Shortage arithmetic: the substitution margin an occupational gap leaves out

Alessandro Conway

First draft 17 March 2026. This version September 2026.

Working paper. The correction is specified and its first term is measured on five occupations, and the adjacency terms wait on the bilateral transition matrix.

Extended summary

An occupational shortage figure is a subtraction. A projection of the positions an occupation will need to fill is set against a count of the people being trained for it, and the difference is published as the number a state is short. The subtraction assumes something that is never written down: that a position in the occupation can be filled only by a person newly trained for exactly that occupation, so that nobody arrives from anywhere else. The occupational flow account built in the companion paper (Conway, 2025) measures where an occupation's people come from, and in every one of the 172 occupations it covers, arrivals from another occupation exceed arrivals from education, by a median factor of 19. The assumption fails across the whole distribution, and the question is what it costs.

This paper writes the correction down and measures its first term. For five healthcare occupations in North Carolina, arrivals from another occupation at the rate the national account observes come to between 49 and 158 per cent of the annual openings the state projection expects: 49.2 per cent for registered nurses with a standard error of 4.1, 123.7 for licensed practical nurses, 91.9 for nursing assistants, 98.9 for medical assistants, and 157.8 for nurse practitioners. Those arrivals are gross and are matched by departures of similar size, and the projection already nets part of that movement out of its replacement figure, so the table cannot be read as a shortage half the size of its headline. The table establishes instead that movement from adjacent occupations supplies these occupations on a scale comparable to the openings themselves, and that a supply count consisting of graduates counts the smallest inflow.

The correction that follows is arithmetic. A raw gap is reduced, tier by tier of skill adjacency, by the workers the adjacent occupations can release into it. Each tier's contribution is the observed transition rate into the occupation times the slack in the adjacent occupation, capped at the largest inflow from that tier the account has recorded. Every term is a measured quantity: adjacency from the skill similarity of occupation pairs, slack from the same projections that produce the raw gap, and transition rates and caps from the flow account. Licensing is a reading of the results rather than a parameter. A pair of occupations that similarity places close and that the account shows nobody crossing is a blocked pair whatever the statute says, and the distribution of those pairs across licensed and unlicensed occupations is the evidence on which a claim about licensing has to be tested. The bilateral matrix that supplies the transition rates is built by the companion account and not yet published, and the correction is set out here with its inputs, its cap, and its readings fixed before it runs.

1 Shortage figures

Every workforce report contains a sentence of the form "the state will be short several thousand nurses by 2033", and the sentence is produced by one operation whatever the occupation. The positions a projection expects the occupation to fill over a horizon, growth and replacement together, are reduced by the people the training system is expected to deliver into it. The Bureau of Labor Statistics and the state workforce agencies publish the first term for every state and occupation, the postsecondary completions survey

supplies the second once a crosswalk assigns each programme’s graduates to the occupations it trains for, and the difference is the number presented to a legislature.

The subtraction is a statement about supply, and supply is understood in it as new entrants. That reading is built into the instruments: a training-capacity decision funds places, a licensure board counts examination passes, a state projection model tracks graduates, and each of these acts on the people entering the occupation for the first time. An occupation, though, is a stock with five flows. People arrive from education, from other kinds of non-employment, and from other occupations, and they leave for other occupations and out of work. The flow account measures all five for the United States, for 230 occupations annually from 2009 to 2025 with a standard error on every figure. Arrivals from another occupation run between four and twelve times arrivals from education, they exceed arrivals from education in every occupation the account covers, and they exceed all arrivals from outside employment together in 170 of 172. Those arrivals occur at every age, and across the whole matrix the lateral arrivals aged thirty-five and over, on their own, outnumber every arrival from education by nearly seven to one.

Economists have understood since Arrow and Capron (1959) that a shortage is a statement about adjustment, about how far the quantity demanded at the going wage exceeds the quantity supplied and how quickly the two close, and that the closing happens through prices, through substitution among inputs, and through the movement of people between occupations. The practical arithmetic dropped the substitution term because nothing measured it. The flow account measures it, and this paper puts it back.

2 The first term

Table 1 takes five healthcare occupations in North Carolina, where the state projection and the employment count both exist at the detailed level, and sets three quantities beside one another. The first is the projection’s own annual openings, growth and replacement together, from the state projections published through Projections Central. The second is the occupation’s employment in the May 2024 Occupational Employment and Wage Statistics, the stock those openings refer to. The third is the rate at which the national flow account observes people arriving in the occupation from another occupation, the mean of the annual rates for 2022 to 2025 with a standard error from the 160 replicate weights the Census Bureau publishes, applied to the state stock. The account is national and the stock belongs to one state. The account’s own test of whether national flow patterns apply to individual states finds no state-specific occupational effect that survives, within a bound of two to three percentage points, and the application here sits inside that bound.

Table 1. Lateral arrivals against projected openings, five healthcare occupations, North Carolina. Rates per hundred prior-year holders are means of 2022 to 2025 from the national flow account, with standard errors in parentheses. Employment is from the OEWS for May 2024, openings are the state projection’s annual growth plus replacement, and replacement per 100 is the projection’s replacement openings over employment. The nursing assistant rate is that of the broader Census aide category, which also covers orderlies, psychiatric aides, and home health aides.

Occupation	Employment	Openings	Replacement per 100	Departures per 100	Lateral arrivals per 100	Lateral over education	Implied arrivals	Share of openings, %
Registered nurses	111,120	7,570	5.3	9.4 (0.5)	3.35 (0.28)	5.1	3,723	49.2 (4.1)
Nurse practitioners	9,800	860	4.0	11.8 (1.7)	13.85 (2.32)	120.4	1,357	157.8 (26.4)
Licensed practical nurses	18,010	1,580	7.3	20.4 (1.6)	10.86 (1.20)	10.0	1,955	123.7 (13.7)
Nursing assistants	64,010	9,110	12.8	22.7 (0.9)	13.09 (0.77)	14.4	8,377	91.9 (5.4)
Medical assistants	23,650	3,340	11.8	23.8 (1.6)	13.97 (1.44)	13.2	3,303	98.9 (10.2)

Read across the registered nurse row. North Carolina employs 111,120 registered nurses, the projection expects 7,570 openings a year, of which 5,940 are replacement, and the account observes 3.35 arrivals per hundred prior-year holders from another occupation. At that rate 3,723 people a year enter the state's registered nurse workforce from a different occupation, 49.2 per cent of the projected openings, against 0.65 per hundred, some 720 people, arriving from education. For licensed practical nurses the lateral rate is 10.86 per hundred and the implied arrivals are 123.7 per cent of projected openings, for nursing assistants 91.9, and for medical assistants 98.9. For nurse practitioners, an occupation entered almost entirely by registered nurses who complete a graduate credential while working, the figure is 157.8, on a rate whose standard error is 2.32 points. Across the five occupations together the projection expects 22,460 openings a year, and lateral arrivals at observed rates would supply 18,715 of them.

The table cannot be read as a shortage cut in half. The arrivals are gross. Registered nurses lose 4.43 per hundred a year to other occupations against the 3.35 who arrive, and licensed practical nurses lose 13.19 against 10.86. At the detailed level only about a tenth of gross lateral movement changes the size of any occupation, and the remainder is two-way traffic between neighbouring occupations. The projection's replacement figure also includes a modelled separation rate that nets some of that traffic out, at 5.3 per hundred for registered nurses against the 9.4 the account observes leaving for any reason, and at 7.3 for licensed practical nurses against 20.4. Where implied arrivals exceed projected openings, this is the reason: the separation rate underneath the projection is well below the rate the account measures at the detailed level, so both the openings and the arrivals that fill them are larger than the headline figure allows for.

The table does establish the order of magnitude of the inflow the arithmetic omits. For every occupation in it, movement from adjacent occupations is of the same size as the openings themselves, and a supply count consisting of graduates counts an inflow between five and 120 times smaller than the one it leaves out. A shortage number built on that count sets one training system's output against the whole of an occupation's demand, and the difference between the two consists mostly of lateral arrivals.

3 The correction

The correction keeps the raw gap and subtracts from it what the adjacent occupations can release. Write G for the raw gap over the horizon, the projection's cumulative openings less the training system's cumulative completions, and index the adjacent occupations j by the tier t of their skill similarity to the target i. Then

$$G^{*} = G - \sum_{t} \min\left(\sum_{j \in t} m_{ji}\, S_{j},\ \bar{m}_{ti}\, E_{i} \right),$$

where m_{ji} is the observed annual rate of transition from j into i per worker in j, S_j is the slack in occupation j, its projected supply in excess of its own projected openings and zero where there is none, E_i is the target occupation's employment, and $\bar{m}_{ti}$ is the largest annual inflow rate from tier t into i that the flow account has recorded. A tier is therefore never asked to supply more than it has been observed to supply. Each tier's term is what the adjacent occupations would release at the rates people already move, limited to the workers those occupations can spare and capped at the rates observed in the past.

The formula contains screens for adjacency, availability, willingness, and permissibility. Adjacency decides which j enter which tier, and it comes from the similarity of occupation pairs on the skills, knowledge, and work activities that the Occupational Information Network records. The composite weights skills by one half and knowledge and activities by a quarter each, and it is held as a table of 798,342 pairs above a similarity of 0.5. The tier thresholds in use in workforce practice place a pair above 0.985 within three months of retraining, between 0.970 and 0.985 within six, and between 0.950 and 0.970 within twelve. The companion paper on skills and moves (Conway, 2026a) tests those thresholds against where observed flows break, and until it reports, the correction runs on the thresholds as used, with their sensitivity as the first robustness table. Availability is S_j, computed from the same projections and completions as the raw gap, so that an occupation counted as short cannot also be counted as a source. Willingness is the direction of the wage differential: a pair enters the sum only where the target pays at least as much as the source, since movement against a wage gradient is small in the account. The gravity estimates on interstate movement put the wage elasticity of flows at 0.114 (Conway, 2026b), so pay moves people weakly towards higher pay and cannot be relied on to move them the other way.

Permissibility is where this paper departs from the version of the correction specified before it. A licensing screen requires a table of which occupations a licence bars entry to, by state. The attempt to build one for the migration paper produced a credentialling rate from the Current Population Survey that is an upper bound on legal licensing, and a compact timeline for one occupation. The screen is not needed as an input, because the transition rate already reflects it. A pair that similarity places in the first tier and that the account shows nobody crossing, at a rate indistinguishable from zero on a stock of a hundred thousand, is a blocked pair whatever the reason, and the formula gives it nothing through m_{ji}. Licensing then becomes a reading of the results. Take the pairs where similarity predicts a flow and the account records none, and ask whether they concentrate in occupations that require a licence. If they do, the claim that licensing is what makes a shortage real, that the overstatement is smallest where public attention and public money are highest, is tested on the whole occupation set, and if they do not, the claim is retired.

The correction produces a range, since the tier thresholds, the transition rates, and the cap each have a standard error or a defensible alternative, and the range is the result. A workforce board that receives it receives the raw gap, the tier by tier subtraction with each term's source named, and the interval, and can rerun the arithmetic with its own thresholds on a spreadsheet.

4 Prespecification

The inputs are four, and each has a source. The bilateral transition matrix, by year, age band, origin, and destination for the occupations the survey supports, is produced by the flow account's own build from the March supplement, on the within-interview design that makes occupational transitions measurable at all. Two independent codings of an unchanged job disagree 48.1 per cent of the time at the detailed level, so a matrix linked across interviews measures coding disagreement, and the supplement's design, with both codes assigned at one sitting from one description, cancels it. The similarity table is the composite described above, built from the Network's importance ratings. Employment and wages are the Occupational Employment and Wage Statistics by state. Openings and completions are the state

projections and the postsecondary completions survey, allocated to occupations by the crosswalk and entry-band rules already used to build the raw gap.

The readings are fixed as well. The correction runs for every occupation with a projection and a positive raw gap in every state, and the object of interest is the distribution of $1 - G^{*} / G$, the share of the headline that the adjacent occupations remove, across occupations and states, with its interquartile range and its tails. The licensing test compares that share, and the count of blocked pairs, between occupations with and without a licence requirement in the Bureau's own entry classification. A distribution of overstatement concentrated below five per cent with no tail would make the correction not worth finishing, and that reading is recorded here before the matrix is published. A median above a fifth would change practice. It is the order the worked case makes plausible, and the arithmetic of Section 2 cannot settle it on its own.

5 Conclusion

A shortage figure is used as an instruction to a training system, and it is wrong in size whenever an occupation takes in people by a route the figure omits. For 170 of the 172 occupations the flow account covers, the omitted route, movement from other occupations, is the largest inflow, and for five healthcare occupations in one state it is of the same order as the openings themselves. The correction that restores it consists of four measured quantities and a subtraction, it produces a range rather than a point, and the people who have to act on it can audit it. Its adjacency terms wait on one table, and its readings are fixed.

References

Arrow, K. J. and Capron, W. M. (1959). Dynamic shortages and price rises: the engineer-scientist case. *Quarterly Journal of Economics*, 73(2), 292–308.

Bureau of Labor Statistics (2025). Occupational Employment and Wage Statistics, May 2024. State occupational projections through Projections Central, projectionscentral.org.

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

Conway, A. (2026a). Do skills predict moves? A gravity test of skill similarity against observed occupational transitions. 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.