expressed in terms of the expected wage returns, under certain assumptions on the amenity cost distribution $F^c$.

In particular, suppose the amenity draws $\varepsilon^c$ are uniformly distributed with with a minimum at 0 and a maximum normalized to 1. And suppose also that very few job offers are accepted at the maximum amenity cost draw: that is, if $\frac{F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\right)}{F^w(\varepsilon)}$ is close to 1 for all values of $\varepsilon$. Under these assumptions, using integration by parts, I show in Appendix G that $\Omega(\varepsilon)$ can be approximated by:

$$\begin{aligned} \Omega(\varepsilon) &\approx \frac{\sigma^w}{\sigma^c}\int_0^\infty (x-\varepsilon)\frac{f^w(\varepsilon+x)}{1-F^w(\varepsilon)}dx \\ &= \frac{\sigma^w}{\sigma^c}\mathbb{E}\left[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon\right] \end{aligned} \tag{24}$$

Notice that $\sigma^w\mathbb{E}[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon]$ is equal to the expected wage return to a new job match. This expected return is simple to identify in the data. The overall change in wages for an individual can be disaggregated into the return to a new job and a contribution from human capital. Using (3):

$$\mathbb{E}\left[w_i'-w_i(\varepsilon)|w_i'\geq w_i(\varepsilon)\right]=\sigma^w\mathbb{E}\left[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon\right]+\gamma'\left(X_i'-X_i\right) \tag{25}$$

This suggests the following empirical specification:

$$\Delta w_{it}=\beta_0+\beta_1 NewJob_{it}+\beta_2' X_{it}+d_i+d_t+\varepsilon_{it} \tag{26}$$

where $\Delta w_{it}$ is the change in worker $i$'s log wage between $t-1$ and $t$, and $NewJob_{it}$ is a dummy taking 1 if the worker began a new job between $t-1$ and $t$. To proxy for human capital, I control for a vector $X_{it}$ of demographic characteristics[14] and, in some specifications, individual fixed effects $d_i$ - to absorb unobserved time-invariant components of human capital. The $d_t$ variable is a time fixed effect. Conditional on the human capital controls, the $\beta_1$ coefficient identifies the expected wage return to a job match - by comparing the wage evolution through a job transition against the counterfactual of remaining in the same job.

It is important to stress that 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 $\beta_1$ identifies. Of course, this conditional mean is driven by selection on job offers - but it is precisely this selection which interests me. In particular, according to Proposition 1, skilled workers should expect larger wage rents conditional on changing job - to the extent that skills are specialized (larger $\sigma^w$). And Proposition 2 predicts that the effect of education should be especially large for younger workers (since they are 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 age groups.

It should be noted that there is already a literature which estimates wage returns to job transitions - using similar specifications to (26). These wage returns are known to be increasing in education (Bartel and Borjas, 1981; Mincer, 1986) and decreasing in age (see also Topel and Ward, 1992 and Chapter 6 of Manning, 2003). In his analysis of lifecycle earnings, Manning (2003) interprets these wage changes in the context of a jobs ladder; and Gottfries and Teulings (2016) consider skill differences within this framework. But to my knowledge, this study is the first to link these effects to geographical mobility.

Before moving to the estimation, it is also worth briefly discussing the choice of a log wage specification. The model above can be interpreted in terms of either log or linear utility: see equation (1). These yield specifications in log and dollar wage changes respectively.[15] It should be emphasized that the log specification is more conservative: any skill gradient in *proportional* rents will be shallower than in *dollar* rents, because better educated workers earn substantially more.

[14] Specifically: age and age squared; four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in age and a time trend; black and Hispanic race dummies and immigrant status; and a gender indicator which is also interacted with all previously mentioned variables.

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

## 5.2 Data

I estimate this specification using the Survey of Income and Program Participation (SIPP). The SIPP offers substantial samples and high-frequency waves, just four months apart. Job status is recorded at the end of each 4-month wave.[16] My sample consists of employees aged 25 to 64 in the SIPP panels beginning 1996, 2001, 2004 and 2008, covering the period between 1996 and 2013. I identify $w_{it}$ with log hourly wages at the end of each 4-month wave $t$.[17]

Of course, the sample is necessarily restricted to individuals who were in employment at the end of wave $t-1$. Since the reservation wage of unemployed workers is unobserved, I do not observe the job rents accruing to their matches. However, this is unlikely to be an important omission in the context of interpreting skill differences in migration rates. Table 2 confirms that migration rates are larger for the unemployed and economically inactive (as the model predicts: the jobless should expect larger rents). But, migration rates among the initially employed are still quite similar to migration rates overall - across education groups.

[Table 2 here]

## 5.3 Empirical estimates

I present estimates of (26) in Table 3. Column 1 reports the basic regression, without controlling for fixed effects: I estimate the expected wage return to changing job as 0.03. In the next four columns, 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 0.02 for high school dropouts (the omitted category) to 0.06 for those with postgraduate qualifications. As columns 3 to 5 show, this education gradient is largely driven by younger workers: among the under-35s, $\beta_1$ ranges from 0.01 (and statistically insignificant) for dropouts to 0.13 for postgraduates. In the remaining five columns, I repeat this exercise controlling for fixed effects: the results are very similar.

[Table 3 here]

[16] Respondents to the SIPP do report their job status at the end of each month (in the four months since the previous wave), but I do not exploit this variation. This monthly data is likely to be subject to large measurement error due to poor recall, as information is only collected at the end of each 4-month wave. In particular, it is known that the SIPP suffers from severe seam bias (see e.g. Marquis and Moore, 2010): monthly changes in individuals' outcomes (whether job status or wages) tend to be larger between months at the seam of two waves than between months within the same wave.

[17] I use hourly wage data for workers paid by the hour, and I impute hourly wages for salaried workers using monthly earnings and hours. To guard against measurement error, I restrict my sample to wage changes where pay duration does not change: that is, the worker is either paid by the hour in both periods or salaried in both periods. I also restrict attention to wage observations between \$2 and \$100 in 2000 prices. And finally, I exclude workers with multiple jobs or business income at the end of a wave.

This suggests that job rents are steeply increasingly in skill, and especially so for younger workers. This is consistent with the prediction of Proposition 2: younger workers are lower down the jobs ladder, so they are the main beneficiaries of the larger rents on offer in skilled markets. This mechanism can be tested directly in the data. Specifically, if age matters only in as much as it affects a worker's rung on the ladder (which can be identified by the initial wage), the effect of skill on rents should be invariant with age if I control for a worker's initial wage. To this end, I re-estimate all the specifications in Table 3 - but this time, controlling also for (i) the worker's log wage in the previous period $w_{it-1}$ and (ii) an interaction between $w_{it-1}$ and the $NewJob_{it}$ dummy.

[Table 4 here]

I report the results in Table 4. It is difficult to interpret the coefficients on $w_{it-1}$ and its interaction. Measurement error and regression to the mean will be a concern, and $w_{it-1}$ may also be picking up unobserved components of human capital. Still, the coefficient on the interaction between $w_{it-1}$ and $NewJob_{it}$ is negative, which is consistent with the first part of Proposition 2: workers can expect larger returns to job matches if they were initially lower down the jobs ladder. The elasticity of the wage returns to the initial wage hovers between -0.15 and -0.3 across the different specifications.

More interesting is what happens to the interactions between $NewJob_{it}$ and the education effects. Notice first that the coefficients on these interactions are much larger than in Table 3: this is because the better educated typically earn more, but those individuals on higher wages expect lower job rents. The coefficients are also very precise, monotonically increasing in education (even for qualifications below college degree) and, most importantly, remarkably similar across age groups. For example, without fixed effects, the effect of a postgraduate qualification on mean wage rents is about 30 log points (relative to high school dropouts) across all age groups, controlling for initial wage. The effect is somewhat smaller (around 0.2) when I control for fixed effects in columns 6-10; but again, it varies little by age. This is strong evidence in favour of the jobs ladder explanation: the large age differences in the education effect on rents in Table 3 are entirely explained by variation in workers' initial wage. And this offers a plausible explanation for why the skill mobility gap is so much larger for younger workers.

## 5.4 Quantifying the effect of job rents on mobility

Finally, it is useful to briefly consider whether these estimates of skilled wage rents are sufficiently large to plausibly explain the observed mobility gap. Assume again that the amenity draws $\varepsilon^c$ are uniformly distributed between 0 and a normalized value of 1; and assume also that very few job offers are accepted at the maximum amenity cost draw. Then, substituting (24) for $\Omega(\varepsilon)$ in (20) gives:

$$\rho_N(\varepsilon) \approx \frac{\pi}{1-\pi}\frac{\sigma^w}{\sigma^c}\mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' \geq \varepsilon\right]\rho(\varepsilon) \tag{27}$$

And integrating (27) over the match quality (distributed $G$) of employed job finders:

$$\frac{\int_\varepsilon \rho_N(\varepsilon)\, dG(\varepsilon)}{\int_\varepsilon \rho(\varepsilon)\, dG(\varepsilon)} = \frac{\pi}{1-\pi}\frac{\sigma^w}{\sigma^c}\frac{\int_\varepsilon \mathbb{E}[\varepsilon' - \varepsilon|\varepsilon' \geq \varepsilon]\,\rho(\varepsilon)\, dG(\varepsilon)}{\int_\varepsilon \rho(\varepsilon)\, dG(\varepsilon)} = \frac{\pi}{1-\pi}\frac{1}{\sigma^c}\beta_1 \tag{28}$$

where $\frac{\sigma^w \int_\varepsilon \mathbb{E}[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon]\rho(\varepsilon)dG(\varepsilon)}{\int_\varepsilon \rho(\varepsilon)dG(\varepsilon)}$ is the expected wage rents across all employed workers (weighted by individual job finding rates $\rho(\varepsilon)$) - which is identified by $\beta_1$ in equation (26).

[Table 5 here]

In Table 5, I set out estimates of $\int_\varepsilon \rho(\varepsilon)\, dG(\varepsilon)$, $\int_\varepsilon \rho_N(\varepsilon)\, dG(\varepsilon)$ and $\beta_1$ for each education group; and based on these, I impute values for $\frac{1-\pi}{\pi}\sigma^c$ - which represents a broad measure of the "costs" inhibiting non-home job finding, due to both the non-home share of job offers $\pi$ and preferences over amenities $\sigma^c$. See the table notes for further details. $\int_\varepsilon \rho(\varepsilon)\, dG(\varepsilon)$ is the overall flow of new jobs among those who were initially employed: that is, the job-to-job finding rate. I identify this using the SIPP data (row 1). $\int_\varepsilon \rho_N(\varepsilon)\, dG(\varepsilon)$ is the average rate of non-home job finding among the initially employed. As I argue above, the non-home finding rate corresponds specifically to those workers leaving their home area (despite a positive amenity cost) for the sake of a job. I identify this with the *job-motivated* annual cross-state migration rate in the CPS illustrated in Figure 2 above (row 2). Since migration rates are similar among the initially employed and the full sample (see Table 2), I approximate $\int_\varepsilon \rho_N(\varepsilon)\, dG(\varepsilon)$ using the average migration rate across all individuals.

The third row of Table 5 reports the ratio of $\int_\varepsilon \rho_N(\varepsilon)\, dG(\varepsilon)$ to $\int_\varepsilon \rho(\varepsilon)\, dG(\varepsilon)$, and the final row derives values for $\frac{1-\pi}{\pi}\sigma^c$ based on (28). It turns out that these cost estimates vary little by education. In other words, given my (uniform) distributional assumption on amenity costs, the skill differences in expected job rents $\beta_1$ explain the bulk of the variation in job-motivated mobility.

# 6 Selection on amenity cost draws

## 6.1 Theoretical results on realized amenity costs

Until now, I have focused exclusively on selection on productivity draws - and the implications for the expected wage rents across all matches. But the model also offers useful predictions on the wage rent accruing to specifically *non-home* matches. These rents depend on the relative importance of selection on productivity and amenity draws. To the extent that workers select into migration because of large productivity shocks (and despite large amenity draws), this will be manifested in relatively large returns to long-distance matching. This offers a useful empirical test to distinguish between different explanations of high skilled mobility.

To address these selection effects, it is useful to study the distribution of *realized* amenity costs - conditional on accepting a non-home job offer. Let $Z(c|\varepsilon)$ be the distribution of realized amenity costs $c = \sigma^c \varepsilon^c$, given initial match quality $\varepsilon$. Conditional on accepting a non-home offer, the probability of having drawn an amenity cost exceeding $c$ is:

$$1 - Z(c|\varepsilon) = \frac{\int_{\frac{c}{\sigma^c}}^{\infty} \left[1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)\right] dF^c(\varepsilon^c)}{\int_0^{\infty} \left[1 - F^w\left(\varepsilon + \frac{\sigma^c}{\sigma^w}\varepsilon^c\right)\right] dF^c(\varepsilon^c)} \tag{29}$$

I make the following claim:

**Proposition 4.** *Given a worker's initial match quality $\varepsilon$, $1 - Z(c|\varepsilon)$ is increasing in both $\sigma^w$ and $\sigma^c$ for any amenity cost $c$.*

That is, a larger $\sigma^w$ and $\sigma^c$ cause a dominating transformation (by the first order stochastic criterion) of the distribution of realized amenity costs. I leave the proof to Appendix G. Intuitively, a larger $\sigma^w$ raises the size of job rents, so workers are more likely to accept non-home offers with high amenity cost draws $c$. And a larger $\sigma^c$ implies larger unconditional amenity cost draws; so (trivially) realized amenity costs will also be larger.

If these realized amenity costs can be observed, this would offer a useful test to discriminate between the rents and costs explanations of the skill mobility gap. If skilled mobility is driven by low costs (i.e. low $\sigma^c$), realized amenity costs should be *decreasing* in education. But if skilled mobility is driven by large wage rents (large $\sigma^w$), realized amenity costs should be *increasing* in education. Intuitively, in the latter case, skilled workers would be moving *because* of large rents - and *despite* the associated costs.

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

**Proposition 5.** *The expected wage return to non-home job finding identifies an upper bound on the expectation of realized amenity costs. And if amenity cost draws are always positive, the differential between the expected return to non-home and home matches identifies a lower bound on the expected realized costs.*

The intuition for the upper bound is simple. Let $\mathbb{E}_N[\varepsilon' - \varepsilon|\varepsilon]$ represent the expected gain in match quality on accepting a *non-home* (subscript $N$) job offer. The associated wage return can then be expressed as:

$$\sigma^w \mathbb{E}_N \left[\varepsilon' - \varepsilon | \varepsilon\right] = \int_c \mathbb{E}\left[\sigma^w \left(\varepsilon' - \varepsilon\right) | \sigma^w \left(\varepsilon' - \varepsilon\right) \geq c\right] dZ(c|\varepsilon) \tag{30}$$

where $\mathbb{E}[\sigma^w(\varepsilon' - \varepsilon)|\sigma^w(\varepsilon' - \varepsilon) \geq c]$ is the expected wage return across all matches, *conditional* on the return exceeding an amenity cost $c$. To derive the expected return to non-home matches, this is integrated over the distribution of realized amenity costs $Z$, as derived in (29). Since $\mathbb{E}[\sigma^w(\varepsilon' - \varepsilon)|\sigma^w(\varepsilon' - \varepsilon) \geq c] \geq c$, it follows that the expected wage return must exceed the expected amenity costs, $\int_0^\infty c dZ(c|\varepsilon)$, associated with those matches. Intuitively, workers will only accept non-home offers if the associated rents exceed the cost of moving.

The lower bound is identified by the differential between the expected wage return to non-home and home job finding:

$$\begin{aligned} & \sigma^w \mathbb{E}_N \left[\varepsilon' - \varepsilon | \varepsilon\right] - \sigma^w \mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon\right] && (31) \\ = \ & \sigma^w \int_0^{\sigma^c} \left\{\mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' - \varepsilon \geq c\right] - \mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' - \varepsilon \geq 0\right]\right\} dZ(c|\varepsilon) \end{aligned}$$

Given my assumption that the match productivity distribution $F^w$ has a monotonically increasing hazard rate, it follows that the term in curly brackets in (31) is less or equal to $c$ for all amenity cost draws $c \geq 0$. See Appendix G for a proof. And if migration costs are positive, it then follows that the non-home/home differential in expected wage returns identifies a lower bound on the expected realized costs, $\int_0^\infty c dZ(c|\varepsilon)$.

The intuition for this result is that the amenity cost is not always binding: there are some local offers which would be sufficient to justify a non-home match. And consequently, the differential in rents should underestimate the magnitude of costs.

## 6.2 Implications for "non-job" migration

Until now, I have restricted my analysis of migration to "non-home" job finding: it is non-home migration which best represents the "job-motivated" moves which drive the skill mobility gap in Figure 2. In this context, amenity cost draws are always positive - since, by definition, workers prefer their home area to anywhere else. However, as Figure 2 shows, a sizable fraction of long-distance moves are motivated by "non-job" reasons (family, housing or other local amenities), which are suggestive of negative costs. Indeed, Kennan and Walker (2011) emphasize the importance of negative costs in explaining many migration decisions. In these cases, in the language of the model, workers are either returning to their home area or perhaps changing their home area.

Selection on amenity costs offers a natural framework for analysing patterns in reasons for moving. As Proposition 4 shows, a low $\sigma^w$ will discourage migration with high associated costs. And this can help explain why lower skilled workers report disproportionately non-job reasons for moving: see Figure 2. Given that non-job reasons (with presumably negative amenity costs) account for a large fraction of low skilled moves, the lower bound result in Proposition 5 may not be robust for that demographic - since that result requires that all cost draws are positive.

## 6.3 Empirical evidence

To test the prediction of Proposition 4, I next offer estimates of the wage returns to cross-state matching. I use the following empirical specification:

$$\Delta \log w_{it} = \beta_0 + \beta_1 NewJob_{it} + \beta_2 NewJob_{it} \cdot Move_{it} + \beta_3 Move_{it} + \beta_4' X_{it} + d_i + d_t + \varepsilon_{it} \quad (32)$$

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 amenity cost of movers can be identified by the coefficient $\beta_2$: this gives the difference in wage returns to non-home and home area job matching. The upper bound can be identified by $\beta_1 + \beta_2$: this is the overall wage return to a non-home match.

To study how $\beta_2$ varies with education, I interact $NewJob_{it} \cdot Move_{it}$ (and the other key variables) with education effects - though in practice, given the specification is more demanding here, I just use a single college graduate dummy (taking 1 for any individual with at least four years in college).

[Table 6 here]

I present estimates of (32) in Table 6. Again, column 1 reports the basic regression with no fixed effects. I estimate $\beta_1$ as 0.02 and $\beta_2$ as 0.07. That is, the average returns to cross-state job finding are 7 percent greater than local job finding. This implies the expected amenity costs of migrants are bounded below by 0.07 and above by 0.09, as a fraction of a worker's initial wage.

It turns out this entire effect is driven by college graduates, and largely by the young among them. In column 2, I interact the variables $NewJob_{it}$, $NewJob_{it} \cdot Move_{it}$ and $Move_{it}$ with a graduate dummy. The coefficient on $NewJob_{it} \cdot Move_{it}$ is now zero, which suggests the average differential between cross-state and local rents is negligible for the low skilled. The implied bounds for low skilled workers' expected realized costs are insignificantly different from zero.

In contrast, the interaction between $NewJob_{it} \cdot Move_{it}$ and the graduate dummy takes a coefficient of 0.16. The implied bounds for skilled workers' expected realized costs, on summing up the basic and interaction coefficients, are 0.14 (= 0.158 - 0.016) and 0.19 (= 0.158 - 0.016 +

0.015 + 0.033). The next three columns show this effect is largely driven by younger workers. These results are very similar when I control for individual fixed effects: see columns 6-10.

Based on Proposition 4, these large amenity costs are indicative of large wage dispersion $\sigma^w$ or strong amenity preferences $\sigma^c$. In other words, skilled workers are moving *because* of large rents and *despite* large costs. This casts heavy doubt on the hypothesis that skilled mobility is driven by low migration costs.

This analysis comparing migrants and stayers builds on previous work by Lkhagvasuren (2014). He compares the wage *levels* of recent cross-state migrants and stayers. Among the college-educated, he shows that recent migrants earn more than stayers; but among the low skilled, the reverse is true. He argues these effects are driven by skill differences in the dispersion of a location-worker productivity match, coupled with a substantial migration cost (so workers only move for a good match).[18] But on studying wage *changes* in longitudinal data, I show these effects are associated with differences in job rents more broadly, irrespective of geography.

## 6.4 Comparing cost estimates to existing literature

There are several studies in the literature which have estimated the cost of migration, and it is worth comparing my estimates to theirs. Previous studies have typically interpreted migration expenses as a one-off cost paid at the point of moving. For the sake of comparability, I convert my amenity cost estimates into one-off cost equivalents using the logic of equation (18).

The estimates suggest the low skilled are typically moving with negligible amenity costs. More interesting is the case of college graduates. The estimated cost bounds are 0.14 and 0.19 (based on column 2 of Table 6), so take a mid-point of 0.165. Average monthly earnings for graduates in my SIPP sample are \$4,032 (in 2000 prices). Taking 16.5 percent of this number yields a monthly amenity cost of about \$665. Equation (18) suggests this cost should be discounted at the sum of the separation rates to both unemployment $\delta$ and new jobs $\rho(\varepsilon)$; the monthly interest rate $r$ is negligible in comparison. I take a value of 0.03 for the monthly separation rate to unemployment and 0.03 for the job-to-job transition rate (see Shimer, 2005*b*), yielding an overall discount rate of 0.06. Dividing \$665 by 0.06 yields an expected one-off migration cost (conditional on moving) of about \$11,000.

How does this compare with existing estimates in the literature? It should be emphasized that earlier estimates do vary substantially partly because they identify different objects. Most studies do not allow for individual heterogeneity in migration costs, which rules out the selec-

[18] To explain the negative effect of recent migration on wages for the low skilled, Lkhagvasuren relies on a (calibrated) positive correlation between ability and migration costs *within* education groups. But I show it is possible to account for skill differences in these results - even if the heterogeneity in migration costs is independent of ability and education. Intuitively, low skilled workers (facing meager job rents) will typically move only if they draw a low (or even negative) migration cost.

tion effects I have described above. Bayer and Juessen (2012) estimate a cross-state migration cost of $34,000, using a dynamic structural model. Lkhagvasuren (2014) calibrates a Roy model and estimates a migration cost of $28,000 to $54,000 between census divisions. And Davies, Greenwood and Li (2001) estimate cross-state migration costs of around $200,000 in a conditional logit framework.

In contrast, Kennan and Walker (2011) allow for large individual heterogeneity in migration costs; and my analysis follows their example. They estimate a much larger average (unconditional) cost of $312,000, though the cost for people who actually move state is typically negative: this is because most moves are motivated by idiosyncratic amenity payoffs, which Kennan and Walker factor into the cost. It should be emphasized that their sample is restricted to high school graduates, who are more likely to move for non-job reasons (see Figure 2). Indeed, my estimates in Table 6 also point to non-positive realized costs (on average) for the low skilled.

# 7 Subjective evidence on migration costs

## 7.1 Imputing amenity costs

In this section, I estimate migration costs more directly using subjective responses to the PSID. I show there is little variation in these costs by education. And reassuringly, the numbers are consistent with the realized costs implied by the estimates in Table 6.

My analysis is based on a unique set of questions on willingness to move for work. In the years 1969-72 and 1979-80, employed respondents to the PSID 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?"[19]

Of course, there may be concerns about the relevance of this data to current trends. But as I show in Appendix F.1, age and education differentials in cross-state mobility in my 1970s PSID sample look very similar to contemporary patterns in Figure 1 above. Also, just like in Figure 2, the skill mobility gap in the PSID sample is entirely driven by individuals who report moving for job reasons. There has been a decline in mobility since the 1980s (Molloy, Smith and Wozniak, 2011), but this effect was fairly uniform across education groups (see Appendix A.4).

In interpreting the responses to these questions, it helps to set out a simple selection model.

[19] Similar questions were also asked of the unemployed. But, there are few unemployed workers in the sample; and since I do not know their reservation wage for a *local* job, it is difficult to impute amenity costs. In any case, I report some results for the unemployed in the footnotes that follow.

Suppose an employed worker is offered a job in another locality. Let:

$$w^R(w_i, c_i) = w_i + c_i \tag{33}$$

be the minimum wage required to tempt a worker $i$ to move, where $w_i$ is the worker's current wage and $c_i$ is the amenity cost. Workers only report being "willing to move" - and disclose their reservation moving wage - if:

$$w^R(w_i, c_i) \leq w_i^{CO} \tag{34}$$

where $w_i^{CO}$ is a cut-off value. Clearly, there is an element of subjectivity in the definition of $w_i^{CO}$.[20] But, one might assume $w_i^{CO}$ approximates the best wage that can be "realistically" attained, so workers with $w^R(w_i, c_i) > w_i^{CO}$ expect only a remote likelihood of moving. For those who satisfy (34) and disclose their reservation, amenity costs $c_i$ can then be imputed as $w^R(w_i, c_i) - w_i$.

## 7.2 Estimates of amenity costs

Throughout, I restrict my sample to household heads as defined in the PSID: these are always male, unless there is no husband (or cohabiting partner) present or the husband is too ill to respond to the survey.

In the first panel of Figure 4, I plot the share of employed heads who are "willing to move" for work. About 50 percent of employed workers respond affirmatively.[21] But remarkably, this does not vary systematically with education. As an aside, notice that older workers do report being less willing to move, and this may help explain part of the age differentials in mobility - together with the differences in job rents estimated in Table 3.

[Figure 4 here]

In Table 7, I disaggregate those unwilling to move by the reasons they give. The most common are family/location ties and financial; and together with age/health reasons, these account for the bulk of the age differences. However, with the exception of health, none of these categories exhibit substantial variation by education.

[Table 7 here]

[20]Of course, if respondents were offered a million dollar salary, the vast majority would move.

[21]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?" 73 percent answer affirmatively: intuitively, the unemployed are more willing to bear the cost of migrating because their outside option is worse (which also reflects the evidence in Table 2). As with the employed, the fraction answering yes varies little with education.

Of course, these subjective responses are only useful if the low skilled do not systematically overstate (in relative terms) their willingness to move for work. And it turns out they are entirely realistic about their meager prospects in this regard. 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 4 plots the share of respondents who claim they might move for work. The results here clearly reflect the familiar age/education mobility patterns from Figures 1 and 2 above.[22] The contrast with the first panel is striking: the fact that low skilled workers expect low mobility is apparently unrelated to their "willingness" to move. This strongly suggests there is some other factor apart from costs at play.

Now, the first panel of Figure 4 tells us nothing about the amenity costs of those individuals who *are* "willing to move". In these cases, based on (33), amenity costs $c_i$ can be imputed as $w^R(w_i, c_i) - w_i$. The distribution of these imputed costs will of course be truncated - since $w^R(w_i, c_i)$ exceeds the cutoff $w_i^{CO}$ for many individuals. But critically, as Figure 4 shows, the fraction of observations which are truncated is invariant with skill: about 0.5 in each education group. That is, migration is a very unrealistic proposition for half the individuals in each group - so selection should not be a concern in comparisons by education. And an analysis of the imputed amenity costs will then be informative about the remainder of the population, the more "marginal" residents.

[Table 8 here]

In Table 8, I report sample means of imputed amenity costs, $w^R(w_i, c_i) - w_i$, in hourly wage terms for employed heads aged 25 to 64 - conditional on expressing willingness to move. I offer estimates using both dollar wage differentials and log differentials: the latter yields a proportionate estimate of the amenity cost, relative to the worker's wage. I proxy $w_i$ with the average hourly wage earned over the previous year. I exclude outliers from the sample - with log differentials below the 1st or above the 99th percentile of the distribution. The standard deviations of the remaining imputed costs are large, reflecting earlier results from Kennan and Walker (2011) which point to considerable heterogeneity in migration costs.

The average dollar cost is \$8.44 in hourly wage terms. These costs vary little (and unsystematically) with education and age. The one exception is the unusually low cost for college graduates aged 45-64: at under \$5, this is markedly less than younger graduates. But in any case, the 45-64 age group account for little of the skill mobility gap overall.[23]

[22]Further, I show in Appendix F.2 that these responses (on whether individuals "might move") have substantial predictive power for individuals' future migration decisions; and this predictive power does not vary significantly across education groups.

[23]The PSID also asks unemployed individuals for the minimum wage offer they would require to move. Using this information, it is possible to estimate $w^R(w_i, c_i) - w_i$ for the unemployed also (again, with $w_i$ representing

Given that the dollar costs are similar, the log gap is unsurprisingly decreasing in education. Across all groups, the average log gap is 0.46 for dropouts and 0.30 for college graduates. But, this difference is quantitatively small. It seems implausible to claim that a cost difference of 16 log points (among the 50 percent of individuals who express willingness to move) can account for the steep mobility gradients in Figure 1.

Reassuringly, the magnitude of the log differentials in Table 8 is consistent with the amenity costs implied by the estimated wage rents above. Using the SIPP data, for college graduate movers (who are mostly moving for job reasons; see Figure 2), I estimated average realized amenity costs of between 0.14 and 0.19 of a job's discounted future wage flows - conditional on moving. This compares to a 0.30 log gap for college graduates in my PSID sample who express a willingness to move for job reasons. The PSID estimate is somewhat larger, and this should be expected - given the PSID offers estimates of ex ante *unconditional* costs; whereas the SIPP estimates are *conditional* on moving, so should be selected from the bottom of the costs distribution.

## 7.3 Predictive power of imputed costs

Given that these results are based on the subjective judgments of respondents, there may be doubts over accuracy. But reassuringly, the cost measures do have significant predictive power for future migration decisions. Let $\rho_N\left(\varepsilon | c_i, \sigma_i^w, X_i\right)$ denote the instantaneous migration probability for some individual $i$ with initial match quality $\varepsilon$, conditional on an amenity cost $c_i$ (for a subsequent move), the dispersion $\sigma_i^w$ of wage offers, and a vector $X_i$ of demographic characteristics. The probability of moving before time $t$ is then:

$$\Pr\left(Mig_i^\tau = 1, t < \tau\right) = 1 - \exp\left(-\rho_N\left(\varepsilon | c_i, \sigma_i^w, X_i\right)\tau\right) \tag{35}$$

and I estimate the migratory response using a complementary log-log model:

$$\Pr\left(Mig_i^\tau = 1, t < \tau\right) = 1 - \exp\left(-\exp\left(\beta_c c_i + \beta_w w_i + \beta_X' X_i\right)\tau\right) \tag{36}$$

where, based on (3), I have expressed match quality $\varepsilon$ as a function of the initial wage $w_i$ and human capital indicators $X_i$. The advantage of this specification is that the $\beta$ parameters can (intuitively) be interpreted as the elasticities of the instantaneous migration rate with respect to its determinants. This interpretation is independent of the time horizon $\tau$ associated with the migration variable, and I effectively normalize $\tau$ to one year (to correspond with the PSID data

the average hourly wage earned over the previous year), though this conflates the amenity cost with the value of employment (relative to unemployment): the average dollar differential is \$8.44 for the employed, compared to just \$2.61 for the unemployed.

interval).[24]

The principle challenge to identification is that the offer dispersion $\sigma_i^w$ facing the individual is unobserved - and may be correlated with the amenity cost $c_i$ and current wage $w_i$. Unfortunately, I do not have convincing instruments, and there is insufficient power to control for individual fixed effects. So instead, I rely on the vector $X_i$ to control for the offer distribution. I include in $X_i$ a set of demographic characteristics[25], 8 occupation and 12 industry fixed effects relating to the individual's initial job, and also a set of year effects.

[Table 9 here]

I report my estimates in Table 9. I restrict my sample to the years 1970-3, which cover those employed individuals who reported reservation wages (for moving) in the previous wave. The first two columns report the elasticity of cross-state migration (in the previous 12 months) to the binary indicator for "willingness to move" (lagged one year). Willingness to move adds 130 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 four columns, I restrict the sample to those who are "willing to move" and estimate elasticities with respect to imputed amenity costs. In columns 3 and 4, I study the response to dollar imputed costs and dollar wages; and in columns 5 and 6, I study log imputed costs and log wages. Column 3 shows that a \$10 reduction in the dollar imputed cost adds 29 log points to the migration rate; and a \$10 reduction in the initial wage adds 42 points. And in column 5, the elasticities of the migration rate to the log imputed cost and initial wage are -0.85 and -0.97 respectively. These estimates are statistically significant. In columns 4 and 6, I allow for skill heterogeneity in the elasticities, but the interaction coefficients (though large in magnitude[26]) are estimated with substantial error. This is perhaps unsurprising, given the number of observations: there are 133 cross-state movers in the sample for the final four columns, of whom just 39 are college graduates.

In any case, the key point to take from this table is that the subjective costs *do* have predictive power - which suggests they are informative about the true costs of moving. And this reinforces the message above that migration costs vary little with education.

[24] The average marginal effects (not reported here) are very similar to those from probit and logit estimates.

[25] Specifically: age and age squared; four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in age; and gender, black and Hispanic dummies.

[26] In principle, a positive interaction (i.e. a smaller elasticity) for college graduates is consistent with the predictions of the model. For a worker earning $w_i$ with amenity cost $c_i$, the non-home job finding rate can be expressed as $\rho\left(\frac{w_i+c_i}{\sigma^w}\right)$. And the elasticity with respect to $w_i$ or $c_i$ is $\frac{1}{\sigma^w}\frac{f^w(\varepsilon)}{1-F^w(\varepsilon)}$ - which is decreasing in match productivity dispersion $\sigma^w$ for given match quality $\varepsilon = \frac{w_i+c_i}{\sigma^w}$. Intuitively, if job rents are larger, a given change in amenity costs or wages should matter less for migration decisions on the margin.

# 8 Conclusion

I have argued that skilled workers are more footloose because they benefit from substantial rents, irrespective of geography, as they climb the jobs ladder. This is particularly so for younger workers, who are just beginning their careers. While these rents are unimportant for local job flows, they play a critical role in driving long-distance mobility - given the typically large cost of these moves.

The job rents explanation is attractive firstly because it is theoretically intuitive: skilled work is necessarily more specialized, which naturally yields larger wage rents on forming a successful match. And second, it has strong empirical foundations: these wage rents are easily observed in the data. Furthermore, my estimates of skilled wage rents are large enough to plausibly explain the mobility gap.

Importantly, this hypothesis makes no claims on the geographical structure of these rents. Though the literature has often emphasized the importance of the worker-location match in explaining skilled mobility, the evidence is not supportive. In particular, I show the skill mobility gap is not driven by large net flows to particular states, even within detailed occupation groups. In my framework, all that matters is the match between workers and *firms* - irrespective of location.

Another popular view is that skilled mobility is driven principally by low migration costs. But this claim is undermined by evidence that wage rents are much larger for skilled workers in cross-state job matches. This suggests skilled workers typically select into migration *because* of large wage rents and *despite* steep migration costs. In contrast, among the low skilled, wage rents are similarly small in both local and cross-state matches: that is, they typically move *because* of a low cost draw and *despite* meager rents. I also offer more direct evidence using subjective data from the PSID, which suggests little difference in (unconditional) migration costs by skill.

While I have focused here on skill differentials in mobility, the model has broader applications. In its most abstract sense, it describes a jobs ladder in two dimensions: productivity and amenities (or the non-productive attributes of jobs). Above, I have considered the implications for residential amenities and residential choices. But there are also consequences for decisions over workplace amenities or commuting requirements. The model will therefore be useful in describing how these vary with the dispersion of match productivity and over the lifecycle.

# References

**Amior, Michael, and Alan Manning.** 2015. "The Persistence of Local Joblessness." CEP Discussion Paper 1357.

**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.

**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.

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

**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.” *The 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.** 2012. “A Spatial Knowledge Economy.” NBER Working Paper No. 18188.

**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.

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

**Glaeser, Edward L., and Joshua D. Gottlieb.** 2009. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature*, 47(4): 983–1028.

**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.

**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, C.A.L., and F. Robert-Nicoud.** 2010. "On the Origins of Land Use Regulations: Theory and Evidence from US Metro Areas."

**Hoynes, Hilary W.** 2002. "The Employment, Earnings, and Income 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.

**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.

**Kaplan, Greg, and Sam Schulhofer-Wohl.** 2012*a*. "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.** 2012*b*. "Understanding the Long-Run Decline in Interstate Migration." NBER Working Paper No. 18507.

**Kennan, John.** 2015. "Spatial Variation in Higher Education Financing and the Supply of College Graduates." http://www.ssc.wisc.edu/∼jkennan.

**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.

**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. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton: Princeton University Press.

**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, J.** 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.

**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 K. Wozniak.** 2011. "Internal Migration in the United States." *Journal of Economic Perspectives*, 25(3): 173–96.

**Molloy, Raven, Christopher L. Smith, and Abigail Wozniak.** 2014. "Declining Migration Within the US: The Role of the Labor Market." NBER Working Paper No. 20065.

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

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

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

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

**Shimer, Robert.** 2005*b*. "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.

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

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

**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.

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

# Tables and figures

Table 1: Net cross-state migration rates by education

| | 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) |
| HS dropout | 1.81 | 0.28 | 0.15 | 1.59 | 0.63 | 0.40 | 1.59 | 0.85 | 0.53 |
| HS graduate | 1.93 | 0.27 | 0.14 | 1.57 | 0.41 | 0.26 | 1.57 | 0.59 | 0.38 |
| Some college | 2.37 | 0.28 | 0.12 | 1.96 | 0.51 | 0.26 | 1.96 | 0.77 | 0.40 |
| Undergraduate | 3.06 | 0.27 | 0.09 | 2.68 | 0.56 | 0.21 | 2.68 | 0.86 | 0.32 |
| Postgraduate | 3.57 | 0.32 | 0.09 | 3.27 | 0.66 | 0.20 | 3.27 | 0.98 | 0.30 |

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 basic estimates, and the final six offer within-occupation estimates - based on 2-digit and 3-digit occupation categories. For each education group, these are constructed by weighting occupation-specific migration rates by occupational employment shares. The sample consists of individuals aged 25 to 64 in the ACS between 2000 and 2009, and this is further restricted to the employed for columns 4-9. Migrants are defined as individuals whose current state of residence differs from state of residence 12 months previously. Employment status and occupation are recorded at time of survey. Occupational codes are based on the census 2000 scheme.

Table 2: Cross-state migration rates by initial employment status: SIPP

| | Cross-state mig rates (%) over 4-month waves | | | |
|---|---|---|---|---|
| | All individuals | By emp status in $t-1$ | | |
| | | Employed | Unemployed | Inactive |
| | (1) | (2) | (3) | (4) |
| HS dropout | 0.31 | 0.27 | 0.40 | 0.33 |
| HS graduate | 0.37 | 0.30 | 0.65 | 0.48 |
| Some college | 0.53 | 0.46 | 0.89 | 0.73 |
| Undergraduate | 0.73 | 0.63 | 1.35 | 1.06 |
| Postgraduate | 0.86 | 0.76 | 2.12 | 1.24 |

This table reports cross-state migration rates across four-month waves, based on the SIPP panels of 1996, 2001, 2004 and 2008, which cover the period between 1996 and 2013. The full sample consists of 1.7m individual-wave observations. Column 1 reports migration rates for all individuals aged 25-64, by education group. Columns 2-4 reports these rates separately by initial employment status.

## Table 3: Estimates of wage returns to job finding

| | No fixed effects | | | | | Fixed effects | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All ages | All ages | 25-34 | 35-44 | 45-64 | All ages | All ages | 25-34 | 35-44 | 45-64 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| New job | 0.025*** | 0.018*** | 0.009 | 0.020* | 0.025** | 0.027*** | 0.018** | 0.009 | 0.027** | 0.019 |
| | (0.003) | (0.006) | (0.010) | (0.011) | (0.013) | (0.004) | (0.008) | (0.014) | (0.013) | (0.013) |
| New job * HS grad | | -0.004 | 0.002 | 0.009 | -0.023 | | -0.005 | 0.004 | -0.007 | -0.014 |
| | | (0.008) | (0.012) | (0.014) | (0.016) | | (0.010) | (0.017) | (0.018) | (0.018) |
| New job * Some coll | | -0.003 | 0.020* | -0.021 | -0.015 | | -0.004 | 0.027 | -0.028 | -0.014 |
| | | (0.008) | (0.012) | (0.014) | (0.016) | | (0.010) | (0.017) | (0.017) | (0.017) |
| New job * Undergrad | | 0.028*** | 0.058*** | 0.041** | -0.022 | | 0.032*** | 0.055*** | 0.046** | -0.004 |
| | | (0.010) | (0.015) | (0.018) | (0.019) | | (0.012) | (0.020) | (0.022) | (0.023) |
| New job * Postgrad | | 0.057*** | 0.125*** | 0.031 | 0.019 | | 0.057*** | 0.129*** | 0.030 | 0.030 |
| | | (0.014) | (0.025) | (0.026) | (0.024) | | (0.018) | (0.032) | (0.035) | (0.029) |
| Demographic controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sample | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 |

This table offers estimates of equation (26), based on four-month wave transitions in the SIPP panels beginning 1996, 2001, 2004 and 2008. I regress log wage changes (within individuals) on a new job dummy, interacted with a set of education effects. I report specifications both without individual fixed effects (columns 1-5) and including them (6-10). Throughout, I control for a full set of wave effects and a detailed set of demographic characteristics, specifically: age and age squared; four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in age and a time trend; black and Hispanic race dummies and immigrant status; and a gender indicator which is also interacted with all previously mentioned variables. I base my wage variable on hourly wage data for workers paid by the hour, and I impute hourly wages for salaried workers using monthly earnings and hours. To guard against measurement error, I restrict my sample to wage changes where pay duration does not change (that is, the worker is either paid by the hour in both periods or salaried in both periods). I also restrict attention to wage observations between \$2 and \$100 in 2000 prices, and I exclude workers with multiple jobs or business income at the end of a wave. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Table 4: Estimates of wage returns to job finding - controlling for initial wage

| | No fixed effects | | | | | Fixed effects | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All ages | All ages | 25-34 | 35-44 | 45-64 | All ages | All ages | 25-34 | 35-44 | 45-64 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| New job | 0.530*** | 0.600*** | 0.513*** | 0.612*** | 0.690*** | 0.375*** | 0.428*** | 0.373*** | 0.421*** | 0.478*** |
| | (0.016) | (0.016) | (0.026) | (0.028) | (0.032) | (0.015) | (0.016) | (0.028) | (0.027) | (0.030) |
| $\log w_{it-1}$ | -0.123*** | -0.121*** | -0.134*** | -0.127*** | -0.111*** | -0.666*** | -0.663*** | -0.684*** | -0.717*** | -0.662*** |
| | (0.001) | (0.001) | (0.002) | (0.002) | (0.002) | (0.004) | (0.004) | (0.008) | (0.008) | (0.006) |
| New job * $\log w_{it-1}$ | -0.211*** | -0.275*** | -0.244*** | -0.277*** | -0.308*** | -0.148*** | -0.192*** | -0.173*** | -0.184*** | -0.214*** |
| | (0.006) | (0.007) | (0.012) | (0.013) | (0.013) | (0.006) | (0.007) | (0.013) | (0.012) | (0.013) |
| New job * HS grad | | 0.032*** | 0.020* | 0.048*** | 0.033** | | 0.019** | 0.015 | 0.013 | 0.032** |
| | | (0.008) | (0.012) | (0.013) | (0.015) | | (0.008) | (0.014) | (0.014) | (0.014) |
| New job * Some coll | | 0.060*** | 0.057*** | 0.051*** | 0.075*** | | 0.032*** | 0.036*** | 0.015 | 0.048*** |
| | | (0.008) | (0.012) | (0.013) | (0.015) | | (0.008) | (0.014) | (0.014) | (0.013) |
| New job * Undergrad | | 0.191*** | 0.177*** | 0.223*** | 0.184*** | | 0.128*** | 0.115*** | 0.148*** | 0.130*** |
| | | (0.011) | (0.017) | (0.020) | (0.021) | | (0.010) | (0.017) | (0.019) | (0.019) |
| New job * Postgrad | | 0.283*** | 0.304*** | 0.277*** | 0.275*** | | 0.190*** | 0.220*** | 0.164*** | 0.192*** |
| | | (0.015) | (0.025) | (0.026) | (0.025) | | (0.014) | (0.025) | (0.025) | (0.024) |
| Demographic controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sample | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 |

This table is identical to Table 3 (see associated notes for further details), except I also control for the lag of the log wage and its interaction with the new job dummy variable. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Table 5: Quantifying the effect of job rents

| | HS dropout (1) | HS grad (2) | Some coll (3) | Undergrad (4) | Postgrad (5) |
|---|---|---|---|---|---|
| Annual overall job-to-job finding rate (SIPP): $\int_{\varepsilon}\rho(\varepsilon)\,dG(\varepsilon)$ | 0.205 | 0.203 | 0.209 | 0.198 | 0.173 |
| Annual job-motivated cross-state mig rate (CPS): $\int_{\varepsilon}\rho_N(\varepsilon)\,dG(\varepsilon)$ | 0.005 | 0.005 | 0.008 | 0.014 | 0.020 |
| Share of job finds which are cross-state: $\frac{\int_{\varepsilon}\rho_N(\varepsilon)dG(\varepsilon)}{\int_{\varepsilon}\rho(\varepsilon)dG(\varepsilon)}$ | 0.022 | 0.027 | 0.036 | 0.072 | 0.116 |
| $\beta_1$ (col 2, Table 3) | 0.018 | 0.014 | 0.015 | 0.046 | 0.075 |
| Calibrated $\frac{1-\pi}{\pi}\sigma^c = \frac{\int_{\varepsilon}\rho(\varepsilon)dG(\varepsilon)}{\int_{\varepsilon}\rho_N(\varepsilon)dG(\varepsilon)}\beta_1$ | 0.81 | 0.52 | 0.41 | 0.64 | 0.64 |

The annual flow of new jobs among the initially employed (column 1) is based on the SIPP; see Section 5.2 for a sample description. I estimate the fraction of individuals aged 25 to 64 (irrespective of initial employment status) at the end of each four month wave who have a new job (which they did not hold in the previous wave). To derive an annual rate, I multiply this number by 3. I report the annual rate at which individuals move state for self-reported job-related reasons in columns 2 and 3, based on the CPS; see notes under Figures 1 and 2. To ensure the SIPP and CPS data cover the same period, I restrict both samples to the years 1999-2013 - the maximum range ensuring coverage of all the studied variables. My $\beta_1$ estimates (column 6) are based on regression estimates from column 2 of Table 3.

## Table 6: Estimates of wage returns to cross-state job finding

| | No fixed effects | | | | | Fixed effects | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All ages | All ages | 25-34 | 35-44 | 45-64 | All ages | All ages | 25-34 | 35-44 | 45-64 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| New job | 0.023*** | 0.015*** | 0.020*** | 0.013** | 0.010* | 0.025*** | 0.015*** | 0.024*** | 0.012* | 0.007 |
| | (0.003) | (0.003) | (0.005) | (0.005) | (0.006) | (0.004) | (0.004) | (0.006) | (0.007) | (0.007) |
| New job * grad | | 0.033*** | 0.051*** | 0.039*** | 0.008 | | 0.037*** | 0.047*** | 0.051*** | 0.020 |
| | | (0.007) | (0.011) | (0.014) | (0.013) | | (0.009) | (0.014) | (0.018) | (0.017) |
| New job * move | 0.070*** | -0.016 | -0.033 | -0.006 | 0.005 | 0.058** | -0.021 | -0.030 | -0.035 | 0.009 |
| | (0.023) | (0.025) | (0.030) | (0.057) | (0.057) | (0.026) | (0.031) | (0.040) | (0.065) | (0.075) |
| New job * move * grad | | 0.158*** | 0.191*** | 0.142 | 0.010 | | 0.139*** | 0.170** | 0.164 | -0.004 |
| | | (0.045) | (0.062) | (0.087) | (0.093) | | (0.052) | (0.073) | (0.103) | (0.118) |
| Move | 0.009 | 0.010 | 0.016 | 0.027 | -0.020 | 0.010 | 0.0120 | 0.020 | 0.031 | -0.017 |
| | (0.008) | (0.012) | (0.017) | (0.020) | (0.022) | (0.011) | (0.016) | (0.024) | (0.028) | (0.036) |
| Move * grad | | 0.000 | 0.007 | -0.046 | 0.039 | | -0.006 | 0.003 | -0.055 | 0.035 |
| | | (0.016) | (0.025) | (0.029) | (0.030) | | (0.022) | (0.035) | (0.039) | (0.044) |
| Demographic controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sample | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 | 780,071 | 780,071 | 191,287 | 237,491 | 351,293 |

This table offers estimates of equation (32), based on four-month wave transitions in the SIPP panels beginning 1996, 2001, 2004 and 2008. I regress log wage changes (within individuals) on a new job dummy, a dummy for a cross-state move and an interaction between the two; and I also include interactions between all those variables and a college graduate dummy. I report specifications both without individual fixed effects (columns 1-5) and including them (6-10). Throughout, I control for a full set of wave effects and a detailed set of demographic characteristics. See the notes under Table 3 for further details on the controls and sample.

Table 7: Disaggregation of those unwilling to move by reported reason

| | Unwilling to move | By reported reason for separation: Family/location ties | Financial | Age/health | Other | Not recorded | Observations |
|---|---|---|---|---|---|---|---|
| *Education* | | | | | | | |
| HS dropout | 50.67 | 29.75 | 10.49 | 5.56 | 3.14 | 1.75 | 7,115 |
| HS graduate | 46.54 | 27.65 | 10.68 | 3.39 | 3.31 | 1.51 | 5,815 |
| Some college | 43.60 | 26.73 | 9.59 | 2.76 | 2.89 | 1.64 | 2,390 |
| College graduate | 46.14 | 27.77 | 9.57 | 1.94 | 4.23 | 2.62 | 2,919 |
| *Age group* | | | | | | | |
| 25-34 | 33.66 | 22.02 | 7.53 | 0.14 | 2.66 | 1.31 | 7,171 |
| 35-44 | 48.35 | 30.67 | 11.18 | 0.89 | 3.85 | 1.75 | 4,465 |
| 45-64 | 62.39 | 33.64 | 12.64 | 10.04 | 3.71 | 2.35 | 6,603 |

In the 1970s, the PSID asked individuals: "Would you be willing to move to another community if you could earn more money there?" In this table, I report the percentage responding negatively, and disaggregate these individuals by their stated reason for being unwilling to move. Statistics in this table are pooled across employed and unemployed individuals. These questions were posed to the employed in the waves of 1970-2 and 1979-80, and to the unemployed in every wave between 1970 and 1980 excluding 1976. The column labelled "not recorded" refers to the small number of individuals who claim to be unwilling to move, but who are coded as N/A for the reason why; these include retirees, students and housewives, among others.

Table 8: Imputed amenity costs (in hourly wage terms), conditional on “willingness to move”

| | Dollar gap ($ 2000) | | | | Log gap | | | | Observations |
|---|---|---|---|---|---|---|---|---|---|
| | All | 25-34 | 35-44 | 45-64 | All | 25-34 | 35-44 | 45-64 | |
| HS dropout | 7.61 | 7.22 | 8.01 | 7.57 | 0.46 | 0.45 | 0.45 | 0.48 | 2,549 |
| | (7.64) | (6.98) | (7.81) | (7.90) | (0.43) | (0.39) | (0.42) | (0.48) | |
| HS graduate | 9.35 | 9.41 | 9.06 | 9.57 | 0.44 | 0.47 | 0.41 | 0.43 | 1,583 |
| | (8.37) | (8.38) | (8.52) | (8.24) | (0.40) | (0.40) | (0.40) | (0.39) | |
| Some college | 8.99 | 9.10 | 7.67 | 10.05 | 0.40 | 0.44 | 0.34 | 0.41 | 697 |
| | (10.11) | (10.60) | (10.24) | (9.13) | (0.40) | (0.42) | (0.36) | (0.39) | |
| Coll graduate | 7.84 | 8.99 | 8.80 | 4.87 | 0.30 | 0.37 | 0.32 | 0.19 | 610 |
| | (10.93) | (8.92) | (11.62) | (11.95) | (0.52) | (0.38) | (0.64) | (0.48) | |

This table reports mean imputed amenity costs by education and age for employed workers (conditional on expressing willingness to move), with standard deviations in parentheses. I present two alternative estimates of the imputed cost. The "dollar gap" is equal to $w^R(w_i, c_i) - w_i$, where $w^R(w_i, c_i)$ is the minimum hourly wage required to tempt a worker $i$ to take a job in another area, and $w_i$ is the worker's average hourly wage in the previous 12 months (with wages expressed in 2000 dollars). The "log gap" is the log difference between the reservation and current wage: $\log w^R(w_i, c_i) - \log w_i$. I pool individuals with undergraduate and postgraduate qualifications because of small samples. The sample consists of employed heads aged 25-64 in the PSID waves of 1969-72. 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. I exclude outliers from the sample - with a "log gap" below the 1st or above the 99th percentile of the distribution. The specific question eliciting the reservation wage is: "How much would a job have to pay for you to be willing to move?"

Table 9: Responses of cross-state migration to lagged cost measures

*Complementary log-log regressions of cross-state move (in previous 12 months) on 1-year lagged costs and wages*

| | Unconditional sample | | Conditional sample (willing to move) | | | |
|---|---|---|---|---|---|---|
| | | | Dollar costs/wages | | Log costs/wages | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Willing to move (binary) | 1.256*** | 1.311*** | | | | |
| | (0.202) | (0.235) | | | | |
| Willing to move (binary) * Grad | | -0.173 | | | | |
| | | (0.446) | | | | |
| Imputed cost | | | -0.029* | -0.041** | -0.851** | -1.135*** |
| | | | (0.016) | (0.020) | (0.354) | (0.350) |
| Imputed cost * Grad | | | | 0.029 | | 0.971 |
| | | | | (0.036) | | (0.928) |
| Previous wage | | | -0.042*** | -0.042** | -0.971*** | -0.901*** |
| | | | (0.015) | (0.020) | (0.291) | (0.339) |
| Previous wage * Grad | | | | 0.010 | | 0.053 |
| | | | | (0.029) | | (0.667) |
| Demographic controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry, occupation FEs (lagged) | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10,721 | 10,721 | 4,366 | 4,366 | 4,366 | 4,366 |
| Cross-state mig rate | 0.021 | 0.021 | 0.035 | 0.035 | 0.035 | 0.035 |

This table reports responses of cross-state migration (in the previous 12 months) to subjective costs and wages (lagged by one year), based on complementary log-log regressions. I study three subjective cost measures: the binary indicator of willingness to move for work, the dollar gap amenity cost, and the log gap amenity cost. The latter two are observable only for individuals who express willingness to move. These measures are described in greater detail in the notes under Table 8. 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 (36). The sample consists of household heads aged 25-64 in the PSID waves of 1970-3, who were employed in the previous wave (when costs and wages are measured). 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 further restricted for the dollar gap and log gap measures, as described in the notes under Table 8. All specifications control for (i) demographic controls, specifically age and age squared, four education indicators (high school graduate, some college, undergraduate and postgraduate), each interacted with a quadratic in age, and gender, black and Hispanic dummies; (ii) fixed effects denoting occupation (8 categories) and industry (12 cateogries) one year ago; and (iii) year fixed effects. Errors are clustered by individual, and robust SEs are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

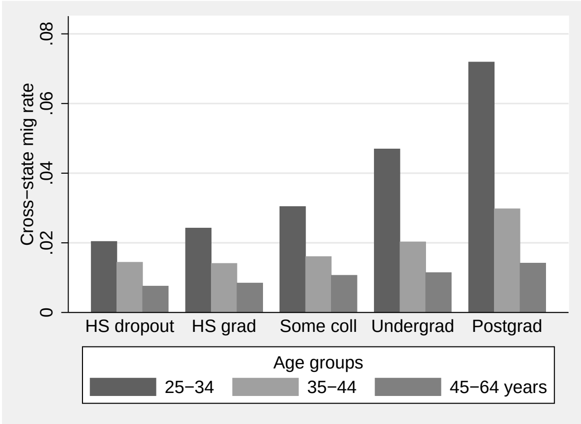


Figure 1: Annual cross-state migration rates (CPS 1999-2015)

This figure reports the fraction of the sample living in a different state 12 months previously, disaggregated by age and education, in CPS March waves between 1999 and 2015. I exclude all individuals living abroad one year previously. I also restrict attention to the top earner in each household. This makes it easier to interpret the evidence below on reported reasons for moving, and it makes little difference to estimates of skill mobility differentials (see Appendix A.2). Finally, Kaplan and Schulhofer-Wohl (2012*a*) show there are inconsistencies in the CPS's procedure for imputing migration status in cases of non-response. The non-response rate for migration status is 14 percent in my sample, and this varies little with education. I choose to drop these observations. See Appendix A.1 for further discussion of all these data issues.

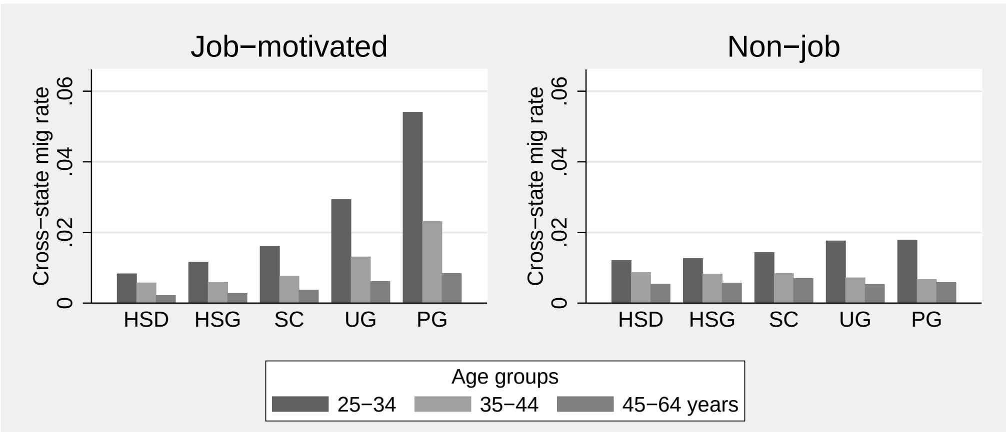


Figure 2: Annual cross-state migration rates by reported reason

The first panel reports the fraction of individuals who moved state primarily for job-related reasons in the previous 12 months; and the second panel does the same for non-job reasons. Data is based on the March CPS between 1999 and 2015. The sample is restricted to top earners in each household, due to concerns that many household dependents simply report the reasons of the breadwinners (see Appendix A.1). See notes below Figure 1 for further details.

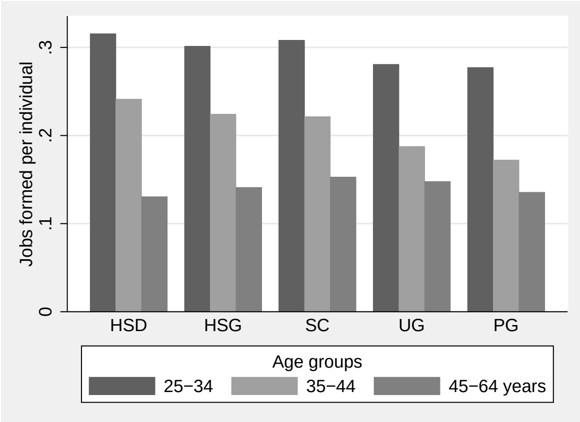


Figure 3: Annual flow of new jobs (SIPP 2004 panel)

This figure reports the flow of new jobs among individuals aged 25 to 64: that is, the average number of job matches formed per individual. These estimates are based on job transitions over four-month waves in the 1996, 2001, 2004 and 2008 panels of the Survey of Income and Program Participation (SIPP), which cover the period between 1996 and 2013. I exclude individuals with multiple jobs or business income. See Section 5.2 for further details on this dataset.

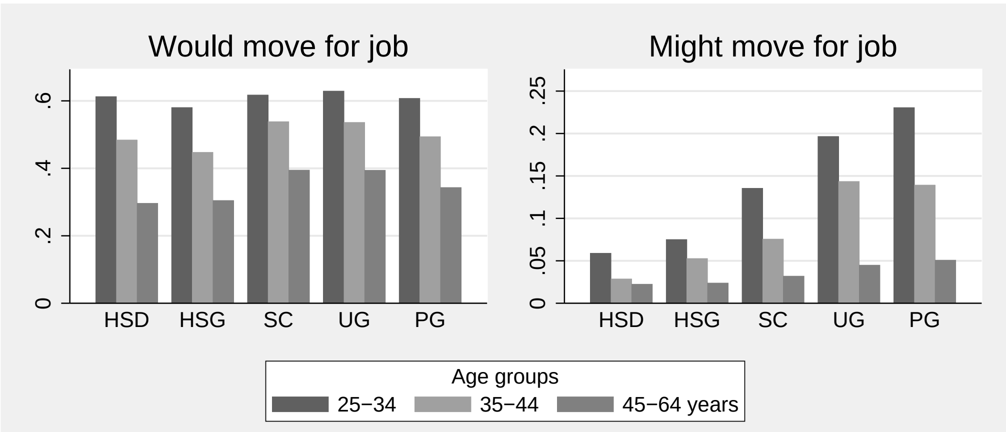


Figure 4: Share who “would move” and “might move” for better job: PSID 1969-80

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 in the years 1969-72 and 1979-80, when both questions were asked. The full sample consists of 18,893 observations.

# Appendices: For online publication

## A Supplementary CPS estimates of basic mobility gap

### A.1 Sample description

In this Appendix, I estimate the skill mobility gap separately for household top earners (which define the sample for Figures 1 and 2 in the main text) and the full sample of individuals. I also report the mobility gap for single-year age groups, and I study changes over time.

I begin with a description of my Current Population Survey (CPS) sample. All estimates from the CPS in this study are based on the March waves, which include the Annual Social and Economic Supplement (ASEC). The ASEC reports whether respondents moved county or state in the previous 12 months. Since 1999, individuals have also given their primary reason for moving. All estimates below are based on pooled cross-sections between 1999 and 2015. I use CPS data organized by IPUMS (King et al., 2010).

I restrict the sample to individuals aged 25 to 64 who lived in the US for the previous 12 months. Focusing on the over-25s helps ensure my results are not conflated by individuals leaving college.

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 has resulted in some inconsistencies in the coding of responses: many household dependents simply report the reasons of the breadwinners.[27] My strategy is to restrict 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 40 percent of the original sample. But as I show below, it makes little difference to estimates of the skill mobility gap.

Finally, Kaplan and Schulhofer-Wohl (2012*a*) show there are inconsistencies in the CPS's procedure for imputing migration status in cases of non-response: 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 by education: 13 percent of college graduates and 14 percent of non-graduates are affected. I choose to drop all these observations.

[27] This is most clearly visible among children: in households with at least one adult moving for job-related reasons, 77 percent of under-16s also report moving for job reasons.

## A.2 Sensitivity to household top earner restriction

I begin by studying the sensitivity of the skill mobility gap to the top earner sample restriction. Table A1 shows the effect is small across all age and education categories - though migration rates for individuals with postgraduate qualifications are affected somewhat more than other education groups.

[Table A1 here]

## A.3 Mobility gap by single-year age

In Figure A1, I report estimates of the annual cross-state migration rate by education and single-year age group, based on my basic sample of household top earners. Consistent with Figure 1 in the main text, the skill mobility gap is largely driven by the under-35s. And this figure makes clear that mobility differentials are also decreasing in age within this group. Among individuals aged 25, the migration rate of college graduates is 8.7 percent compared to 4.6 percent for non-graduates; and these numbers decline (and converge) to 3.5 and 2.3 percent respectively for those aged 34.

[Figure A1 here]

## A.4 Historical changes in mobility gap

The CPS analysis in the main text is restricted to the period 1999-2015, when I have information of reasons for moving. But the skill mobility gap is by no means a new phenomenon. In Figure A2, I plot annual cross-state migration rates using the March waves of the CPS from 1963 to 2015.[28]

[Figure A2 here]

For this exercise, I do not restrict my sample to household top earners - so my results are not conflated with changes in household composition over time. But I continue to exclude individuals living abroad 12 months previously and those with imputed migration data. Following Kaplan and Schulhofer-Wohl (2012*b*), I also omit households with members in the military: military households are unusually mobile, and the period saw a decline in military employment.

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

As is well known, migration rates have declined overall: see e.g. Molloy, Smith and Wozniak (2011). Kaplan and Schulhofer-Wohl (2012*b*) argue this was driven by the declining geographical specificity of occupational returns, coupled with improvements in communications technology. Molloy, Smith and Wozniak (2014) explain it by a declining rate of labor market transitions.

In any case, the key point from the perspective of this paper is the persistence of the skill mobility gap, as measured by the ratio of graduate to non-graduate mobility. This ratio did decline in the 1960s and 1970s from about 2.2 to 1.7, but it has changed little since then. Having said that, care must be taken in interpreting these trends, because the period experienced a large expansion of higher education - so there may have been important changes in the composition of these education groups.

# B Supplementary CPS estimates on reasons for moving

## B.1 Breakdown of migration by reported reasons for moving

This Appendix presents supplementary estimates of migration rates by reported reasons for moving, based on the CPS sample described in Section A.1. First, I present a detailed disaggregation of cross-county and cross-state migration in the CPS by reported reason. Second, I test the robustness of the results in Figure 2 (on the skill gradients in job-related and non-job migration) to individual demographic controls. Third, I disaggregate the skill gradients in job-related and non-job migration into finer reason categories.

[Table A2 here]

Table A2 disaggregates cross-county migration by primary reason for moving, separately for cross-state and cross-county moves. The first column gives the percentage of the full sample who changed state for each recorded reason, and the second column reports the percentage of cross-state migrants who moved for each recorded reason. The final two columns repeat this exercise for cross-county moves within states.

The bottom row shows that, each year, about 2 percent of the sample move across states and another 2 percent switch county within states. About half of cross-state moves are job-motivated, compared with a third of within-state moves. Job-motivated moves are almost always driven by the needs of a *specific* job. Usually, this is due to a job change or transfer; and among within-state moves, commuting reasons also feature prominently. The commuting motivation can easily be interpreted in the context of a cross-state match: after accepting a distant job (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 3 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.

## B.2 Robustness of Figure 2 to individual controls

Next, I show the patterns in job-motivated and non-job migration in Figure 2 are robust to individual demographic controls, within each age group. Specifically, I estimate complementary log-log models for annual incidence of cross-state migration of the form of equation (36).

[Table A3 here]

The $\beta$ estimates for the education effects are presented in Table A3. I report results both with and without a detailed range of demographic controls: specifically age, age squared, black and Hispanic race dummies, immigrant status, marital status, a range of indicators for number of own children, and a gender indicator which is also interacted with all previously mentioned variables. The reported coefficients give the log point effect of a particular level of education, relative to high-school dropout (the omitted category), on the instantaneous migration rate (for the specified motivation).

With regards to job-motivated migration (columns 1-3), Table A3 shows there are positive and strongly significant education effects within each age group. The coefficients change little after controlling for demographic characteristics. Among those aged 25-34, a postgraduate education adds 179 log points to the migration rate (controlling for individual characteristics), relative to high school dropouts. This effect comes to 129 log points among the 35-44s, and 134 among the 45-64s.

Columns 4-6 report the effects on non-job migration. As in Figure 2, the education effects are increasing for the 25-34s without demographic controls (though much more slowly than in columns 1-3) and somewhat decreasing for the 35-44s. It turns out the positive slope for the under-35s is entirely explained by individuals moving to attend or leave college. This is clear from columns 7-9, where the dependent variable now takes 1 for any non-job move which is not motivated by attending or leaving college. Controlling for demographic characteristics makes little difference to all these results.

## B.3 Disaggregation of job-motivated and non-job skill gradients

In Table A4, I offer a finer disaggregation of the skill mobility gradient by reasons for moving. Again, I estimate complementary log-log models for migration by reported reason, of the form of equation (36), on education effects and a range of demographic controls. In each row of the table, I report education slopes for the individual motivations. The first four columns report estimates for the incidence of cross-state moves, and the final four columns for within-state cross-county moves. I pool all age groups together in each specification.

[Table A4 here]

The first row reports effects for all motivations combined. Interestingly, the positive education gradient is only present for cross-state moves and not within-state. Mechanically, this is for two reasons. First, the (positive) education slope of job-motivated migration is much steeper for cross-state than within-state moves (see the second row). This result is consistent with the model, to the extent that cross-state migration is more costly - in which case skill differences in job rents matter more (see Proposition 1). Second, there is a strong negative slope in non-job migration for within-state moves.

Among job-related moves, the positive skill gradient is driven by motivations relating to a *specific* job - whether moving for a new job or commuting reasons. The new job motivation has a stronger skill gradient for cross-state moves, and the commuting motivation is stronger within-state. In contrast, better educated workers are significantly *less* likely to move speculatively - to look for work. A postgraduate education reduces the speculative migration rate by 48 log points across states, relative to dropouts, and by 84 points across counties (within states).

The negative skill slope in cross-county non-job migration is driven by a broad range of motivations: "other family reasons", cheaper housing, "other housing reasons", better neighborhood, climate, health and retirement. Of course, it is possible that lower skilled workers are simply subject to more family, housing and amenity shocks. For example, they tend to be more credit constrained, so housing costs may be a more important contributor to migration decisions. Or they may suffer more from family instability (see e.g. McLanahan, 1985). But, meager job rents offer an alternative hypothesis. If job rents are smaller, a given "non-job" shock (which increases the value of moving away) is more likely to break a worker's current job match. That is, workers value their jobs less - so they are happier to give them up to move elsewhere.

There are just three non-job motivations with significant positive skill slopes: the desire to purchase a home, attending or leaving college, and the residual "other reasons" category (across states only). The first two are of course entirely intuitive.

# C Net and gross migratory flows within occupations

In this Appendix, I break down gross and net migration rates by age group and by individual occupation. First, in Table A5, I replicate the results in Table 1 in Section 3 - but this time separately for individuals aged 25-34 and 35-64. For the under-35s at least, there is a discernible effect of education on net migration rates (especially for the within-occupation estimates), but only for the postgraduate-educated. And for all age groups and occupation schemes, the ratio of net to gross migration is still strongly decreasing in education. I conclude from this that the skill mobility gap is not driven by large net flows to particular states, even within detailed occupation categories and within distinct age groups.

[Table A5 here]

Table A5 (and Table 1 in the main text) offer averaged within-occupation migration rates by education. But it is also useful to study migration rates within individual occupation groups. Occupations are themselves a useful proxy for skill, and offer much greater variation than the five (education) data points I use above. It is worth emphasizing again (as I do in Section 3) that occupations are recorded at the time of survey, immediately *after* the period in which migration occurs. There may be a concern here that occupation and the migration decision are then simultaneously determined. But arguably, this is a useful time to measure occupation in this particular exercise - since an individual's ex post occupation is a good indicator of the job market in which they were searching.

[Figure A3 here]

In Figure A3, I plot annual gross (O markers) and net (X markers) migration rates within each occupation group against its skill percentile, where skill is identified with an occupation's college graduate share. I use the same occupation sample as in Tables 1 and A5 (based on the 2000 census scheme), but I exclude individuals in the armed forces: they are unusually mobile, with a cross-state migration rate of 19 percent, compared to 2 percent for other workers. I report estimates separately for 2-digit (left column) and 3-digit (right) occupations, and separately for individuals aged 25-64 (i.e. full sample, top row), 25-34 (middle) and 35-64 (bottom). The size of each marker is proportional to the occupation's employment sample.

In each case, the skill gradient in net migration is remarkably flat (even within 3-digit occupations), despite a steep gradient (especially among the under-35s) in gross migration rates.[29] This strongly reinforces my message in Section 3 in the main text.

[29]In terms of gross migration rates, there is one sizable outlier - with a skill percentile of around 0.9 but an unusually low migration rate. These are school teachers, who are constrained in their mobility by state licensing laws (see Kleiner, 2000).

# D Contribution of returning students

In this section, I check whether returning students may be contributing to the skill mobility gap. Throughout my analysis of reasons for moving in the CPS, I have excluded those individuals aged under 25. And in Table A3, I also exclude those who explicitly report moving either to attend or leave college. 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 youngest age group in the analysis above: those aged 25-34. Indeed, Kennan and Walker (2011) emphasizes that a large fraction of long-distance moves in the US involve people returning to former locations; and Kennan (2015) considers in particular how individuals return home to begin work after studying in another state.

The contribution of this return migration can easily be assessed in the PSID. In this exercise, I restrict attention to heads aged 25-34 in the annual PSID waves between 1990 and 1997. I exclude PSID waves after 1997 because these are biennial: in those years, it is not possible to keep track of return migration at annual frequencies. The first row of Table A6 reports the fraction of heads in each age group who were recently students (either in the current or previous annual wave). Since I exclude under-25s from my sample, the numbers are small and lie below 4 percent in each education group.

[Table A6 here]

The remaining rows report annual cross-state migration rates by education. The second row of the table gives the migration rates for the full sample, illustrating the familiar positive skill gradient. I exclude recent students in the third row, but this makes little difference since they comprise such a small fraction of the sample.

However, excluding recent students does not address the concerns entirely, because ex-students may yet return to their home town 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 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 skill gradient is clearly positive for both return and non-return rates, and the gradient is not steeper for the former. This demonstrates that returning students (and return migration in general) cannot account for the mobility gap.

# E Estimates of job finding and separation rates

Table A7 reports job transition rates based on the SIPP panels of 1996, 2001, 2004 and 2008, which cover the period between 1996 and 2013. The model in Section 4.5.3 predicts a theoretically ambiguous effect of skill on the job finding rate from unemployment; and indeed, column 1 shows the effect (though positive) is relatively small. Having said that, the skill gradient in job finding is steeper once the economically inactive are included (column 3). Separation rates $\delta$ to unemployment (and non-employment) are steeply decreasing in education (columns 2 and 5); and this explains the bulk of skill differentials in unemployment rates (see e.g. Mincer, 1991). Finally, the job-to-job transition rate is decreasing in education (column 3).

[Table A7 here]

# F Supplementary estimates for PSID sample

## F.1 Estimates of skill mobility gap

In this Appendix, I offer supplementary estimates for my 1970s PSID sample. First, I show that patterns in cross-state migration rates by age and education are comparable to those from my 1999-2015 CPS sample. And second, I show that subjective expectations are strong predictors of future mobility - and equally strong across education groups.

[Figure A4 here]

Figure A4 reports the share of household heads[30] between 1970 and 1980 who moved state in the previous 12 months, by age and education. The patterns are very similar to those in Figure A2, though the migration rates are somewhat higher in all groups. This is unsurprising, given the time trends documented in Figure A2 above.

[Figure A5 here]

In Figure A5, I break down these migration rates by reported reason, specifically job-related or non-job. For 18 percent of cross-state moves, the PSID reports the reason to be "ambiguous" or "mixed" or simply unknown. I allocate these cases to the job-motivated and non-job categories within age-education cells, according to the proportions in the non-ambiguous data. The patterns look very similar to those in Figure 2 in the main text: just as in the CPS sample, the skill mobility gap is almost entirely driven by job-motivated migration.

[30] 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.

## F.2 Predictive power of subjective expectations

The analysis in Section 7 is predicated on the predictive power of subjective expectations of future mobility. This can be investigated empirically. The PSID asks: “Do you think you might move in the next couple of years?” and “Why might you move?” Based on the responses, the second panel of Figure 4 in the main text shows that better educated workers are much more likely to report they “might” move for work; and the effect is strongest for the young. This reflects the patters in Figure A5 above.

In this section, I exploit the survey’s longitudinal dimension to check how the predictive power of these subjective expectations varies by education. Table A8 reports the fraction of the sample (household heads over 1970-80) who moved residence (over any distance) for job-related reasons in the previous 12 months, in cells defined along two dimensions: (i) education (college graduate or non-graduate) and (ii) whether they reported one year previously that they “might move” for specifically job-related reasons.

[Table A8 here]

For both education groups, individuals who claimed they “might move” were about 20 percentage points more likely to do so in the subsequent year; and this number does not differ significantly by education. This suggests that there is no systematic difference by education in the accuracy of these subjective expectations.

# G Supplementary theoretical derivations and results

In this Appendix, I derive three theoretical results to support the argument in the main text. The first is to show that $\Omega(\varepsilon)$ can be expressed in terms of the *expected* rents - under certain distributional assumptions. This is equation (24) in Section 5.1 in the main text. The second is to show that realized amenity costs, conditional on accepting a non-home job offer, are increasing in both $\sigma^w$ and $\sigma^c$. And the third is to show the differential in expected wage rents between non-home and home area matches can serve as a lower bound on the expected realized amenity costs - as I argue in Section 6.1 in the main text.

## G.1 Derivation of equation (24) in Section 5.1 : expected rents

The aim here is to derive the following approximation for $\Omega(\varepsilon)$:

$$\Omega(\varepsilon) \approx \frac{\sigma^w}{\sigma^c}\mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' \geq \varepsilon\right]$$

where the operator $\mathbb{E}$ denotes the conditional expectation of the match productivity draw from the distribution $F^w$. To derive this expression, I first assume that $\varepsilon^c$ is distributed uniformly with a minimum at 0 and maximum normalized to 1. Based on (21), this implies that:

$$\begin{aligned}
\Omega(\varepsilon) &= \int_0^1 \left[\frac{1-F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1-F^w(\varepsilon)}\right] d\varepsilon^c \\
&= \frac{\sigma^w}{\sigma^c}\int_0^1 \left[\frac{1-F^w(\varepsilon+x)}{1-F^w(\varepsilon)}\right] dx
\end{aligned} \tag{A1}$$

where I have defined $x \equiv \frac{\sigma^c}{\sigma^w}\varepsilon^c$, so $d\varepsilon^c = \frac{\sigma^w}{\sigma^c}dx$. Now, suppose that very few job offers are accepted at the maximum amenity cost draw. That is, $\frac{F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\right)}{F^w(\varepsilon)}$ is close to 1 for all values of $\varepsilon$. (A1) can then be approximated as:

$$\begin{aligned}
\Omega(\varepsilon) &\approx \frac{\sigma^w}{\sigma^c}\int_0^\infty \left[\frac{1-F^w(\varepsilon+x)}{1-F^w(\varepsilon)}\right] dx \\
&= \frac{\sigma^w}{\sigma^c}\int_0^\infty x\frac{f^w(x)}{1-F^w(x)}dx
\end{aligned} \tag{A2}$$

where the second line follows from integration by parts. And $\int_0^\infty x\frac{f^w(x)}{1-F^w(x)}dx$ is equal to the conditional expectation $\mathbb{E}[\varepsilon'-\varepsilon|\varepsilon'\geq\varepsilon]$.

## G.2 Impact of $\sigma^w$ and $\sigma^c$ on $Z(c|\varepsilon)$, the distribution of realized amenity costs (Section 6.1)

To show there is a dominating transformation of $Z(c|\varepsilon)$ by the first order stochastic criterion, it is sufficient to demonstrate dominance by the hazard rate criterion. For given initial match quality $\varepsilon$ and amenity cost $c$, the hazard rate of $Z(c|\varepsilon)$ is:

$$\frac{z(c|\varepsilon)}{1-Z(c|\varepsilon)} = \left[\sigma^c\int_{\frac{c}{\sigma^c}}^\infty \frac{1-F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1-F^w\left(\varepsilon+\frac{c}{\sigma^w}\right)}dF^c(\varepsilon^c)\right]^{-1} \tag{A3}$$

The assumption that $F^w$ has a monotonically increasing hazard rate ensures that $\frac{1-F^w\left(\varepsilon+\frac{\sigma^c}{\sigma^w}\varepsilon^c\right)}{1-F^w\left(\varepsilon+\frac{c}{\sigma^w}\right)}$ is increasing in $\sigma^w$, conditional on $\varepsilon$ and $\varepsilon^c$. If follows that $\frac{z(c|\varepsilon)}{1-Z(c|\varepsilon)}$ is decreasing in $\sigma^w$ at every $c$, given $\varepsilon$. That is, there is a hazard rate dominating transformation of the $Z(c|\varepsilon)$ distribution.

It is also clear by inspection that $\frac{z(c|\varepsilon)}{1-Z(c|\varepsilon)}$ is decreasing in $\sigma^c$ at every $c$, given $\varepsilon$. Again, this represents a hazard rate dominating transformation.

## G.3 Lower bound on expected amenity costs (Section 6.1)

Following the argument given in Section 6.1, it suffices to show that the expression in the curly brackets in equation (31) is less than $c$, i.e.:

$$\mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' - \varepsilon \geq c\right] - \mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' - \varepsilon \geq 0\right] \leq c \tag{A4}$$

conditional on the initial match quality $\varepsilon$ - where the operator $\mathbb{E}$ denotes the conditional expectation of the match productivity draw from the distribution $F^w$. This can usefully be rearranged as:

$$\mathbb{E}\left[\varepsilon' - \varepsilon - c | \varepsilon' - \varepsilon - c \geq 0\right] \leq \mathbb{E}\left[\varepsilon' - \varepsilon | \varepsilon' - \varepsilon \geq 0\right] \tag{A5}$$

Since I have assumed the amenity draw $c$ always exceeds zero, it is sufficient to show that:

$$\frac{d}{dx} \log \mathbb{E}\left[\varepsilon' - x | \varepsilon' - x \geq 0\right] \leq 0 \tag{A6}$$

for all $x \equiv \varepsilon + c$. Writing this in terms of the match distribution $F^w$:

$$\begin{aligned} \frac{d}{dx} \log \mathbb{E}\left[\varepsilon' - x | \varepsilon' - x \geq 0\right] &= \frac{d}{dx} \log \frac{\int_x^\infty \varepsilon f^w(\varepsilon) d\varepsilon}{1 - F^w(x)} \\ &= \frac{d}{dx} \log \frac{\int_x^\infty [1 - F^w(\varepsilon)] d\varepsilon}{1 - F^w(x)} \\ &= -\frac{1 - F^w(x)}{\int_x^\infty \left[\frac{1 - F^w(\varepsilon)}{f^w(\varepsilon)}\right] f^w(\varepsilon) \varepsilon} + \frac{f^w(x)}{1 - F^w(x)} \end{aligned} \tag{A7}$$

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

$$\frac{d}{dx} \log \mathbb{E}\left[\varepsilon' - x | \varepsilon' - x \geq 0\right] \leq -\frac{1 - F^w(x)}{\int_x^\infty \left[\frac{1 - F^w(x)}{f^w(x)}\right] f^w(\varepsilon) \varepsilon} + \frac{f^w(x)}{1 - F^w(x)} = 0 \tag{A8}$$

so equation (A6) is satisfied.

# Appendix references

**Kaplan, Greg, and Sam Schulhofer-Wohl.** 2012*a*. "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.** 2012*b*. "Understanding the Long-Run Decline in Interstate Migration." NBER Working Paper No. 18507.

**Kennan, John.** 2015. "Spatial Variation in Higher Education Financing and the Supply of College Graduates." http://www.ssc.wisc.edu/∼jkennan.

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

**King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick.** 2010. "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0 [Machine-readable database]." Minneapolis: University of Minnesota.

**Kleiner, Morris M.** 2000. "Occupational Licensing." *The Journal of Economic Perspectives*, 14(4): 189–202.

**McLanahan, Sara.** 1985. "Family Structure and the Reproduction of Poverty." *American Journal of Sociology*, 90(4): 873–901.

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

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

**Molloy, Raven, Christopher L. Smith, and Abigail Wozniak.** 2014. "Declining Migration Within the US: The Role of the Labor Market." NBER Working Paper No. 20065.

# Appendix tables and figures

Table A1: Migration rates (%) for all individuals and household top earners

| | 25-34s | | 35-44s | | 45-64s | |
|---|---|---|---|---|---|---|
| | All indiv's | Top earners | All indiv's | Top earners | All indiv's | Top earners |
| HS dropout | 2.08 | 2.04 | 1.38 | 1.44 | 0.80 | 0.76 |
| HS graduate | 2.50 | 2.42 | 1.44 | 1.41 | 0.87 | 0.85 |
| Some college | 3.01 | 3.04 | 1.67 | 1.60 | 1.10 | 1.07 |
| Undergrad | 4.53 | 4.69 | 2.06 | 2.03 | 1.08 | 1.14 |
| Postgrad | 6.77 | 7.19 | 2.82 | 2.98 | 1.28 | 1.42 |

This table reports annual cross-state migration rates (by age and education) separately for all individuals and household top earners, based on CPS March waves between 1999 and 2015. See Section A.1 for further sample details.

Table A2: Breakdown of migration motivations for household top earners aged 25-64

| | State moves | | County moves (within states) | |
|---|---|---|---|---|
| Primary reason | % full sample | % state migrants | % full sample | % county migrants |
| JOB-MOTIVATED | 0.92 | 51.75 | 0.56 | 30.94 |
| New job or job transfer | 0.69 | 38.84 | 0.30 | 16.52 |
| Easier commute | 0.05 | 2.93 | 0.17 | 9.11 |
| Looking for work | 0.09 | 5.01 | 0.05 | 2.52 |
| Other job-related reasons | 0.09 | 4.96 | 0.05 | 2.79 |
| NON-JOB | 0.86 | 48.25 | 1.26 | 69.06 |
| Family | 0.40 | 22.71 | 0.49 | 26.69 |
| Change in marital status | 0.08 | 4.29 | 0.15 | 8.21 |
| Establish own household | 0.05 | 3.00 | 0.13 | 6.89 |
| Other family reasons | 0.27 | 15.42 | 0.21 | 11.60 |
| Housing | 0.21 | 11.81 | 0.57 | 31.18 |
| Want to own home | 0.04 | 2.03 | 0.15 | 8.22 |
| New or better housing | 0.05 | 2.67 | 0.17 | 9.63 |
| Cheaper housing | 0.05 | 2.82 | 0.10 | 5.26 |
| Other housing reasons | 0.08 | 4.30 | 0.15 | 8.08 |
| Amenities | 0.16 | 7.41 | 0.16 | 6.41 |
| Better neighborhood | 0.03 | 1.83 | 0.08 | 4.21 |
| Climate, health, retirement | 0.10 | 5.58 | 0.04 | 2.2 |
| Attend/leave college | 0.05 | 3.00 | 0.03 | 1.75 |
| Other reasons | 0.06 | 3.32 | 0.06 | 3.03 |
| ALL REASONS | 1.78 | 100 | 1.82 | 100 |

This table presents migration rates for household top earners aged 25-64, by primary reason in CPS March waves between 1999 and 2015. See Section A.1 for further sample details. 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 gives the percentage of *state-movers* reporting each reason. The final two columns repeat the exercise for cross-county moves within states. 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.

Table A3: Log point responses of job-motivated and non-job migration rates

| | JOB-MOTIVATED | | | NON-JOB | | | NON-JOB EXCL. COLLEGE | | |
|---|---|---|---|---|---|---|---|---|---|
| | 25-34 | 35-44 | 45-64 | 25-34 | 35-44 | 45-64 | 25-34 | 35-44 | 45-64 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Specification 1: no demographic controls | | | | | | | | | |
| HS graduate | 0.335*** | 0.016 | 0.255* | 0.050 | -0.060 | 0.070 | 0.016 | -0.070 | 0.064 |
| | (0.109) | (0.140) | (0.142) | (0.091) | (0.104) | (0.093) | (0.092) | (0.104) | (0.093) |
| Some college | 0.667*** | 0.290** | 0.554*** | 0.189** | -0.029 | 0.282*** | 0.070 | -0.046 | 0.271*** |
| | (0.106) | (0.130) | (0.140) | (0.089) | (0.105) | (0.091) | (0.091) | (0.106) | (0.091) |
| Undergrad | 1.275*** | 0.840*** | 1.065*** | 0.403*** | -0.174 | 0.017 | 0.146 | -0.213* | 0.007 |
| | (0.102) | (0.127) | (0.139) | (0.089) | (0.111) | (0.100) | (0.092) | (0.112) | (0.100) |
| Postgrad | 1.910*** | 1.423*** | 1.373*** | 0.435*** | -0.236* | 0.092 | 0.145 | -0.296** | 0.061 |
| | (0.104) | (0.128) | (0.139) | (0.102) | (0.127) | (0.105) | (0.109) | (0.129) | (0.106) |
| Specification 2: demographic controls | | | | | | | | | |
| HS graduate | 0.269** | -0.055 | 0.162 | -0.060 | -0.176 | 0.030 | -0.090 | -0.186* | 0.024 |
| | (0.113) | (0.146) | (0.143) | (0.095) | (0.111) | (0.096) | (0.096) | (0.112) | (0.096) |
| Some college | 0.604*** | 0.235* | 0.481*** | 0.047 | -0.155 | 0.249*** | -0.066 | -0.173 | 0.239** |
| | (0.112) | (0.132) | (0.141) | (0.097) | (0.114) | (0.094) | (0.100) | (0.114) | (0.094) |
| Undergrad | 1.110*** | 0.721*** | 0.982*** | 0.220** | -0.246** | 0.066 | -0.017 | -0.286** | 0.056 |
| | (0.111) | (0.130) | (0.141) | (0.099) | (0.120) | (0.104) | (0.104) | (0.121) | (0.104) |
| Postgrad | 1.794*** | 1.292*** | 1.343*** | 0.342*** | -0.274** | 0.162 | 0.052 | -0.336** | 0.131 |
| | (0.114) | (0.133) | (0.140) | (0.113) | (0.136) | (0.109) | (0.121) | (0.138) | (0.110) |
| Observations | 207,706 | 264,521 | 429,193 | 207,706 | 264,521 | 429,193 | 207,706 | 264,521 | 429,193 |
| Mig rate (%) | 2.144 | 1.016 | 0.438 | 1.487 | 0.788 | 0.598 | 1.286 | 0.771 | 0.591 |

Each column reports education effects from complementary log-log regressions on job-motivated (columns 1-3) and non-job migration incidence (columns 4-6) across states. I also study the effect on an indicator for any cross-state non-job move *excluding attending/leaving college* (columns 7-9). I report results separately for three age groups. Coefficients should be interpreted as the log point effect of a particular level of education (relative to high-school dropout, the omitted category) on the instantaneous migration rate, conditional on the empirical model described by equation (36). The sample consists of household top earners aged 25 to 64 in CPS March waves between 1999 and 2015; see Section A.1 for further sample details. I include results for specifications both with and without detailed demographic controls: age, age squared, black and Hispanic race dummies, immigration status, marital status, a range of indicators for number of own children, and a gender indicator which is also interacted with all previously mentioned variables. All specifications control for a set of year fixed effects (for the individual CPS cross-sections). Robust SEs in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Log point responses of migration rate, by detailed reported reason

| | CROSS-STATE | | | | CROSS-COUNTY WITHIN STATE | | | |
|---|---|---|---|---|---|---|---|---|
| Primary reason | HS grad | Some coll | Coll grad | Post grad | HS grad | Some coll | Coll grad | Post grad |
| All reasons | 0.003 | 0.203*** | 0.443*** | 0.841*** | -0.097** | 0.044 | -0.015 | 0.000 |
| | (0.046) | (0.045) | (0.046) | (0.048) | (0.038) | (0.038) | (0.040) | (0.046) |
| JOB-RELATED | | | | | | | | |
| All job-related reasons | 0.132* | 0.441*** | 0.918*** | 1.491*** | 0.067 | 0.361*** | 0.550*** | 0.726*** |
| | (0.076) | (0.073) | (0.072) | (0.074) | (0.081) | (0.079) | (0.080) | (0.087) |
| New job/transfer | 0.342*** | 0.797*** | 1.364*** | 2.008*** | 0.238* | 0.652*** | 1.009*** | 1.264*** |
| | (0.106) | (0.104) | (0.103) | (0.104) | (0.137) | (0.135) | (0.133) | (0.142) |
| Commute | -0.244 | 0.017 | 0.193 | 0.543* | 0.197 | 0.407*** | 0.463*** | 0.562*** |
| | (0.274) | (0.269) | (0.271) | (0.289) | (0.140) | (0.137) | (0.142) | (0.154) |
| Look for work | -0.168 | -0.407*** | -0.336** | -0.479** | -0.343 | -0.444** | -0.807*** | -0.844*** |
| | (0.166) | (0.156) | (0.170) | (0.235) | (0.216) | (0.206) | (0.249) | (0.296) |
| Other job reasons | 0.155 | 0.327* | 0.598*** | 0.990*** | -0.295 | 0.074 | 0.014 | -0.008 |
| | (0.201) | (0.197) | (0.197) | (0.208) | (0.233) | (0.222) | (0.236) | (0.269) |
| NON-JOB | | | | | | | | |
| All non-job reasons | -0.065 | 0.049 | 0.012 | 0.068 | -0.154*** | -0.071 | -0.258*** | -0.349*** |
| | (0.058) | (0.058) | (0.061) | (0.067) | (0.044) | (0.044) | (0.047) | (0.055) |
| Change in marital status | 0.004 | 0.133 | -0.180 | 0.014 | 0.447*** | 0.548*** | 0.137 | 0.209 |
| | (0.214) | (0.215) | (0.231) | (0.246) | (0.157) | (0.154) | (0.160) | (0.182) |
| Establish own household | -0.011 | 0.179 | 0.118 | 0.023 | -0.004 | 0.105 | -0.193 | -0.326* |
| | (0.243) | (0.252) | (0.257) | (0.293) | (0.139) | (0.140) | (0.155) | (0.193) |
| Other family reasons | -0.112 | -0.087 | -0.223** | -0.411*** | -0.320*** | -0.292*** | -0.559*** | -0.829*** |
| | (0.095) | (0.098) | (0.105) | (0.123) | (0.097) | (0.101) | (0.109) | (0.138) |
| Want to own home | 0.170 | 0.431 | 0.700** | 0.887** | 0.002 | 0.294** | 0.372*** | 0.305** |
| | (0.328) | (0.328) | (0.334) | (0.358) | (0.137) | (0.136) | (0.137) | (0.152) |
| New or better housing | 0.059 | -0.038 | 0.153 | 0.007 | -0.021 | 0.005 | 0.041 | 0.032 |
| | (0.244) | (0.247) | (0.248) | (0.269) | (0.119) | (0.120) | (0.126) | (0.141) |
| Cheaper housing | -0.011 | -0.233 | -0.797*** | -0.521* | -0.349** | -0.445*** | -0.783*** | -1.142*** |
| | (0.226) | (0.232) | (0.265) | (0.307) | (0.143) | (0.150) | (0.167) | (0.239) |
| Other housing reasons | -0.397** | -0.283 | -0.217 | 0.140 | -0.388*** | -0.430*** | -0.677*** | -0.626*** |
| | (0.187) | (0.189) | (0.199) | (0.212) | (0.131) | (0.133) | (0.149) | (0.161) |
| Better neighborhood | 0.343 | 0.389 | 0.060 | 0.007 | -0.115 | -0.069 | -0.307* | -0.604*** |
| | (0.257) | (0.259) | (0.281) | (0.328) | (0.161) | (0.160) | (0.179) | (0.213) |
| Climate, health, retirement | -0.111 | 0.201 | -0.089 | 0.034 | -0.393* | -0.424** | -0.764*** | -0.379 |
| | (0.155) | (0.153) | (0.167) | (0.179) | (0.203) | (0.209) | (0.251) | (0.260) |
| Attend or leave college | 2.353** | 3.462*** | 4.189*** | 4.614*** | 0.135 | 1.628*** | 1.700*** | 1.631*** |
| | (1.017) | (0.998) | (0.997) | (1.001) | (0.554) | (0.493) | (0.496) | (0.525) |
| Other reasons | -0.069 | 0.082 | 0.364* | 0.497** | -0.410** | -0.321* | -0.605*** | -0.847*** |
| | (0.210) | (0.211) | (0.218) | (0.234) | (0.187) | (0.191) | (0.209) | (0.257) |

This table reports education effects from complementary log-log regressions on annual migration incidence, estimated separately for (i) cross-state moves and (ii) cross-county moves within states. Each row reports the effects on moving for the motivation specified, with the first row presenting education effects on the overall migration incidence (all reasons). The first four columns gives results for cross-state migration and the final four for cross-county migration within states. Coefficients should be interpreted as the log point effect of a particular level of education (relative to high-school dropout, the omitted category) on the instantaneous migration rate, conditional on the empirical model described by equation (36). The sample consists of household top earners aged 25 to 64 in CPS March waves between 1999 and 2015; see Section A.1 for further sample details. The sample size in each regression is 901,420. Each regression controls for a detailed set of individual characteristics: age, age squared, black and Hispanic race dummies, immigration status, marital status, a range of indicators for number of own children, a gender indicator which is also interacted with all previously mentioned variables, and a set of year fixed effects (for the individual CPS cross-sections). 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. Robust SEs in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

## Table A5: Net cross-state migration rates by age and education

| | 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 aged 25-34* | | | | | | | | | |
| HS dropout | 2.71 | 0.42 | 0.15 | 2.50 | 1.24 | 0.49 | 2.50 | 1.60 | 0.64 |
| HS graduate | 3.27 | 0.37 | 0.11 | 2.82 | 0.96 | 0.34 | 2.82 | 1.36 | 0.48 |
| Some college | 3.84 | 0.34 | 0.09 | 3.29 | 1.11 | 0.34 | 3.29 | 1.64 | 0.50 |
| Undergraduate | 5.67 | 0.42 | 0.07 | 5.02 | 1.31 | 0.26 | 5.02 | 1.96 | 0.39 |
| Postgraduate | 8.09 | 0.78 | 0.10 | 7.62 | 1.97 | 0.26 | 7.62 | 2.87 | 0.38 |
| *Individuals aged 35-64* | | | | | | | | | |
| HS dropout | 1.49 | 0.24 | 0.16 | 1.23 | 0.57 | 0.46 | 1.23 | 0.74 | 0.60 |
| HS graduate | 1.51 | 0.24 | 0.16 | 1.16 | 0.36 | 0.31 | 1.16 | 0.51 | 0.44 |
| Some college | 1.84 | 0.27 | 0.15 | 1.46 | 0.47 | 0.32 | 1.46 | 0.69 | 0.48 |
| Undergraduate | 2.02 | 0.24 | 0.12 | 1.69 | 0.47 | 0.28 | 1.69 | 0.70 | 0.42 |
| Postgraduate | 2.45 | 0.26 | 0.11 | 2.16 | 0.56 | 0.26 | 2.16 | 0.80 | 0.37 |

This table reports annual gross and net cross-state migration rates within education groups, separately for individuals aged 25-34 and 35-64. See the notes under Table 1 for details on samples and construction of variables.

## Table A6: Cross-state migration rates for 25-34s: students and return moves

| | HS dropout | HS grad | Some coll | Undergrad | Postgrad |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| % recent students by education | 3.92 | 1.96 | 3.53 | 1.61 | 2.76 |
| Migration rate: all cross-state moves (%) | | | | | |
| Full sample | 2.63 | 2.74 | 4.51 | 6.45 | 8.56 |
| Excl. recent students | 2.68 | 2.7 | 4.56 | 6.36 | 8.24 |
| Migration rate: return moves (%) | 1.17 | 1.45 | 1.74 | 2.56 | 3.31 |
| Migration rate: non-return moves (%) | 1.46 | 1.29 | 2.77 | 3.89 | 5.25 |
| Observations | 1,711 | 3,170 | 1,840 | 1,054 | 362 |

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 includes all household heads aged 25-34 residing in the US in the previous wave. 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.

Table A7: Job finding and separation rates: SIPP

| | Labor force participants | | | All individuals aged 25-64 | |
|---|---|---|---|---|---|
| | No job to job | Job to no job | Job to new job | No job to job | Job to no job |
| | $\rho(\varepsilon_R)$ | $\delta$ | $\int_\varepsilon \rho(\varepsilon)\, dG$ | $\rho(\varepsilon_R)$ | $\delta$ |
| | (1) | (2) | (3) | (4) | (5) |
| HS dropout | 39.49 | 3.78 | 5.58 | 7.46 | 7.24 |
| HS graduate | 42.41 | 2.61 | 4.58 | 9.97 | 5.11 |
| Some college | 44.21 | 2.40 | 4.68 | 12.14 | 4.64 |
| Undergraduate | 49.34 | 1.82 | 4.26 | 14.00 | 3.68 |
| Postgraduate | 53.71 | 1.41 | 3.61 | 14.99 | 2.97 |

This table reports four-month job transition rates, based on the 1996, 2001, 2004 and 2008 panels of the SIPP, which cover the period between 1996 and 2013. Column 1 gives the percentage of unemployed workers at the end of wave *t-1* who are employed at the end of wave *t* (4 months later); and vice versa for column 2. Column 3 reports the percentage of employed workers at the end of *t-1* who have a new job at the end of *t*. Columns 4-5 report transition rates from joblessness to employment and vice versa among all individuals - i.e. including the economically inactive. Throughout, I exclude workers with multiple jobs or business income at the end of each wave. The full sample consists of 1.4m observations.

Table A8: Difference-in-difference for job-motivated moving rate

*Share who moved residence for job reasons in previous 12 months*

| | Might move for job (lagged 1 year) | | Difference |
|---|---|---|---|
| | No | Yes | |
| Non-graduate | 0.017 | 0.214 | 0.197 |
| | | | (0.013) |
| College graduate | 0.029 | 0.238 | 0.209 |
| | | | (0.017) |
| Difference | 0.012 | 0.024 | 0.012 |
| | (0.003) | (0.021) | (0.022) |

This table reports sample shares who moved residence (any distance) for self-reported job reasons in the previous 12 months, by (i) education and (ii) whether they reported one year previously that they "might move" for specifically job-related reasons. Robust standard errors are in parentheses, clustered by individual. The full sample, based on household heads in the PSID between 1970 and 1980, consists of 42,287 observations.

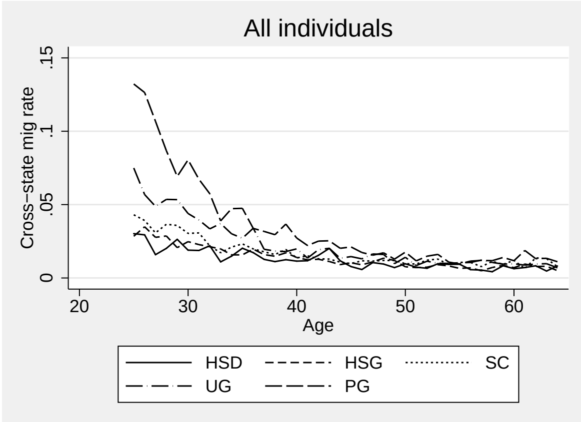


Figure A1: Cross-state migration rate by single-year age (CPS 1999-2015)

This figure reports annual rates of cross-state migration for household top earners, by single-year age and education: high school dropout (less than 12 years of schooling), high school graduate (12 years of schooling), some college (between 1 and 3 years of college), undergraduate (4 years of college) and postgraduate (5 or more years of college). See Appendix A.1 for further details on sample.

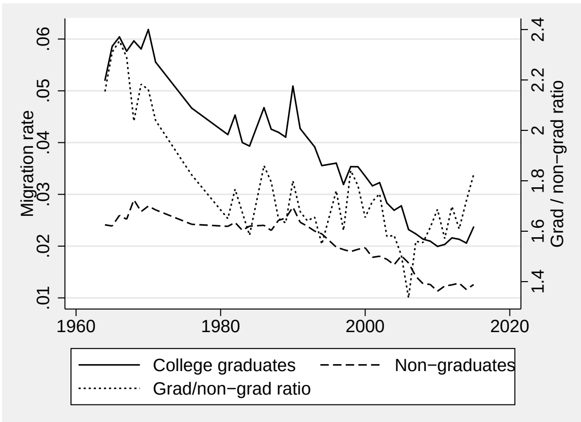


Figure A2: Annual cross-state migration rates by education: 1963-2015

This figure reports annual rates of cross-state migration over time using the CPS, separately for college graduates and non-graduates. The right-hand scale gives the ratio of the two. The sample is based on all individuals aged 25-64, excluding military households. I also omit observations with imputed migration observations.

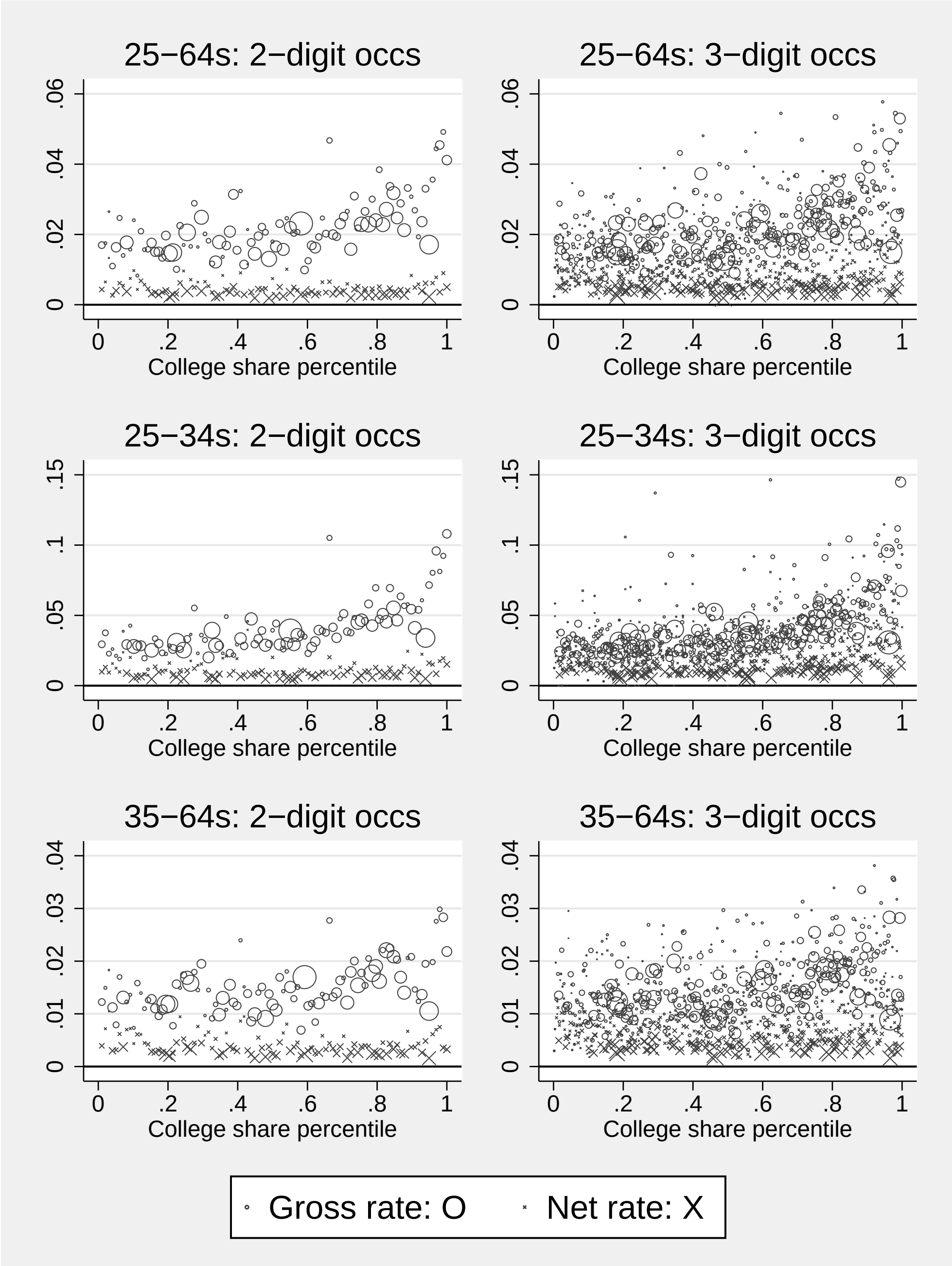


Figure A3: Annual gross and net cross-state migration by occupation

This figure reports annual gross and net cross-state migration rates within detailed occupation groups, based on employed civilians in the ACS between 2000 and 2009. Within each occupation group, 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 $j$. Occupations are based on the 2000 census scheme. I report estimates separately for 2-digit (left column) and 3-digit (right) occupations, and separately for individuals aged 25-64 (i.e. the full sample, top row), 25-34s (middle) and 35-64s (bottom). The size of each marker is proportional to the occupation's employment sample.

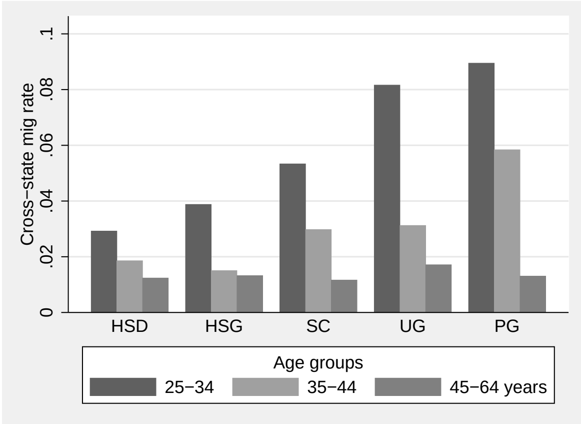


Figure A4: Annual cross-state migration rates (PSID 1970-80)

This figure reports the share of household heads who moved state in the previous 12 months. 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.

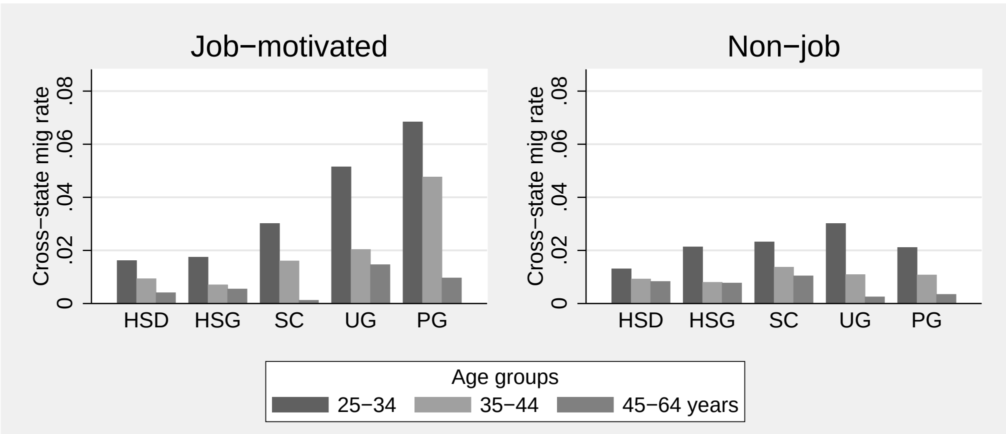


Figure A5: Annual cross-state migration rates by reported reason (PSID 1970-80)

The first panel reports the fraction of household heads who moved state primarily for job-related reasons in the previous 12 months; and the second panel does the same for non-job reasons. For 18 percent of cross-state moves, the PSID reports the reason to be "ambiguous" or "mixed" or simply unknown. I allocate these cases to the job-motivated and non-job categories within age-education cells, according to the proportions in the non-ambiguous data.