Why don't less educated workers move? The role of job search in migration decisions

Maria Balgova*

October, 2018

Click here for most recent version

Abstract

I establish a new stylised fact showing that less educated workers are not only less mobile, they are also significantly less likely to move with a job in hand. Using evidence based on US panel data, I argue that a large portion of the observed differences in migration behaviour is driven by the differences in workers' employment options in other regions. Compared to college graduates, less educated workers find job search in more distant regions much more difficult. This limits their options to move for guaranteed employment, forcing them to move speculatively, and thereby reducing their overall mobility. I develop these results in two stages. First, I adapt the recent literature from empirical IO on discrete choice models with heteregeneous option sets to isolate the impact of differences in employment opportunities on migration decisions. Second, I extend the standard model of job search (McCall (1970)) to multiple residential locations. I estimate this model to quantify the size of cross-regional job search frictions, finding that they can explain approximately half of the migration propensity gap between the more and the less educated. This result opens a new policy channel in addressing regional differences and those left behind: the importance of the ability to find a job before moving suggests a large social return to improving regional search and matching for less educated groups.

*maria.balgova@ccc.ox.ac.uk. Department of Economics, University of Oxford. I am indebted to my supervisors Margaret Stevens and Abi Adams for their guidance and encouragement. I would also like to thank Steve Bond, Tony Venables, Peter Neary, Simon Quinn, Ian Crawford, Jean Flemming, Alex Plekhanov, and the conference and seminar participants at the Search and Matching Conference (University of Cambridge), Oxford Applied Micro Seminar, Royal Economic Society Symposium (University of Sussex), and ZEW Summer School (Mannheim) for advice and comments. This work would not be possible without the generous support of ESRC and Wolf Scholarship (Corpus Christi College).

1 Introduction

Less educated individuals move less: compared to their college-educated counterparts, they are up to three times less likely to move into another region of the same country[1] (Greenwood (1997)). This is despite the evidence that benefits to migration can be large. In the short run, moving to another region makes it possible for an individual to weather negative local shocks, while in the long run, migration can significantly improve the lifetime earnings of the movers and their offspring alike (Bound and Holzer (1996), Kennan and Walker (2011), Chetty and Hendren (2018)). Both of these effects have been found to be larger for the less educated: they are more exposed to regional economic fluctuations, and the impact of where a child grows up on her future earnings is particularly strong for low-income households (Gregg et al. (2004), Hoynes (1999), Bartik (2018)).This makes the low migration propensity of the less educated a puzzle.

The existing literature explains this low mobility as a reflection of small net expected gains to moving. When deciding whether to move, the worker weighs up expected utility in different regions, which depends on factors such as expected income, location preferences, and migration costs. The existing empirical evidence shows that these are less favourable for the less educated: they tend to have lower and less dispersed wages, different amenity preferences, and larger migration costs. As a result, their expected returns to moving are relatively small, and they move relatively infrequently.

In this paper, I argue that the literature has overlooked a crucial mechanism: cross-regional job search frictions. They not only change the expected returns to moving, but also the way migration decisions are made. I show that college graduates usually move for a specific job, while the less educated tend to move speculatively, without a job offer in hand. Using reduced-form and structural evidence, I argue that the less educated workers find job search in more distant regions much more difficult. This diminishes their options to move for a specific job, forcing them to move under greater uncertainty, and thereby reducing their overall mobility.

The analysis is based on a monthly panel of a sample of US inhabitants from 1996 to 1999 (SIPP). The monthly data on employment and location, combined with the fact that the survey follows respondents when they move, allows me to infer whether a particular move was for a specific job or speculative. I analyse these outcomes by drawing on other labour market information about the workers and matching it with proxy data on recruitment strategies of firms. The paper proceeds in two steps.

First, I extend the growing literature on discrete choice models with heterogeneous choice sets to test whether the less educated have fewer opportunities to migrate with a job in hand. Speculative migration is independent of employers' decisions, but migra-

[1]Based on inter-state migration of a sample of US adults, 1996-1999.

tion for a specific job is impossible without first receiving a job offer. This means that an individual's choice set is determined by the recruitment behaviour of her employers. However, because these choice sets are unobserved, I estimate their impact by embedding them in a model of joint location and employment choice. Workers' choice sets are identified using sectoral and occupational variation in cross-regional hiring; I also control for inter-personal variation in wages, living costs, and migration costs, which were shown to drive the migration and employment decisions. The model is estimated as a conditional logit with unobserved heterogeneous choice sets using maximum likelihood.

I find that the less and the more educated face significantly different opportunities to move for a specific job. The more educated are more likely to work for large and more spatially concentrated companies, which tend to advertise their vacancies more broadly; as a result, the more educated are more likely to be offered a job in distant region. Overall, less than a half of all high school dropouts can choose to migrate for a specific job, but this increases to 95% for college graduates. This translates into large differences in migration propensity: high school dropouts would be about 40% more likely to move if they were given the same job search opportunities as the more educated.

Moreover, controlling for differences in opportunity is crucial for correct interpretation of the estimates of sensitivity to income variation and migration costs. A standard conditional logit suggests that the less educated are not as responsive to income as college graduates; letting choice sets vary virtually eliminates this gap. Intuitively, a model with heterogeneous choice sets allows for the possibility that the choice probability of employment away is low because that option is not available, not because the less educated don't care for the better paid employment in other regions. A standard logit implicitly assumes that all workers have the same choice set, which does not allow us to distinguish between workers' preferences and their options.

In the second part of this paper, I build a formal model of cross-regional job search, which allows me to estimate the frictions that drive the different employment opportunities of those with and without college degrees, as well as to quantify their importance alongside other explanations of migration. My point of departure is the seminal partial equilibrium model of job search (McCall (1970)), in which workers wait to receive random wage offers, and follow an optimal stopping rule in deciding whether to accept them or not. I add a location dimension, so that workers may receive job offers from other regions, as well as migrate speculatively. Because the model allows for different job-finding probabilities depending on the region of origin and destination of the vacancy, it can capture the idea that it is harder to search across regions than locally. I estimate the model using the method of simulated moments, using location-employment flows between regions to identify the model parameters.

The structural estimates of the model reveal considerable cross-regional search frictions. When searching on the job, the probability that a college graduate would receive a job offer from another region is about 30% smaller than the probability of receiving a local job offer. This gap is more than 60% for those without a college degree. Cross-regional job search when unemployed is even harder, with individuals of any education seeing less than 1% of the vacancies in distant regions.

The estimated cross-regional search frictions are also economically significant. The greater ability of the more educated to move for a specific job, as opposed to moving under income uncertainty, can explain almost 50% of the gap in migration propensity and three quarters of the difference in the type of migration. While the lower migration propensity of those without a college degree partly reflects their lower wages, higher migration costs, and different location preferences, it is also a consequence of a labour market that makes it much more difficult for them to find a job in another region without having to move there first.

Related literature The paper's main contribution is in modifying the way we model within-country migration. The standard assumption in the literature is that migration is a decision in which income is uncertain. The new stylised fact established in this paper demonstrates that the majority of moves are in fact for a specific job – after this uncertainty has been resolved. I show that not all workers are equally able to move this way, and that this difference in employment opportunity matters for migration. My theoretical model explicitly incorporates the process of job search alongside the choice of optimal location, allowing for both types of migration.

This is not the first paper to point to the two different ways of migrating, but it is the first paper to take them seriously. Possibly the first paper to discuss speculative vs "contracted" migration was a theoretical paper by Silvers (1977); different types of migration were also touched on in Amior (2015), Herzog et al. (1993) and Lutgen and der Linden (2015). In terms of the causal mechanism in question, my paper is closest to the literature on the role of information and networks in migration decisions (Mckenzie and Rapoport (2007), Patacchini and Zenou (2012), Wilson (2017)); also discussed in Gregg et al. (2004). However, to the best of my knowledge, this is the first paper to systematically study the types of regional migration and what they imply about the workers' decision-making process.

The findings of this paper help to bridge the literatures on local labour demand shocks and regional adjustment. There is substantial evidence on the role of mobility as a buffer to local demand shocks (Bartik (2018), Blanchard and Katz (1992)), but research also shows that the less educated, low-income households who are more exposed to these shocks are also least likely to move, exacerbating regional disparities

(Bound and Holzer (1996), Wozniak (2010), Hoynes (1999)). This paper explains that this is partly because the frictions between regional labour markets make it difficult for the less educated to migrate.

Understanding why and how the less educated move is essential for policy. Governments intervene in regions that have experienced prolonged periods of high unemployment and low growth, often by implementing job-creation policies such as hiring subsidies and special tax arrangements for new businesses. As the literature shows (Kline and Moretti (2013), Glaeser and Gottlieb (2008)), the welfare impact of such an intervention largely depends on workers' migration response. This paper deepens our understanding of that response. My results also suggest that, if the aim of place-based policies is in fact to improve welfare of the most vulnerable groups, improving the cross-regional labour market is a viable alternative to subsidies and taxation.

This paper also makes two contributions to methodology. First, the reduced-form model in the first part of the paper applies the concept of unobserved variation in choice sets, which was developed in empirical IO literature (Abaluck and Adams (2017), Goeree (2008)). I adapt the method to labour market behaviour: instead of capturing the set of products a consumer is aware of, I use it to model the sequencing of migration and job search decisions.

Second, the structural model in this paper is the first to combine both search-move and move-search strategies in one framework[2]. The existing literature assumes that individuals either only search after they migrate (Kennan and Walker (2011), Epifani and Gancia (2005), Kline and Moretti (2013)), or that migration can only happen as a result of a specific employer-employee match (Beaudry et al. (2014), Lutgen and der Linden (2015), Amior (2015)). Incorporating both enables a better understanding of the the interaction between job search and migration decisions.

2 Data

One of the main points of this paper is to distinguish between two types of migration: before and after job search. This is possible by using monthly panel data that follows the individuals when they move. This section describes this dataset, and defines the education categories and regions used in the analysis. I also describe the data on choice set variation.

The SIPP My main dataset comes from the 1996 wave of the Survey of Income and Program Participation (for a discussion of why I use this particular wave, see section

[2]Molho (2001) does include both, but his work focuses on analysing spatial equilibrium instead of migration behaviour, and he assume the migration strategy is a choice made by the individual as opposed to something partially determined by the labour market

A.1 in the Appendix). It tracks a nationally-representative sample of US inhabitants over the period of 4 years, providing monthly data on their income, employment, and residence, among others. I also use four supplementary datasets (County Business Patterns, IPUMS, the Bureau of Labor Statistics, FRED) to capture firms' recruitment behaviour and regional variables such as rents and unemployment rates. The full list of variables and their sources is summarised in Table 12 in the Appendix.

I use the Survey of Income and Program Participation (SIPP) because of its three key features: it is (i) a panel with (ii) monthly data that (iii) follows its respondents when they move. These three features are necessary to analyse the relationship between cross-regional job search and migration. I need panel data in order to observe the outcomes of the worker's job search, and I need information on her geographic movement to be able to study migration. Finally, the data needs to be of sufficiently high frequency to allow me to link the two, i.e. to be able to see the worker's labour market outcomes just after the move. SIPP is, to my best knowledge, the only major dataset that combines these three features (for a detailed comparison of the different data sources, see Hernández-Murillo et al. (2011)).

I restrict my attention to a subsample of adult working men. The aim of this paper is to understand the role of job search in migration decisions, which is why I exclude individuals below the age of 24 (and college-related migration) and above 50 (to avoid migration motivated by retirement), as well as those adults with no employment spell over the duration of the panel. I focus on men to reflect the fact that 75% of respondents in the data live in households where the primary earner is male and is thus likely to have a disproportionate influence on the migration decision.

The descriptive statistics on the subsample used in my analysis are summarised in Table 10 in the Appendix. There are 16,720 working-age males in my sample, about 13% of which did not finish high school, while a quarter of the sample has completed a four-year college degree or more. The table shows that there is considerable variation between the education groups in terms of individual and household income, employment, and migration propensity. At the same time, however, the three groups are similar in terms of their age, racial and urban profile, making it easier to abstract from these variables.

Throughout the paper, the outcome of primary interest is the location-employment combination (based on the four large US regions) chosen by the workers each month. Unsurprisingly, employment is more popular than unemployment, and the most frequent location choice corresponds to the most populous region of the US (the South). Cross-regional migration is a relatively rare event, with 2.5% of the workers choosing to become employed in an away region and 0.5% of opting for unemployment away, during the 4 years of the sample. A full breakdown of the distribution of individuals' choices can be seen in Table 11 in the Appendix.

Data on recruitment strategies In the absence of a detailed information on the geographic radius of recruitment, I draw on the findings of the human resources literature to construct three proxies: firm size, spatial concentration, and online vacancy posting.

The first two variables, sector-specific average firm size and geographic concentration, are constructed using firm data from the County Business Patterns dataset for the years 1996-1999. I calculate spatial concentration as a Gini index based on Ellison and Glaeser (1997). The third variable looks at what proportion of all vacancies, within each occupation, are advertised on the Internet (Hershbein and Kahn (2018)), looking at a more recent period (2007, 2010-2014).

I match these proxies on firm recruitment with the worker dataset. The SIPP provides detailed data on industry and occupation for each employed worker, distinguishing between 409 occupation categories and 233 different industries. I aggregate these into 21 occupation and 13 industry groups, using the internal grouping in the census codes for the occupation and industry variables. These are then matched with the occupation and industry data used in the CBP and online vacancy data. The result is a merged dataset which records recruitment proxies for each worker, based on her industry and occupation group. For more information, see section A.4 in the Appendix.

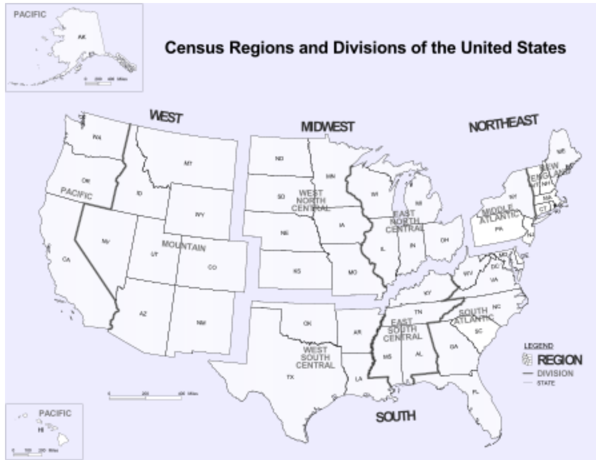
Defining education categories This paper works with three education categories: high school dropouts, high school graduates, and college graduates. They are based on SIPP's "eeducate" variable which captures the highest degree or grade completed. The mapping is the following:

- high school dropouts: individuals that have not received a high school diploma or equivalent
- high school graduates: individuals who have graduated from high school and may have vocational degrees or some college
- college graduates: individuals that hold a bachelor's degree (graduated from a four-year college) or more

In some parts of the paper, I merge the first two categories to only distinguish between the less educated (i.e. less than a four-year college degree) and the more educated (i.e. college graduates).

Defining migration The SIPP contains several levels of geographic information. Each household is located in a state and metropolitan area (if applicable), alongside a

Figure 1: The 4 large census regions of the USA.



Source: US Census Bureau, Geography Division.

within-household address index which changes every time a household moves, even if it is within the same neighborhood. The majority of moves fall into the latter category - 63 % of moves are within the same urban or rural area.[3]

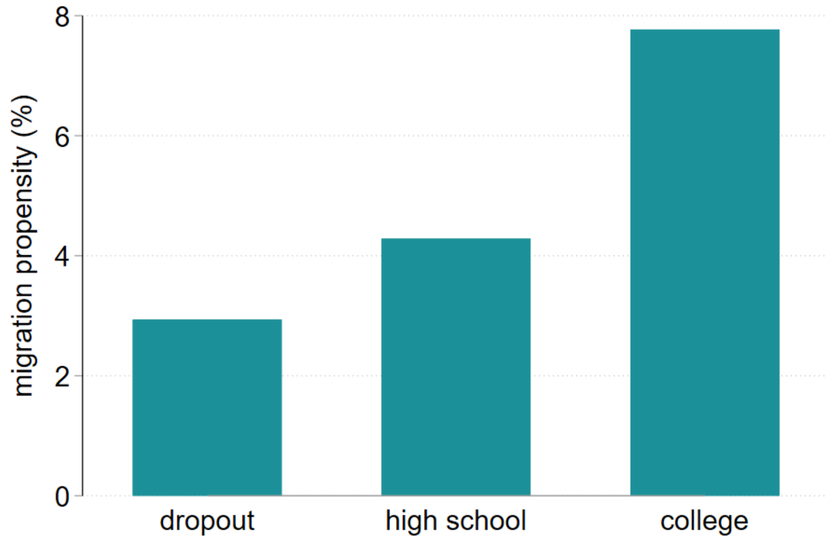
For the purposes of my analysis, I work with two definitions of migration: inter-state and cross-regional. Inter-state migration is widely used in literature when more detailed commuting zones are not available. In my dataset, about 13% of all moves are between states. Most of the analysis is conducted at the level of regions (Figure 1). Because of the larger geographic areas, this type of migration is somewhat rarer: just below 50% of all the inter-state moves are moves between regions.

3 Stylised facts about regional migration and education

This paper is motivated by two stylised facts: the less educated move less, and they are less likely to move for a specific job. The education differences in migration propensity are well-known in the literature (Greenwood (1997)), but the pattern in the type of migration has not been established before. In this section, I document both stylised facts in my dataset and explain the categorisation of migration into speculative and job-specific. I also discuss the wider implications of the different migration behaviour.

[3]The dataset distinguishes between urban and rural households, and identifies the metropolitan area of the urban households. A "rural area" is thus defined as the area of each state that is not urban. The 63% hence refers to moves within the same metropolitan area, plus moves between different parts of the same state which are rural.

Figure 2: Propensity to migrate into another state, US, 1996-1999.



Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Sample: adults (anyone over the age of 18). States refer to the 50 US states.

3.1 The propensity to migrate

Migration is a rare event. The probability that an adult in my sample of US inhabitants moves into a different state in the time period 1996-1999 is about 4.5%. In general, the inter-state migration propensity reported in the literature ranges between 3 and 7%, and has been in decline since the 1980s (Molloy et al. (2011)).

Despite that, a permanent feature of the migration statistics is the large gap in migration propensity between the more and the less educated (Greenwood (1997)). In my sample, college graduates are up to 3 times more likely to move into another states than high-school dropouts (Figure 2). These numbers are similar to the estimates in other studies: for example, Hernández-Murillo et al. (2011) report the migration propensity to be 2.6% for high school dropouts and 5.7% for college graduates.

This effect is robust to demographic characteristics, household composition, employment status, and distance migrated. The first two columns in Table 1 present the results of a binary logit model of the migration decision, in which the coefficient on education remains positive and statistically significant even after including a set of controls; further robustness checks can be found in Figure 17 in the Appendix.

3.2 Speculative migration vs. migration for a specific job

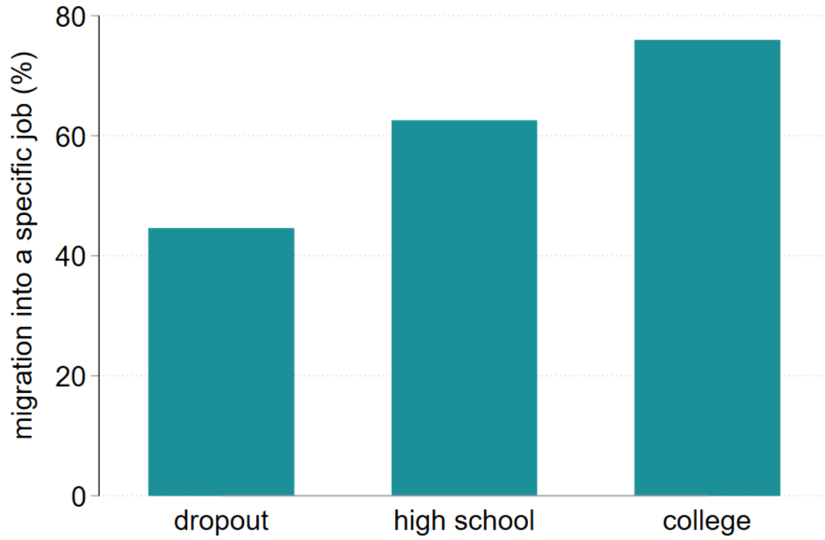
The second stylised fact motivating this paper is about *how* individuals move. My panel data allows me to distinguish between migration followed by employment (*migration*

Table 1: Binary logit models of propensity to migrate and the type of migration, between regions

	(1) migration propensity	(2) migration propensity	(3) speculative migration	(4) speculative migration
college	1.1003***	1.1191***	-0.7856**	-0.7035**
	(0.0969)	(0.0974)	(0.3222)	(0.3325)
employed in previous month	-1.0738***	-0.9871***	-1.7744***	-1.8898***
	(0.1368)	(0.1382)	(0.3894)	(0.4073)
married		-0.1835		-0.5011
		(0.1216)		(0.3824)
kids		-0.1064		-0.5243
		(0.1177)		(0.4043)
young		-0.5989***		0.4678
		(0.0980)		(0.3247)
unemployment rate			1.6547***	1.6317**
			(0.6373)	(0.6549)
constant	-6.4955***	-6.0552***	-0.4140	-0.1427
	(0.1297)	(0.1417)	(0.4101)	(0.5125)
N	16720	16720	428	428

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
The sample is restricted to working men between the age of 25 and 50. Columns (1) and (2) model the probability of moving into a different Census region. Columns (4) and (5) explain the likelihood that, conditional on moving, the migration is speculative (as opposed to for a specific job). "Employed in previous month" is a dummy variable equal to 1 if the worker was in employment the month before migration. "Unemployment rate" refers to individual-specific probability of being unemployed in any given month.

Figure 3: Probability of moving into a specific job (conditional on moving), US, 1996-1999.



Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Sample: adults (anyone over the age of 18). Migration for a specific job refers to migration followed by employment. States refer to the 50 US states.

for a specific job) and migration followed by unemployment (*speculative migration*) by looking at the timing of their migration and employment. I categorise all individuals that became employed within one month of their move migrated with a job in hand, while all migrants that were unemployed for at least a month after moving moved speculatively and had to search upon arrival.

The data shows that conditional on migration, college graduates are up to 3 times more likely to move into a specific job than high school dropouts (Figure 3). While more than half of high school dropouts that move across states have no job lined up, the number is only about a fifth for college graduates.

Table 1 establishes this stylised fact formally. I run a binary logit model of whether a migrant moves speculatively or not. The estimates show that the negative relationship between education and speculative migration is a robust one (see also figures 18, 19 and 20 in the Appendix). In particular, the less educated are less likely to move into a specific job (i.e. be employed after migration) even after controlling for their greater unemployment rate overall[4].

[4] In fact, there are two different effects that I have to control for: employment before moving, and overall unemployment rate. The unemployed move more (Herzog et al. (1993)), which is why the binary regressions in Table 1 include employment status in previous month. Moreover, since the unemployment rate of the less educated tends to be higher in general, I also have to control for this when estimating their propensity to migrate speculatively. This is captured by the individual-specific unemployment rate before moving ("unemployment rate").

Table 2: The list of options in the logit model of the employment-location choice.

option	employment status	region
1	employed	Northeast
2	unemployed	Northeast
3	employed	Midwest
4	unemployed	Midwest
5	employed	South
6	unemployed	South
7	employed	West
8	unemployed	West

An individual can decide to be employed or unemployed in each of the census regions of the US (conditional on that option being in her choice set). The dependent variable in the model is a dummy vector, the elements of which are equal to 1 if a particular option was selected, and 0 otherwise.

4 A descriptive model of migration

Are the two stylised facts – the relatively low migration propensity and the low probability of migrating into a specific job – about the migration of the less educated related? To answer this question, I estimate a descriptive, reduced-form model of migration and employment choice. Each worker makes optimal decisions about where and whether to work, but her options are limited by whether the firms in her occupation and industry hire cross-regionally. At the same time, employment-location options may differ in the offered income, living costs, and migration costs. This set-up allows me to decompose the observed behaviour into variation in the attributes of the different employment-location options, and variation in the availability of these options. The results show that, among other reasons, the less educated move less because they receive relatively fewer opportunities to do so. Moreover, I demonstrate that controlling for choice set heterogeneity is is important for producing unbiased estimates of workers' preferences.

4.1 Discrete choice model of location and employment

Typically, empirical literature on migration focuses on two decisions: whether to migrate, and where to migrate to. Conditional on location decision, models of the labour market then estimate the worker's decision whether to accept a job or remain unemployed.

In this paper, I nest these three choices within a single model of employment-location choice. Every period, the worker decides where to live and whether to work, which determines her employment status and may result in migration. Using the four census regions of the US as locations, the model works with eight outcome variables, listed in Table 2. The dependent variable is a 8x1 vector, the elements of which are equal to 1 if the given option was selected, and 0 otherwise.

This discrete choice model seeks to decompose the observed choice probabilities of these eight options into variation in the attributes of these options, and differences in the choice sets available to individuals. The interpretation of the model is based on the random utility model (RUM) framework. RUM assumes that the decision-maker is a utility-maximiser, so that the observed choice probabilities are proportional to the utility differences between options, once choice set heterogeneity is controlled for. As a result, the estimated coefficients are informative both about workers preferences over income, migration costs, etc., and how likely they are able to choose a given option.

Modelling utility: conditional choice probability First, I define the choice probability conditional on the option being available to the individual. It is a function of option attributes, X_{it}: monthly income, local rent and house prices (to capture living costs and amenities), and migration dummy, which is my catch-all variable for all types costs associated with moving. I also include a dummy for unemployment, to reflect the fact that just like migration, unemployment also often costs the worker more than just the pure financial loss of foregone income. This specification is based on the seminal model of migration by Sjaastad (1962), which argues that a worker decides to move if the the benefits (income, amenities) outweigh the costs (living costs and migration costs). Note that the model set-up is relevant even if not all the migration in my dataset is job-related – although Amior (2015) shows that more than a half of cross-state moves are primarily for employment reasons. For a detailed discussion of the incentives to migrate, see Appendix A.3.

The optimal decision is also based on factors unobservable by the economist, such as individual-specific location history, family ties, amenity preferences, and disutility of work. I denote these factors ϵ_{it}.

The worker chooses the option that gives her the highest utility from among the options available to her:

$$P_{is} = \text{Prob}(\beta X_{is} + \epsilon_{is} > \beta X_{it'} + \epsilon_{it'} \ \ \forall \ \ t' \neq s) \tag{1}$$

where P_{is} denotes the conditional probability of choosing a particular option s, and β captures the worker's preferences over the components of X_{is}. Assuming that ϵ_{is} is type I extreme value distributed, the conditional choice probability P_{is} has the standard

Table 3: Full and restricted choice sets. Example for a worker resident in the West.

options in...	
full choice set	restricted choice set
1 [employed, Northeast]	-
2 [unemployed, Northeast]	2 [unemployed, Northeast]
3 [employed, Midwest]	-
4 [unemployed, Midwest]	4 [unemployed, Midwest]
5 [employed, South]	-
6 [unemployed, South]	6 [unemployed, South]
7 [employed, West]	7 [employed, West]
8 [unemployed, West]	8 [unemployed, West]

The left hand column lists all the employment-location options available to a worker that does receive job offers from other regions. The right hand side column lists the options available to a worker that only receives job offers from her home region (is not capable of cross-regional job search).

logit closed-form solution (McFadden et al. (1973)):

$$P_{is} = \frac{exp(\beta X_{is})}{\sum_{t=1}^{8} exp(\beta X_{it})} \tag{2}$$

Modelling choice sets: recruitment behaviour The second part of the overall (unconditional) choice probability is the likelihood that a given option lies in the individual's choice set. This augmentation of the standard multinomial logit draws on the empirical IO literature (Abaluck and Adams (2017), Goeree (2008)), which explicitly recognises that the consumer's optimal choice may be skewed by the set of goods she pays attention to. Allowing for choice set heterogeneity makes sense in my model because a worker can only choose to be employed if she has received a job offer. If some workers are more likely to receive job offers than others, they will be able to choose employment options more often. And if workers differ in their probability of receiving a job offer from a distant region, this will impact their ability to migrate for a specific job.

I impose a specific structure on the choice set heterogeneity in this model. A worker may face one of two choice sets: a full choice set with all eight options, or a restricted one which does not include employment in regions other than her home one. This means that a worker can always choose to be unemployed in any region, because that decision does not depend on the behaviour of employers. She can also always choose to be employed in her home region, because I hypothesise that it is easier to find a job in a familiar labour market compared to searching in a distant region. Table 3 lists an example of the options in these two choice sets for a worker that resides in the West.

A worker's choice set depends on the recruitment practices in her sector and occupation, Z_i. I use three variables to proxy for these: average firm size and spatial concentration (by industry), and future online vacancy posting (by occupation). The first two variables[5] are selected based on findings in the human resources literature, which show that firms hire within a broader geographic radius if they are looking for skill or quantity of labour that is relatively harder to find in their local labour market (for a full discussion, see Appendix A.4). This holds true for firms that are large or in spatially concentrated sectors. The data on online vacancy posting[6] from late 2000s is used to pin down occupations for which the employers are willing to search more broadly. As not all vacancies are posted online, and there is significant variation across education groups and occupations, this variable is likely to capture some of the relevant variation in recruitment strategies.

The empirical link between migration and recruitment proxies are documented in Figure 4. Its three panels plot the average values of firm size, spatial concentration, and online vacancy posting, separately for migrants and stayers, and for the two different types of migration. It shows that migrants on average work in more spatially concentrated sectors, and in occupations that are more likely to be advertised online. However, this relationship is mostly driven by those that move into a specific job: moving with a job in hand is more likely for workers working for larger and more spatially concentrated companies that tend to post their vacancies online.

I model choice set heterogeneity as a binary variable equal to 1 if a worker can choose from all options, and 0 if her choice set is restricted. Assuming that the unobserved part of the choice set variation is also type I extreme value distributed, the probability that worker i has a full choice set, μ_i, is a logit function[7] of Z_i:

$$\text{Prob (full choice set} = 1) = \mu_i = \frac{exp(\gamma Z_i)}{1 + exp(\gamma Z_i)} \tag{3}$$

where γ stands for the coefficients of Z_i.

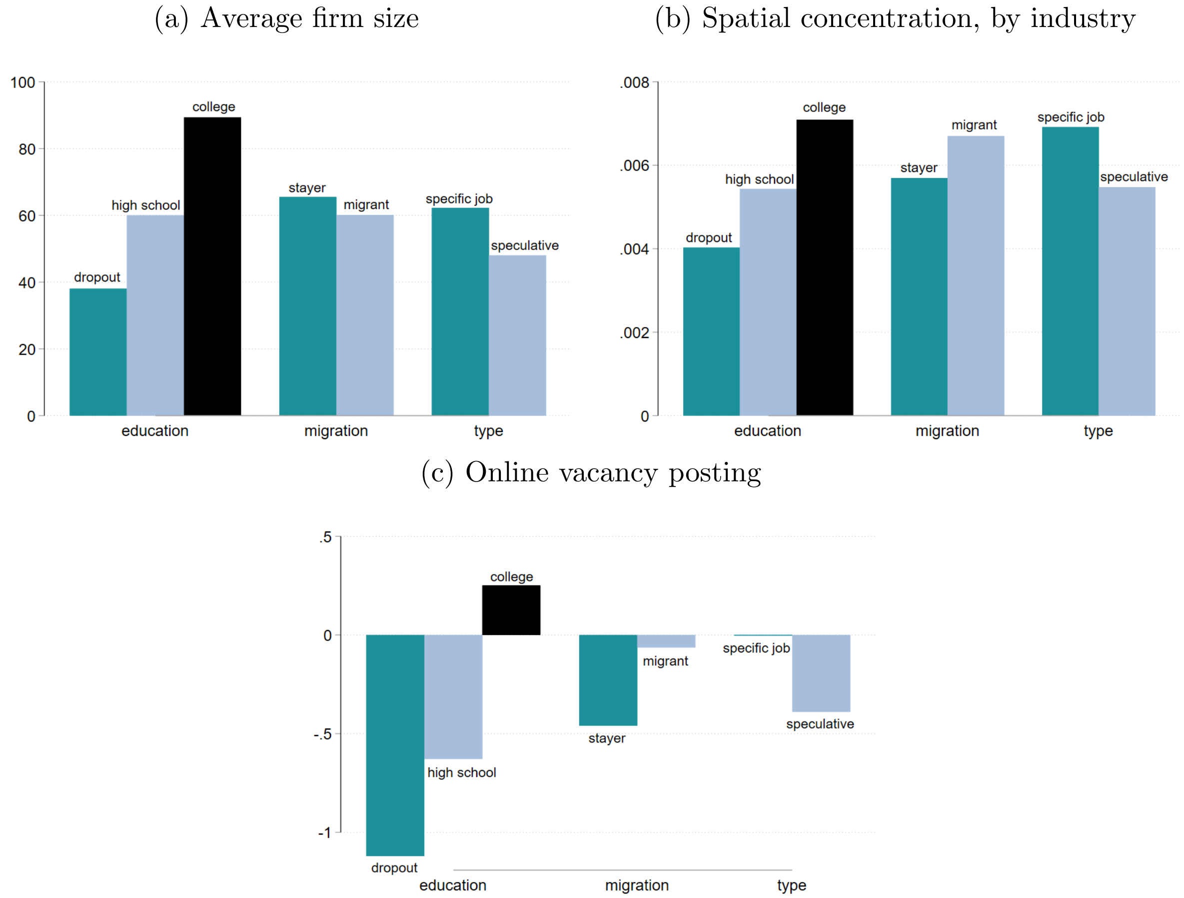
The overall probability that the worker chooses option s is a combination of the conditional choice probability P_{is} and the probability that it is in her choice set, which is a function of μ_i. The choice set itself depends on the worker's home region, and so the exact unconditional probabilities will be individual-specific. Let's return to the example of a worker residing in the West of the US, whose two possible choice

[5] They are constructed from the Country Business Patterns dataset; see section 2 for more detail.

[6] Taken from Hershbein and Kahn (2018). See section 2 for more detail.

[7] The RUM interpretation does not apply in this case. Choice set is not something the worker can choose, it is given by factors exogenous to her.

Figure 4: Recruitment proxies and migration



Recruitment behaviour and worker characteristics. Panel (a) summarises average firm size by education, migration propensity, and type of migration (conditional on migration). Panel (b) plots average index of spatial concentration by education, migration propensity, and type of migration (conditional on migration). The higher the value of the index, the more spatially concentrated the industry. Panel (c) summarises the gap between all vacancies and those posted online. The variable is calculated as log odds ratio: it is the log of odds ratio of the probability density mass of each occupation category in the two datasets. Negative number means that the vacancies are under-represented in the online data, and vice versa. Data sources: SIPP (1996-2000), County Business Patterns (1996-1999), JOLTS (2007-2014), Burning Glass (2007-2014).

sets were described in Table 3. She can only choose to be employed in the Northeast (option 1) if she receives cross-regional job offers, that is, if she has a full choice set. This will happen with probability μ_i. Conditional on option 1 being in her choice set, she will choose it with the probability P_{i1}. The unconditional probability of choosing employment in the Northeast for this worker, R_{i1}, is thus:

$$R_{i1} = \mu_i P_{i1} = \frac{exp(\gamma Z_i)}{1 + exp(\gamma Z_i)} \frac{exp(\beta X_{i1})}{\sum_{t=1}^{8} exp(\beta X_{it})} \tag{4}$$

The unconditional choice probability for the other employment-away choices (in this case, options 3 and 5) can be derived analogously.

In contrast, the option to be unemployed anywhere, or employed locally, is available under both choice sets (see Table 3). The unconditional choice probability for options $s = \{2, 4, 6, 7, 8\}$ will be thus a weighted average of the conditional choice probabilities under those two scenarios. If the worker faces a full choice set (with probability μ_i), she will choose to stay home as unemployed (option 8 in this case) with the probability P_{i8}, P_{i8} being the conditional probability of choosing $s = 8$ from among all the eight options. With probability $1 - \mu_i$, the worker receives no away employment offers, and the conditional probability of choosing unemployment at home is S_{i8}. S_{i8} is different from P_{i8}, because it calculates how optimal option 8 is among five, not eight, options. For the worker in our example, the unconditional choice probability for option 8, R_{i8}, is:

$$R_{i8} = \mu_i P_{i8} + (1 - \mu_i) S_{i8} \tag{5}$$

where

$$S_{i8} = \frac{exp(\beta X_{i8})}{\sum_{t=\{2,4,6,7,8\}} exp(\beta X_{it})} \tag{6}$$

The unconditional choice probabilities for the other options in the restricted set, $s = \{2, 4, 6, 7\}$ can be derived similarly.

4.2 Estimation

Identification The goal of the this model is to decompose the variation in workers' choices into variation in the utility of the options, and variation in their availability.

The workers' preferences are identified from the variation in attributes across options. The coefficients on the various attributes reveal how much does the probability of choosing an option changes when the value of the attribute changes.[8] I use seven

[8]In general, random utility models like this one require normalisation of the coefficients, because only differences in utility matter for identifying the optimal choice. This is not necessary here, because the conditional choice probabilities are modelled as conditional logit. Normalisation of the scale of utility is done implicitly by assuming that the unobserved part of utility, ϵ_{it}, is type I extreme distributed with

option attributes: income, regional unemployment rate, unemployment dummy, migration dummy, switching cost dummy, regional rent, and regional house price. The data sources on these variables are described in section 2 and Table 12.

The challenge in estimating the variation of the choice sets lies in the fact that they are unobservable to the researcher. In the previous section, I build a model where the probability of full choice set was denoted as the latent variable μ_i. We learn about it indirectly, by embedding it in the unconditional probabilities (4) and (5).

Identification of μ_i rests on two components. It is identified from the variation in cross-regional recruitment practices, as proxied by variables Z_i. It also depends on the specific choice set structure I imposed in the model, which manifests itself in the functional forms of the unconditional probabilities. In particular, I assume that unemployment anywhere and employment at home are always in the worker's choice set, but employment away may not be.

The exclusion restrictions for the identification of choice sets are that firms recruit all the workers in the same sector and occupation the same way; and that workers do not choose their sector and industry because of the migration opportunities it offers – worker's choice of industry and occupation is taken as given at the point of her employment-location choice.

Imputing income The model includes monthly income because it is an important part of the worker's migration and employment decision. However, some parts of the income variable have to be imputed before the full model can be estimated.

For options that include employment, the income corresponds to the specific worker's monthly wage offer, which will be unobserved unless it is accepted. For unemployment options, income equals the worker's reservation wage, which is also unobserved.[9]

Reservation wages are calculated as the smallest observed wage in each worker-region bin. Workers are sorted into bins by occupation category, experience (captured as 5-year age intervals), education, year, and region of residence. The smallest observed wage in each bin is a super-consistent estimate of the reservation wage (Keane et al. (2011)), assuming that the reservation wages are uncorrelated with any relevant unobservables, such as differences in propensity to migrate or location preferences.

Monthly wage offers are extrapolated from the observed monthly earnings. I need to control for two sources of selection bias: the probability that a worker would choose to be employed, and that she would choose the given region. To do so, I follow the method proposed by Dahl (2002), which expands the seminal work by Heckman (1979) to multi-regional labour markets. I estimate a wage offer equation separately for each

variance equal to $\pi^2/6$.

[9] Reservation wage reflects non-labour income (social security, savings) as well as the discounted expected benefit of job search, i.e. the future earnings.

of the four regions, which is then used to impute individual-specific average monthly wage offer (see Appendix A.5).

The likelihood function The likelihood function that rationalises the data follows the standard form for a discrete choice model:

$$L(\beta, \gamma) = \prod_i \prod_t (R_{it})^{d_{it}} \tag{7}$$

the log-likelihood of which is

$$logL(\beta, \gamma) = \sum_i \sum_t d_{it} log(R_{it}) \tag{8}$$

R_{it} denotes the unconditional probability of worker i choosing option t, and d_{it} is a dummy variable which equals 1 for the selected t and 0 otherwise. This expression assumes that the unobserved components of utility are independent across options, and that the workers' decisions are independent of each other.

The assumption that both parts of R_{it}, conditional choice probability and the choice set probability, are logit, means that the log-likelihood has a closed-form solution. The model is estimated by maximum likelihood.

4.3 Results

This section presents estimates of the discrete choice model from section 4.1. It explains the worker's choice over employment and location as a function of the attributes of these options and the worker's choice set, proxied by data on firm recruitment. The dependent variable is a dummy signalling the workers' choice over eight options, employment or unemployment in each of the four large US Census regions. The are two main results: workers' migration opportunities increase in education, and controlling for these differences is vital for obtaining unbiased preference estimates.

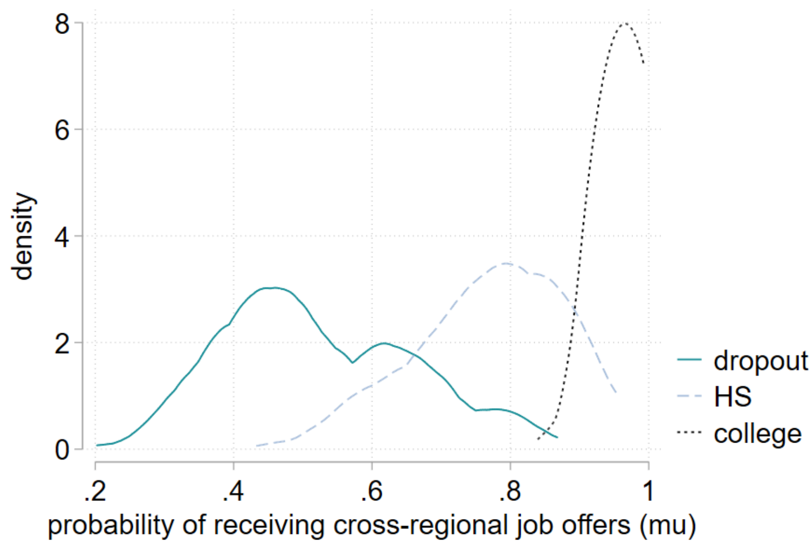
The results of the model are summarised in Table 4. The first column is a standard conditional logit. It serves as a benchmark for the augmented model estimated in columns (2) and (3). It explains the observed choices as a function of the option attributes, implicitly assuming that all workers always have full choice sets (so there is no choice set heterogeneity). It shows that workers prefer options that give them higher income (although the income coefficient is not statistically significant), better amenities (as captured by house prices), and with lower living costs (rents). Employment options are preferred to unemployment. The coefficients on the education-specific migration dummies show that migration costs are greatest for the least educated. The model also estimates a relatively large and significant switching cost. Such strong persistence is

Table 4: Multinomial logit model of the choice of employment or unemployment in one of the 4 large US regions, allowing for choice set heterogeneity.

	full choice set	heterogenous choice set	
	(1)	(2)	(3)
utility			
Dependent variable: probability of choosing one of the 8 options			
income	0.032	0.143***	0.152***
	(0.054)	(0.029)	(0.029)
regional unemp. rate	0.941***	-0.322***	-0.318***
	(0.336)	(0.061)	(0.061)
unemployment dummy	-0.821***	-1.243***	-1.226***
	(0.192)	(0.076)	(0.076)
regional house price	8.836	10.042***	10.043***
	(8.098)	(1.467)	(1.470)
regional rent	-0.008**	-0.003***	-0.003***
	(0.004)	(0.001)	(0.001)
mig. cost (dropouts)	-2.333***	-2.138***	-1.809***
	(0.300)	(0.284)	(0.280)
mig. cost (high school)	-1.390***	-1.200***	-1.143***
	(0.129)	(0.084)	(0.085)
mig. cost (college)	-0.369***	-0.300***	-0.345***
	(0.130)	(0.084)	(0.084)
switching cost	-2.742***	-2.737***	-2.740***
	(0.103)	(0.053)	(0.053)
choice set			
Dependent variable (latent): probability of receiving cross-regional job offer			
employees per firm		1.926*	2.210***
		(0.858)	(0.814)
spatial index (industry)		0.068	0.339
		(0.169)	(0.6746)
online vacancies (occupation)		0.927***	0.403**
		(0.192)	(0.214)
constant		1.691***	
		(0.370)	
dropout			-0.334
			(0.737)
HS			0.570**
			(0.366)
college			2.254*
			(1.292)
N	3997	3997	3997

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ *All models, including (1), also contain option-specific constants. (1) assumes all workers face the same choice set.*

Figure 5: Estimated distribution of the probability of choosing from a full choice set (the probability of receiving a cross-regional job offer), by education.



Kernel density plot of estimated μ_i, by education. Solid line: dropouts. Dashed line: high school graduates. Dotted line: college graduates. Based on specification (6) in Table 4.

expected when looking at month-to-month decisions.

Columns (2) and (3) in Table 4 allow choice sets to vary. These results are presented in two panels: the upper summarises workers' preferences, and the bottom the variables that determine their choice set.

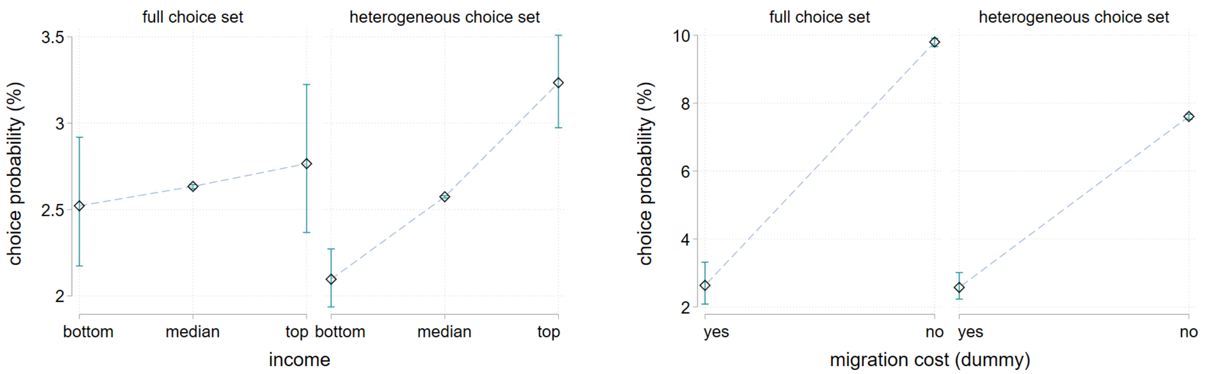
Starting from the bottom, the statistical significance of the recruitment proxies tells us that choice set variation matters. Workers coming from industries and occupations that are more spatially concentrated, with larger firms, and where vacancies are more likely to appear online, are more likely to receive a cross-regional job offer, and choose to be employed away. At the same time, the education-specific constants in column (3) show that the average probability of having access to the full choice set increases in education, suggesting there are other factors behind choice set variation not covered by my model.

Based on these estimates, Figure 5 plots the probability that a worker receives cross-regional job offers, μ_i. As education increases, the average probability of receiving cross-regional job offers increases, and the variance of this probability falls. The median dropout worker has a 48% chance of being able to select employment away. The median μ_i rises to 78% for high school graduates, and to almost 97% for college graduates. Fewer than 1% of college graduates face a less than 90% chance of being able to choose from the full choice set.

The upper panel of Table 4 shows that allowing for choice set heterogeneity changes

Figure 6: Marginal effects of income and migration costs on the probability that an average high school graduate in the Northeast chooses employment in the South (option 5).

(a) income (b) migration costs



The figures plot the estimated probability that average high school graduate in the Northeast chooses employment in the South. Full choice set plots the probabilities for a standard logit model, which assumes that all workers always choose from all options. Heterogeneous choice set plots the probabilities for a conditional logit in which the probability that a worker receives a job in another region varies with her education, sector, and occupation. Left hand panel shows how this choice probability changes as the income offered in this option increases. Bottom = bottom 1% of the wage distribution in the South. Median = median wage. Top = top 1% of the wage distribution. Right hand side panel shows how this choice probability changes between a world with 0 migration costs, and a world where migration costs exist (the average migration costs estimated in Table 4). The diamond represent the point estimates; the solid lines depict 95% confidence intervals.

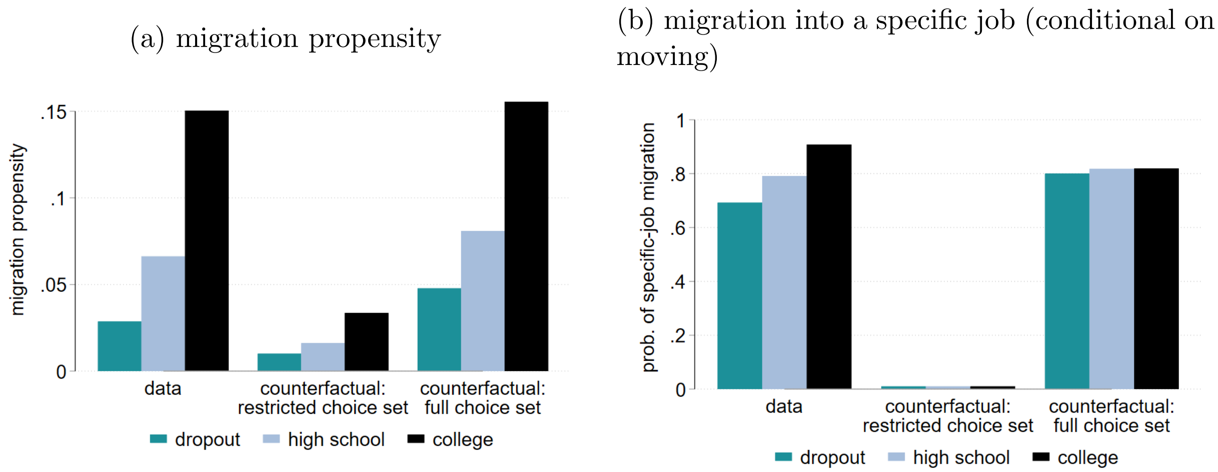
my estimates of workers' preferences. Comparing coefficients in columns (1)-(3), we can see that income and unemployment dummy increase in magnitude, while migration costs become smaller. The sign on regional unemployment rate changes from positive to negative, indicating that workers indeed do prefer locations where the labour market is tighter[10]. These coefficients become more statistically significant, because varying choice sets explain some of the behaviour a standard conditional logit could not explain. The magnitude and sign of the variables changes because allowing for choice set heterogeneity is similar to including an omitted variable, which reduces the omitted variable bias in the estimated coefficients[11].

Take the example of income and migration costs, the estimates of which figure prominently in the migration literature. Figure 6 plots how the predicted choice probability of a particular option changes when income and migration costs change. Each panel contains two lines: one representing the standard conditional logit (where all

[10]This change in the sign suggests that workers stay in regions with relatively higher unemployment rate because speculative migration is not optimal, and they lack the option to move for a specific job.

[11]The bias discussed here pertains to my interpretation of the logit coefficients as estimating workers' *preferences*; the coefficients are not biased as such. If we ignore the RUM framework and instead interpret the logit coefficients as the workers' *responsiveness* to option attributes, then controlling for choice set heterogeneity provides more detail on this responsiveness.

Figure 7: Impact of choice set variation: counterfactual migration behaviour if all workers received cross-regional job offers vs. if none of them did.



Left hand side: probability of regional migration. Right hand side: probability of moving for a specific job, conditional on migration. The first cluster of bars plots the observed migration behaviour. The other clusters capture counterfactuals calculated from the estimated reduced-form model. The middle cluster plots migration behaviour when all workers must choose from the restricted choice set, i.e. never receive cross-regional job offers. The values of job-specific migration are 0 for all education groups. The last cluster plots migration behaviour when all workers can always choose from the full choice set, i.e. employment in other regions is always available. All predicted probabilities are based on the coefficients from my preferred specification, column (3) in Table 4.

workers have a full choice set), the other based on the model with heterogeneous choice sets. Looking at income, we can see that workers are estimated to be much more sensitive to income variation when I allow their choice sets to vary. Intuitively, in order for the standard logit to explain the low probability of selecting employment in distant regions, workers cannot care much for the higher income these options offer. By letting choice sets to vary, I bring in another explanation: workers do not choose employment away because they may not have received a job offer from another region. For migration costs, the effect is the opposite: allowing for choice set heterogeneity actually decreases the estimated migration costs, resulting in a flatter line on the right hand side panel of Figure 6. However, the intuition is the same: some of the low choice probability of employment away is now explained by the lack of cross-regional job offers rather than high migration costs. Looking at column (3) in Table 4, we can see that this effect is greatest for the high school dropouts, who also face the most severe choice set restrictions.

How much does choice set heterogeneity matter? To answer this question, I use the model estimates from Table 4 to construct two counterfactuals[12]: a world where all workers have a restricted choice set, and a world where all workers have a full choice set. The corresponding predicted migration propensities and predicted probabilities of migration into a specific job are plotted in Figure 7. The differences in migration behaviour are stark. Effectively ending any cross-regional hiring, as I do in the middle set of bars, shows that removing the possibility to move into a specific job would have a large and negative impact on migration propensity of the workers. This effect is the strongest for the most educated group, whose migration propensity would decline by more than 10 percentage points, compared to about 2 percentage points decline for high school dropouts. On the other hand, allowing all workers to always choose from among all options doubles the migration propensity of the less educated. It would also almost completely eliminate any education-based differences in the type of migration, suggesting that the relatively low share of migration for a specific job among the less educated is more due to their restricted choice sets rather than their preferences or the attributes of their options.

To investigate the differences between education groups further, I estimate the descriptive model of migration separately for those with and without college degree[13]. In the model estimated in Table 4, I allow migration costs to vary by education; here I allow other preferences to vary, too. If the differences in education are primarily driven by variation in workers' tastes, rather than different job search opportunities and option attributes, we should see different preference coefficients for the less and the more educated, and the choice set variables should lose their significance.

The results in Table 5 show that this is not the case. If we compare the coefficients in the two rightmost columns, we can see that some differences in preferences exist – the more educated seem to be more responsive in general to the different option attributes – but they are small. Furthermore, the variables capturing choice set variation remain statistically significant. They are different from the coefficients estimated in Table 4, but that is because here they relate to the within-group variation, whereas the general model captures the between-group variation.

Moreover, controlling for choice set heterogeneity reduces the bias in the education-specific preference estimates. Similarly to the general results, I estimated the standard conditional logit alongside the logit with heterogeneous choice sets. These results are given in the first two columns of Table 5. I use these estimates to predict choice probability of a particular option as the income offered in that option increases. As Figure

[12]I do this because discrete choice models do not have a straightforward measure of fit like linear regressions. A multinomial choice model like the one estimated in this section will by construction set the predicted choice probability of each option equal to its sample average probability.

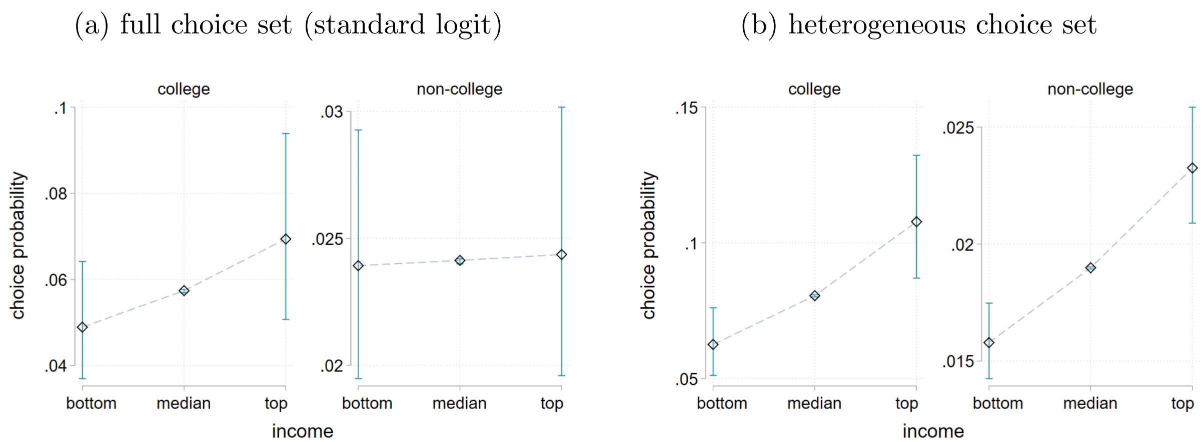
[13]I pool the high school graduates and dropouts in this way because the dropout subsample is not large enough to identify the model.

Table 5: Multinomial logit model of the choice of employment or unemployment in one of the 4 large US regions, allowing for choice set heterogeneity, by education.

	full choice set		heterogenous choice set	
	less than college	college	less than college	college
preferences				
Dependent variable: probability of choosing one of the 8 options				
income	0.006	0.095	0.127***	0.159***
	(0.072)	(0.081)	(0.036)	(0.061)
regional unemp. rate	0.669	0.767	-0.320***	-0.475***
	(0.442)	(0.488)	(0.085)	(0.095)
unemployment dummy	-1.313***	-1.542***	-1.093***	-2.130***
	(0.230)	(0.312)	(0.087)	(0.191)
regional house price	4.815	12.600	8.902***	9.218***
	(12.025)	(12.500)	(2.071)	(2.270)
regional rent	-0.006	-0.009	-0.003***	-0.003***
	(0.005)	(0.006)	(0.001)	(0.001)
migration cost	-1.575***	0.057	-1.426***	1.399***
	(0.132)	(0.285)	(0.0803)	(0.097)
switching cost	-2.668***	-3.185***	-2.636***	-3.671***
	(0.110)	(0.281)	(0.067)	(0.094)
choice set				
Dependent variable (implied): probability of receiving cross-regional job offer				
employees per firm			5.262***	0.785**
			(1.777)	(0.436)
spatial index (industry)			0.164	0.747
			(15.033)	(4.789)
online vacancies (occupation)			0.744***	0.506***
			(0.271)	(0.202)
constant			0.546	-0.342**
			(0.410)	(0.147)
N	2840	1157	2840	1157

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ *All models contain option-specific constants.*

Figure 8: Marginal effects of income on the probability that an average worker chooses employment in the South (option 5), by education.



The figures plot the estimated probability that average worker chooses employment in the South, if she currently resides in the Northeast, as the income offered in this option increases. The probabilities are calculated separately for workers with and without college education, using estimates from Table 5. Left hand side panel: full choice set plots the probabilities for a standard logit model, which assumes that all workers always choose from all options. Right hand side panel: heterogeneous choice set plots the probabilities for a conditional logit in which the probability that a worker receives a job in another region varies with her education, sector, and occupation. Bottom = bottom 1% of the wage distribution in the South. Median = median wage. Top = top 1% of the wage distribution.The diamond represent the point estimates; the solid lines depict 95% confidence intervals.

8 shows, the standard logit estimates the less educated to be much less responsive to income changes than the more educated. This difference virtually disappears when I control for differences in choice sets faced by these two groups.

Overall, the descriptive model of migration showed that job search matters. The results presented in this section showed that this is true for two reasons. First, firm size, spatial concentration and online vacancy posting are statistically significant additions to the location-employment choice. Second, controlling for choice set heterogeneity improves our interpretation of the estimates of workers' preferences.

5 A theoretical model of cross-regional job search with frictions

The empirical results in section 4 demonstrated that different workers have different opportunities to migrate into a specific job, and that these differences matter for their overall migration behaviour. To analyse the role of cross-regional job search formally, I build a model of frictional regional labour markets. Every period, a worker decides where to live and whether to work there, responding to job offers from various regions and to the utility of the different options. The probability of receiving an offer depends both on the region of origin of the vacancy and the worker, which allows me to model cross-regional job search frictions. The model pins down equilibrium employment, location, and wage, of the workers, which will be estimated in section 6.

5.1 The setup

The point of departure is a standard partial equilibrium model of job search (McCall (1970)), where workers wait to receive random wage offers and follow an optimal stopping rule in deciding whether to accept them or not. The same mechanism lies at the heart of this model, but I add a location dimension, so that workers may receive wage offers from different regions, and may choose to move to another region even in the absence of a specific job offer. This means there are 3 endogenous variables in this model: a worker's employment status, her wage, and her location.

The focus of the model, as well as the whole paper, is equilibrium analysis. In this model, I assume a steady-state equilibrium exists, and then describe and determine its properties. Although the model analyses dynamic and forward-looking behavior, the environment is stationary. All the model parameters and distributions (of location preferences and wage offers) are constant over time, and so are the values of employment and unemployment in different regions.

I define region to correspond to local labour market, i.e. a commuting zone. This means that a move to a different region necessarily implies a move to a new labour market. Furthermore, all the workers within the same region are assumed to be subject to the same labour market conditions. There are J regions in the model, denoted $j = 1, 2, ..., J$.

Regions can differ in their labour market characteristics, as well as on features unrelated to the labour market. Local (regional) labour markets may differ in wage offer distributions ($F_j(z)$), job offer arrival rates (θ_j, λ_j), and the probability of exogenous job destruction (δ_j). The attractiveness of a region also depends on weather, landscape, local amenities, geographic links to other regions, etc. This is captured by a vector of individual location preference parameters $\boldsymbol{\gamma} = \begin{bmatrix}\gamma_1 & \gamma_2 & ... & \gamma_J\end{bmatrix}$. Each γ_j is a random variable drawn from a fixed multivariate distribution G with variance g^2 and region-specific means $\bar{\gamma}_j$.

Migration is costly. Every time a worker moves, she pays a fixed cost K. It captures the financial, social and psychological burden of moving.

Job search happens both when unemployed and on the job. All workers draw from the same wage distributions, so the employed are not more or less likely to receive high wage offers than the unemployed. The job offer arrival rate follows a Poisson process, in which the probability of receiving more than one offer is equals to 0. I denote θ as the probability of receiving a job offer when unemployed, and λ as the job offer rate for the employed. There is no explicit job search decision: all workers search.

The probability of receiving a job offer may not only vary by employment status but also by geography. First, the job offer rates may be higher in one region than another, which is captured by the J subscript in θ_j, λ_j. Second, it may be easier to receive job offers from one's current region of residence than from other regions. To capture this variation, I define two job search wedges, ζ_1 and ζ_2, for search on the job and when unemployed, respectively. They denote how much more difficult it is to find a vacancy in another region compared to searching locally. The probability of receiving a job offer from region k when the worker resides in region m can then be defined as

$$\theta_k^m = \zeta_2 * \theta_m^m \tag{9}$$

$$\lambda_k^m = \zeta_1 * \lambda_m^m \tag{10}$$

The implicit assumption in this setup is that the wedge between local and cross-regional job search is the same in all regions.

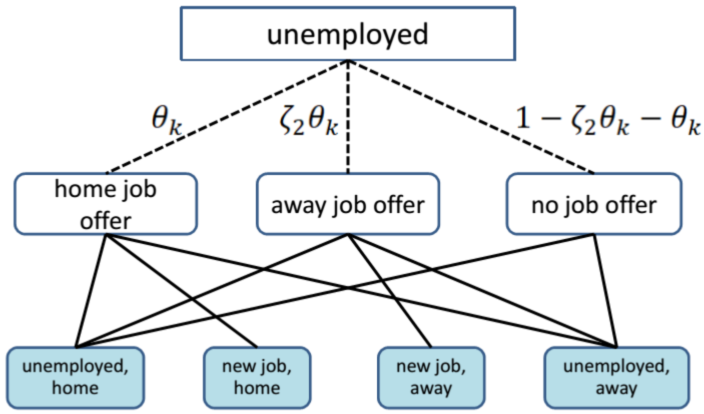
There are three reasons why a worker would change her location or employment. Each period, an employed worker may be fired; this happens with region-specific probability δ_j. Each period, a worker (unless she has just been fired) may receive a job offer. The probability of such an event depends on θ_j and λ_j, and the wage will be drawn from the relevant region-specific wage offer distribution $F_j(z)$. Third, the worker makes a draw of her location preferences $\boldsymbol{\gamma}$ from distribution G. As a result of these stochastic events, the worker re-optimizes her location and employment choice and may choose to migrate.

Timing of events and decisions Upon learning about her job options and location preferences at the start of the current period, the worker makes the joint decision about where to live and whether to work (if available). Her decisions will come into force at the start of the next period, when she also makes a new draw of the random variables.

The timing of the model is as follows:

1. The worker starts the period living in region j. She may be employed or unemployed.
2. The worker draws her location preferences for the next period.
3. If the worker is employed, her job may be exogenously destroyed, or she may receive at most one job offer from one of the J regions of the economy.
4. If the worker is unemployed, she may receive at most one job offer from one of the J regions of the economy.

Figure 9: Choices and possible outcomes for an unemployed worker



The scheme links the functioning of the labour market with workers' options and their potential outcomes. For instance, with the probability θ^k, the worker receives a job offer in her home region. In that case, she can decide between three outcomes: accept the offer, reject it and stay unemployed at home, or reject it and move into unemployment in another region. On the other hand, migration for a specific job is only possible if she receives a job offer from there first.

5. The worker makes the best choice for the next period. She compares the value of her current state with all her other options, taking into account her location preferences, and choose that location-employment combination that maximises her expected lifetime utility.

6. The worker's decisions will come into force at the start of the new period.

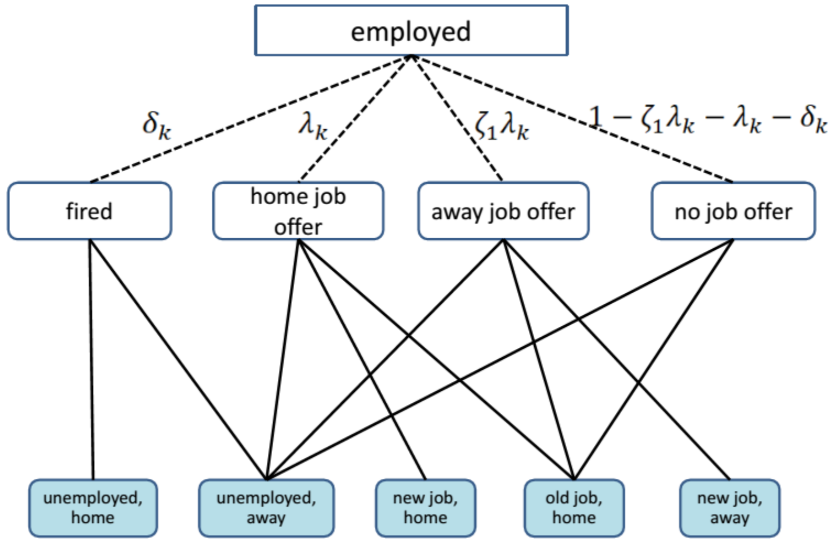
Utility maximisation The worker is a rational utility maximiser. She decides where to live and which job offers to accept based on all the information available to her so that her expected lifetime utility is maximised. All workers are infinitely-lived, risk-neutral, and homogenous up to their idiosyncratic location preferences. The worker has perfect information about the aggregate characteristics of the local labour markets and of her own location preferences, although she does not know what wage offers she will receive in a given period (if any), and what her draw of location preferences will be.

The worker's optimal decision is based on the values of employment and unemployment in different regions. There are J Bellman equations that describe the discounted expected value of employment, E_j, and J Bellman equations for unemployment, U_j.

The worker's choice also depends on her options. Figures 9 and 10 visualise the links between a worker's options and her potential choices. This structure tells us which utilities the worker compares when making her optimal choice, and is embedded in the recursive forms of the values of employment and unemployment defined below.

I start with defining *the best unemployment option*, U_m^*, as the highest net utility

Figure 10: Choices and possible outcomes for an employed worker



The scheme links the functioning of the labour market with workers' options and their potential outcomes. For instance, with the probability λ_k, the worker receives a job offer in her home region. In that case, she can decide between three outcomes: accept the offer, reject it and stay employed at home, or reject it and move into unemployment in another region.

a worker in region m can receive as unemployed. U^*_m is equal to the highest value of unemployment less migration costs, given the vector of idiosyncratic preference draws $\boldsymbol{\gamma}$:

$$U^*_m(\boldsymbol{\gamma}) = \max_j[U_j + \gamma_j - K] \tag{11}$$

The value of unemployment (equation (12)) has two parts: the current period payoff, and the expected utility of future optimisation. For every period the worker spends as unemployed in region m, she receives non-labour income (net of living costs) denoted b_m. The second part of (12) describes the future payoffs. With probability θ_m, the worker receives a local wage offer z; with probability $\zeta_2\theta_m$ she is offered a job in region j. She compares either offer against her best unemployment option U^*_m, taking into account her location preferences and migration costs. Alternatively, with probability $1-(J-1)\zeta_2\theta_m-\theta_m$, she receives no job offers and picks her best unemployment option. The expectations of the future payoffs are taken over both the values of wages z and idiosyncratic location preferences $\boldsymbol{\gamma}$, since these are the the elements unknown to the worker when making a choice today. The utility flow is discounted at the rate i.

$$U_m = \frac{1}{1+i}\left[b_m + \theta_m E_{z,\gamma}\max[V_m(z)+\gamma_m, U_m^*] + \sum_{j\neq m}^{J}\zeta_2\theta_m E_{z,\gamma}\max[V_j(z)+\gamma_j - K, U_m^*]\right]$$
$$+ \frac{1}{1+i}\left[(1-(J-1)\zeta_2\theta_m - \theta_m)\, E_\gamma U_m^*\right] \tag{12}$$

The value of employment in region m, as described in equation (13), has also two parts although the future payoff is more complex, capturing the possibilities described in Figure 10.

$$\begin{aligned} V_m(w) = {} & \frac{1}{1+i}\left[w + \lambda_m E_{z,\gamma}\max[V_m(w)+\gamma_m, V_m(z)+\gamma_m, U_m^*]\right] \\ & + \frac{1}{1+i}\left[\sum_{j\neq m}^{J}\zeta_1\lambda_m E_{z,\gamma}\max[V_m(w)+\gamma_m, V_j(z)+\gamma_j - K, U_m^*]\right] \\ & + \frac{1}{1+i}\left[\delta_m E_\gamma U_m^* + (1-(J-1)\zeta_1\lambda_m - \lambda_m - \delta_m)\, E_{z,\gamma}\max[V_m(w)+\gamma_m, U_m^*]\right] \end{aligned} \tag{13}$$

Vector $\mathbf{V_j}$ stacks all the possible values of employment, as a function of wage, for a given region j. Assuming that the regional wage offer distribution $F_j(z)$ is discrete, the length of $\mathbf{V_j}$ corresponds to the length of the support of $F_j(z)$.

When making her optimal choice, the worker compares the values of employment and unemployment available to her plus her current draw of idiosyncratic location preferences $\boldsymbol{\gamma}$, and chooses the one that gives her the highest utility. This choice pins down her employment status, wage, and location. For example, if the worker is unemployed and has not received any job offers, her optimal choice will be unemployment in her best region:

$$U_m^*(\boldsymbol{\gamma}) = \max_j[U_j + \gamma_j - K]$$

On the other hand, if she is employed and receives a job offer from another region, her best choice is defined as:

$$\max[V_m(w)+\gamma_m, V_k(z)+\gamma_k - K, U_m^*(\boldsymbol{\gamma})]$$

where m stands for home region and z is the new job offer from region k.

5.2 Equilibrium

To close the model, I define equilibrium as one where the endogenous variables of the model, regional unemployment rates ($\{\mu_j\}_1^J$), population shares ($\{\alpha_j\}_1^J$), and accepted wage distributions $\{H_j(z)\}_1^J$, are constant. Throughout the paper, the total population is normalised to 1, so per-person and aggregate outcomes are the same[14].

A general J-region model requires $3J$ equilibrium conditions to pin down the J local unemployment rates μ_j, the J regional population shares α_j, and the $J\omega$ equilibrium wage distributions $H_j(z)$, where ω denotes the size of the discrete wage support. They are all based on imposing an equilibrium of flows between different parts of the labour market or different regions.

Equilibrium in local labour markets Equilibrium in the local labour market requires a steady unemployment rate, which means that the flows in and out of the stock of unemployed must be balanced.

$$\textit{flows into local unemployment in } k = \textit{flows out of local unemployment in } k \tag{14}$$

Equilibrium in regional population shares To find equilibrium regional population shares α_j, the net migration into every region has to be zero. I can set up a similar balance-of-flows expression as for local unemployment rate that will pin down a particular regional population variable α_k.

$$\textit{migration into region } k = \textit{migration out of region } k \tag{15}$$

Given that the population shares α_j must sum up to one, the Jth condition is

$$\sum_1^J \alpha_j = 1 \tag{16}$$

which implicitly pins down α_J.

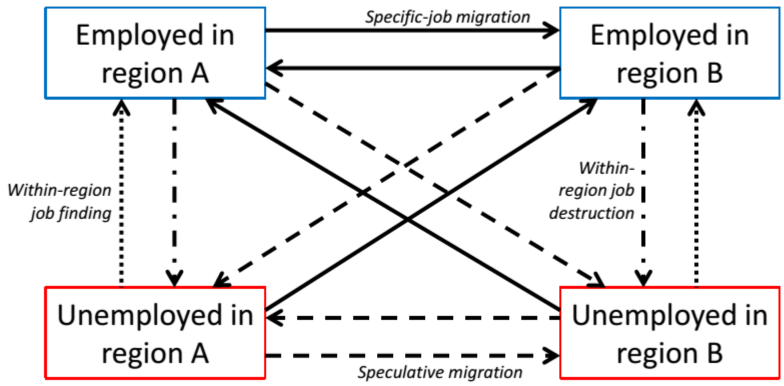
Wage distribution $H_j(z)$ is the equilibrium distribution of accepted wages. It has the same support as the distribution of wage offers, $F_j(z)$. It is derived by setting equal the flows of workers paid a particular wage w, which is why this needs to be specified for each point of the wage support, giving us $J\omega$ equations in total.

$$\textit{new workers that are paid } w \textit{ in region } k = \textit{workers that used to be paid } w \textit{ in } k \tag{17}$$

To understand how these flows equations link to the model developed in the previous section, consider a two-region example with a single wage. Figure 11 illustrates all the possible stocks and flows in a $J = 2$ economy. The boxes represent stocks of workers, employed and unemployed, in each region, while arrows depict the flows between these

[14]This holds true as long as population stays constant.

Figure 11: Equilibrium in a 2-region model.



The arrows represent flows between the different employment-region stocks. Solid line: migration for a specific job. Dashed line: speculative migration. Dotted line: flows from unemployment to employment withing the same region. Dotdashed line: flows from employment to unemployment within the same region.

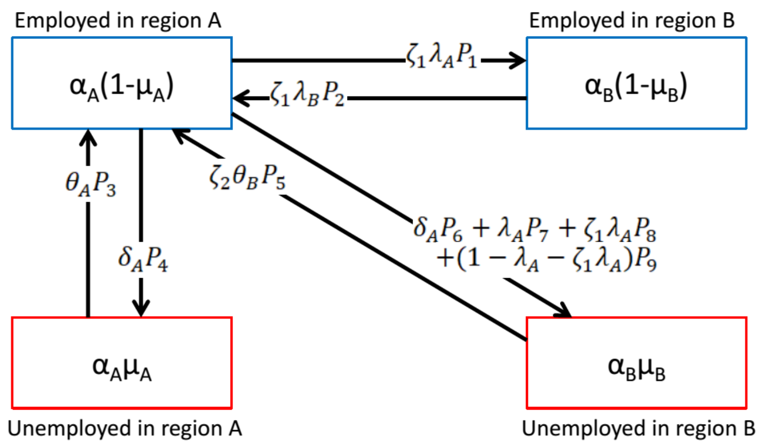
stocks, in and out of employment and between and within regions. The economy is in equilibrium when the size of the flows is such that the size of the stocks stays the same. For example, the equilibrium condition for regional population shares would specify that the magnitude of the flows in and out of each region must perfectly offset each other. Figure 11 makes it clear that this implies equality between the flows from employment to employment, employment to unemployment, unemployment to employment, and unemployment to unemployment from B to A, and the same flows in the opposite direction, from A to B.

Figure 12 shows how the model pins down the size of these flows. For example, the flow of employed in A to employment in B is $\alpha_A(1-\mu_A)\zeta_1\lambda_A P_1$. There are $\alpha_A(1-\mu_A)$ employed workers in region A. Fraction $\zeta_1\lambda_A$ of them receive a job offer in the other region, which is necessary in order to move for a specific job. Conditional on that, P_1 denotes the probability that the worker accepts the offer:

$$\begin{aligned} P_1 &= \Pr(V_B \text{ is the optimal choice out of } \{V_A, U_A, V_B, U_B\}, \text{ given preferences } \boldsymbol{\gamma}) \\ &= f_B(z)\left[\Pr(V_B(z)+\gamma_B-K > \{V_A(w)+\gamma_A, U_A+\gamma_A, U_B+\gamma_B-K\})\right] h_A(w) \end{aligned} \tag{18}$$

where $f_j(z)$ and $h_j(z)$ are probability density functions of $F_j(z)$ and $H_j(z)$, respectively. The expression of the overall flows thus depends on the optmisation of the worker, which is in turn the function of employment and unemployment values in different locations, and her choice set.

Figure 12: Equilibrium in a 2-region model: constant stock of employed workers in region A.



The four boxes represent the 4 different stocks of workers. The arrows demonstrate the possible flows in and out of the stock of employed in A. In equilibrium, these flows must perfectly offset each other. There are 3 inflows (from unemployment in A, unemployment in B, and employment in B) and 3 outflows (into unemployment in A, unemployment in B, and employment in B); the expressions on each arrow denotes the share of the corresponding stock that flows that way, i.e. the probability that a worker in a given stock would move into a different stock.

Finding the equilibrium The model is solved numerically.

I use value function iteration to calculate the values of employment and unemployment across regions, $\mathbf{V_j}$ and U_j. The challenging part of this procedure is to find the expected maximum value of future decisions, such as $E_{z,\gamma}\max[V_j(z) + \gamma_j - K, U_m^*]$. This is a core part of the Bellman equations, in which expectations need to be taken over both wage offer distributions and location preferences. The assumption that γ_j are type-I extreme distributed simplifies this step considerably. As Rust (1987) demonstrated, under this assumption the expected future utility has a closed-form solution, which I simply plug into the Bellman equation:

$$E_{z,\gamma}\max[V_j(z) + \gamma_j - K, U_m^*] = f_j(z)^T[ln(exp(\mathbf{V_j}(z) - K) + exp(U_m^*))] \tag{19}$$

Having solved for values of employment and unemployment as a function of the model parameters, I can solve for conditional choice probabilities such as P_1 in equation (18). These probabilities are necessary in order to construct the equlibrium conditions of the model.

The final step is solving the system of flow equations for the equilibrium endogenous variables $\{\mu_j\}_1^J$, $\{\alpha_j\}_1^J$, and $\{H_j(z)\}_1^J$. I do this using value function iterations in two stages. First, I use the wage offer distribution $F_j(z)$ instead of the endogenous distribution of wages $H_j(z)$ and solve for equilibrium population shares and unemployment rates. I then plug in these first-stage solutions to calculate the equilibrium wage distribution. I repeat the iteration until the values from the two stages converge.

6 Structural estimation

Section 4 presents evidence that differences in cross-regional job search - captured as variation in workers' choice sets - matter for the optimal decision about where and how to migrate. Section 5 subsequently embeds this hypothesis into a formal theoretical model of employment and location choice in a world with frictional job search. In this section, I estimate the model parameters by matching the migration and employment flows generated by the theoretical model with those observed in the US between 1996 and 1999. The results show that cross-regional search frictions are substantial. Receiving a job offer from another region is especially difficult for the less educated, and for those searching as unemployed.

6.1 Identification of parameters

Returning to the four large Census regions of the US as my unit of analysis, the structural model contains 20 parameters to be estimated:

- 4 domestic job offer probabilities for search when unemployed, θ_j
- 4 domestic job offer probabilities for search on the job, λ_j
- 2 job offer wedges, ζ_1 and ζ_2, capturing the cross-regional job-finding frictions relative to home job-finding frictions
- 4 job-destruction probabilities δ_j
- 1 migration cost parameter K
- 4 mean location preference $\bar{\gamma}_j$
- 1 variance of the location preference distribution g^2

I set non-labour earnings, b_j to 0 for all regions. The variation in non-labour income and savings will be instead estimated as a part of the region-specific location preferences, $\bar{\gamma}_j$. I set the distribution of these preferences, γ_j, to be type-I extreme value distribution. As in other parts of this paper, it makes it possible to find closed-form solutions for workers' choice probabilities, as well as the moments I base my estimation on. The parameters of the regional wage distributions are taken from the data, given the partial equilibrium model presented here does not explain how are wages set (for more detail see the next section).

Table 6: Data moments: matrix of transition probabilities for the more educated (in %)

	e,1	u,1	e,2	u,2	e,3	u,3	e,4	u,4
e,1	19.24984	0.16173	0.00682	0.00038	0.00909	0.00076	0.00492	0.00038
u,1	0.13370	1.05521	0.00000	0.00152	0.00076	0.00189	0.00000	0.00076
e,2	0.00682	0.00000	24.32591	0.20150	0.01477	0.00038	0.00720	0.00076
u,2	0.00000	0.00114	0.15112	1.13361	0.00038	0.00152	0.00076	0.00114
e,3	0.00682	0.00000	0.00985	0.00076	28.57478	0.26361	0.01212	0.00038
u,3	0.00000	0.00038	0.00114	0.00076	0.20188	1.71690	0.00038	0.00152
e,4	0.00682	0.00000	0.00909	0.00076	0.00909	0.00000	20.78645	0.24695
u,4	0.00038	0.00038	0.00000	0.00038	0.00038	0.00227	0.19127	1.47980

The transition probabilities are calculated from the 1996-1999 SIPP data. The sample is working adult males between the age 25 and 50. e = employment. u = unemployment. The numbers denote one of the four census regions of the US (see table 2).

The model is identified from data moments on the worker movement between regions and in and out of employment. In this section, I show how the theoretical expressions for these flows link to data to help uncover the model parameters.

A worker can be in one of eight states (employment or unemployment in one of the 4 census regions of the US), which gives me $8 * 8 = 64$ transition probabilities. However, the probabilities in a transition matrix must sum up to 1, giving us 63 free moments. These moments are calculated from the 1996 SIPP panel, separately for the more and the less educated. They are presented in Tables 6 and 7.

The model is estimated separately for the less and the more educated. The "more educated" category includes workers with a college degree or more; the "less educated" are high school graduates, high school dropouts, and workers with less than four years of college education (section 3.1 describes my this categorisation).

The first step in identification is to normalise the model. Because in the worker's optimisation only differences in utility matter, I set the average location preference for the West to 0, which reduces the number of parameters to be estimated to 19.

The analytical expressions for the moments are backed out by considering under what circumstances would a worker find it optimal to make a particular transition (I used the same procedure to find the model equilibrium in section 5.2). Take the transition rate from unemployment to employment from region m to region k, $T_{ue,mk}$. Using the notation from section 5, there are $\alpha_m \mu_m$ unemployed residing in region m. In order for some of these unemployed to move into employment in region k, they need to receive such a job offer. This happens with probability $\zeta_2 \theta_m$. However, the move is

Table 7: Data moments: matrix of transition probabilities for the less educated (in %)

	e,1	u,1	e,2	u,2	e,3	u,3	e,4	u,4
e,1	14.90083	0.23907	0.00048	0.00000	0.00556	0.00097	0.00145	0.00000
u,1	0.20692	1.91083	0.00000	0.00024	0.00000	0.00073	0.00000	0.00000
e,2	0.00073	0.00024	21.64856	0.33382	0.00580	0.00193	0.00338	0.00048
u,2	0.00000	0.00000	0.27508	2.34255	0.00073	0.00169	0.00000	0.00097
e,3	0.00193	0.00024	0.00411	0.00048	31.29649	0.53929	0.00604	0.00024
u,3	0.00000	0.00024	0.00024	0.00097	0.42640	3.87870	0.00097	0.00024
e,4	0.00193	0.00000	0.00218	0.00024	0.00508	0.00121	18.86438	0.36549
u,4	0.00024	0.00000	0.00024	0.00193	0.00000	0.00193	0.32391	2.39162

The transition probabilities are calculated from the 1996-1999 SIPP data. The sample is working adult males between the age 25 and 50. e = employment. u = unemployment. The numbers denote one of the four census regions of the US (see table 2).

only going to happen if it is better than all other options available to the worker, in this case unemployment in one of the 4 regions of the US. This conditional probability depends on the utility of the employment and unemployment options, the worker's idiosyncratic location preferences, and the relevant migration costs. The probability has to be calculated for all possible wage offers, and the overall expression needs to be weighted by the probability of receiving the particular wage offer, $f_k(z)$. This is expressed in the second line of the transition rate (20). Assuming that the idiosyncratic location preferences are type-I extreme value distributed, this probability has a closed-form, logit solution. The resulting transition rate is given in expression (20).

$$
\begin{aligned}
T_{ue,mk} &= \text{job-specific migration from unemployment in } m \text{ to } k \\
&= \alpha_m \mu_m \zeta_2 \theta_m f_k(z)^T \begin{bmatrix} \Pr(V_k(z_1) + \gamma_k - K > \{U_m + \gamma_m, ..., U_J + \gamma_J - K\}) \\ \Pr(V_k(z_2) + \gamma_k - K > \{U_m + \gamma_m, ..., U_J + \gamma_J - K\}) \\ ... \\ \Pr(V_k(z') + \gamma_k - K > \{U_m + \gamma_m, ..., U_J + \gamma_J - K\}) \end{bmatrix} \\
&= \alpha_m \mu_m \zeta_2 \theta_m^m \int \frac{exp(V_k(z) - K)}{exp(V_k(z) - K) + exp(U_m) + \sum_{j \neq m}^{J} exp(U_j - K)} \mathrm{d}F_k(z)
\end{aligned} \tag{20}
$$

The 19 model parameters are determined jointly by all the data moments, but moment expressions such as (20) enable me to see the links between the model and the data explicitly. In this particular case, knowing the equilibrium population distribution and unemployment rates, α_m and μ_m, along with the rest of the model parameters, would allow me to uniquely identify the unemployed job search wedge ζ_2. ζ_1 is identified from the between-region employment-to-employment transition rates. The region-specific domestic job-finding probabilities for the employed and the unemployed are backed out from the within-region employment-to-employment and unemployment-to-employment transitions, respectively. δ_j, the regional job-destruction probability, is determined by the local employment-to-unemployment moves. Overall migration rates pin down migration costs, while the region-specific flows identify average location preferences.

With 19 unknowns and 63 equations, the order condition is satisfied. This is particularly helpful in my case because migration is a relatively rare event. Some off-diagonal elements of the moment matrices in Tables 6 and 7 are small and close to 0, so having multiple observations on between-region unemployment-to-employment flows allows me to estimate parameters such as ζ_2 more accurately.

6.2 Method: estimation by simulated moments

The model parameters are recovered using the method of simulated moments (McFadden (1989)). An algorithm finds the set of parameter values such that the moments generated by the data match those generated by the model.

In practice, SMM estimation involves minimising the squared distance between the data moments (as given in Tables 6 and 7) and the corresponding moment expressions as derived from the structural model (an example of such expression is given in equation (20)). Define $\mathbf{D}$ to be a 63x1 vector of data moments and $\mathbf{M}(\rho)$ the 63x1 vector of moment expressions, where ρ is the set of the 19 unknown parameters:

$$\rho = \{\delta_1, \delta_2, ..., \delta_4, \theta_1, ..., \theta_4, \lambda_1, ..., \lambda_4, \zeta_1, \zeta_2, K, \bar{\gamma_1}, ..., \bar{\gamma_3}, g\} \tag{21}$$

The SMM estimates of ρ, $\hat{\rho}$, is then a 19x1 vector of variables that can be defined as:

$$\hat{\rho} = \text{argmin}(\mathbf{M}(\rho) - \mathbf{D})^T \mathbf{W}(\mathbf{M}(\rho) - \mathbf{D}) \tag{22}$$

$\mathbf{W}$ is the weighting matrix. It determines how much "attention" to pay to individual moments in the minimisation process. Usually, $\mathbf{W}$ is the inverse of the variance-covariance matrix of the data moments. However, I simply use an identity matrix, giving an equal weight to all moments. Altonji and Segal (1996) argue that variance-covariance matrices suffer from a serious downward bias both in small samples and when the distribution is poorly behaved. In this case, even though the sample contains almost 4,000 individuals, the moments that describe migration are often based on a much smaller number of observations. As a result, a variance-covariance matrix would offer relatively little benefit over a simple identity matrix.

The first step in the estimation procedure is to guess some parameters and use them to evaluate employment and unemployment Bellman equations and workers' conditional choice probabilities.

The model moment expressions also depend on the equilibrium population distribution and local unemployment rate, α_j and μ_j. I take these directly from the data. Ideally, I would solve the model for every set of parameter values and use the corresponding α_j and μ_j; this would considerably increase the computation burden of estimation. Given that the analysis in this model is based on the assumption that the economy is in the equilibrium, the observed population shares and unemployment rates should be equal to these values.

I make a similar argument about the equilibrium wage distribution, which is also taken from the data. I approximate it using a 2-point support separately for each region: log wages at 25th and 75th percentile, attributing a 50% probability density to both. To recover the distribution of wage offers, I plug the guess model parameters, along with the endogenous variables calculated above, into the equilibrium condition for accepted wage distribution. Inverting the set of equations that define accepted wages in equilibrium allows me to solve for the distribution of wage offers, which can then be plugged into the moments. In the baseline specification, results of which are presented in the next section, I skip this step and use the approximation for accepted wage distribution as approximation for wage offers, too. An extension using 3-point wage support and estimate of wage offer distribution calculated as described here can be found in Appendix as a robustness check.

Once the model moments for a particular guess of parameter values are calculated (vector $\mathbf{M}(\rho)$), I compare this against data moments (vector $\mathbf{D}$). I repeat these steps until the squared difference of the model and data moments is sufficiently close to 0. The corresponding parameter values are the SMM estimate $\hat{\rho}$.

6.3 Results

The goodness of the model's fit is summarised in Table 8. Instead of comparing the model predictions for all 63 moments, they are aggregated into 8 transition rates, between employment and unemployment and home and away regions. The fit is good: for most moments, the difference between data and the model prediction is less than two decimal points.

Table 8: The SMM estimates of the structural model: goodness of fit.

	less educated		more educated	
	data	model	data	model
employed to …				
… employed home	86.71%	87.89%	92.94%	93.61%
… employed away	0.04%	0.04%	0.10%	0.10%
… unemployed home	1.48%	1.47%	0.87%	0.86%
… unemployed away	0.01%	0.00%	0.00%	0.00%
unemployed to …				
… employed home	1.23%	1.23%	0.68%	0.68%
… employed away	0.00%	0.00%	0.00%	0.00%
… unemployed home	10.52%	10.81%	0.68%	0.68%
… unemployed away	0.01%	0.01%	0.01%	0.01%

The actual results of the estimation are presented in Table 9, separately for the less and the more educated. The main parameters of interest are the cross-regional job search wedges, ζ_1 and ζ_2. These variables capture the proportion between away and domestic job offers. A number close to 1 means that receiving a cross-regional job offer is as easy as finding a job locally, while a number close to 0 means that cross-regional job offers are virtually impossible compared to local ones.

I find that the wedges are sizable and vary by education. An employed college is 35% less likely to receive a job offer from another region, compared to local offers. The wedge is even greater for those without a college degree, who receive on average four cross-regional offers for each ten local ones. There are also differences between search on-the-job and when unemployed. The latter is virtually infeasible: workers of either education receive less than one cross-regional offer for a hundred of local offers.

The differences in the estimated migration costs are similarly large. Converted to US dollars, the total moving costs for a college graduate are almost \$100,000; they rise to \$300,000 for those without a college degree. Despite the large range, these estimates are in line with the literature: Kennan and Walker (2011) estimate the migration cost of high-school graduates at more than \$300,000, while Amior (2015) arrives at a much more modest \$13,000.

The education variation in the rest of the model parameters are much smaller. In general, in line with the greater churn in the labour markets for less skilled labour, the job-finding and job-destructing variables are somewhat larger for those without a college degree. On the other hand, the distributions of their location preferences are

Table 9: The SMM estimates of the model parameters, by education.

description	parameter	less educated	more educated
job offer arrival rate, on-the-job, Northeast	λ_1	0.857	0.660
job offer arrival rate, on-the-job,Midwest	λ_2	0.907	0.514
job offer arrival rate, on-the-job, South	λ_3	0.869	0.644
job offer arrival rate, on-the-job, West	λ_4	0.998	0.619
job offer arrival rate, unemployed, Northeast	θ_1	0.096	0.110
job offer arrival rate, unemployed, Midwest	θ_2	0.102	0.113
job offer arrival rate, unemployed, South	θ_3	0.096	0.102
job offer arrival rate, unemployed, West	θ_4	0.116	0.110
job destruction probability, Northeast	δ_1	0.016	0.008
job destruction probability, Midwest	δ_2	0.015	0.008
job destruction probability, South	δ_3	0.017	0.009
job destruction probability, West	δ_4	0.019	0.012
job search wedge, on-the-job	ζ_1	0.399	0.674
job search wedge, unemployed	ζ_2	0.0001	0.001
migration cost (log)	K	12.626	11.485
std. dev. of location preferences	g	1.609	1.627
mean location preference, Northeast	$\bar{\gamma}_1$	-0.162	-0.037
mean location preference, Midwest	$\bar{\gamma}_2$	0.004	0.108
mean location preference, South	$\bar{\gamma}_3$	0.289	0.240

very similar, suggesting that variation in location preferences is unlikely to account for much of the migration differences. The next section explores this question in greater depth.

7 Explaining the migration-education stylised facts

Estimates of the structural model in section 6 support my hypothesis that cross-regional job search is more frictional than domestic search, and that these frictions are larger for the less educated. However, there are other factors that shape migration behaviour, such as wages, migration costs, and preferences. In this section, I use the structural estimates from section 6 to show that job search frictions can account for almost half of the observed gap in migration propensity between the less and the more educated, and 75% of the differences in the type of migration.

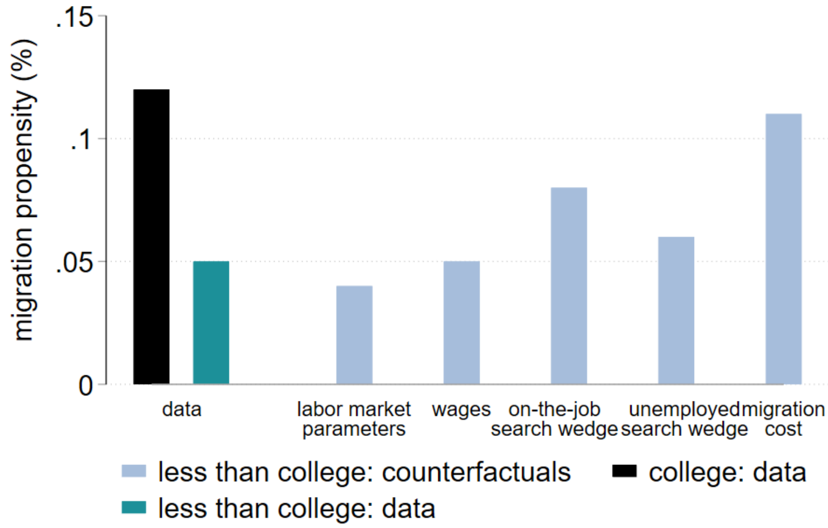
7.1 Decomposing the migration propensity gap between the more and the less educated

The less and the more educated vary along many dimensions: labour market variables, wage distribution, location preferences, migration costs, and cross-regional job search frictions. They combine to give rise to the migration propensity and migration type gaps described in section 3 of this paper. I use the structural estimates (Table 9) to decompose these gaps into the differences in the individuals components.

The decomposition of the relative difference in migration propensity by education is presented in Figure 13. It plots the changes to the migration propensity of the less educated when I change specific parameters so that they equal those of the more educated. For example, the “labour market parameters” bar shows that the less educated would move less if their job-finding and job-destruction probabilities were the same as that of the more educated, keeping all the other parameters fixed. In general, the decomposition shows that job search frictions are as important as migration cost in explaining the migration propensity differences. On the other hand, wage differences play an insignificant role. This is partly because higher mean and higher variance have opposing effects on migration propensity, effectively cancelling each other out. However, it also reflects that the observed regional wage differences are not sufficiently large to motivate migration flows.

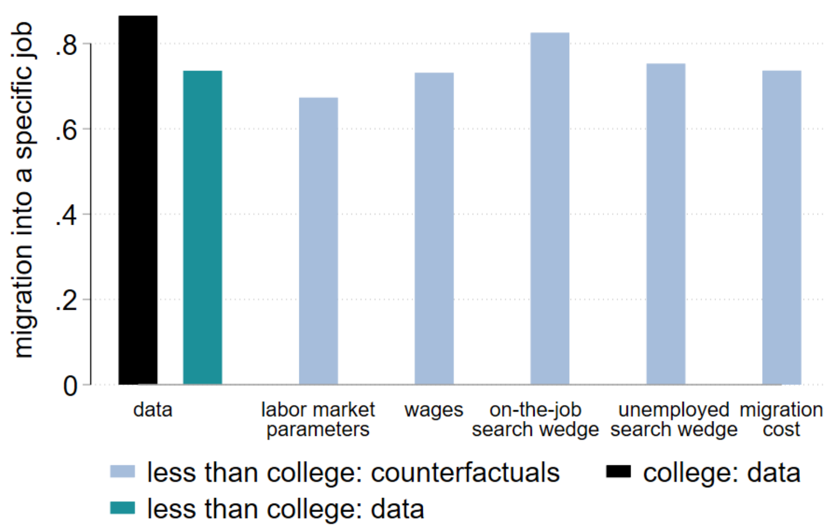
Figure 14 repeats the exercise for the propensity to migrate for a specific job, conditional on moving. Here, cross-regional job search frictions are the primary explanation, combining to wipe out most of the higher share of job-specific migration

Figure 13: Decomposition of the education differences in migration propensity.



Calculated using the structural estimates in Table 9. The two leftmost bars represent observed behaviour. The other five bars plot the change in the migration behaviour of the less educated if the given parameters were equal to that of the more educated. For example, the rightmost bar in panel (a) shows that if the less educated faced the same migration costs as the more educated, their propensity to migrate would be almost as high as that of the more educated.

Figure 14: Decomposition of the education differences in migration for a specific job.



Calculated using the structural estimates in Table 9. The two leftmost bars represent observed behaviour. The other five bars plot the change in the migration behaviour of the less educated if the given parameters were equal to that of the more educated. For example, the rightmost bar in panel (a) shows that if the less educated faced the same migration costs as the more educated, their propensity to migrate would be almost as high as that of the more educated.

among the more educated. Migration costs, a prominent factor behind their lower migration propensity, play a minor role, because they have to be paid regardless of which way the worker migrates. The education differences in wage distributions are virtually insignificant, for the same reason as above.

The impact of search frictions is large because they change the sequencing of workers' migration and employment decisions. In the absence of cross-regional job offers, all migration is by definition speculative – a decision in which income is uncertain. Receiving job offers in other regions makes it possible for the worker to decide about migration *after* the income uncertainty has been resolved[15]. These two strategies may deliver very different utility, resulting in starkly different migration patterns. Importantly, this mechanism does not rely on risk aversion.[16] To see why, consider this stylised example of a two-region economy. Let's say that the expected income of staying in the worker's home region gives her income of \$600. If she moves to the other region, she may either be unemployed and earn \$0, or get a job that pays \$1,000. If she has a 50% probability of finding a job, the expected income of moving into the other region is \$500. If the only way to migrate is speculatively, she would never make the move (\$600 > \$500). If, on the other hand, she is able to search for jobs in the other region *before* she moves, there is a 50% chance of migration: she would move if she was offered a job in the other region ($\$600 < \$1,000$) but not if she didn't ($\$600 > \$0$). This is why the estimated cross-regional job search frictions result in stark differences in the migration propensity of college graduates versus those without a degree.

7.2 Policy implications: the role of job search frictions

Understanding why and how the less educated move is vital for addressing both regional inequalities and worker mobility. There is evidence that migration is an important buffer to local demand shocks (Bartik (2018), Blanchard and Katz (1992)), but research also shows that the less educated, low-income households that are more exposed to these shocks are also least likely to move (Bound and Holzer (1996), Wozniak (2010), Hoynes (1999)). The implications are potentially large: Amior and Manning (2018) argue that this lack of mobility leads to the persistent and substantial regional disparities in unemployment and welfare. Governments do intervene in these regions, but the impact of their policies is often ambiguous and largely depending on workers' migration response (Kline and Moretti (2013), Glaeser and Gottlieb (2008)).

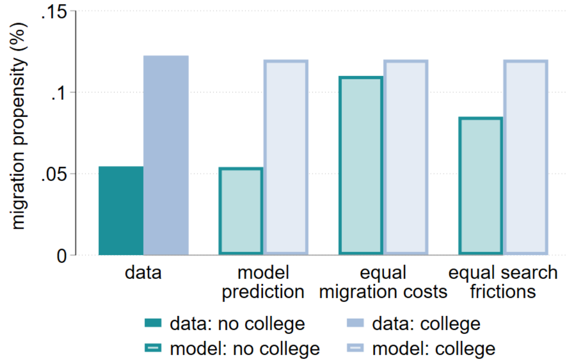
In this section, I construct a series of counterfactuals evaluating the possible impact of reducing job search frictions on migration behaviour. I compare it to the effect of

[15]However, this does not mean that the decision is made under certainty. Factors such as location preferences and future employment opportunities are still unknown to the migrant.

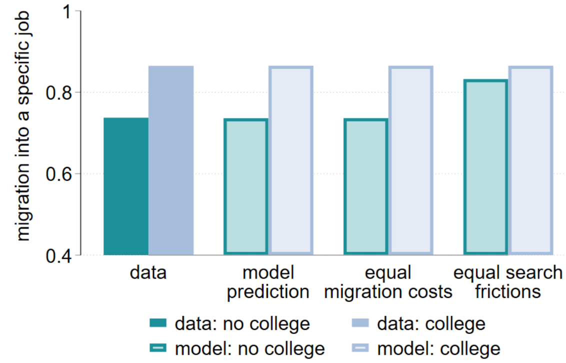
[16]The theoretical model presented in this paper assumes all workers are risk-neutral.

Figure 15: Counterfactuals comparing the role of migration costs and cross-regional search frictions in migration behaviour.

(a) migration propensity



(b) migration for a specific job (conditional on moving)



Calculated using the structural estimates in Table 9. The two leftmost bars represent observed behaviour. The next set from the left plot the values predicted by the model. The other bars plot the counterfactual migration behaviour of the less educated if the given parameters were equal to that of the more educated. For example, the rigthmost bars show that if the less educated faced the same search frictions in their cross-regional job search as the more educated, about half of the migration propensity gap, and almost all of the migration type gap, would disappear. The search frictions scenario changes the search wedge for both on-the-job and unemployed search.

reduction in migration costs, since subsidies to moving are often suggested as a solution to low mobility. I then discuss the potential policy lessons of this paper.

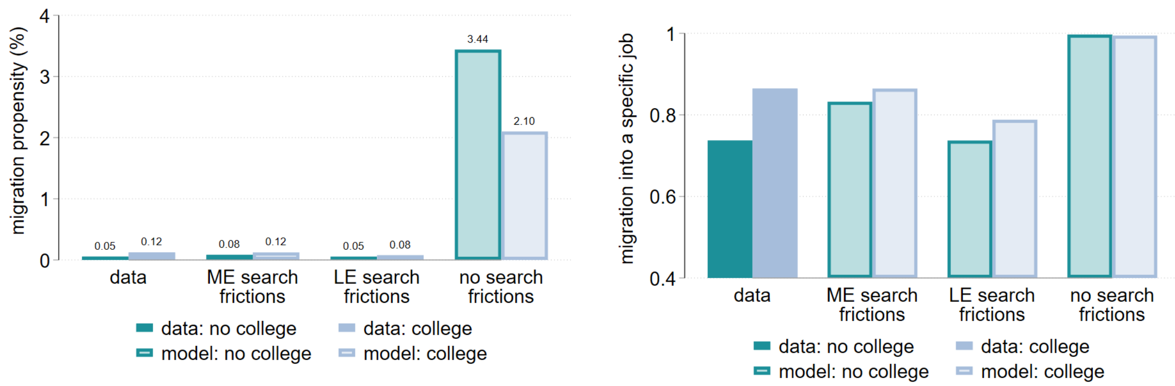
First, I look at what would happen if cross-regional search frictions were equalised across education, and compare it to an alternative policy in which I eliminate the differences in migration costs instead. The first set of bars in both panels in Figure 15 plots the observed migration behaviour (migration propensity and the share of migration that is into a specific job). The last set depicts a scenario in which I make cross-regional job search as easy for the less educated as it is for the college graduates (changing the job search wedges for both the employed and the unemployed). The central set does the same, but with migration costs, lowering them to the level of the college graduates. Lowering migration costs has a bigger impact on closing the education gap in migration propensity, but a large portion of the less educated are still forced to move speculatively. Improving cross-regional job search, on the other, closes about 75% of the gap in the type of migration, and almost a half of the difference in migration propensity.

What would happen if we could remove cross-regional search frictions entirely? I answer this question with a second set of counterfactuals, depicted in Figure 16. It allows me to see how would migration behaviour of the less and the more educated change under search frictions of the other group, and if cross-regional job search be-

Figure 16: Counterfactuals evaluating the role of cross-regional search frictions on migration behaviour.

(a) migration propensity

(b) migration for a specific job (conditional on moving)



Calculated using the structural estimates in Table 9. The two leftmost bars represent observed behaviour. The other bars plot migration behaviour when cross-regional search frictions change. For example, the "ME search frictions" cluster represents the migration behaviour if both education groups faced the cross-regional search frictions of the more educated group. The rightmost bars consider what would happen if there were no cross-regional frictions, i.e. if it was as easy to find a job elsewhere as it is to find a job locally. The job-finding probabilities (and other labour market variables) remain education-specific.

came as easy as the local one. The main takeaway, in line with the decomposition in the previous section, is that migration behaviour would be radically different if the cross-regional search wedge was eliminated. Not only would virtually no workers move speculatively, but the less educated would in fact be more mobile than the more educated. This answers the puzzle that motivates this paper: the less don't move more because they get few opportunities to move for a specific job.

Figure 16 also shows how much the migration behaviour of the college graduates depends on their relatively frictionless cross-regional job search. Their migration propensity would fall by a third, and the rate of their speculative migration would be almost in line with the less educated, if they faced the same cross-regional search frictions as the less educated. The difference in the overall migration propensity between the two education groups would fall to 3 percentage points.

Of course, these counterfactuals estimate the upper bound of the search friction effect. The analysis ignores any negative spillover and general equilibrium effects, such as equilibration of rents and wages, changes in firms' vacancy posting strategies, and negative search externalities among the workers. All of these would likely reduce the overall effect estimated here.

There are several reasons why a policymaker wishing to improve mobility may want to consider addressing the job search frictions rather than simply subsidising migration costs. First, even though levelling migration costs between the education groups has a bigger impact on the propensity to migrate than equalising cross-regional search frictions, the latter has a bigger impact on the differences in the type of migration (Figre 15). This matters because the descriptive model in section 4 shows that migration for a specific job is preferred to speculative migration, so allowing workers to move more *and* move more into a specific job would result in greater welfare gains. By changing migration costs, the policymaker is changing the cost-benefit balance of migration, but by addressing search frictions the policymaker improves the functioning and the efficiency of the market, increasing rather merely re-distributing welfare.[17] Second, the difference in the migration costs between the more and the less educated estimated in the structural model is large, in the ballpark of \$200,000, making investments into frictionless cross-regional job search a potentially cheaper alternative.

Improving cross-regional job search for those without a college degree would allow economic growth to spread more evenly. A policymaker may want to focus on sustaining strong growth in the better-off regions, rather than trying to relocate business into the depressed regions. In that case, a well-working inter-state job market would be a key part of the equilibrating channel.

An active labour market policy aimed at improving posting vacancies across regions would probably focus on two mechanisms. First, there is the cost of vacancy posting across large geographic areas. It may seem that this would be lowered sufficiently with the advancement of internet, but the actual impact on the labour market is ambiguous. Data on online vacancy posting (Hershbein and Kahn (2018)) reveals there still are significant differences in the likelihood of different jobs being posted. However, advertisement of vacancies is only a small portion of the "black box" frictions discussed in this paper. The human resources literature points out that larger, more specialised firms not only advertise more widely, but they also interview more candidates and use different selection procedures. The second part of labour market intervention may thus look at the factors that determine the firm's recruitment strategy. For example, interview transportation subsidies for the unemployed could make it feasible for workers to not only apply, but actively pursue job search in distant regions.

8 Conclusion

The main message of this paper is that job opportunities matter for understanding migration behaviour. I use a combination of reduced-form and structural evidence to show that the differences in the ease of finding a job in another region *before* the worker moves can explain up to a half of the gap in migration propensity between the more

[17] Of course, the matter is not as simple. Kline and Moretti (2013) demonstrate that in a second-best world and/or in the presence of agglomeration forces, place-based policies may in fact improve overall welfare.

and the less educated workers. The less educated are more likely to move speculatively, and move less overall, because they are about half as likely to receive a job offer from another region.

I motivate the paper by establishing a new, previously undocumented stylised fact about the relative timing of migration and job search. The less educated are not only less likely to move, they are significantly less likely to move into a specific job. I then analyse what drives this behaviour first by estimating a conditional logit of employment-location with an unobserved choice set heterogeneity. In the second step, I model search frictions formally by adding a location dimension to a standard partial equilibrium model of search. The estimates of both models lend support to the hypothesis that cross-regional search is more difficult for the less educated. The structural model also allows me to simulate a set of counterfactuals looking at possible impact of different policies focused on improving mobility.

Regardless of my results, the stylised facts introduced in this paper have implications for how we model within-country migration. I show that 65% of migration is into a specific job, which is at odds with the standard assumption in the literature that migration is a decision in which income is uncertain. First, any new model needs to distinguish between moving into a specific job (income) and moving for expected income. Second, the model must take into account that migration for a specific job is conditional on how firms advertise their vacancies. The reduced-form model in the first part of this paper shows how these two features can be embedded in a standard discrete choice model of migration, while the structural model includes them in a standard model of job search.

The policy implications of this paper are two-fold. The reduced-form results show that, in order to correctly interpret workers' responsiveness to income, living costs, and migration costs, we need to control for the fact that not all of them choose from the same set of options. The counterfactuals based on the structural model suggest that improving the efficiency of cross-regional labour market could not only have a potentially large impact on overall mobility, but also improve individuals' welfare by allowing them to migrate after the job search uncertainty has been resolved.

References

Jason Abaluck and Abi Adams. What do consumers consider before they choose? identification from asymmetric demand responses. Working Paper 23566, National Bureau of Economic Research, June 2017.

Joseph G Altonji and Lewis M Segal. Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business & Economic Statistics*, 14(3):353–366, July 1996.

Michael Amior. Why are higher skilled workers more mobile geographically? the role of the job surplus. Discussion Paper 1338, Centre for Economic Performance, 2015.

Michael Amior and Alan Manning. The persistence of local joblessness. *American Economic Review*, 108(7):1942–70, July 2018.

Alexander Bartik. Moving costs and worker adjustment to changes in labor demand: Evidence from longitudinal census data. *Manuscript, University of Illinois at Urbana-Champaign*, 2018.

Paul Beaudry, David A. Green, and Benjamin M. Sand. Spatial equilibrium with unemployment and wage bargaining: Theory and estimation. *Journal of Urban Economics*, 79(C):2–19, 2014.

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

George J. Borjas, Stephen G. Bronars, and Stephen J. Trejo. Self-selection and internal migration in the United States. *Journal of Urban Economics*, 32(2):159–185, September 1992.

John Bound and Harry J Holzer. Demand shifts, population adjustments, and labor market outcomes during the 1980s. Working Paper 5685, National Bureau of Economic Research, July 1996.

Raj Chetty and Nathaniel Hendren. The impacts of neighborhoods on intergenerational mobility ii: County-level estimates*. *The Quarterly Journal of Economics*, 133(3): 1163–1228, 2018.

David E. Clark and William J. Hunter. The impact of economic opportunity, amenities and fiscal factors on age-specific migration rates. *Journal of Regional Science*, 32 (3):349, 1992. ISSN 00224146.

Gordon B. Dahl. Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica*, 70(6):2367–2420, 2002. ISSN 1468-0262.

Glenn Ellison and Edward L Glaeser. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5):889–927, October 1997.

Paolo Epifani and Gino A. Gancia. Trade, migration and regional unemployment. *Regional Science and Urban Economics*, 35(6):625 – 644, 2005. ISSN 0166-0462. doi: http://dx.doi.org/10.1016/j.regsciurbeco.2004.09.003. URL `http://www.sciencedirect.com/science/article/pii/S0166046204000869`.

Edward L Glaeser and Joshua D Gottlieb. The economics of place-making policies. Working Paper 14373, National Bureau of Economic Research, October 2008.

Michelle Sovinsky Goeree. Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074, 2008. ISSN 1468-0262.

Michael J. Greenwood. Internal migration in developed countries. In M. R. Rosenzweig and O. Stark, editors, *Handbook of Population and Family Economics*, volume 1 of *Handbook of Population and Family Economics*, chapter 12, pages 647–720. Elsevier, 1997.

Paul Gregg, Stephen Machin, and Alan Manning. Mobility and joblessness. In *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, pages 371–410. National Bureau of Economic Research, Inc, 2004.

John R. Harris and Michael P. Todaro. Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, 60(1):126–142, 1970. ISSN 00028282. URL `http://www.jstor.org/stable/1807860`.

James J. Heckman. Sample selection bias as a specification error. *Econometrica*, 47 (1):153–161, 1979.

Rubén Hernández-Murillo, Lesli S. Ott, Michael T. Owyang, and Denise Whalen. Patterns of interstate migration in the United States from the Survey of Income and Program Participation. *Federal Reserve Bank of St. Louis Review*, 93(3):169–186, May 2011.

Brad Hershbein and Lisa B. Kahn. Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, 108(7):1737–72, July 2018.

Henry W. Herzog, Alan M. Schlottmann, and Thomas P. Boehm. Migration as spatial jobsearch: A survey of empirical findings. *Regional Studies*, 27(4):327–340, 1993.

Hilary Hoynes. The employment, earnings, and income of less skilled workers over the business cycle. Working Paper 7188, National Bureau of Economic Research, June 1999.

Michael P. Keane, Petra E. Todd, and Kenneth I. Wolpin. *The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications*, volume 4 of *Handbook of Labor Economics*, chapter 4, pages 331–461. Elsevier, 2011.

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

Patrick Kline and Enrico Moretti. Place Based Policies with Unemployment. *American Economic Review*, 103(3):238–43, May 2013.

Vanessa Lutgen and Bruno Van der Linden. Regional equilibrium unemployment theory at the age of the internet. *Regional Science and Urban Economics*, 53:50 – 67, 2015.

Ofer Malamud and Abigail Wozniak. The impact of college on migration: Evidence from the vietnam generation. *The Journal of Human Resources*, 47(4):913–950, 2012.

J. J. McCall. Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126, 1970.

Daniel McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.

Daniel McFadden et al. Conditional logit analysis of qualitative choice behavior. 1973.

David Mckenzie and Hillel Rapoport. Network effects and the dynamics of migration and inequality: Theory and evidence from mexico. *Journal of Development Economics*, 84(1):1 – 24, 2007.

Ian Molho. Spatial search, migration and regional unemployment. *Economica*, 68(270): 269–283, 2001.

Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Internal migration in the united states. *Journal of Economic Perspectives*, 25(3):173–96, September 2011.

Enrico Moretti. Chapter 14 - local labor markets*. volume 4, Part B of *Handbook of Labor Economics*, pages 1237 – 1313. Elsevier, 2011. doi: http://dx.doi.org/10.1016/S0169-7218(11)02412-9. URL `http://www.sciencedirect.com/science/article/pii/S0169721811024129`.

Eleonora Patacchini and Yves Zenou. Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, 42(6):938–949, 2012.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033, 1987.

Arthur L. Silvers. Probabilistic income-maximizing behavior in regional migration. *International Regional Science Review*, 2(1):29–40, 1977.

Larry A. Sjaastad. The costs and returns of human migration. *Journal of Political Economy*, 70(5):80–93, 1962.

Riley Wilson. Moving to jobs: The role of information in migration decisions. unpublished manuscript, November 2017.

Abigail Wozniak. Are college graduates more responsive to distant labor market opportunities? *The Journal of Human Resources*, 45(4):944–970, 2010.

A The Appendix

A.1 Data

I use the Survey of Income and Program Participation (SIPP) because of its three key features: it is a (i) panel with (ii) monthly data that (iii) follows its respondents when they move. These three features are necessary to analyse the relationship between cross-regional job search and migration. I need panel data in order to observe the outcomes of the worker's job search, and I need information on her geographic movement to be able to study migration. Finally, the data needs to be of sufficiently high frequency to allow me to link the two, i.e. to be able to see the worker's labour market outcomes just after the move. SIPP is, to my best knowledge, the only major dataset that combines these three features (for a detailed comparison of the different data sources, see Hernández-Murillo et al. (2011)). High-level monthly data on employment and wages is often a snapshot of the economy, missing the panel dimension. On the other hand, the available panel data at a similarly high frequency (such as the Current Population Survey) does not track respondents when they move; or, when it does (such as the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics), the data is collected at low frequency (annual, bi-annual, or every ten years in the case of census data), making it hard for me make reasonable inference about whether the worker migrated with a job in hand or not. The SIPP combines all these features at a large enough sample size to allow me to conduct meaningful empirical analysis.

The data collection for SIPP occurs every four months, when the respondents provide monthly retrospective data. There are several consecutive SIPP panels, each lasting between three and four years, so that there is potentially up to four years of monthly data on each participating individual. The first SIPP panel started in 1984, but I focus on the data collected after the major re-design in 1996. There are five such panels: 1996, 2001, 2004, 2008, and 2014 (ongoing).

I use the 1996 panel, which covers data from December 1995 to February 2001, for two reasons. First, it covers a "normal" period with no large downturns or recessions. Second, it follows a large redesign of the survey aiming to improve the data quality and extend the length of the survey, and manages to avoid some of the budget cuts and complications that plagued the later surveys (a 2000 panel had to be cancelled after 8 months due to budget restrictions, while the 2004 panel ran for less than 3 years).

Each SIPP panel contains between 14,000 and 52,000 interviewed households, but like in all panel datasets, this number changes over time due to sample attrition. In the SIPP, individuals may exit the sample by becoming unresponsive, through entering the army or an institution (such as prison), or migrating overseas. At the same time, the panel sample can grow: because the SIPP is based on households, new individuals

Table 10: Descriptive statistics for the sample of working men between the age of 25 and 50

	dropout	high school	college	total
age	38.60	38.37	39.26	38.65
	(7.427)	(7.221)	(7.475)	(7.327)
white	84.6%	86%	89.3 %	86.8 %
	(0.361)	(0.347)	(0.309)	(0.339)
employed	97.3%	98.4%	99.3%	98.5%
	(0.161)	(0.125)	(0.0833)	(0.120)
monthly wage ($)	1829.7	2677.1	4703.2	3156.8
	(1207.1)	(1987.5)	(4422.3)	(3019.0)
total household income ($)	3282.4	4550.4	7074.0	5124.1
	(2319.8)	(3060.3)	(5449.4)	(4043.1)
urban population	79.1%	77.8%	88.1%	80.9%
	(0.406)	(0.415)	(0.324)	(0.393)
migration propensity (state)	0.0499%	0.117%	0.286%	0.158%
	(0.0223)	(0.0343)	(0.0534)	(0.0397)
migration propensity (region)	0.0268%	0.0553%	0.156%	0.0806%
	(0.0164)	(0.0235)	(0.0394)	(0.0284)
total	2,093	10,431	4,196	16,720

Summary descriptive statistics for the sample used in the empirical sections of this paper. The values are averaged over the whole duration of the survey, 1996-1999. Migration propensities are calculated monthly.

Table 11: Observed choices

option	selected	%
employment, Northeast	1948	16.36
unemployment, Northeast	139	1.17
employment, Midwest	2616	21.98
unemployment, Midwest	227	1.91
employment, South	3846	32.31
unemployment, South	339	2.85
employment, West	2492	20.93
unemployment, West	297	2.49
employment, home	10612	89
unemployment, home	968	8.1
employment, away	292	2.5
unemployment, away	53	0.5
total	11925	100

Summary of individual's choices. The upper panel presents distribution of choices from the 8 options in the logit model. The lower panel aggregates these options based on employment/unemployment and migration.

may enter the survey any time by becoming a member of the respondent household, e.g. via marriage. This results in a degree of flux in and out of the panel sample. Of the total of 115, 996 individuals surveyed in the 1996 panel, a fifth (21%) remained in the survey for a year or less.[18] Just over a half of the sample (52,558 individuals), remained in the dataset for the full 4 years. The descriptive statistics on the subsample used in my analysis are summarised in Table 10. The dsitribution of people's choices is given in Table 11.

Data on regional house prices, rents, and unemployment rates are drawn from supplementary datasets. The rent data is taken from the 5% sample of the 2000 census. I calculate median rent by state and then aggregate it up to the census regions used in the model. I also include data on quarterly data on median house price sales by region, for two reasons. First, rent and house price data correlate with both living

[18]This may be due to inflow of new respondents as well as sample attrition.

costs and local amenities, which is why it makes sense to include two variables. Second, unlike the rent data, house prices vary over time. The third variable, monthly regional unemployment rate, is taken from the Bureau of Labor Statics.

The full list of variables and their sources is summarised in Table 12 below.

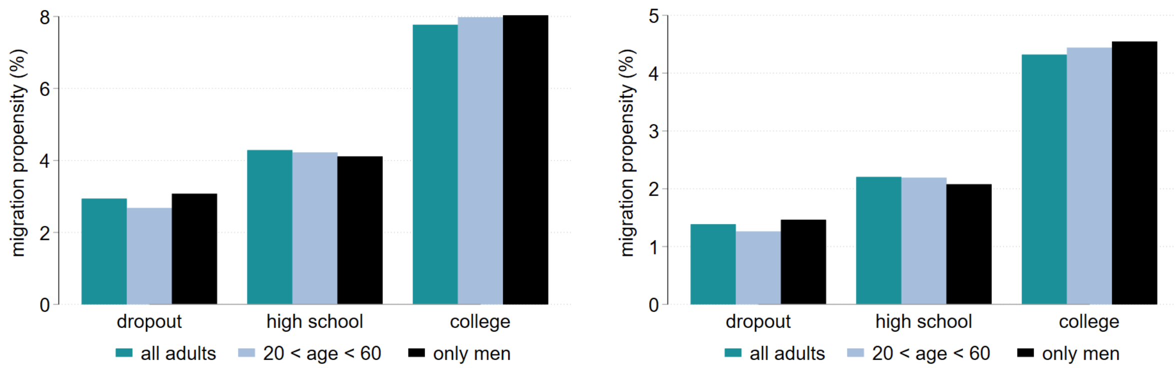
Table 12: Definitions and sources of variables.

variable	definition	source
migration	dummy equal to 1 if person changes region of residence	SIPP
speculative migration	dummy equal to 1 if person migrates and is unemployed in the following month	SIPP
income	accepted wage offer: observed monthly wage rejected wage offer: imputed average monthly wage offer reservation wage: imputed average monthly lowest wage	SIPP
unemployment	dummy equal to 1 if person is unemployed in a given month	SIPP
unemployment rate, regional	monthly state-level unemployment rate, aggregated to regions, 1996-1999	Bureau of Labor Statistics
unemployment rate, individual	individual-specific probability that the worker is unemployed in a given month. Excludes months of and after migration.	SIPP
regional rent	median rent calculated from the 5% sample of 2000 Census	IPUMS
regional house price	quarterly median house sale price, by census region, 1996-2000	Federal Reserve Bank of St. Louis
migration cost	dummy equal to 1 if person changes region of residence	SIPP
switching cost	dummy equal to 1 if this month's employment or location differs from previous month	SIPP
young	dummy equal to 1 if individual is less than 35 years old	SIPP
kids	dummy equal to 1 if individual is in a family with at least 1 child	SIPP
married	dummy equal to 1 if individual currently married	SIPP
education	highest achieved education	SIPP
online vacancies	log odds ratio of the probability density mass of a given occupation in online postings and full universe	JOLTS, Burning Glass
spatial concentration	Gini index of distribution of sector-specific employment relative to regional population	County Business Patterns
employees per firm	average number of employees per firm, per sector, 1996-1999	County Business Patterns

A.2 Stylised facts about regional migration and education

Figure 17: Propensity to migrate, US, 1996-1999.

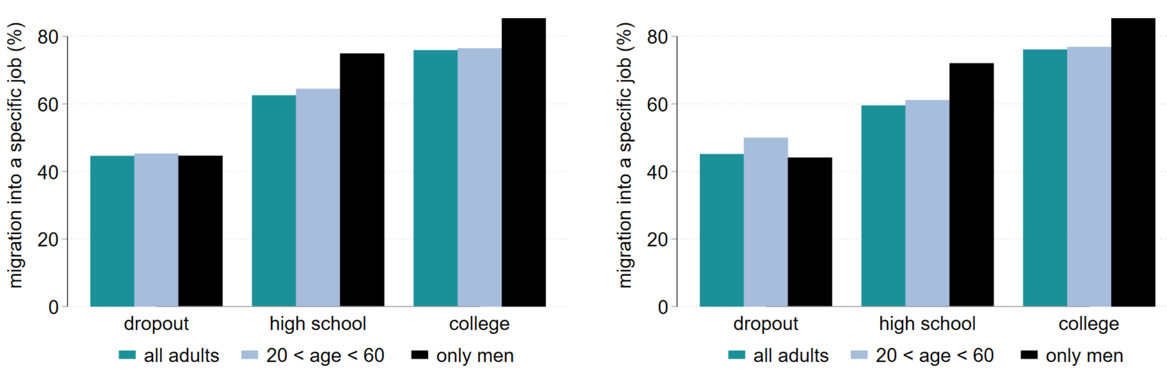
(a) Inter-state migration (b) Regional migration



Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Adult is defined as anyone over the age of 18. States refer to the 50 US states. Regions refer to the 4 Census regions (see section 3.2 for definition).

Figure 18: Probability of moving into a specific job (conditional on moving), US, 1996-1999.

(a) Inter-state migration (b) Regional migration

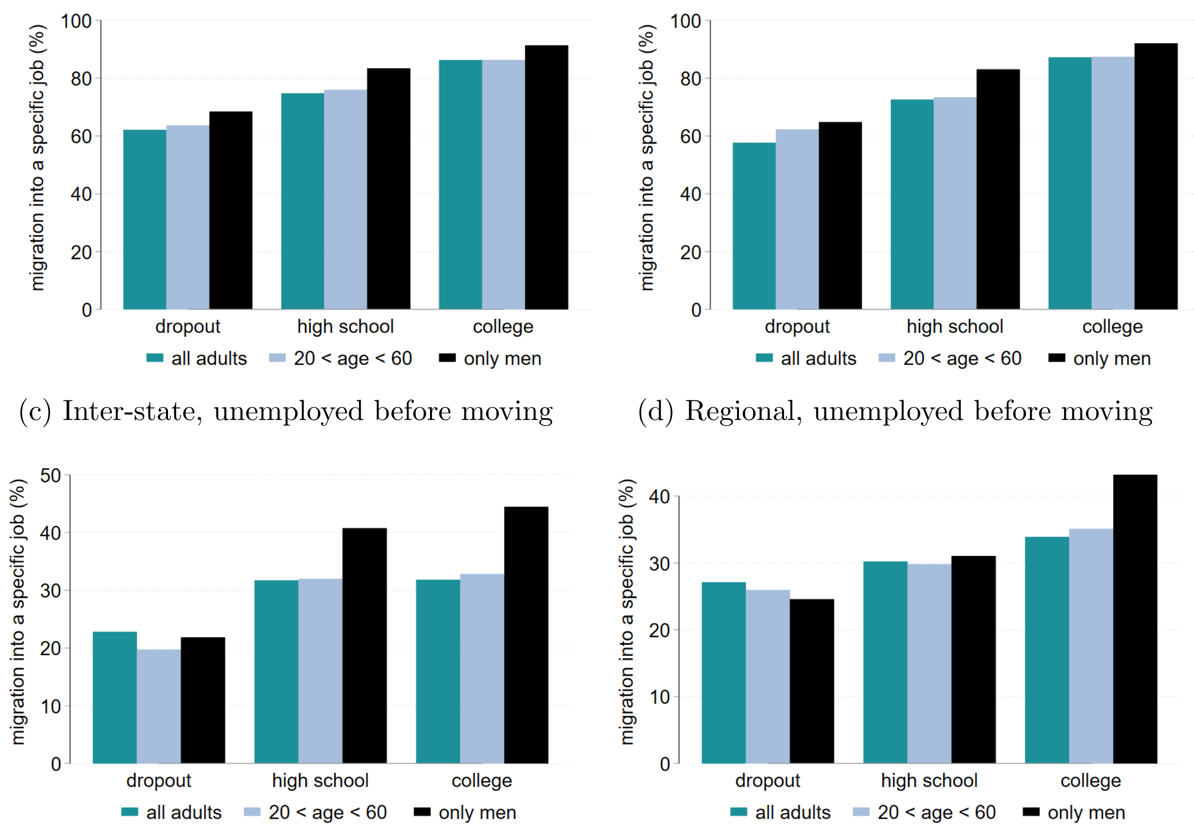


Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Adult is defined as anyone over the age of 18. Migration for a specific job refers to migration followed by employment. States refer to the 50 US states. Regions refer to the 4 Census regions (see section 3.2 for definition).

Figure 19: Probability of moving into a specific job, conditional on moving and employment state before migration, US, 1996-1999.

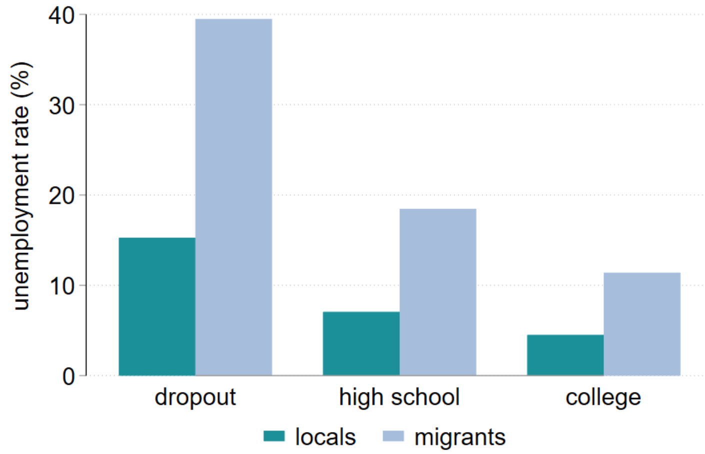
(a) Inter-state, employed before moving (b) Regional, employed before moving

(c) Inter-state, unemployed before moving (d) Regional, unemployed before moving



Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Adult is defined as anyone over the age of 18. Migration for a specific job refers to migration followed by employment. States refer to the 50 US states. Regions refer to the 4 Census regions (see section 3.2 for definition).

Figure 20: Unemployment rates for stayers and migrants, US, 1996-1999.



Calculated from the Survey of Income and Program Participation, 1996-1999 panel. Sample: adults (anyone over the age of 18). The bars represent average state-level unemployment rate for the given group.

A.3 Incentives for cross-regional migration

The first step in explaining the education differences in migration propensity is to understand why people migrate in general.

Income, and especially earnings differences across regions, is one of the most important drivers of migration both in theoretical models and in terms of the focus of the existing literature. Most the existing models of migration follow in the vein of Sjaastad (1962) and Harris and Todaro (1970), in which mobility is the response to income and job differences between regions. This has been confirmed in the empirical literature: a seminal paper by Kennan and Walker (2011) demonstrates the importance of expected income for migration decisions of young men, while Borjas et al. (1992) highlights the role state-level differences in the return to skill.

Given than earnings vary significantly with education, income is one of the leading explanations for the education migration gap. In general, average earnings rise with education, making it more likely that the more educated can afford to pay migration costs. Higher returns to skill (Borjas et al. (1992)) and larger regional dispersion of wages (Amior (2015)) have also been suggested as explanations. However, several studies have pointed out that regional differences in employment rates in particular are in fact larger for the less educated (Hoynes (1999), Gregg et al. (2004)), making their small migration propensity something of a puzzle.

The differences in wages and unemployment rates across US regions are considerable (Moretti (2011)). Table 13 documents these differences for the individuals in my data. A high school graduate may gain on average 300$ in her monthly nominal wage by

Table 13: Regional variation in monthly wages and non-employment rates, by education

education	Northeast	Midwest	South	West
		wages		
dropouts	1,704	1,637	1,504	1,549
high school graduates	2,351	2,247	2,079	2,390
college graduates	4,145	3,691	3,629	4,066
		non-employment rate (%)		
dropouts	19.86	20.53	18.05	23.34
high school graduates	10.71	9.41	10.83	12.27
college graduates	6.27	5.97	5.72	6.95

Calculated from the Survey of Income and Program Participation, 1996-1999 panel.
The sample are working males between the age of 25 and 50.

Table 14: Across-state variation in monthly wages, by education

	average wage		maximum wage	
	dropouts	college graduates	dropouts	college graduates
mean	1,913	4,361	6,307	22,469
standard deviation	338	824	4,164	10,097
coeff. of variation	0.1769	0.1891	0.6603	0.4494

Calculated from the Survey of Income and Program Participation, 1996-1999 panel.
The sample are working males between the age of 25 and 50.

moving from the lowest-wage to the highest-wage region (South and West, respectively). The average gain for a college graduate is even higher at 500$ a month (for a move from the South to the Northeast). Differences in non-employment rates[19] are also non-negligible. The difference between the best and worst performing region is more than 5 percentage points for a high school dropout and and almost 3 percentage points for a high school graduate.

At the same time, the results of direct comparisons of income variation by education are mixed. While the absolute wage differences between regions are larger for the more educated, Table 14 shows that this difference relative to average wage (as measured by coefficient of variation) is comparable, if not greater for the less educated. The opposite is true for unemployment rates (Table 15): the regional variation (as measured by standard deviation) is larger for the less educated, although the coefficient of variation is actually larger for college graduates.

[19] I define non-employment rate as $100 - $ employment rate.

Table 15: Across-state variation in monthly non-employment rates, by education

	non-employment rate (%)	
	dropouts	college graduates
mean	19.3%	5.8%
standard deviation	0.07	0.03
coeff. of variation	0.37	0.51

Calculated from the Survey of Income and Program Participation, 1996-1999 panel. The sample are working males between the age of 25 and 50.

Expected income is not everything. There are also amenities (climate, culture, quality of schools, etc.), living and housing costs, retirement and other lifecycle considerations, and proximity to friends and family. In a seminal paper by Kennan and Walker (2011), about half of the moves in their sample are subsequently reversed by return migration, highlighting the importance of residence history and social networks. Moving for college is another widely-studied motive for migration: Gregg et al. (2004) hypothesise that this experience of moving makes college graduates more mobile for the rest of their lives. Clark and Hunter (1992) and Whisler et al. (2008) show how individuals' changing preference for amenities over their lifetime shapes migration behaviour. A 2008 study by Rappaport estimates that 44% of the variation in population growth in metropolitan areas over the past 50 years is due to winter weather alone.

Similarly, the education gap in migration may exist because individuals with less education they face higher migration costs (Greenwood (1997), Moretti (2011)), or because they have different preferences over locations (Diamond, 2016). Researchers hypothesised that the stylised fact is simply a spurious correlation arising from those that self-select into college education also self-selecting into migration. Malamud and Wozniak (2012) bring evidence that the effect from education to migration propensity is causal, but the exact mechanism remains unresolved.

The job search explanation presented in this paper would be questionable if a significant number of moves was due to reasons unrelated to the labour market. However, since Roback (1982), models of local labour markets assumed that workers' utility depends on a combination of all of these factors. In this framework, the worker weighs up expected income, local amenities, migration costs, social networks, etc. to maximise her utility. This means that the migration decision then depends on labour markets even if employment is not the primary reason for moving.

Indeed, survey data suggests that job-related reasons are one of the main factors behind migration decisions. Using the Current Population Survey, which asks respondents about their reason to migrate, Amior (2015) documents that employment is a significant reason for moving across education and age categories. Almost a third of

Table 16: Impact of migration on wages, within-person comparison

	monthly wage premium		
	$ (average)	% (average)	fraction that is positive
dropouts	84	19%	73%
high school graduates	292	29%	67%
college graduates	1096	46%	78%

Calculated from the Survey of Income and Program Participation, 1996-1999 panel.
The sample are working males between the age of 25 and 50.

Table 17: Impact of migration on unemployment

	non-employment rate (%)		duration of unempl. spells		prob. of spell>6 months (%)	
	stayers	migrants (after move)	stayers	migrants (after move)	stayers	migrants (after move)
dropouts	18	26	5.5	4.2	29	20
high school	9	12	4.8	3.8	24	19
college	4	6	4.5	3.7	21	15

Calculated from the Survey of Income and Program Participation, 1996-1999 panel.
The sample are working males between the age of 25 and 50.

all cross-county moves, and just over a half of all cross-state moves was primarily for job-related reasons. This number increases in distance migrated, which is encouraging given that this paper use data on migration across the four census regions of the US.

Data on post-migration outcomes also confirms that a large majority of migrants benefit from their move to a new labour market. Because this paper uses panel data, I can calculate the impact of migration on wages by comparing an individual's wage before and after migration. The results are summarised in Table 16. The average increase in monthly wages varies between 19% to 46%, with only about 30% of workers taking a pay cut. In this calculation, I do not control for self-selection into migration, but the existing literature has in general documented positive effect of migration on worker outcomes. A 2000 paper by Rodgers and Rodgers finds positive returns to moving for every postmove year, estimating the wage premium at 20% on average. The data on unemployment incidence and duration paints a similar picture, as shown in Table 17. Migrants are more likely to be unemployed following migration, but they stay unemployed for shorter, compared to stayers.

A.4 Capturing variation in choice sets: firm recruitment strategies

In order to identify the importance of choice set heterogeneity in workers' employment-location decisions, I need to collect variables that capture this heterogeneity. Ideally, I would use detailed occupation- and industry-specific information on the geographic radius within which the representative employer advertises their vacancies and searches for employees.

To my best knowledge, no such large-scale recruitment dataset exists. Instead, I draw on the findings of the human resources literature and then use firm data to construct proxies to recruitment behaviour.

The literature suggests that firms that hire across a wide geographic radius are likely to fall into one of two categories. They may be looking for workers of specific skill that may be hard to find in their local labour market. The studies by Barron, Bishop and Dunkelberg (1985) and Russo, Nijkamp and Rietveld (1996) have demonstrated that firms that are looking for more educated workers search across greater distance, using more resources. Second, the search radius is also going to increase with the size of the firm itself. Tardos and Pedersen (2011) and Russo et al. (1996) show that larger firms spend more money on recruitment, interview more applicants, and reach out into more distant regions. Moreover, these studies point out that larger companies are more likely to use formal recruitment channels, such as job board or newspaper postings, and professional recruitment agencies, that are more likely to reach outside of the local labour market. Small companies, on the other hand, are much more likely to rely on informal channels, such as referrals and word of mouth, as well as using only one recruitment channel (Barber et al, 1999).

Based on these findings, I construct three proxies for the geographic radius of recruitment. Using firm data from the County Business Patterns dataset for the years 1996-1999, I calculate the average company size (as measured by the number of employees) by industry. I use the firm location information in this dataset to calculate an index of spatial concentration of each industry. The index (based on Ellison and Glaeser (1997)) compares the total employment in a sector in a county against the total population in each county of the US. The higher the number, the more clustered the given industry is. Finally, I use a more recent data on online vacancy posting that looks at what proportion of all vacancies, within each occupation, are advertised on the Internet (Kahn and Hershbein, 2016). Even though the data in question is from a more recent period (2007, 2010-2014), this variable provides a more direct evidence on the willingness of a firm to disseminate their vacancy information as widely as possible.

A.5 Imputing wage offers

The mean of the wage offer distribution is smaller than the mean of the observed wage because workers do not accept wages that are not large enough. This is a well-known problem in labour economics that is usually solved following a selection control approach introduced by Heckman (1979). It consists of first parametrically estimating the probability that a given worker would be employed, and then adding this selection correction term into a regression that explains observed wages as a function of worker characteristics. The coefficients from this second-stage regression then allow me to calculate average wage offer for a worker, conditional on her characteristics.

For the purposes of this model, however, I also have to control for the probability of observing wage from a particular region. Because the worker's decision is along two dimensions (employment and location), I have to explicitly control for both workers choosing highest possible wages, and workers rejecting high wages in favour of living in anther region. One way to solve this problem is to add variables to the Heckman control function that determine workers' migration decision. The disadvantage of this approach is that, by relying on a rich set variables to capture both types of bias, the overall wage regression become difficult to estimate and convergence of results is challenging.

My preferred way of estimating wage offers follows Dahl (2002). In his estimates of state-level education premium, Dahl faces the same multinomial selection bias. His solution is to calculate, for each region, the probability that a local worker becomes employed there, and the probability that a worker moves into employment from elsewhere. He then uses a polynomial of these probabilities as a proxy for the correction function, adding it to the wage regression in the same way as the Heckman correction term.

The results of the Heckman wage regressions are presented in Table 18; Dahl estimates are summarised in Table 19. The wage regression, using the given correction function, is estimated separately for each region, using the wage data on workers working and living in the region at the time. I use race as exclusion restriction in the Heckman specification. Overall, the coefficients across the two regressions are quite similar, as is the (unreported) standard deviation of the error term, suggesting that both models can explain similar mount of variation in the observed wages. The Dahl specification is my preferred one because, thanks to the lower demands on estimation, it allows me to add more worker characteristics in the wage regression.

Table 18: Heckman selection model for monthly regional wages.

	(1) Northeast	(2) Midwest	(3) South	(4) West
wage				
age	280.6687***	278.3296***	219.8698***	263.7825***
	(8.0228)	(5.6201)	(4.9114)	(6.7049)
education	1116.2263***	1021.3983***	1000.2749***	966.8457***
	(23.3704)	(16.8157)	(13.5038)	(18.2518)
_cons	2245.7523***	1741.0617***	2044.4947***	2462.5954***
	(60.8001)	(43.4558)	(36.1614)	(49.1740)
select				
age	0.0319***	0.0228***	0.0257***	0.0052
	(0.0056)	(0.0051)	(0.0041)	(0.0047)
education	0.3023***	0.3506***	0.1887***	0.2267***
	(0.0155)	(0.0138)	(0.0109)	(0.0121)
race	-0.1585***	-0.2641***	-0.1362***	-0.0898***
	(0.0105)	(0.0116)	(0.0089)	(0.0068)
_cons	1.5561***	1.5149***	1.7357***	1.4580***
	(0.0468)	(0.0405)	(0.0347)	(0.0371)
N	78530	108208	144165	95882
ll	-7.152e+05	-9.692e+05	-1.286e+06	-8.572e+05

Standard errors in parentheses

Both stages also include occupation groups and year dummy.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 19: Dahl selection model for monthly regional wages

	(1) Northeast	(2) Midwest	(3) South	(4) West
age	279.8684***	270.4605***	217.3884***	258.5599***
	(8.2174)	(5.7405)	(5.0433)	(6.8946)
education	1091.7691***	961.5562***	996.4935***	948.8347***
	(24.6079)	(17.5926)	(14.0792)	(19.1049)
year	73.9644***	82.4742***	88.6739***	113.9386***
	(12.0316)	(8.3644)	(7.3824)	(10.0632)
home_prob	-2421.5960*	-3933.6999***	-1485.0401*	262.5190
	(1005.0397)	(727.8249)	(668.3174)	(799.1905)
mig_prob	-4.065e+04***	-1.387e+04***	356.7142	-2.473e+04***
	(5571.2857)	(3179.5752)	(3265.3795)	(4907.0853)
home_prob2	10403.0351***	-901.3914	6360.9951***	-746.5879
	(1842.0758)	(1183.6315)	(1235.7193)	(1155.8361)
mig_prob2	38877.6788***	14905.4679***	-968.5881	23634.7809***
	(5710.3835)	(3239.1183)	(3312.2007)	(4903.5391)
cross_prob1	55462.4511***	14841.3148***	6372.0374	32798.8037***
	(6795.9589)	(4093.9859)	(4091.4949)	(6184.1172)
cross_prob2	-7915.8938***	4587.8734***	-4672.2692***	1460.6144
	(1774.5541)	(1099.2049)	(1186.0104)	(1083.5815)
cross_prob3	-7807.7315*	4440.2111	-4814.0601*	-4116.8270
	(3847.9920)	(2300.8156)	(2205.1413)	(3446.3211)
_cons	2231.6960***	2152.4964***	1894.9275***	1540.6217***
	(408.2322)	(292.2394)	(273.3837)	(323.9317)
N	72852	101648	134390	87900
ll	-6.856e+05	-9.368e+05	-1.241e+06	-8.204e+05

Standard errors in parentheses

The regression also includes occupation groups and year dummy.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$