Do skills predict moves? A gravity test of skill similarity against observed occupational transitions

Alessandro Conway

First draft 29 January 2026. This version September 2026.

Working paper. The model, its inputs, and its readings are fixed, and the estimation waits on the bilateral transition matrix.

Extended summary

A whole advisory industry now tells workers and workforce boards which occupations they can move to, and it does so with a similarity score. Two occupations whose skill profiles overlap at 98.5 per cent are presented as a few months of retraining apart, a machinist is directed towards industrial maintenance, and public money for bridge programmes and reskilling grants goes where the score points. The score is a claim about feasibility, that a person with one profile could do the other job. Whether anyone makes the move is a different claim, and the advice depends on it, yet it has not been tested at scale, although the data to test it are public and the test costs little.

This paper specifies the test and fixes its inputs. Observed flows between occupations come from the March supplement of the Current Population Survey, which records in one interview what a person does now and what they did last year. The companion flow account (Conway, 2025) shows that this is the only design in that survey which measures occupational transitions at all. Two independent codings of a job the survey itself records as unchanged disagree 48.1 per cent of the time at the detailed level, so a matrix linked across interviews measures coding disagreement. The disagreements also fall along boundaries in the classification, such as elementary against secondary teachers or supervisors against the staff they supervise, and those are the pairs a similarity score ranks highest. A linked panel would therefore find similarity predicting moves whether or not anyone moved. The within-interview design cancels that error, and it is the one used here. Similarity is a composite over the skills, knowledge, and work activities the Occupational Information Network records, held as 798,342 occupation pairs above a similarity of 0.5, and the tier thresholds used in workforce practice sit at 0.985, 0.970, and 0.950.

The model is a gravity equation estimated by Poisson pseudo-maximum likelihood on gross flows, with origin and destination effects that absorb the size of each occupation. The elasticity of flows to similarity is therefore identified from how flows vary across the destinations of one origin, holding the wage differential. Similarity is symmetric in a pair and the wage differential is antisymmetric, which divides the test in two. With origin and destination effects, the level of traffic between a pair identifies the similarity elasticity; with pair effects added, similarity drops out and the wage differential is identified from the direction of that traffic alone. The flow account has already shown that between a third and two fifths of gross lateral movement has a direction at the pair level and that a tenth changes the size of any occupation. The two halves of the test therefore separate what a similarity score can be expected to explain, which is traffic, from what it cannot, which is where the traffic nets to. A flexible version of the model enters similarity in bins instead of as one elasticity, locates where observed flows break, and sets the breaks against the three thresholds in use. Licensing is read from the residuals rather than entered as a regressor: the pairs that similarity places close and

that nobody is observed crossing are a map of what the score misses, and the residuals show whether that map concentrates in licensed occupations.

1 The claim under test

Skill-similarity scores are built from occupational profiles. Each occupation in the Network has importance ratings on thirty-five skills, thirty-three knowledge areas, and forty-one work activities, and a composite similarity between two occupations is the cosine of their two rating vectors, weighted one half on skills and a quarter each on knowledge and activities. The scores are used in career tools, in workforce board planning, in the design of bridge programmes, and in the shortage arithmetic of the companion paper (Conway, 2026a), where a tier of adjacency decides which occupations can release workers into a short one. In every one of those uses a high score is read as evidence that people move.

The score establishes something narrower. It measures whether two jobs draw on the same abilities in similar proportions, which bears on what a worker could do. A worker's decision to move also depends on a wage differential, on where the destination jobs are, on whether a licence stands in the way, on what the worker knows about the destination, and on whether the destination is hiring. Gathmann and Schönberg (2010) showed that workers do move along task-distance gradients in German administrative data, and Kambourov and Manovskii (2008) showed that occupation-specific human capital is large enough to keep people where they are. Neither result gives the realised movement associated with a point of similarity in the United States at the detailed level, and neither tested the thresholds that a practitioner's tool now draws between an easy move and a career change. That elasticity and those thresholds are what this paper estimates.

2 The transition matrix

The natural object is a transition matrix, the number of people who were in occupation i last year and are in occupation j now. The survey has two ways of building one, and the flow account measures the difference between them. Following a person across two interviews a year apart, with occupation coded independently on each occasion, gives an annual mobility of 50.1 per cent at the detailed level. Among the 114,491 people whom the survey itself records as holding the same occupation in both years and working all fifty-two weeks, the two codings still disagree 48.1 per cent of the time. The March supplement instead asks in one interview what a person does now and what they did last year, both codes are assigned at one sitting from one respondent's description, and the annual mobility it gives on the same linked people is 10.3 per cent. The errors of the two codes are correlated, and the correlation cancels in the difference.

For this paper the choice matters for a reason beyond precision. The disagreements between independent codings follow the boundaries of the classification: the largest single confusion is elementary against secondary school teachers, followed by supervisors against the staff they supervise and by generic against specific managerial categories. Those pairs have near-identical skill profiles, so a matrix built from linked interviews contains spurious transitions concentrated where similarity is highest, and a gravity model estimated on it would return a large positive elasticity that measures the coding process. The within-interview matrix is the flow account's own construction, by year, age band, origin, and destination, for the 230 occupations the survey supports, with a standard error on every cell from the 160 replicate weights the Census Bureau publishes. It is built and not yet published, and this paper runs on it once it is.

3 The model

Write f_{ij} for the weighted number of people observed moving from occupation i to occupation j over the pooled years, s_{ij} for their composite similarity, and w_i and w_j for the mean wages of the two occupations in the Occupational Employment and Wage Statistics. The estimating equation is

$$\mathbb{E}\left[f_{ij} \right] = \exp\left(\eta \ln s_{ij} + \beta \ln \frac{w_j}{w_i} + \alpha_i + \delta_j \right),$$

estimated by Poisson pseudo-maximum likelihood, which keeps the pairs where nobody moves and handles the heteroskedasticity that a log-linear specification mishandles (Santos Silva and Tenreyro, 2006). The origin effect α_i absorbs the size of the origin, its overall propensity to lose people, and everything else that moves its departures as a whole, and the destination effect δ_j does the same for arrivals. The elasticity η is then identified from the variation in similarity across the destinations of one origin, and across the origins of one destination, holding the wage differential between them.

The migration paper (Conway, 2026b) shows why the pooled model is the one to estimate. There a wage elasticity estimated one occupation at a time was identified only from year-to-year movement in a survey cell mean, of which a reliability calculation attributed 54 per cent to sampling noise, and the per-occupation estimates were a null produced by the design. Pooling identified the elasticity from cross-sectional differences at a reliability of 0.76, and the estimate came out at 0.114, weak and precisely measured. The same holds here: the pooled model identifies η from variation across pairs, and any per-occupation version is reported as a distribution rather than as the estimate.

Two properties of the regressors divide the test. Similarity is symmetric, $s_{ij} = s_{ji}$, and the wage differential is antisymmetric, since $\ln(w_j / w_i) = -\ln(w_i / w_j)$. Adding a pair effect μ_{ij}, symmetric by construction, absorbs similarity entirely, together with every symmetric friction between the pair, and leaves the wage differential identified from the asymmetry of the pair’s traffic alone, which is how many more people go from i to j than from j to i. The first specification, with origin and destination effects, therefore measures what similarity explains, which is the level of traffic between a pair, and the second, with pair effects, measures what the wage differential explains, which is its direction. The flow account gives reason to expect the two to differ. Across the whole matrix, gross lateral movement runs at 19.0 million people a year in 2022 to 2025 at the detailed level, of which 42.1 per cent has a direction at the pair level and 10.4 per cent changes the size of any occupation. Over 2009 to 2025 the pair share is 32.3 per cent. Most of what a similarity score would predict is two-way traffic between neighbouring occupations, while a workforce board wants to know whether a move will fill a vacancy that stays filled, which is a question about the smaller antisymmetric part.

The thresholds are tested by relaxing the functional form. In place of $\eta \ln s_{ij}$ the model takes indicators for bins of similarity of width 0.005 from 0.90 to 1, with everything below 0.90 as the base, and the sequence of bin coefficients traces flow against similarity net of size and wages. A threshold is confirmed where the curve steps by more than the standard errors of the adjacent bins, and the three cutoffs in use, 0.985, 0.970, and 0.950, are set against the steps the data contain. A smooth curve with no step anywhere would mean that the tiers are a convenience of presentation, uninformative about where a move becomes hard.

4 Licensing

The specification that preceded this one entered a licensing indicator at the destination as a regressor. It is dropped, for a reason the migration paper established. The survey's only credentialling variable records any active certification or licence and is an upper bound on legal licensing, the compact timelines that would identify a licensing effect exist for one occupation, and a regressor that measures the wrong thing precisely is worse than none.

Its place is taken by the residual map. The fitted model predicts a flow for every pair from similarity, wages, and the two size effects. Where the observed flow falls far below the prediction, at a rate indistinguishable from zero on stocks large enough to show one, something outside the score blocks the move. The map is reported in full, as the pairs sorted by the ratio of observed to predicted flow with the standard error of the observed cell, and it is read three ways. Whether the blocked pairs concentrate in destinations that the Bureau of Labor Statistics classifies as requiring a licence is the licensing test, run on every occupation the matrix covers instead of on the nurse example alone. Whether they concentrate in pairs whose destinations are geographically concentrated is the location test, using the interstate migration measures of the companion paper. The pairs that remain after both are the information gaps, moves that are feasible, permitted, and available and that people do not make, and those are where a bridge programme has something to do.

5 Prespecification

The inputs are four, and each has a source. The transition matrix is the flow account's within-interview build from the 2009 to 2025 supplements, pooled with year effects, on the Census 2010 occupation codes. Similarity is the composite over the Network's importance ratings, keyed to the Standard Occupational Classification and mapped to Census codes by the account's bridge, with the employment-weighted mean similarity taken where a Census code maps to several classification codes. Wages are the national means of the May 2024 Occupational Employment and Wage Statistics, mapped to Census codes the same way. Employment enters only through the fixed effects.

The readings are fixed as well. The primary estimate is η from the first specification with standard errors clustered by pair, reported with β from the second. The tier test is the bin curve and the location of its steps against 0.985, 0.970, and 0.950, and the residual map is the full list. An η within two standard errors of zero on the pooled matrix would end the paper, since it would mean that a similarity score contains no information about who moves and that advice built on one has no empirical basis. The practice would be vindicated by an η large enough that a move from the first tier to the third cuts the predicted flow by more than half, with steps in the bin curve at or near the cutoffs in use. Whichever result it is, the estimated thresholds replace the assumed ones in the shortage arithmetic of the companion paper, which takes them as given until this one reports.

6 Conclusion

Similarity scores have been used as evidence about movement without being tested against movement, because no transition matrix measured moves instead of coding noise. The flow account supplies the matrix and the Network supplies the score, and a gravity model on public data then gives the elasticity that every skill-based recommendation currently implies without stating it. The model, its inputs, and its readings are fixed here, and the estimation follows the matrix.

References

Conway, A. (2025). Where workers come from: an occupational flow account for the United States. Working Paper AC-WP-2025-01.

Conway, A. (2026a). Shortage arithmetic: the substitution margin an occupational gap leaves out. Working paper.

Conway, A. (2026b). Where occupations move: interstate migration by occupation in the United States, 2005 to 2024. Working Paper AC-WP-2026-02.

Gathmann, C. and Schönberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, 28(1), 1–49.

Kambourov, G. and Manovskii, I. (2008). Rising occupational and industry mobility in the United States: 1968–97. *International Economic Review*, 49(1), 41–79.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4), 641–658.