

Distance and Occupation in Interstate Migration: A Gravity Measure for 209 Occupations in the United States, 2005 to 2024

Alessandro Conway

Working Paper AC-WP-2026-02

Version 1.5 (revised October 2026)

Not peer reviewed. Comments welcome.

Abstract

Employers who cannot fill jobs locally must know how far the labour market of each occupation extends, a question left open by published migration statistics, which describe persons. I build a measure of it, the distance elasticity of interstate moves into each of 209 occupations, from gravity models fitted with zeros to 531,349 movers in the American Community Survey for 2005 to 2024. The elasticity averages -0.77 , about two thirds of its variance between occupations is real, and for major groups it hardly changes when movers are classified by the occupation of the previous year. I then put it to four uses. Markets are widest for graduates: the elasticity runs from -0.97 for movers without a high school diploma to -0.58 for holders of graduate degrees, and education and age account for 57 to 63 per cent of its variance between occupations net of sampling noise. After 2020, distance did not come to matter less for teleworkable occupations. For an employer recruiting from other states, the geography of migration common to all occupations gives the best out-of-sample forecast of where an occupation's arrivals come from, occupation-specific information adds about one per cent per arrival, and a rule with one parameter does as well as the best model, with occupations differing in how strongly arrivals come from places that specialise in them. Finally, workers leave states that specialise in their occupation somewhat less often, and arrivals spread across states about as incumbents do, though these, like every other estimate here, are descriptive.

Keywords: interstate migration, occupational labour markets, distance elasticity, gravity model, forecasting, American Community Survey

JEL classification: J61, R23, J62

Status of this version. Working paper, not peer reviewed.

Suggested citation: Conway, Alessandro (2026). "Distance and Occupation in Interstate Migration: A Gravity Measure for 209 Occupations in the United States, 2005 to 2024." Working Paper AC-WP-2026-02, Version 1.5.

© Alessandro Conway 2026. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0).

Distance and Occupation in Interstate Migration: A Gravity Measure for 209 Occupations in the United States, 2005 to 2024

Alessandro Conway*

First draft March 2026

This version October 2026

Abstract

Employers who cannot fill jobs locally must know how far the labour market of each occupation extends, a question left open by published migration statistics, which describe persons. I build a measure of it, the distance elasticity of interstate moves into each of 209 occupations, from gravity models fitted with zeros to 531,349 movers in the American Community Survey for 2005 to 2024. The elasticity averages -0.77 , about two thirds of its variance between occupations is real, and for major groups it hardly changes when movers are classified by the occupation of the previous year. I then put it to four uses. Markets are widest for graduates: the elasticity runs from -0.97 for movers without a high school diploma to -0.58 for holders of graduate degrees, and education and age account for 57 to 63 per cent of its variance between occupations net of sampling noise. After 2020, distance did not come to matter less for teleworkable occupations. For an employer recruiting from other states, the geography of migration common to all occupations gives the best out-of-sample forecast of where an occupation's arrivals come from, occupation-specific information adds about one per cent per arrival, and a rule with one parameter does as well as the best model, with occupations differing in how strongly arrivals come from places that specialise in them. Finally, workers leave states that specialise in their occupation somewhat less often, and arrivals spread across states about as incumbents do, though these, like every other estimate here, are descriptive.

Keywords: interstate migration, occupational labour markets, distance elasticity, gravity model, forecasting, American Community Survey.

JEL classification: J61, R23, J62.

1 Introduction

An employer opening a plant, a hospital system short of nurses, and a state agency planning for a new industry all require the same thing, some idea of where the workers they cannot find locally would come from, and so of how far the labour market for each occupation extends beyond the state. Published statistics on migration describe persons and places, with rates of moving between

*Working paper, not peer reviewed. A companion paper, *An Occupational Flow Account for the United States, 2009 to 2025: Mobility, the Sources and Destinations of Each Occupation's Workers, and the Industry Shock of 2020 and 2021*, builds the national account of arrivals into and departures from each occupation. The author works in workforce analytics in the private sector. The paper uses only public data and received no funding.

states and counties, and contain little on occupations. The nearest occupational figure that can be drawn from them, the share of an occupation's workers who lived in another state a year earlier, counts arrivals but gives no sense of how far they came.

In this paper I build a measure of how far each occupation's labour market extends and then put it to use. It comes from the gravity equation of trade and migration (Anderson and van Wincoop, 2003; Head and Mayer, 2014; Beine, Bertoli, and Fernández-Huertas Moraga, 2016), fitted separately to the moves of the people who end up in each of 209 occupations in the United States, using the American Community Survey between 2005 and 2024. Its distance elasticity, which I call the reach of an occupation's labour market, measures how quickly the number of movers falls with the distance between the state they left and the state they moved to. For instance, an elasticity of -1 means that, of two destinations that do not border the origin, the one ten times as far away has a tenth as many movers, while an elasticity nearer zero, a longer reach, means that moves decline more slowly with distance. As the survey records occupation once, after the move, the model describes where the people who were nurses or welders in a state at the survey date had lived a year earlier, whatever work they did there.

In Section 4 I show that the measure can be read for single occupations, as about two thirds of its variance between occupations is real and the ordering survives changes of sample and specification and, for major groups, of the year when occupation is recorded. I then put it to the questions a planner or an economist would ask of it. Which occupations have national labour markets, and why? Most of the differences go with the education and age of the workforce, and the part that belongs to the occupation is largest among holders of graduate degrees (Section 5). Did remote work widen the markets of teleworkable occupations after 2020? Their arrivals rose above trend in 2021 and 2022, distinctly so only among non-graduates, and distance did not come to matter less (Section 6). Where would an employer's out-of-state recruits come from? Here I fit a second form of the model, with an effect for each pair of states common to all occupations and an exponent on the occupation's specialisation of each origin, and the exercise doubles as a test of the measure: the geography of migration common to all occupations matters far more than the differences in reach, and a rule built on the origins of all movers forecasts held-out years almost as well as the best model (Section 7). Do workers move to, and stay in, the places that specialise in their occupation? Arrivals spread across states about as incumbents do, and workers leave specialised states somewhat less often (Section 8).

All estimates are descriptive. In the applications wages and housing are held fixed, and I describe where recruits would come from under the pattern of past moves without claiming anything about how many would come. Before the measure (Section 4) come the sources that already exist (Section 2) and the data (Section 3), and after the applications in Sections 5 to 8 I close on the limits of the data (Section 9).

2 Existing measures

Most studies of internal migration in the United States are about persons, with little on occupations (Molloy, Smith, and Wozniak, 2011; Kennan and Walker, 2011), and Schwartz (1973) is the usual

early reference on how the fall of migration with distance varies with education. The most recent published estimates for groups of workers that I know of are those of Cooke, Ellis, and Wright (2022), whose Poisson gravity model of interstate flows in the American Community Survey for 2014 to 2016 implies elasticities of about -0.75 for the labour force aged 25 to 64, -0.62 for holders of a four-year degree, and -0.91 for other workers. By contrast, the distance decay of applications that Marinescu and Rathelot (2018) find among job seekers is measured between zip codes, so it cannot be compared with mine. I have not found published distance elasticities of migration for detailed occupations in the United States, nor an occupational gravity model tested on how well it forecasts the origins of later arrivals. Furthermore, a single elasticity fitted to all movers hides precisely the differences between occupations that Section 4 measures.

3 Data

I use an IPUMS USA extract of the American Community Survey one-year files for 2005 to 2024 (Ruggles et al., 2025), with 59.8 million person records. The 2020 file is left out, as the Census Bureau released only experimental estimates for that year, so the panel has nineteen vintages and the 2021 file follows directly on from that of 2019.

The universe consists of civilian employed people living in households in the fifty states and the District of Columbia. Members of the armed forces are removed by their detailed employment status, because about half of them have a civilian occupation code and their moves are postings. People living in group quarters are removed in every year, because the 2005 survey did not sample them, and as from 2006 they are 0.8 per cent of employment but 4.1 per cent of interstate movers, with a mean age of 25, the moves of the youngest workers are understated. In all, the universe holds 27,026,795 employed person records, of whom 531,349 are interstate movers.

A worker is an interstate mover if the state of residence one year earlier is one of the 51 jurisdictions and differs from the current one. Respondents are asked where they lived a year ago and nothing about the route or the reason. In other words, a move is a change of state between two dates. Some moves go unreported: Foster, Fiorio, and Ellis (2024), in a Census Bureau working paper that is not peer reviewed, find that around 68 per cent of householders who filed a permanent change of address reported living elsewhere a year earlier. The counts are therefore a lower bound on moves, and if the unreported moves differed in distance from the reported ones the elasticities would be affected, something I cannot check.

Occupation is coded on the OCC2010 basis, the IPUMS harmonisation of the Census occupation codes, and the universe contains 475 codes. Like age, education, and class of worker, occupation is recorded at the survey date, so every mover is assigned to the occupation held after the move. In 2019 the Census Bureau changed the layout and examples of the occupation questions (U.S. Census Bureau, n.d.), and the occupation series is discontinuous between 2018 and 2021, as Appendix A describes. The estimates of reach pool all nineteen vintages and are little affected, while Section 6, where years are compared, includes checks with fixed occupation weights.

Where a respondent leaves an item blank the Census Bureau allocates a value from a donor, and Kaplan and Schulhofer-Wohl (2012) found that allocation inflated interstate migration in the Cur-

rent Population Survey between 1999 and 2005. Here, records with an allocated occupation, prior residence, or migration status are 12.1 per cent of employed records and 17.4 per cent of interstate movers, the pattern I document for changes of occupation in the companion paper (Conway, 2026), and I repeat the main estimates on the records where all three items were reported.

Distance is the great-circle distance between the centres of population of two states, and the states form 2,550 ordered pairs. Contiguity is taken from the Census county adjacency file, which counts counties facing each other across water as adjacent. For Section 5 I compute a set of traits of occupations on the same universe: the shares with a bachelor’s degree and with a graduate degree, the share under 30, the mean log hourly wage, the shares self-employed and in government, and the concentration of the occupation across states. The credentialling rate is the share holding any professional certification or licence in forty monthly samples of the Current Population Survey from 2015 to 2024, and teleworkability is the classification of Dingel and Neiman (2020).

4 The measure

The gravity equation

For each occupation k I count the weighted number of interstate movers F_{odk} who lived in state o a year earlier and hold occupation k in state d at the survey date, summed over the nineteen vintages, and fit

$$E[F_{odk}] = \exp(\alpha_{ok} + \gamma_{dk} + \beta_k \ln D_{od} + \delta_k C_{od}), \quad (1)$$

where D_{od} is distance, C_{od} indicates contiguous states, and α_{ok} and γ_{dk} are origin and destination effects specific to the occupation. The model is estimated by Poisson pseudo-maximum likelihood (Santos Silva and Tenreyro, 2006), and β_k is the reach of occupation k . Origin and destination effects absorb the size of the occupation in each state, its wages, the cost of housing, and anything else that makes a state a large source or destination of its movers (Anderson and van Wincoop, 2003; Beine, Bertoli, and Fernández-Huertas Moraga, 2016). As the flows are pooled over years they absorb wage differentials as well, so the model contains no response to wages. It also assumes an elasticity constant in log distance, an assumption I have not tested.

Contiguity enters separately because many moves between neighbouring states are short moves across a border, which the distance between state centres overstates. Its coefficient averages 0.67, so a bordering state sends or receives about 1.95 times as many movers as its distance alone would imply, and for this reason every conversion of an elasticity into a share of movers compares two states that do not border the origin. Dropping the contiguity term moves the mean elasticity to -1.07 , while fitting the model only to pairs of states that do not touch leaves the ranking nearly the same.

Zeros, sample size, and inference

All 2,550 pairs of states are kept for every occupation, including the many between which no sampled worker moved, as estimation in logarithms is biased under heteroskedasticity and the Poisson estimator accommodates zeros (Santos Silva and Tenreyro, 2006). This matters for a ranking. A

bias that depended on the number of movers would turn a ranking by reach into a ranking by sample size, so I thinned nine large samples at random to between 300 and 4,000 movers, 150 times each. With zeros kept, the mean estimate differs from the full-sample estimate by less than 0.006 from 500 movers upwards. The model on positive flows only, by contrast, gives -0.16 on all movers thinned to 300 against -0.76 on the full sample.

I report an occupation if it has at least 500 sampled interstate movers, which leaves about five for each of the 102 parameters of a fit. The 209 occupations that qualify hold 90.8 per cent of sampled movers, and I name an occupation in the text only if it has at least 1,000, with one exception that is marked. For a single estimate I report a standard error clustered by origin and destination state. To compare occupations, however, I use the sampling error of the survey alone, because the clustered error also contains misfit common to all occupations. I measure it by refitting every occupation on alternate vintages, which contain different people, so that the covariance of the two sets of estimates is the variance of the true differences and their correlation, stepped up by the Spearman-Brown formula, is the reliability of the full-sample estimate. As a check I also resample movers, which treats members of one household as independent and so gives an upper bound.

The measure across occupations

Across the 209 occupations in Table 1 the mean elasticity is -0.77 , close to the -0.76 obtained when all movers are pooled in one fit, with a standard deviation of 0.15. To put that in terms of places: at the mean elasticity, and comparing states that do not border the origin, a destination twice as far away has 59 per cent as many movers and one ten times as far away has 17 per cent. At the 5th percentile of occupations, where the elasticity is -1.00 , those figures are 50 and 10 per cent, and at the 95th percentile, where it is -0.52 , they are 70 and 30 per cent. Every one of the 209 estimates differs from zero at the five per cent level.

Estimates from alternate vintages have a correlation of 0.49 across occupations, which implies a reliability of 0.66 for the estimates from the full sample, with a standard error of 0.06, and resampling of movers gives the upper bound of 0.72. In other words, about two thirds of the variance in Table 1 is difference between occupations and about a third is noise, and the standard deviation of the true differences is 0.12 to 0.13. Reliability rises with the size of the occupation, to 0.78 among the 133 occupations with at least 1,000 movers and 0.88 among the 71 with at least 2,000, while the mean and the spread of the true differences do not change with the threshold.

Under the variants in panel D the ordering of occupations changes little, with rank correlations of 0.87 to 0.97, and on fully reported records the mean elasticity is -0.777 against -0.766 .

Among the 133 occupations with at least 1,000 movers, the five with the longest reach are the residual category of physical scientists (-0.38), writers and authors (-0.42), software developers (-0.43), the residual category of business operations specialists (-0.43), and artists (-0.49). Post-secondary teachers (-0.51) and physicians and surgeons (-0.52) follow closely. At the other end, the five with the shortest reach are insurance sales agents (-1.06), industrial truck operators (-1.04), plumbers and pipefitters (-1.02), welders (-1.01), and carpenters (-0.99). Registered nurses, at -0.64 , have a longer reach than 79 per cent of the 209 occupations, whereas elementary and middle school teachers, at -0.82 , have a longer reach than only 37 per cent of them, although

Table 1: The distance elasticity of interstate moves across occupations. Occupations with at least 500 sampled movers, 2005 to 2024 pooled. Standard errors in parentheses: clustered by origin and destination for the pooled fit, and from resampling occupations for the reliability.

<i>A. Distribution across 209 occupations</i>	
All movers pooled in one fit	−0.756 (0.062)
Mean of the occupations	−0.766
Standard deviation	0.151
5th and 95th percentiles	−0.997, −0.518
25th percentile, median, 75th percentile	−0.860, −0.778, −0.668
<i>B. Signal and reliability</i>	
Correlation of estimates from alternate vintages	0.489
Standard deviation of true differences, alternate vintages	0.123
Reliability, alternate vintages	0.657 (0.060)
Standard deviation of true differences, resampling of movers	0.127
Reliability, resampling of movers (upper bound)	0.715
<i>C. Reliability by minimum number of movers, alternate vintages</i>	
At least 500 movers (209 occupations)	0.657
At least 1,000 movers (133 occupations)	0.777
At least 2,000 movers (71 occupations)	0.880
<i>D. Rank correlation with the baseline</i>	
Pairs of states that do not touch	0.931
Records with occupation, prior residence, and migration status all reported	0.970
Vintages 2005 to 2019 only	0.947
Sampled counts in place of weighted counts	0.921
All pairs without the contiguity term	0.871

nearly all hold a degree.

Reach has a correlation of 0.63 with the arrival share, meaning that occupations that draw more of their workforce from other states tend to draw it from further away. Being conditional on crossing a state line, it leaves out a long move within a large state.

Occupation before and after the move

Because the survey records only the occupation held at the survey date, a mover counted as a cashier in the destination state may have been a cashier at the origin, a student, or a machinist. The March supplement of the Current Population Survey helps here, as it records the state of residence a year earlier and the occupation of the longest job held in the previous calendar year, a job that for a mover may come before or after the move. With 29,031 interstate movers aged 18 to 64 who worked in the previous year over 2005 to 2024, it supports comparisons of the 22 major occupation groups only.

Movers do change work more often. Among workers employed in both years, 15.9 per cent of interstate movers are in a different major group from the year before, against 4.1 per cent of those who stayed in their state, and 26.7 against 10.8 per cent change detailed occupation. Both rates include the coding error that I document for these supplements in the companion paper (Conway, 2026), but the gap between movers and stayers is large.

However, classifying movers by the group held after the move makes little difference to the elasticity. Fitted to flows by state pair and major group in the supplement, with origin and des-

termination effects for each group, the model gives -0.771 pooled over groups whether the movers are classified by the group held now or by the group held a year earlier, close to the -0.756 of the American Community Survey. Across the 16 groups with at least 500 movers, the two sets of elasticities have a correlation of 0.95, close to mechanical as 84 per cent of these movers are in the same group in both years, and differ by -0.008 on average (standard deviation 0.05). The informative comparison is actually among the 3,786 movers who changed group, whose pooled elasticity is -0.82 by the group held now and -0.83 by the group held in the previous year. Thus, at least for major groups, the ordering by reach is not an artefact of the timing of the occupation question, and computer and mathematical occupations (-0.48 and -0.49) and building and grounds work (-1.00 and -1.04) stand at the two ends under either classification.

The measure as a table

All 209 elasticities, each with its standard error clustered by origin and destination, its standard error from resampling movers, and a reliability computed from the variance of the true differences between occupations, are in a file that accompanies the paper. In Section 7 I add a second form of the model, built for forecasting, with an effect for each pair of states common to all occupations and an exponent on the occupation's specialisation of each origin, and with it I test how much the differences in the elasticity matter for where an occupation's recruits come from.

5 National and local markets

I first use the measure to ask which occupations draw on a national labour market and which on a local one, and what goes with the difference.

The education of the mover

The occupations with a long reach are graduate occupations, which raises an obvious question: does reach belong to the occupation, or to the education of the people in it? Fitted to movers grouped by their own attainment, as in Table 2, the elasticity rises in steps from -0.97 for movers who did not finish high school to -0.58 for holders of graduate degrees, a difference of 0.39 with a standard error of 0.04, so that a non-bordering destination ten times as far away has 11 per cent as many movers in the first group and 26 per cent in the second. Fitted to all graduates and all non-graduates, the model gives -0.65 and -0.87 , close to the -0.62 and -0.91 implied by Cooke, Ellis, and Wright (2022) for a shorter period and a narrower age range.

The same difference appears within occupations: in 93 of the 102 occupations with at least 300 sampled movers with a degree and 300 without, the graduates have the longer reach, by a mean of 0.16, and the difference varies between occupations (standard deviation 0.15, interquartile range 0.08 to 0.25), although much of that variation is sampling noise.

Table 2: The distance elasticity by the education of the mover. Panel A is one pooled model on flows by state pair, major occupation group, and education, with origin and destination effects for each group and level. Panel B takes the 102 occupations with at least 300 sampled movers among graduates and among non-graduates. Standard errors in parentheses: clustered by origin and destination in panel A, and over occupations in panel B.

<i>A. By the education of the mover</i>	
Less than high school	−0.970 (0.042)
High school	−0.898 (0.053)
Some college or an associate’s degree	−0.837 (0.054)
Bachelor’s degree	−0.696 (0.072)
Graduate degree	−0.580 (0.063)
Graduate degree minus less than high school	0.390 (0.042)
<i>B. Graduates and non-graduates in the same occupation, 102 occupations</i>	
Mean elasticity of graduates	−0.688 (0.014)
Mean elasticity of non-graduates	−0.849 (0.011)
Mean difference within an occupation	0.161 (0.014)
Standard deviation of the differences	0.145
Occupations in which graduates have the longer reach	93
Occupations in which equality is rejected at five per cent	35

The share of the variance associated with education

Across the 209 occupations the correlation between reach and the share of the workforce with a degree is 0.61, with a standard error of 0.04. Corrected for the unreliability of the elasticity, it is between 0.72 and 0.76, depending on which measure of reliability is used. Figure 1 shows the relation, with a line fitted on the degree share alone that runs from −0.88 for an occupation with no graduates to −0.59 for one composed entirely of graduates.

A regression of the elasticity on the share with a degree, the share with a graduate degree, and the share under 30 has an R-squared of 0.42, held down by the noise in the dependent variable. Net of that noise the three regressors account for 57 per cent of the variance of the true differences when the noise is measured by resampling, and for 63 per cent when it is measured from alternate vintages.¹ Age matters too: the coefficient on the share under 30 is positive, at 0.20 with a standard error of 0.06, so occupations with a younger workforce have a longer reach given their education. Computed from panel A of Table 2, the elasticity that each occupation would have if it depended only on the education of its own movers accounts for 53 to 59 per cent of the true variance. In essence, roughly three fifths of the variance in reach between occupations is associated with education and age, and roughly two fifths is not.

The occupational component within levels of education

To see where the remaining variance lies, I estimated the reach of each occupation again using only movers of one level of education. As Table 3 shows, the standard deviation of the true differences between occupations rises steadily with education, from about 0.04 to 0.05 among movers with

¹As the dependent variable is an estimate, the regression is fitted by feasible generalised least squares with weights that combine the sampling variance of each estimate and the variance left unexplained, following Lewis and Linzer (2005). The share of the true variance accounted for is one minus the ratio of the unexplained variance, net of sampling variance, to the variance of the true differences.

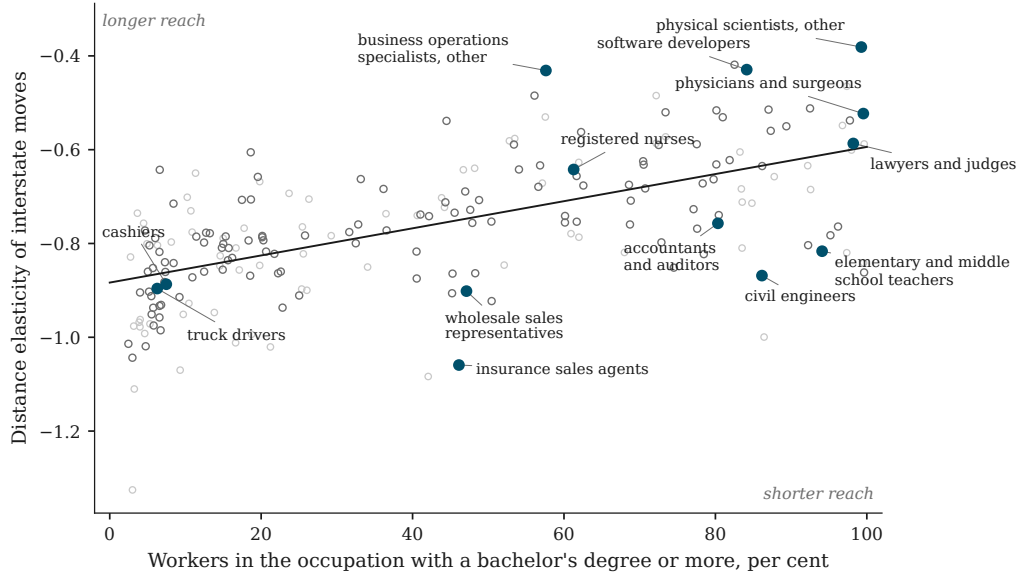


Figure 1: The distance elasticity of interstate moves against the share of the occupation's workforce with a bachelor's degree or more, for the 209 occupations with at least 500 sampled movers. Dark open circles have at least 1,000 movers and pale open circles fewer. The filled circles are the occupations named by a label, all with at least 1,000 movers. The line is fitted on the degree share alone by feasible generalised least squares.

high school only to 0.06 to 0.08 with some college, 0.09 with a bachelor's degree only, and 0.14 with a graduate degree, and the correlation between estimates from alternate vintages rises in the same order, from 0.08 to 0.22, 0.23, and 0.54.

As the rows of Table 3 cover different sets of occupations, I compare standard deviations, and the last two rows hold the set fixed: in the 102 occupations with enough movers of both kinds, the standard deviation of the true differences is 0.02 to 0.05 among non-graduates and 0.10 among graduates. Below a degree, then, the occupation adds little to reach beyond the education and age of its movers, and a planner can take the elasticity for the mover's education as the guide. Among holders of graduate degrees, by contrast, the occupation's own estimate is the better one.

Major groups differ too. In a pooled model with distance interacted with education and with major group, 11 of 21 groups differ from office and administrative work by more than 1.96 standard errors, the largest gaps being for arts and media and for computer and mathematical occupations (+0.17 each).

Departures of single occupations from the line fitted on education and age indicate what the occupational component consists of. Using the resampling error, which understates the noise, 37 of the 209 occupations lie more than 1.96 standard errors from that line, where about 10 would be expected by chance, and 9 lie beyond the Bonferroni bound. Six of the nine have a shorter reach than fitted, by the amounts in brackets: elementary and middle school teachers (0.21), insurance sales agents (0.31), civil engineers (0.27), wholesale and manufacturing sales representatives (0.16), special education teachers (0.36), and accountants and auditors (0.13). Special education teachers have 847 sampled movers and are the one occupation named in this paper with fewer than 1,000. The other three have a longer reach than fitted: software developers (0.19), the residual

Table 3: Differences in reach between occupations within the education of the mover. Each row uses the occupations with at least 300 sampled movers of that education. The standard deviation of the true differences is measured from resampling of movers and from the covariance of estimates on alternate vintages, and the last column is the correlation of the estimates on alternate vintages. A negative covariance is shown as a standard deviation of zero.

Education of the mover	Occupations	S.d. of true differences		Correlation
		Resampling	Vintages	
Less than high school	20	0.069	0.000	−0.018
High school only	64	0.038	0.050	0.083
Some college or an associate’s degree	117	0.058	0.078	0.216
Bachelor’s degree only	126	0.092	0.089	0.230
Graduate degree	73	0.141	0.146	0.539
Non-graduates, the occupations of Table 2	102	0.048	0.023	0.021
Graduates, the same occupations	102	0.100	0.105	0.352

category of business operations specialists (0.28), and the residual category of physical scientists (0.22). Secondary school teachers, social workers, and psychologists are also below the line, and postsecondary teachers and physicians are above it, all within the bound.

To be fair to the occupations below the line, it holds only education and age fixed, and school teaching and nursing are mostly done by women, who may more often move as partners of another mover. Part of a departure from the line may therefore belong to the household circumstances of the people in an occupation, which the survey does not let me separate. But, for an employer, the consequence is the same either way: a school district or an accountancy firm recruiting from other states draws on a nearer field than the education of its staff would suggest.

Credentialling

Several occupations below the line are ones where most workers hold a credential. The credentialling rate, the share reporting any professional certification or licence, contains no information on whether a credential transfers between states, and on its own it is unrelated to reach (−0.01, standard error 0.04). Given the share with a degree, its coefficient is −0.20 (0.04), which comes from graduate occupations (−0.25, 0.09, against −0.07, 0.07, among the others, a difference of −0.19 with a standard error of 0.12) and falls to −0.11 (0.07) under employment weights. Johnson and Kleiner (2020), by contrast, study licensing itself, with a measure of state-specific examinations, and find lower interstate migration where such examinations exist.

Age

Across five age bands, reach is longest at 25 to 34 (−0.70, against −0.83 at 55 and over, a difference of 0.13 with a standard error of 0.05), in agreement with the coefficient on the share under 30 above. Age is measured after the move, and the youngest band includes many moves from study to work.

Concentration across states

Whether an occupation is concentrated in a few states has little to do with how far its market extends. Its concentration across states has a correlation of only 0.17 with the distance elasticity, or 0.21 corrected for unreliability, and given the share with a degree its coefficient is 0.17 with a standard error of 0.15.

6 Remote work after 2020

Did remote work widen the labour markets of the occupations that can be done from home? Any change after 2020 has to be read against the trend before it, and so I start with how arrivals and the distance elasticity had been moving since 2005.

Change since 2005

Across the 209 occupations the interstate arrival rate was 2.25 per cent of employment in 2005 to 2009, 2.08 in 2010 to 2014, 2.24 in 2015 to 2019, and 2.35 in 2021 to 2024, and with each occupation's weight held fixed it was flat, at 2.29 in the first period and 2.27 in the last. The fifth of occupations with the highest arrival shares, among them physicians, scientists, and pilots, had about twice the rate of the others throughout (4.09 and 4.32 per cent in the first and last periods). These are moves of employed people from 2005, so they begin after much of the long decline in interstate migration that Molloy, Smith, and Wozniak (2011) and Kaplan and Schulhofer-Wohl (2017) study.

The distance elasticity moved towards zero over the same years. Fitted to flows by occupation and period with origin and destination effects for each, it was -0.80 in 2005 to 2009, -0.75 in 2010 to 2014, -0.74 in 2015 to 2019, and -0.72 in 2021 to 2024, a change of 0.09 with a standard error of 0.02. With a distance slope for each occupation the change within occupations is 0.05 (0.02), and it had appeared by 2010 to 2014, well before the pandemic. By education, the change was 0.04 among graduates and 0.08 among non-graduates, and the pooled change of 0.09 exceeds both because the graduate share of movers rose from 43 to 56 per cent. With origin and destination effects and a distance slope for each cell of five levels of education and five age bands, the change is 0.06 (0.02). For the 71 occupations with at least 2,000 movers, the change from the decade to 2014 to the years from 2015 averages 0.03, and the changes measured on odd and even vintages have a correlation of only 0.09, so the shift was common to occupations and, beyond sampling noise, little distinguishes one occupation's change from another's.

Arrivals in teleworkable occupations

If remote work weakened the tie between where people live and where their employer is, the occupations that can be done from home might show more arrivals, and a longer reach, after 2020. I grouped occupations by the measure of Dingel and Neiman (2020), under which 119 occupations with 30 per cent of employment can be done entirely at home and 274 with half of employment cannot. In the first group the share who worked at home rose from 9 per cent in 2019 to 30 per cent in 2021 and stood at 22 per cent in 2024, against 3, 8, and 6 per cent in the second. Standardised

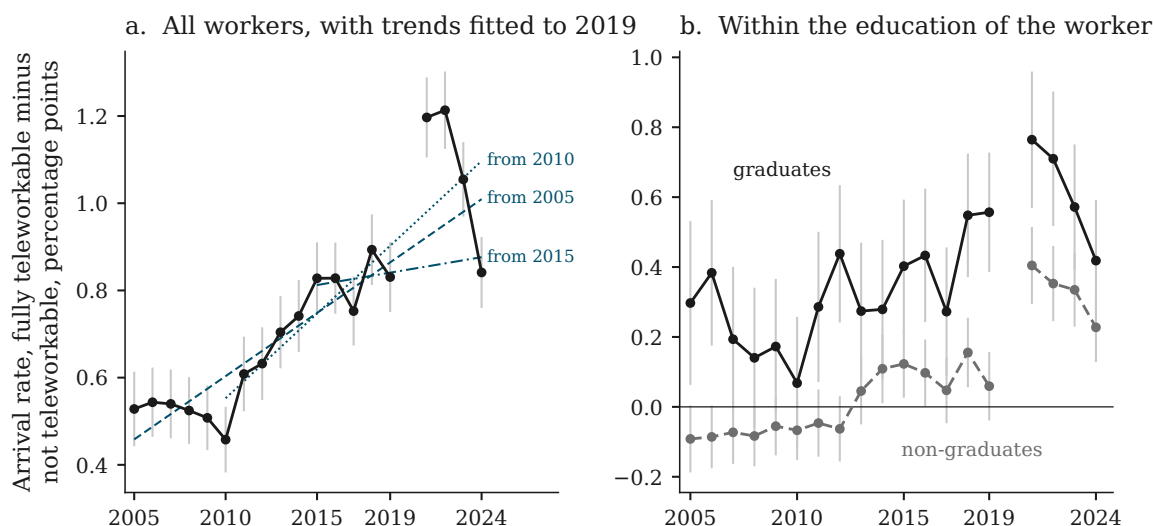


Figure 2: The interstate arrival rate of fully teleworkable occupations minus that of occupations that cannot be done at home, in percentage points, standardised to the age distribution of the pooled universe. Bars are 95 per cent intervals that treat records as independent. Panel a: all workers, with linear trends fitted to the years from 2005, 2010, and 2015 to 2019 and extended to 2024. Panel b: workers with and without a bachelor's degree. There is no 2020 vintage, and the 2021 file includes moves made during 2020.

for age, the interstate arrival rate of the fully teleworkable group exceeded that of the other by 0.46 points in 2010 and 0.83 in 2019, so the gap was already widening before 2020. It then stood at 1.20 in 2021, 1.21 in 2022, 1.05 in 2023, and 0.84 in 2024 (Figure 2, panel a).

Set against a linear trend fitted to the years before 2020, the gap in 2021 and 2022 was between 0.19 and 0.35 points above trend, depending on whether the trend starts in 2005, 2010, or 2015, and in 2024 it was between 0.26 and 0.04 points below, no more than the standard error of a single year's gap. To see whether departures of that size are unusual, I moved the end of the trend back to each year from 2010 to 2016 and measured the gap two and three years later, all before 2020. With trends of up to ten years, 3 of these 14 placebo departures are as large as those of 2021 and 2022, and with trends of five years none is. Among non-graduates one of 14 matches the rise with the longer trends and none with the shorter, and none does with each occupation's weight held fixed, which guards against the change to the occupation questions in 2019. Among graduates, by contrast, 10 or 12 of the 14 do, and with the weights held fixed for all workers 7 of 14 do with the longer trends and none with the shorter. As the 14 departures overlap and come from one series, they are not independent draws and indicate the size of ordinary movement. On that reading, the rise of 2021 and 2022 is distinct among non-graduates and within ordinary movement among graduates.

Teleworkable occupations are graduate occupations, so in panel b I make the comparison within education. Among graduates the gap was 0.56 points in 2019, 0.76 in 2021, and 0.42 in 2024, 0.14 below its level of 2019 (standard error 0.12), and among non-graduates 0.06, 0.40, and 0.23, still 0.17 above it (0.07). With the weight of every occupation present in all nineteen vintages fixed at its pooled employment, the gap was 0.45 points in 2019, 0.72 in 2021, 0.73 in 2022, and 0.41 in 2024.

Studies of moves during the pandemic include Haslag and Weagley (2024), on some 360,000

Table 4: Change in the distance elasticity of a fully teleworkable occupation relative to one that cannot be done at home, against 2005 to 2019. Poisson pseudo-maximum likelihood on flows by state pair, occupation, and period, with origin and destination effects for each occupation and period. Standard errors clustered by origin and destination in parentheses.

	2021 to 2022	2023 to 2024
All movers	−0.030 (0.016)	−0.027 (0.030)
All movers, with a distance slope for each occupation	−0.042 (0.017)	−0.033 (0.029)
All movers, with the degree share interacted in the same way	−0.008 (0.023)	0.019 (0.034)
Graduates only	−0.011 (0.030)	0.014 (0.030)
Non-graduates only	−0.053 (0.031)	0.001 (0.046)

household moves recorded by one moving company, Althoff, Eckert, Ganapati, and Walsh (2022), on the geography of workers whose jobs could be done remotely, and Ramani and Bloom (2021), a working paper that is not peer reviewed, who find on postal records that households left the centres of large cities for their suburbs with little reallocation between metropolitan areas. The small and brief rise in moves across state lines that I find here is consistent with a reallocation that was mostly local.

The distance elasticity

More arrivals need not mean arrivals from further away. I fitted the gravity model to flows by occupation and period, with origin and destination effects for each occupation in each period, and interacted distance with the teleworkability of the occupation and with the periods 2021 to 2022 and 2023 to 2024. Before 2020 a fully teleworkable occupation had an elasticity 0.15 nearer zero than one that cannot be done at home, and Table 4 gives the change in that difference after 2020, where a positive entry would mean that distance came to matter less.

No entry in the table is positive and distinguishable from zero. For 2021 to 2022 all five are negative, and the specification with a distance slope for each occupation, which uses only the change within an occupation, gives −0.04 with a standard error of 0.02. Within education the changes are smaller and less precise, so part of the negative estimate for all movers may be a change in who moved. Thus, distance did not come to matter less for teleworkable occupations after 2020, and their additional arrivals in 2021 and 2022 came, if anything, from slightly nearer.

7 Recruiting from other states

Forecasting the origins of arrivals

A planner requires something more direct than an elasticity. Given that a number of people will move from other states into an occupation in a destination, where will they have lived? I call this the origin composition of a given number of interstate arrivals, with occupation recorded after the move, and compare ways of forecasting it, each fitted on the moves of 2005 to 2014 and scored on the held-out arrivals of 2015 to 2024.

Four rules need no model. Origin shares can be set in proportion to the occupation’s employ-

ment in each other state, or to the realised origin shares of the same occupation and destination in the earlier decade, smoothed towards the employment shares with the weight of λ sampled movers. They can also ignore occupation and use the realised origin shares of all interstate movers into the destination, or multiply those shares by $(E_{ok}/E_o)^c$, the occupation's share of the origin's employment raised to a power. I chose λ and c on the earlier half of the fitting decade, fitting on 2005 to 2009 and scoring on 2010 to 2014, which gives $\lambda = 50$ from a grid of 1 to 1,000 and $c = 0.5$ from a grid of 0 to 1.5, so nothing is tuned on the held-out years.

Six models are fitted by Poisson pseudo-maximum likelihood with destination effects for each occupation, so that the origin shares they imply are those of a multinomial choice of origin (Guimarães, Figueiredo, and Woodward, 2003). Two are structural gravity with origin effects for each occupation (Anderson and van Wincoop, 2003; Head and Mayer, 2014), with distance and contiguity coefficients common to all occupations or for each, the latter being the measure of Section 4, and a third replaces the origin effect by the log of the occupation's employment in the origin. The last three replace origin effects and distance by an effect for each pair of states, estimated jointly on all 209 occupations, which holds the geography of migration common to them, and add the occupation's share of the origin's employment with an exponent common to all, an exponent for each, or an exponent for each with the occupation's own departure from the common effect of distance:

$$E[F_{odk}] = \exp\left(\pi_{od} + \gamma_{dk} + a_k \ln \frac{E_{ok}}{E_o} + d_k \ln D_{od}\right). \quad (2)$$

My main score is the mean log probability that a forecast gives the origin of each later arrival, weighted by the survey weights, which is a proper scoring rule (Gneiting and Raftery, 2007), and every forecast is mixed with equal shares in the proportion 0.99 to 0.01 so that no origin has probability zero. Only 0.3 per cent of later arrivals came from an origin that sent no mover of the occupation in the earlier decade. The second score is the total variation between forecast and realised shares, half the sum of their absolute differences, averaged over destinations by their arrivals.

Table 5 gives the scores. The rule that ignores occupation altogether, the origin shares of all movers into the destination, raises the probability given to the actual origin of an arrival by a factor of 1.41 per arrival over employment shares and 1.18 over the past shares of the occupation itself, and it is better than each in 209 and 206 occupations (1.42 and 1.25 in the average occupation). Every model with origin effects or origin employment does worse than this rule, the mass form in 189 of 209 occupations, as a decade of one occupation's movers is too few to learn a geography that all movers share. The best model, with pair effects, an exponent for each occupation, and an occupation's own departure from the common effect of distance, improves on the all-mover rule by 0.009 per arrival, a factor of 1.009, in 130 of 209 occupations. But it does no better than the rule that multiplies all-mover shares by the square root of the occupation's specialisation, which it beats in only 90. At destinations with at least 50 sampled arrivals the best forecasts close about two thirds of the distance between employment shares and the total variation that sampling alone would leave if a forecast were exact, which is at least 0.21.

In essence, the geography of migration is largely common to all occupations. On held-out years, where all movers to a state came from forecasts almost as well where the people who end up in a

Table 5: Forecasting the origins of the interstate arrivals of 2015 to 2024 from the moves of 2005 to 2014, 209 occupations, 251,165 sampled arrivals. Panel A, weighted by arrivals: log score, the mean log probability of the observed origin per arrival (higher is better), and total variation between forecast and realised origin shares, averaged over destinations (lower is better), for all destinations and for those with at least 50 sampled arrivals. Panel B, unweighted across occupations: mean difference in log score per arrival, standard error across occupations in parentheses, and the number of occupations in which the first scores higher. The best model is the one with pair effects, own exponent, and own distance.

<i>A. Scores</i>			
	Log score	Total variation	
		All	50 or more
Employment shares	−3.583	0.522	0.394
Past shares of the occupation, $\lambda = 50$	−3.404	0.460	0.317
All movers into the destination	−3.238	0.405	0.286
All movers $\times$ specialisation, $c = 0.5$	−3.230	0.400	0.280
Origin effects, common distance	−3.326	0.441	0.323
Origin effects, own distance	−3.322	0.439	0.319
Mass form, own distance	−3.319	0.448	0.327
Pair effects, common exponent	−3.231	0.401	0.280
Pair effects, own exponent	−3.231	0.401	0.280
Pair effects, own exponent and own distance	−3.229	0.399	0.277
<i>B. Differences in log score per arrival</i>			
	Difference	Occupations	
All movers over employment shares	0.351 (0.008)	209	
All movers over past shares of the occupation	0.222 (0.007)	206	
Best model over all movers	0.009 (0.002)	130	
Best model over all movers $\times$ specialisation	−0.000 (0.001)	90	
Origin effects, own over common distance	0.002 (0.001)	116	

given occupation there came from, and the information specific to the occupation, its specialisation across origins and its own sensitivity to distance, adds about one per cent to the probability of each arrival's origin. In total variation at well-sampled destinations it takes the all-mover rule from 0.286 to 0.277, about 12 per cent of the distance that rule stood above the sampling floor. As the rule that multiplies all-mover shares by the square root of specialisation, with its one parameter, ties the best model, a planner can use that rule. The differences in reach measured in Sections 4 and 5 are real, but beside the common geography they matter little for accuracy out of sample (0.001 per arrival in the best model and 0.002 among the models with origin effects).

Specialisation of the origin

Occupations differ in the exponent on specialisation. Refitted on all years, the model gives a mean of 0.51 (10th and 90th percentiles 0.16 and 0.84): given the movers that a pair of states exchanges, an origin where the occupation's share of employment is twice as high sends about 1.4 times as many of the occupation's arrivals. Estimated separately on odd and even vintages, the exponents have a correlation of 0.42 across occupations, a reliability of 0.59 for the full sample, so their differences are real. They are highest for postsecondary teachers (1.31), physicians (1.24), and accountants (1.06) and near zero for secretaries (0.09), personal care aides (0.08), and hand labourers (0.04), with a correlation of 0.31 with the share holding a degree. Because occupation is recorded after the

move, an exponent near zero means that the occupation's arrivals came from places regardless of how much the occupation is practised there, as when people take the work on arrival. An exponent near one or above means that they came from where the occupation is practised, as when people move within it. The high values for physicians and university teachers may also reflect where they trained, a move from study to work that the survey does not separate from a move between jobs (Section 9).

The area of arrivals

Averaged over destinations by their arrivals, the distance within which half the modelled arrivals start has a median of 1,073 km and runs only from 964 to 1,199 km between the 10th and 90th percentiles of occupations. Among those with at least 1,000 movers it is longest for pilots (1,404 km) and shortest for industrial truck operators (829 km).

A new employer

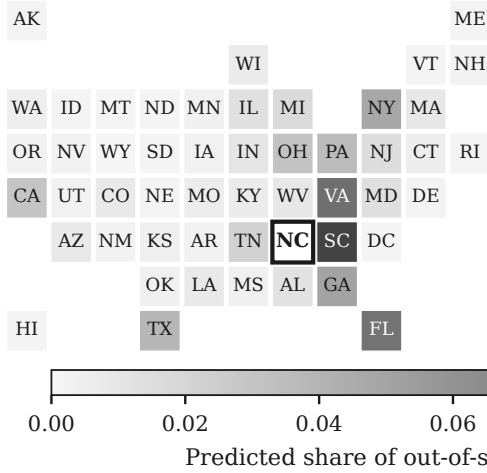
To bring this down to particular places, I apply the model in Table 6 to six cases, each a destination where an employer recruits an occupation from other states, and Figure 3 shows two of them. The cases give shares of out-of-state recruits, on the assumption that recruits follow the pattern of past movers and that wages and housing do not change. How many of its workers an employer finds within the state depends on the local labour market and its wage offer, neither of which the model contains. The first column is the number of interstate arrivals in a year as a share of the occupation's workers in the state, over 2015 to 2024, a measure of how much the occupation there already depends on them: 1.5 per cent for welders in North Carolina and 6.3 per cent for software developers there. In the next three columns the shares by distance band come from the model fitted on all years, and the comparison with the observed shares of 2005 to 2024 is in sample. For the out-of-sample check I score the model fitted on 2005 to 2014 against the arrivals of 2015 to 2024. Over all pairs of occupation and destination with at least 50 sampled arrivals in those years, weighted by the arrivals of the destination, its error in the share of a distance band has a root mean square of 0.057 to 0.063, about one standard deviation of the error of a single forecast, of which sampling alone accounts for about 0.046 to 0.048 at the median of 80 sampled arrivals. On average the model understated the share from beyond 1,600 km by 0.012 and overstated the share from within 800 km by 0.015. The margin comes from cells with at least 50 arrivals and is wider for thinner ones.

A North Carolina plant recruiting welders from other states would, on the pattern of past moves, find nearly three fifths of them within 800 km, the most from South Carolina (12 per cent), Virginia (9), and Florida (9). The observed shares for this destination rest on 35 sampled movers, 19 of them after 2014, and the distance between forecast and record out of sample (0.47) is no more than sampling alone would produce (0.50). For a single occupation in a single state, then, the record is too thin to use directly, and a model fitted to all of them is the better guide. By contrast, a Texas employer recruiting software developers would find three fifths of them beyond 1,600 km, a fifth of them from California. Out of sample the excess of the total variation over sampling is largest for truck drivers in Missouri (0.32 against 0.21) and software developers in Texas (0.18 against 0.10),

Table 6: Where out-of-state recruits would come from, six cases. First column: interstate arrivals in a year as a share of the occupation’s workers in the destination, 2015 to 2024. Next three: forecast shares of out-of-state arrivals by distance from the destination, model with pair effects fitted to 2005 to 2024, with the observed shares of 2005 to 2024 in brackets (in sample). The out-of-sample root mean square error of a band share, about one standard deviation, is 0.057 to 0.063 in cells with at least 50 arrivals. Last two columns: total variation between the model fitted on 2005 to 2014 and the arrivals of 2015 to 2024, beside the value expected from sampling alone, with the number of sampled arrivals in 2015 to 2024.

	Newcomers	Under	800 to	Over	Out of sample	
	per cent	800 km	1,600 km	1,600 km	Variation	Sampling (<i>n</i>)
Welders, North Carolina	1.5	0.57 (0.51)	0.26 (0.29)	0.17 (0.20)	0.47	0.50 (19)
Assemblers, North Carolina	2.0	0.57 (0.58)	0.26 (0.27)	0.17 (0.15)	0.36	0.33 (50)
Software developers, N. Carolina	6.3	0.49 (0.48)	0.25 (0.24)	0.26 (0.27)	0.20	0.14 (274)
Software developers, Texas	5.7	0.07 (0.05)	0.34 (0.31)	0.59 (0.64)	0.18	0.10 (639)
Registered nurses, Arizona	4.4	0.28 (0.23)	0.18 (0.15)	0.54 (0.62)	0.22	0.16 (221)
Truck drivers, Missouri	1.9	0.57 (0.55)	0.30 (0.31)	0.13 (0.14)	0.32	0.21 (116)

a. Welders arriving in North Carolina



b. Software developers arriving in Texas

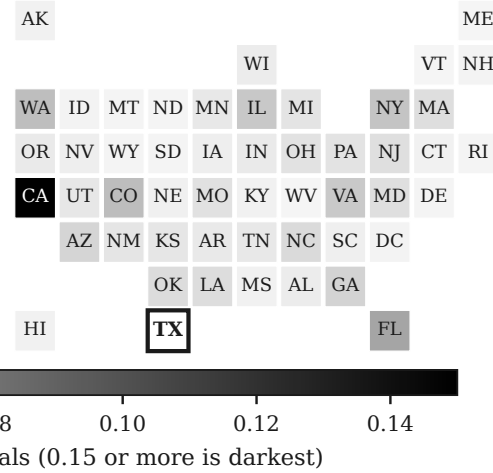


Figure 3: Forecast shares of out-of-state arrivals by state of origin, model with pair effects fitted to 2005 to 2024, one square for each state placed roughly where it lies. The outlined square is the destination. Panel a: welders arriving in North Carolina. Panel b: software developers arriving in Texas.

and for Texas the model fitted on the first decade put too few of the later arrivals beyond 1,600 km (0.57 against 0.63).

A fall in the cost of distance

If remote work lowered the cost of living far from an employer, the distance elasticity of teleworkable occupations would move towards zero. I multiply the model’s shares for the 93 occupations whose teleworkability is at least one half by D_{od}^{Δ} , which changes their distance elasticity by Δ , under three scenarios: $\Delta = 0.161$, the mean difference between graduates and non-graduates within an occupation (Table 2), an elasticity raised to -0.52 , the 95th percentile of the estimates of Section 4, and an elasticity of zero, a ceiling where distance no longer deters movers. The median distance within which half their arrivals start lengthens from 1,087 km to 1,230 km under the first, 1,265 km

under the second, and 1,774 km under the ceiling, and the share from beyond 1,600 km rises from 0.32 to 0.36, 0.38, and 0.53. These remain scenarios, however, and as Section 6 reports, after 2020 the elasticity of teleworkable occupations did not actually move towards zero.

8 Places specialised in an occupation

A question often put is whether workers move to the places that specialise in their occupation, so that technology workers go to California as production workers go to North Carolina, and whether those places then keep them. I answer both with the gravity model, replacing its destination effect by the size of the occupation at the destination.

Arrivals

For each occupation I fitted

$$E[F_{odk}] = \exp(\alpha_{ok} + \ln T_{od} + \theta_k \ln Q_{dk} + \beta_k \ln D_{od} + \delta_k C_{od}), \quad (3)$$

where T_{od} , the flow of all interstate movers between the two states, enters with a coefficient of one, and Q_{dk} is the location quotient of occupation k among workers who did not move between states, its share of the destination's employment divided by its national share. If movers simply took up the occupations that the destination employs, F_{odk} would be T_{od} times the occupation's share of the destination's employment, and θ_k would be one. A value above one means that the arrivals in an occupation are more concentrated in the states specialised in it than its incumbents are, and a value below one that they go disproportionately to states where it is rare. Because occupation is recorded after the move, a coefficient of one cannot separate a choice of place from taking the work a place offers, but differences between occupations in θ_k are informative, as migration that ignored occupation would give every occupation the same value. When I redraw each mover's occupation from the mix of non-movers at the destination, as a placebo, the estimates from odd and even vintages have a correlation of -0.01 .

In the data that correlation is 0.55, so θ_k has a reliability of 0.71 and the standard deviation of the true differences between occupations is 0.29, as Table 7 shows. Its mean is 0.87, or 0.94 once each estimate is corrected for error in the location quotient, which is built on few records for small occupations in small states. On average, then, arrivals are spread across destinations about as incumbents are, and occupations depart from that in both directions. As for the occupations named in the question, they are close to one, with software developers at 0.90 (standard error 0.04), assemblers and fabricators at 0.94, and welders at 0.99, so in these data arrivals in software are no more concentrated in the states that specialise in it than arrivals in factory work are. Physicians (1.31) and postsecondary teachers (1.50) have values above one, and healthcare support (0.51) and health practitioners (0.56) have the lowest group means, with registered nurses at 0.19, so arrivals in nursing are less concentrated than nurses already are in the states that specialise in nursing. The extremes deserve some caution, as they are occupations whose share varies little across states (a median standard deviation of the log quotient of 0.18 for the ten furthest from one, against 0.27 for

Table 7: Places specialised in an occupation. Panel A: the coefficient θ_k of equation (3) across the 209 occupations, in the data and in a placebo in which each mover’s occupation is drawn from the mix of non-movers at the destination. Panel B: linear probability of moving to another state, in percentage points, on the log location quotient of the major group of the previous year’s job in the state of a year earlier, March supplement of the Current Population Survey 2005 to 2024, with effects for group, state, and year, age band, and education, and standard errors clustered by state in parentheses.

<i>A. Arrivals relative to incumbents, 209 occupations</i>		
	Data	Placebo
Mean of θ_k	0.865	1.012
Mean corrected for error in the quotient	0.944	
Correlation of estimates from alternate vintages	0.548	−0.011
Standard deviation of true differences	0.286	0.000
<i>B. Leaving the state, 1,668,909 workers, mean rate 1.78 per cent</i>		
All workers	−0.926	(0.176)
All workers, 2007 to 2024	−0.889	(0.152)
Without the District of Columbia	−0.783	(0.105)
The 25 largest states	−0.670	(0.099)
Unweighted	−0.564	(0.101)
Groups concentrated across states, above the median	−0.881	(0.195)
Groups below the median	−1.055	(0.213)
Graduates	−1.510	(0.312)
Non-graduates	−0.624	(0.139)
Aged 25 to 34	−1.419	(0.302)
Aged 45 to 54	−0.435	(0.246)

all 209), so their estimates rest on small differences. Furthermore, when the coefficient on the flow of all movers is left free, and not fixed at one, it averages 0.99 but is 0.86 for postsecondary teachers, whose value falls to 1.17, while physicians (1.26) and nurses (0.26) change little. The value of θ_k is unrelated to the share with a degree, teleworkability, the concentration of the occupation, and the distance elasticity, with correlations between -0.08 and -0.01 , so it describes a second and separate dimension of an occupation’s market.

Leavers

For leavers I return to the supplement of the Current Population Survey used in Section 4, which records the occupation of the previous calendar year and so can show who leaves the states that specialise in an occupation, classified mostly by the work they did there. I regressed an indicator of having moved to another state on the log location quotient of the major group of the job held in the previous calendar year in the state of residence a year earlier, with effects for the group, the state, and the year, and controls for age and education. Of the coefficients in panel B of Table 7, the first, -0.93 , means that one standard deviation of the log quotient across workers (0.16) goes with a probability of leaving 0.15 points lower, against a mean rate of 1.78 per cent, so a state that specialises in an occupation keeps its workers somewhat better than one where the occupation is rare. Almost no worker is in a state with a quotient above 2.7. The variants in panel B, after the first two supplements, where Kaplan and Schulhofer-Wohl (2012) found that allocation inflated interstate moves, without the District of Columbia, in the 25 largest states, and without weights, give coefficients between -0.89 and -0.56 .

In percentage points, graduates respond to specialisation more than twice as strongly as other workers, and workers aged 25 to 34 more than three times as strongly as those aged 45 to 54. The response is similar for occupations that are concentrated in a few states and for those spread evenly. These estimates fit the view of Manning and Petrongolo (2017) that labour markets are local in the sense that job search falls off quickly with distance, so a worker in a place with many jobs in an occupation has less reason to leave it. Of course, the specialisation of a state is not assigned at random, so the estimates describe who leaves and who stays. Even so, for a state that already specialises in an occupation they mean that its workers are a little less likely to be lost to other states, most of all its young and its graduates.

9 Limits of the data

Occupation after the move

For major groups the elasticities hardly change when movers are classified by the group of the previous year (Section 4), but detailed occupations cannot be checked in this way. Moves from study to work raise a related difficulty, as the origin of a newly qualified physician or a new university teacher is where they trained, and the survey does not separate such moves from moves between two jobs.

The model and the survey

Distance is measured between state centres of population, so a move across a state line within one metropolitan area is given the distance between two states, which the contiguity term absorbs only in part. The model also assumes one elasticity over all distances, and its standard errors ignore the strata and clusters of the survey design. If the moves the survey misses, or those of people in group quarters, differ in distance from the reported ones, the elasticities are affected in a direction I cannot determine. Lastly, as the 2021 file includes moves made during 2020, Section 6 cannot date the rise more closely than that.

10 Conclusion

Published statistics on moves between states describe persons and not occupations, and I have built from them a measure of how far the labour market of each occupation extends. Fitted to 209 occupations, the distance elasticity of interstate moves averages -0.77 , about two thirds of its variance between occupations is real, and for major groups it hardly changes when movers are classified by the occupation of the previous year.

Most of the difference between national and local markets goes with education and age, which account for 57 to 63 per cent of the variance of the true differences between occupations, net of sampling noise (the raw R-squared is 0.42). The occupational component is largest among holders of graduate degrees, where arrivals in school teaching and civil engineering came from nearby and arrivals in software and physical science from far away. After 2020 the elasticity of teleworkable

occupations did not move towards zero.

For a place recruiting from other states, the geography of migration common to all occupations matters far more than anything specific to an occupation, which adds about one per cent per arrival. Occupations differ, however, in how strongly their arrivals come from places that specialise in them, strongly where people move within an occupation and weakly where they take the work on arrival. A rule with one parameter, the origin shares of all movers into the destination times the square root of the occupation's specialisation of each origin, does as well as the best model. For a single occupation and state, where the survey's own record is thin, the model gives a better guide than that record, with an error measured on better-sampled cells, as it does for a North Carolina plant recruiting welders, nearly three fifths of whom would come from within 800 km.

A state that specialises in an occupation keeps its workers somewhat better than one where the occupation is rare, and arrivals in most occupations are spread across destinations about as incumbents are, with physicians and postsecondary teachers drawn more than their incumbents to the states that specialise in them and registered nurses less.

The measure can be estimated for any occupation with enough movers, and its reliability reaches 0.78 among the 133 occupations with at least 1,000. A source that recorded the occupation before and after a move for detailed occupations would separate the movement of an occupation's own workers from that of the people who enter it on arrival, so that the same model could describe both.

References

- Lukas Althoff, Fabian Eckert, Sharat Ganapati, and Conor Walsh. The geography of remote work. *Regional Science and Urban Economics*, 93:103770, 2022.
- James E. Anderson and Eric van Wincoop. Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192, 2003.
- Michel Beine, Simone Bertoli, and Jesús Fernández-Huertas Moraga. A practitioners' guide to gravity models of international migration. *The World Economy*, 39(4):496–512, 2016. doi: 10.1111/twec.12265.
- Alessandro Conway. An occupational flow account for the United States, 2009 to 2025: Mobility, the sources and destinations of each occupation's workers, and the industry shock of 2020 and 2021. Working paper, not peer reviewed, 2026.
- Thomas J. Cooke, Mark Ellis, and Richard Wright. Occupational licensing and interstate migration in the United States. *Monthly Labor Review*, August 2022. doi: 10.21916/mlr.2022.22.
- Jonathan I. Dingel and Brent Neiman. How many jobs can be done at home? *Journal of Public Economics*, 189:104235, 2020.
- Thomas B. Foster, Lee Fiorio, and Mark Ellis. Internal migration in the U.S. during the COVID-19 pandemic. Working Paper CES-24-50, Center for Economic Studies, U.S. Census Bureau, 2024. Working paper, not peer reviewed.

- Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- Paulo Guimarães, Octavio Figueiredo, and Douglas Woodward. A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204, 2003.
- Peter H. Haslag and Daniel Weagley. From L.A. to Boise: How migration has changed during the COVID-19 pandemic. *Journal of Financial and Quantitative Analysis*, 59(5):2068–2098, 2024.
- Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. In Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, editors, *Handbook of International Economics*, volume 4, pages 131–195. Elsevier, 2014.
- Janna E. Johnson and Morris M. Kleiner. Is occupational licensing a barrier to interstate migration? *American Economic Journal: Economic Policy*, 12(3):347–373, 2020.
- Greg Kaplan and Sam Schulhofer-Wohl. Interstate migration has fallen less than you think: Consequences of hot deck imputation in the Current Population Survey. *Demography*, 49(3):1061–1074, 2012.
- Greg Kaplan and Sam Schulhofer-Wohl. Understanding the long-run decline in interstate migration. *International Economic Review*, 58(1):57–94, 2017.
- John Kennan and James R. Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011.
- Jeffrey B. Lewis and Drew A. Linzer. Estimating regression models in which the dependent variable is based on estimates. *Political Analysis*, 13(4):345–364, 2005.
- Alan Manning and Barbara Petrongolo. How local are labor markets? evidence from a spatial job search model. *American Economic Review*, 107(10):2877–2907, 2017.
- Ioana Marinescu and Roland Rathelot. Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics*, 10(3):42–70, 2018. doi: 10.1257/mac.20160312.
- Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Internal migration in the United States. *Journal of Economic Perspectives*, 25(3):173–196, 2011.
- Arjun Ramani and Nicholas Bloom. The donut effect of COVID-19 on cities. Working Paper 28876, National Bureau of Economic Research, 2021.
- Steven Ruggles, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rogers, Jonathan Schroeder, and Kari C. W. Williams. IPUMS USA: Version 16.0 [dataset], 2025.
- J. M. C. Santos Silva and Silvana Tenreyro. The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658, 2006.

Aba Schwartz. Interpreting the effect of distance on migration. *Journal of Political Economy*, 81 (5):1153–1169, 1973.

U.S. Census Bureau. About occupation. <https://www.census.gov/topics/employment/industry-occupation/about/occupation.html>, n.d.

A Arrival shares and the national share

The arrival share is the proportion of an occupation’s workforce that lived in another state a year earlier. Across the 397 occupations with at least 5,000 person records it has a median of 2.10 per cent and a 95th percentile 3.7 times the 5th, and sampling variance is 2.4 per cent of its variance across occupations. The national share was 2.38 per cent in 2005 and 2.13 per cent in 2024. On the 432 occupation codes present in every vintage, the part of the change due to the mix of occupations is +0.124 points to 2024, with a jackknife standard error of 0.006, and the part due to arrival shares within occupations has no stable sign.

The occupation series is discontinuous after 2018, in that the share of employment that changes code between consecutive vintages is 4.1 per cent between 2018 and 2019 and 4.3 per cent between 2019 and 2021, against 1.5 to 2.5 per cent at every other link. The Census Bureau changed the occupation questions in 2019 and the second link spans the pandemic, and I have not established which of the two produced the discontinuity.