nature cities

Article

https://doi.org/10.1038/s44284-023-00009-1

Network constraints on worker mobility

Received: 10 May 2023

Accepted: 17 November 2023

Published online: 11 January 2024

Check for updates

Morgan R. Frank [1,2,3], Esteban Moro[3,4,5], Tobin South[3], Alex Rutherford[6], Alex Pentland[3], Bledi Taska[7] & Iyad Rahwan[6]

How do skills shape career mobility and access to cities' labor markets? Here we model career pathways as an occupation network constructed from the similarity of occupations' skill requirements within each US city. Using a nationally representative survey and three resume datasets, skill similarity predicts transition rates between occupations and predictions improve with increasingly granular skill data. Thus, a measure for skill specialization based on a workers' position in their city's occupation network may predict future career dynamics. Job changes that decrease workers' network embeddedness also increased wages, and workers tend to decrease their embeddedness over their careers. Further, city pairs with dissimilar job embeddedness have greater census migration and increased flows of enplaned passengers according to the US Bureau of Transportation Statistics. This study directly connects workplace skills to workers' career mobility and spatial mobility, thus offering insights into skill specialization, career mobility and urbanization.

As technology[1], pandemics[2,3] and climate change[4] shape the future of work, how can workers maximize employment opportunities in their communities? In addition to workforce development, policy makers must also identify and mitigate the structures within the labor market that hinder workers' mobility and relevance within their local economies[5,6]. To this end, recent studies have used skills to explain US job polarization as a divide between high-skill and low-skill workers[7,8]. However, these broad labor categories—or even refinements to cognitive and physical or routine and nonroutine labor[9]—are too broad to capture skill specialization[10,11] and can obfuscate job seeker dynamics[12–15]. For example, civil engineers and medical doctors are both highly educated, well-paid, cognitive, nonroutine occupations, but the skills required by each occupation are largely nontransferable, thus limiting individuals' job transitions from one occupation to the other. The differences between occupations are most easily understood when we consider labor markets as a multilayer, complex system made of occupations modeled as abstract bundles of skills and abilities[12,16–18].

Although urban labor markets are typically modeled according to their employment by occupation, cities may too be characterized according to their workers' workplace skills[19,20]. For example, the economic agglomeration of similar firms within the same city has long been linked to higher earnings and productivity[21–24]. Cities with workers who leverage many different skills tend to offer greater wealth and more diverse employment opportunities[25,26]. However, studying only the distribution of skills or occupations obscures economic insights. For example, skill specialization typically refers to workers or occupations leveraging a small set of rare skills, but a skill's rarity varies across cities. Thus, relative skill specialization may also be informative. Relative skill specialization could be observed by modeling occupations' skill similarity to other occupations in the same local labor market. Existing studies observe that cities with a densely connected workforce (that is, supporting many occupations requiring similar skills) tend to be more productive, innovative and resilient to economic shocks[13,27,28], which challenges the presumably beneficial role of skill specialization and the division of labor in cities[29].

Likewise, city pairs with many shared employment opportunities may have greater potential for worker mobility[30–32]. For example, workers may maximize wages by relocating to labor markets where their skill specialization is in demand but locally in short supply.

[1]Department of Informatics and Networked Systems, University of Pittsburgh, Pittsburgh, PA, USA. [2]Digital Economy Lab, Institute for Human-Centered Artificial Intelligence, Stanford University, Stanford, CA, USA. [3]Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]Department of Mathematics and GISC, Universidad Carlos III de Madrid, Leganes, Spain. [5]Network Science Institute, Northeastern University, Boston, MA, USA. [6]Center for Humans and Machines, Max Planck Institute for Human Development, Berlin, Germany. [7]Burning Glass Technologies, Boston, MA, USA. e-mail: mrfrank@pitt.edu

However, while employment opportunities and skill demands shape worker relocation[33,34], they remain absent from traditional migration models[35,36] except through measurements of an economy's size. Since skills are critical to the job-matching process[37] and workers' pursuit of earnings[14,15], workers may make career moves based on their relative position in various urban labor markets, thus shaping spatial mobility[30,31,38]. If so, then policymakers could use employment and skills data to identify synergistic or competitive labor markets as they seek to maintain or grow local employment.

In this Article, we consider the following research questions:

- (RQ1) Do broad skill categories explain modeling career mobility as well as detailed skill categories? We hypothesize that occupation similarity based on detailed skills data improves models of job transitions.
- (RQ2) Does a worker's relative skill specialization and position in their city's skill landscape impact their careers and wages? We hypothesize that local labor markets can be modeled as occupation networks and that workers' position in these networks relates to their earnings and career mobility.
- (RQ3) Do cities' occupation networks inform inter-city mobility as workers increase their skill specialization over their careers? We hypothesize that workers will relocate to new cities if corresponding career moves tend to increase their relative skill specialization.

We compare occupation pairs based on skill requirements reported by the US Bureau of Labor Statistics (BLS). Occupations with shared skill requirements exhibit greater flows of workers between them. Thus, we model each US city as an occupation network based on skill similarity. With three large-scale resume datasets, we investigate how individual workers move within and between these networks. Finally, we study US Census migration and airline traffic flows to compare workers' embeddedness across city pairs to inter-city spatial mobility.

Results

Skill similarity predicts job transition rates

We expect that workers will have greater transition rates between pairs of occupations that share skill requirements[16,39,40]. However, abstract descriptions can obfuscate key differences between occupations. As an illustrative example, we compare the skill requirements of surgeons and air traffic controllers in Fig. 1a. The word clouds highlight the most important skills to each occupation using font size. Text color indicates if the skill is social-cognitive or sensory-physical according to ref. 41. Although the specific skills are different, the occupations have the same cognitive skill fraction[41] that is calculated according to $\left(\sum_{s\in \text{cognitive}} \text{onet}_{j,s}\right) / \left(\sum_{s\in S} \text{onet}_{j,s}\right)$. Here, $\text{onet}_{j,s} \in [0,1]$ denotes the real-valued importance of skill $s \in S$ to occupation $j \in J$ according to the Occupational Information Network (O*NET) database from the US BLS (see 'O*NET skills data' section in Methods).

To resolve these differences, we quantify the Jaccard similarity of the skills required by occupations j and j' according to

$$\text{skillsim}(j, j') = \frac{\sum_{s\in S} \min(\text{onet}_{j,s}, \text{onet}_{j',s})}{\sum_{s\in S} \max(\text{onet}_{j,s}, \text{onet}_{j',s})}. \quad (1)$$

Note that similarity scores are real-valued on the interval $\text{skillsim}(j, j') \in [0,1]$. Although occupations that require many different skills are penalized by the denominator in equation (1), we performed an additional analysis in Supplementary Materials Table S1 (model 4), where we control for ubiquitous skills using revealed comparative advantage[41]. We highlight some of the mutually exclusive and shared skills for surgeons and air traffic controllers in Fig. 1b. In our example, although air traffic controllers and surgeons are comparably cognitive, their $\text{skillsim}(j,j')$ is relatively low at $\text{skillsim}(\text{air traffic controller}, \text{surgeon}) = 0.54$ in 2018 (see Supplementary Fig. 2 for the full similarity distribution).

Does $\text{skillsim}(j,j')$ correspond to job transitions? We investigate with four separate datasets. First, we consider job transition rates between occupation pairs according to the nationally representative US Community Population Survey (CPS) from 2012 through 2018 (see 'Tracking career mobility through the community population survey' section in Methods for more details). Additionally, we consider datasets provided by Burning Glass Technologies (BGT), FutureFit AI (FF) and Revelio Labs (Revelio), each containing millions of resumes from individual workers in the US (see Supplementary Sections 6–8 and 'Resume data' section in Methods). We present our analysis of CPS data in Table 1 using ordinary linear regression to model job transition rates between occupations pairs (see Supplementary Tables 19, 22 and 24 for comparable analyses of the resume data). First, we consider a random mixing model based on occupations' national employment according to $\log_{10}(T(j, j')) = \log_{10}(\text{employment}_j \cdot \text{employment}_{j'})$ (model 1) using employment data from the US BLS's Occupation Employment and Wage Statistics (OEWS). Next, we combine national employment with the occupations' Jaccard similarity of skills, aggregated into cognitive, physical, routine and nonroutine skill categories[9] (denoted $\text{CP}(j, j')$) in model 2 (see 'Aggregations of workplace skills' section in Methods). This approach improves predictions of job transition rates, but the best performance is achieved by including $\text{skillsim}(j,j')$ (models 3 and 4) applied to more granular O*NET skills. Interestingly, $\text{CP}(j, j')$ and $\text{skillsim}(j,j')$ each interact positively with employment $\log_{10}(T(j, j'))$, suggesting that $\text{skillsim}(j,j')$ may be more important to career mobility in more common occupations.

Does $\text{skillsim}(j,j')$ perform well because of only a few key O*NET skills? Figure 1c demonstrates the performance of model 3 as we vary the number of O*NET skills included in the calculation of $\text{skillsim}(j,j')$. For each number of included skills, we perform 50 independent trials where O*NET skills are randomly selected to recalculate $\text{skillsim}(j,j')$; specifically for a random subset of skills $S^* \subseteq S$, we calculate

$$\text{skillsim}^*(j,j') = \frac{\sum_{s\in S^*} \min(\text{onet}_{j,s}, \text{onet}_{j',s})}{\sum_{s\in S^*} \max(\text{onet}_{j,s}, \text{onet}_{j',s})} \quad (2)$$

and use skillsim^* in place of skillsim in model 3. Note that this recalculation of skill similarity does not change the regression's degrees of freedom for model 3 and so the predictive performance gains in Fig. 1c are not the result of overfitting the data (that is, including fixed effects, there are nine variables in regression model 3 used to predict a data set of over 12,000 observations). Predictive performance sharply increases with the inclusion of the first 40 O*NET skills and steadily increases with each additional skill thereafter. Although pairs of O*NET skills may contain redundant information about worker transition rates, these results highlight the predictive value of granular workplace skills and suggest that even better model performance may be obtained with increasingly refined skill data.

Urban occupation networks and workers' embeddedness

Skill similarity plays a crucial role in career mobility[37] (Fig. 1). Therefore, we hypothesize that workers' career advancement strongly depends on their skill set[42], the occupations that are supported in their local labor market and the structure determined by the $\text{skillsim}(j,j')$ among local occupations. We model this by constructing a national occupation network with edge weights determined by occupation pairs' $\text{skillsim}(j,j')$ (see Supplementary Section 1 for more details). We test our hypothesis with millions of individuals' job transitions in three large-scale resume datasets from BGT, FF and Revelio. With each career move, a worker may relocate between cities and

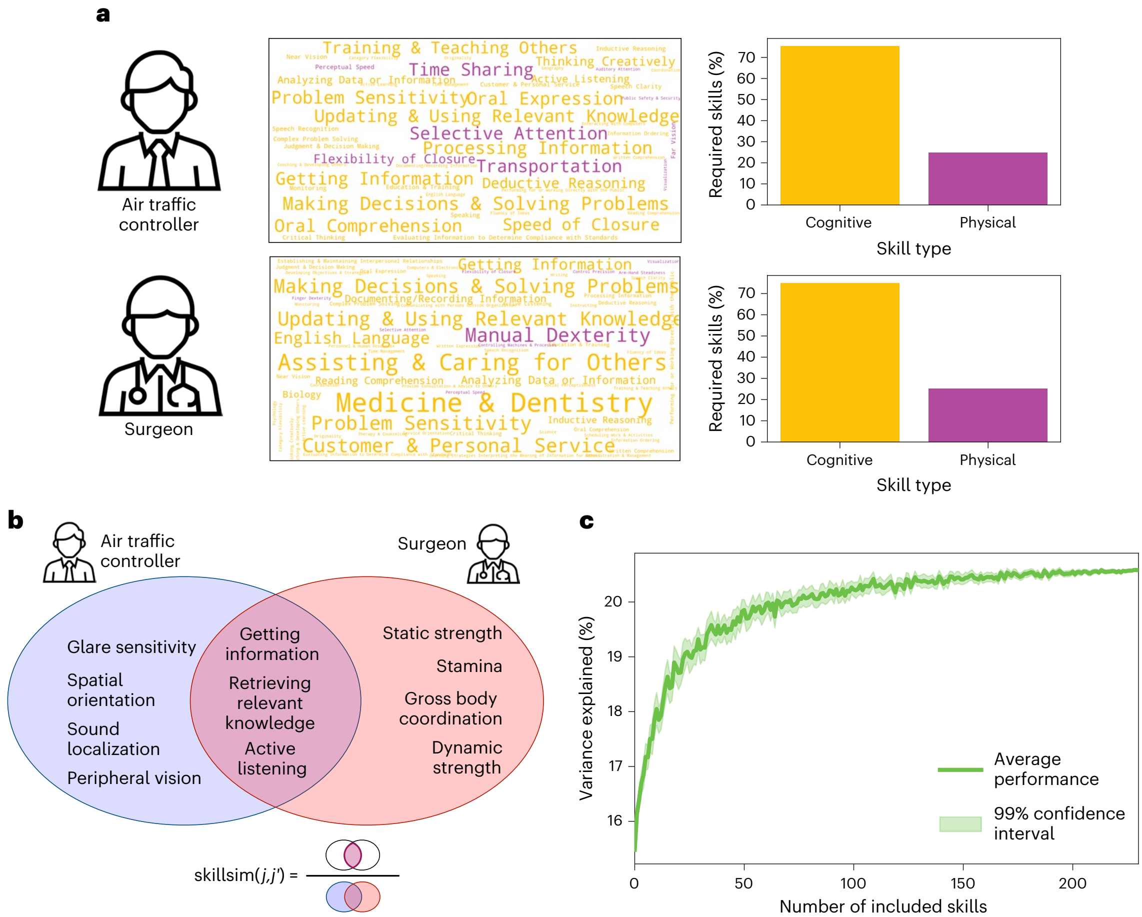

Fig. 1 | Occupations' skill similarity predicts job transition rates. a, Comparing the 2018 O*NET skill requirements of air traffic controllers (SOC: 53-2021) and surgeons (SOC: 29-1067) reveals a dissimilar set of specific skills despite identical levels of cognitive skill requirements on aggregate (right). Word clouds (center) highlight the relative importance of each skill using font size, while colors correspond to skill classification as cognitive or physical according to ref. 41. **b**, A schematic for the Jaccard similarity of the O*NET skill requirements for a pair of occupations. Examples of shared and mutually exclusive O*NET skills are provided for our example occupations. **c**, Recalculating skill similarity using random subsets of O*NET skills and retraining model 3 from Table 1 yields steady improvements in the model's predictions of worker flows between job titles according to the variance explained across 50 independent trials for each number of included skills. Icons are from ref. 60.

may—or may not—change occupations (see Fig. 2a for an example resume and corresponding career moves). Both BGT and Revelio map raw job title strings onto the Standard Occupational Classification (SOC) occupations, and we employed an 'off-the-shelf' pre-trained language model to map FF job titles to SOC occupations using natural language processing (see 'Mapping occupations from FF resumes' section in Methods and Supplementary Section 7). Supplementary Materials Sections 6–8 provide basic statistics on the data coverage of US cities and the national labor market. Our analysis includes 3.2 million job transitions in the FF data, 1.2 million job transitions in the BGT data and 8 million job transitions in the Revelio data between 2012 and 2018.

To describe a worker's position in their local labor market on the basis of their skills, we represent each US city's local labor market as the subnetwork determined by the occupations with nonzero OEWS employment in that city (that is, metropolitan statistical area) and the connections between those locally supported occupations (see examples in Fig. 2b). We hypothesize that a worker of an occupation with high skillsim(j,j') to many other locally supported occupations will have more career mobility without re-skilling. However, saturated demand for similar occupations may dampen the mobility implied by skillsim(j,j'). Therefore, we introduce occupation embeddedness based on both the shared skill requirements among occupations and employment share among similar occupations; specifically, we measure occupation i's employment-weighted embeddedness in city c according to

$$e_{c,j} = \sum_{j' \in \text{Jobs}} \text{skillsim}(j,j') \cdot \text{share}(c,j'), \tag{3}$$

where share(c,j') is the employment share of occupation j' in c (see Fig. 2b for a schematic). Note that $e_{c,j}$ is calculated for each occupation in each city in each year and we use $e_{c,j} = 0$ if share$_{c,j} = 0$, according to OEWS data. $e_{c,j}$ captures the similarity of an occupation's skill requirements to the rest of their local labor market, and so we use this measure to approximate j's relative skill specialization in c. Standard occupation employment metrics, including city fixed effects, year fixed effects and local employment share, as well as an occupation's page rank in their city's occupation network do not explain $e_{c,i}$ (Supplementary Tables 5 and 6). Beforehand, it is not clear if $e_{c,i}$ measures the complementarity (for example, positive economic agglomeration[43]) or the competition (for example, workers with similar skills competing for finite employment opportunities) around an occupation in a city.

Do workers change their relative skill specialization over their careers? If so, then workers should, in general, decrease $e_{c,j}$ with each career transition. We investigate by measuring

$$\Delta \log_{10} e = \log_{10}\left(e_{c_2,j_2}\right) - \log_{10}\left(e_{c_1,j_1}\right), \tag{4}$$

Table 1 | Ordinary least squares regression analysis of job transition rates between pairs of occupations according to the US CPS from 2012 through 2018

Dependent variable: $\log_{10}$ job transitions				
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
$\log_{10}$ total employment ($T(j,j')$)	0.392 (0.008)	0.403 (0.008)	0.416 (0.008)	0.424 (0.008)
	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$
	(0.376, 0.408)	(0.387, 0.419)	(0.401, 0.432)	(0.409, 0.439)
Cognitive/physical similarity ($\mathrm{CP}(j,j')$)		0.172 (0.008)		−0.005 (0.012)
		$P\leq1\times10^{-3}$		$P=0.669$
		(0.156, 0.188)		(−0.028, 0.018)
$\mathrm{skillsim}(j,j')$			0.271 (0.008)	0.324 (0.012)
			$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$
			(0.255, 0.288)	(0.300, 0.348)
$\mathrm{skillsim}(j,j')\times\mathrm{CP}(j,j')$				0.104 (0.008)
				$P\leq1\times10^{-3}$
				(0.087, 0.120)
$\mathrm{skillsim}(j,j')\times T(j,j')$				0.064 (0.013)
				$P\leq1\times10^{-3}$
				(0.039, 0.089)
$\mathrm{CP}(j,j')\times T(j,j')$				0.040 (0.013)
				$P\leq1\times10^{-2}$
				(0.015, 0.065)
Year fixed effects	Yes	Yes	Yes	Yes
R^2	0.155	0.183	0.219	0.237
Adjusted R^2	0.154	0.182	0.219	0.237

Regression data includes 12,648 observations that vary by year and occupation pair. Model 1 represents a random mixing model based on the national total employment of each occupation ($\log_{10}(T(j,j'))$). Subsequent models augment the employment model based on the occupation pair's similarity of skills using canonical skill aggregations (model 2) and the skill similarity using the complete O*NET taxonomy (models 3 and 4). All variables were centered and standardized before analysis. Regression coefficients are reported, followed by coefficient estimate standard error in parentheses, estimate P value and a 95% confidence interval. Similar analyses using resume data are presented in Supplementary Tables 8 and 11.

where $(c_1,j_1)\to(c_2,j_2)$ is a career transition observed in an individual worker's resume such that $c_1\neq c_2$ and/or $j_1\neq j_2$. In general, $\Delta\log_{10}e$ was slightly positive in the BGT resumes ($\langle\Delta\log_{10}e\rangle=0.0011$) and slightly negative in the FF resumes ($\langle\Delta\log_{10}e\rangle=-0.0009$), but workers tended to significantly decrease $e_{c,i}$ over their careers according to Spearman rank correlation (Fig. 2c). Both averages were significantly different from zero as a one-tailed t-test yielded $P\approx0$. The average $\Delta\log_{10}e$ differs from the decreasing trend over a career because the distribution of the number of career moves per resume is skewed. Note that a career transition may be to the same occupation in the same city; these transitions account for 14% of the transitions in the FF resume data and 22% in the BGT resume data.

Since workers decrease $e_{c,i}$ over their careers, how does $\Delta\log_{10}e$ relate to wages? We investigate using the estimated average annual wage associated with each occupation in each city in each year according to OEWS data from the US BLS. Workers rapidly decrease their $e_{c,i}$ in the early parts of their career (Fig. 2c), and career transitions that decrease employment-weighted embeddedness (that is, $\Delta\log_{10}e<0$) correspond to increased wages (that is, $\Delta\log_{10}\mathrm{wage}>0$; Fig. 2d). Thus, we hypothesize that the wage increase associated with decreased $e_{c,i}$ may be explained by trends in workers' spatial mobility (for example, urbanization of high-skilled workers[44]). We investigate using BGT and FutureFit resume datasets in Table 2 while controlling for year fixed effects as well as fixed effects for workers' source and destination city. Additionally, we include a dummy variable (J) indicating if a job transition was between two different occupations and a dummy variable (r) indicating if a job transition was a relocation between two different cities. These fixed effects control for many potential confounds including city size of the source and destination city, location-based economic factors (for example, access to a port or natural resources), and each cities' cost of living. As a baseline, we first ignore cities' occupation networks and analyze the relationship between changes in occupations' educational requirements and wage changes while controlling for occupation changes (J) or a relocations (r). Career moves to occupations with a larger share of local workers with a bachelor's degree according to OEWS data corresponded to wage increases in both resume datasets (models 1 and 4).

Next, we introduce information about a worker's position in their local labor market by considering their occupations' centrality page rank (denoted $P_{c,i}$) in their city's occupation network. An occupation with large $P_{c,i}$ has high $\mathrm{skillsim}(j,j')$ with many other occupations that themselves have high $\mathrm{skillsim}(j,j')$ to other occupations. Including the change in page rank associated with each career move (denoted $\Delta\log_{10}P$) yields mixed results across resume data sets and does not substantially improve the predictions of wages (models 2 and 5). Occupations' page rank uses the topology of cities' occupation networks but only minimally uses cities' employment distributions. That is, occupations are included in a city's network if they have nonzero employment according to OEWS data, but information is omitted about each occupation's local employment share. As we noted above, workers may more easily transition between occupations with large $\mathrm{skillsim}(j,j')$, but only if demand for these similar occupations is not already saturated. $e_{c,i}$ accounts for both the network topology in each city, as well as the employment share among similar occupations.

https://doi.org/10.1038/s44284-023-00009-1

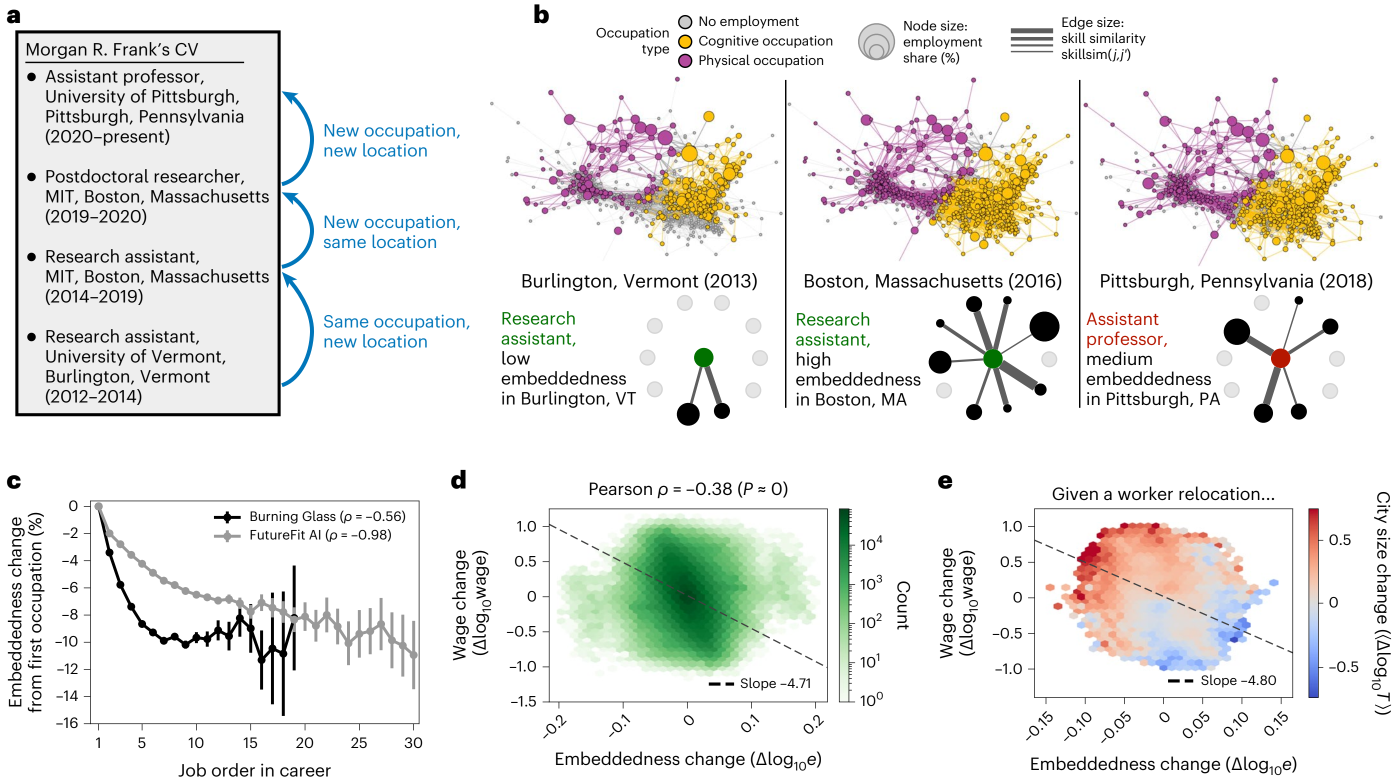

Fig. 2 | Workers decrease their employment-weighted embeddedness, $e_{c,i}$, throughout their careers, and career moves that decrease $e_{c,i}$ correspond to higher wages according to resume data. a, An example career trajectory from the resume of one of the authors. A career move is a transition to the same or a different occupation and may correspond to a relocation between cities. **b**, A city's labor market represented as a subnetwork of the complete occupation network according to the local occupations with nonzero employment and the pairwise skill similarity among those occupations. Red and green text highlight a change in occupation. A worker changes between networks when they relocate between cities, and $e_{c,i}$ may change as a result. **c**, Workers decrease $e_{c,i}$ throughout their careers according to Spearman correlations ρ ($P < 0.02$ for both datasets). Data are presented as mean values with vertical lines representing 95% confidence intervals. The FF analysis includes 6.8 million resumes, while the BGT analysis includes 879,276 resumes. **d**, Using wage estimates from the US BLS data by occupation, city and year, individual workers increased wages when a career move decreased $e_{c,i}$. **e**, Given a relocation associated with a career transitions, changes in $e_{c,i}$ were a stronger predictor of wage gains than changes in city size. Color represents the average wage change in each bin for bins containing at least 20 observations. **d** and **e** use resume data from FF. See Supplementary Fig. 10 for a similar analysis using resumes from BGT.

Even after controlling for occupation changes, relocations, occupations' educational requirements, occupations' network centrality and several fixed effects, we find again that career moves that decreased embeddedness (that is, $\Delta\log_{10}e < 0$) corresponded to increased wages (models 3 and 6). Our regressions in Table 2 include fixed effects for workers' starting city, ending city and year of job transition, as well as dummy variables indicating when a job transition was a relocation (r) or an occupation change (J). As a robustness check, we provide similar analyses restricted to transitions that were an occupation change (Supplementary Table 7), a relocation (Supplementary Table 8), an occupation change without a relocation (Supplementary Table 9) or a relocation without an occupation change (Supplementary Table 10). In general, we find that job transitions that lowered $e_{c,i}$ were associated with increased wages (Supplementary Section 4).

Unlike BGT and FF resume data, Revelio resume data estimates the annual salary of each job in each workers' resume (Supplementary Section 8). Using these salaries in a cross-sectional analysis, we find again that lower $e_{c,j}$ is associated with higher earnings (Supplementary Table 26). These findings are robust to the inclusion of page rank, the national share of workers with a bachelor's degree in an occupation (B) as well as city, year and major SOC code (that is, two-digit SOC code) fixed effects. These variables control for regional and temporal variation in salaries and even for differences in average earnings across major occupations. Combined with Table 2, these results demonstrate both cross-sectional and longitudinal evidence that lower $e_{c,i}$ corresponds to higher wages.

Embeddedness and workers' spatial mobility

Although decreasing $e_{c,i}$ corresponds to increased wages, our observational analysis does not identify the underlying causal mechanisms. However, we can consider potential mechanisms and look for their symptoms. On one hand, low $e_{c,i}$ identifies occupations with relative skill specialization, which suggests workers with those skills may enjoy increased negotiating power within their local labor market. On the other hand, employers may offer higher wages to lure workers to their city to provide skills that are rare among local workers. In the latter case, offers of higher wages may incentivize a worker relocation. Although a random control trial (for example, randomizing $e_{c,i}$ or a city's occupation network) is not possible, we can test this hypothesis using data on spatial mobility between cities. This section provides evidence to support embeddedness as a factor in inter-city mobility.

Traditional models for spatial mobility (for example, the gravity model) ignore labor characteristics[35,36]—except for labor market size. Alternatively, we hypothesize that workers with large $e_{c,i}$ in their current city, c, may relocate to another city, c', where their occupation has a lower $e_{c,i}$ (that is, $e_{c',i} < e_{c,i}$). If so, then larger average combined embeddedness according to

$$\hat{e}_{c,c'} = \sum_{j \in \mathrm{Jobs}_c \cap \mathrm{Jobs}_{c'}} 1/\left(e_{c,j} e_{c',j}\right) \quad (5)$$

should correspond to greater spatial mobility flows between c and c'. Note that lower $e_{c,i}$ increases $\hat{e}_{c,c'}$.

 https://doi.org/10.1038/s44284-023-00009-1

Table 2 | Workers who decrease their employment-weighted embeddedness increase their wages—especially following a change in occupation

	Dependent variable: wage change ($\Delta\log_{10}$ wage)					
	FF resumes			BGT resumes		
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
Relocation dummy (r)	0.000 (0.000)	−0.000 (0.000)	−0.001 (0.000)	0.002 (0.001)	0.002 (0.001)	0.001 (0.001)
	$P=0.940$	$P=0.961$	$P\leq1\times10^{-1}$	$P\leq1\times10^{-1}$	$P\leq1\times10^{-1}$	$P=0.243$
	(−0.001, 0.001)	(−0.001, 0.001)	(−0.002, −0.000)	(0.000, 0.003)	(0.000, 0.003)	(−0.001, 0.002)
Occupation change dummy (J)	0.018 (0.000)	0.018 (0.000)	0.018 (0.000)	−0.026 (0.001)	−0.026 (0.001)	−0.022 (0.000)
	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$
	(0.018, 0.019)	(0.017, 0.019)	(0.018, 0.019)	(−0.027, −0.025)	(−0.027, −0.024)	(−0.023, −0.022)
Bachelor's degree change (ΔB)	0.629 (0.000)	0.622 (0.001)	0.563 (0.001)	0.691 (0.001)	0.692 (0.001)	0.560 (0.001)
	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$
	(0.628, 0.630)	(0.621, 0.623)	(0.562, 0.564)	(0.689, 0.693)	(0.691, 0.694)	(0.558, 0.562)
Page rank change ($\Delta\log_{10}P$)		−0.100 (0.002)	0.974 (0.004)		0.425 (0.004)	1.875 (0.007)
		$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$		$P\leq1\times10^{-3}$	$P\leq1\times10^{-3}$
		(−0.104, −0.095)	(0.966, 0.981)		(0.418, 0.432)	(1.861, 1.890)
Embeddedness change ($\Delta\log_{10}e$)			−0.092 (0.003)			−0.224 (0.006)
			$P\leq1\times10^{-3}$			$P\leq1\times10^{-3}$
			(−0.097, −0.087)			(−0.235, −0.212)
$\Delta\log_{10}e\times r$			0.003 (0.001)			0.012 (0.001)
			$P\leq1\times10^{-3}$			$P\leq1\times10^{-3}$
			(0.002, 0.004)			(0.010, 0.014)
$\Delta\log_{10}e\times J$			−0.222 (0.002)			−0.206 (0.006)
			$P\leq1\times10^{-3}$			$P\leq1\times10^{-3}$
			(−0.226, −0.217)			(−0.217, −0.195)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Source city FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination city FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.409	0.410	0.441	0.487	0.496	0.550
Adjusted R^2	0.409	0.409	0.441	0.487	0.495	0.550

Career transitions between employment opportunities may or may not capture a change in occupation (that is, J) and may or may not accompany a relocation to a new city (that is, r) in the period of 2012 through 2018. Models 1–3 use 3.2 million career transitions in the FF resume data. Models 4–6 use 1.2 million career transitions in the BGT resume data. All variables were centered and standardized before analysis. Ordinary least squares regression coefficients are reported, followed by coefficient estimate standard error in parentheses, estimate P value and a 95% confidence interval. Standard errors are clustered by the major occupation (that is, two-digit SOC code) of the starting occupation. See Supplementary Sections 6 and 7 for additional information about resume data. FE, fixed effect.

We test our hypothesis using inter-city migration according to the US Census Bureau and the number of enplaned passengers flying between city pairs according to the US Bureau of Transportation Statistics ('Migration and enplaned passengers' section in Methods). The combined embeddedness of a city pair measures the similarity of their labor markets beyond simple alternatives, such as similarity of size or similarity of employment by occupation. While typical mobility models, including the gravity model and radiation model, only account for size and distance, combined embeddedness represents similar economic factors encoded in cities' labor markets. If combined embeddedness predicts inter-city mobility, then policymakers' should include this information to invest strategically in their economy rather than only focusing on growth (that is, increasing size). Further, these two datasets containing comprehensive federal data offer a compelling alternative to the potentially biased resume datasets and, thus, enable another form of potential validation for the concept of embeddedness.

Does $\hat{e}_{c,c'}$ predict spatial mobility? As a baseline, we first consider the gravity model[35,45–47], which combines city size (that is, the combined total employment in c and c' denoted N_c and $N_{c'}$) with spatial distance $D_{c,c'}$. With year fixed effects, the gravity model accounts for roughly one-third of mobility in both migration and flight data (Table 3, models 1 and 4). Augmenting the gravity model with $\hat{e}_{c,c'}$ improves predictions to 43% of migration variance explained and 37% of flight variance explained (models 2 and 5). However, the best predictive performance is observed when we include interaction terms between $\hat{e}_{c,c'}$ and the variables from the gravity model (models 3 and 6). As a robustness check, we find that simpler measures based on the similarity of urban employment distributions over occupations do not perform as well nor do they diminish the significance of $\hat{e}_{c,c'}$ (Supplementary Tables 11 and 12). We provide a similar analysis of direct inter-city mobility and a robustness check using an alternative gravity model with city fixed effects in Supplementary Section 10.

https://doi.org/10.1038/s44284-023-00009-1

Table 3 | For both inter-city census migration and enplaned passengers according to the US Bureau of Transportation, the gravity mobility model (model 1 and 4) is improved when combined employment-weighted embeddedness is included (models 2 and 5) but the best performance is achieved when combined embeddedness moderates city size (models 3 and 5)

Dependent variable:	$\log_{10}$ total migration			$\log_{10}$ total enplaned passengers		
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Total employment ($\log_{10}(T_c T_{c'})$)	0.517 (0.002)	0.339 (0.002)	0.411 (0.002)	0.578 (0.005)	0.417 (0.007)	0.529 (0.006)
	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$
	(0.514, 0.521)	(0.335, 0.343)	(0.407, 0.416)	(0.568, 0.589)	(0.403, 0.430)	(0.516, 0.541)
Distance ($\log_{10} D_{c,c'}$)	−0.385 (0.002)	−0.394 (0.003)	−0.394 (0.002)	−0.069 (0.005)	−0.101 (0.005)	−0.101 (0.005)
	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times0^{-3}$	$P \leq 1\times10^{-3}$	$P \leq 1\times0^{-3}$
	(−0.390, −0.381)	(−0.399, −0.389)	(−0.397, −0.390)	(−0.079, −0.060)	(−0.112, −0.090)	(−0.112, −0.091)
Combined embeddedness ($\hat{e}_{c,c'}$)		0.348 (0.003)	0.223 (0.002)		0.307 (0.008)	0.205 (0.007)
		$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$		$P \leq 1\times10^{-3}$	$P \leq 1\times10^{-3}$
		(0.343, 0.353)	(0.219, 0.228)		(0.292, 0.321)	(0.190, 0.220)
$\hat{e}_{c,c'} \times \log_{10} T_{c,c'}$			0.186 (0.002)			0.235 (0.006)
			$P \leq 1\times10^{-3}$			$P \leq 1\times10^{-3}$
			(0.183, 0.190)			(0.223, 0.246)
$\hat{e}_{c,c'} \times \log_{10} D_{c,c'}$			−0.022 (0.002)			0.000 (0.004)
			$P \leq 1\times10^{-3}$			$P=0.930$
			(−0.026, −0.018)			(−0.008, 0.009)
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.358	0.425	0.449	0.326	0.370	0.417
Adjusted R^2	0.358	0.425	0.449	0.325	0.370	0.416

The migration analysis is based on 172,000 observations, while the enplaned passenger analysis is based on 25,000 observations. Ordinary least squares regression coefficients are reported, followed by coefficient estimate standard error in parentheses, estimate P value and a 95% confidence interval. All variables were centered and standardized before analysis. See Supplementary Section 10 for an analysis of directed mobility with additional controls for labor market tightness and Supplementary Section 5.2 for analysis with a fixed effects gravity model.

Larger $\hat{e}_{c,c'}$ corresponds to greater inter-city mobility and interacts significantly with combined total employment (that is, $\log_{10}(N_c N_{c'})$). Examining this relationship more closely, we find that $\hat{e}_{c,c'}$ moderates the relationship between combined city size and spatial mobility (Fig. 3). City pairs with larger $\hat{e}_{c,c'}$ and larger combined employment exhibited greater spatial mobility. We perform a mediation analysis in Supplementary Section 11, which suggests that embeddedness may mediate 34.5% of the effect of city size on total migration and 27.9% of the effect of city size on total enplaned passengers. We observe the same results when using the radiation mobility model[36] instead of the gravity model. These results are not an artifact of these data as these relationships disappear after data randomization (Supplementary Section 5). Our main results hold under additional analyses of directed mobility and controls for labor market tightness (that is, a city's unemployment divided by its job vacancies, according to the BLS Job Openings and Labor Turnover Survey) in Supplementary Section 10.

Discussion

US job polarization is described as a division between low- and high-skill work[48], but workers must match more-specific skill requirements when pursuing employment opportunities (that is, skill matching[14,15,37,49]). Accordingly, more-specific labor categories, such as cognitive/physical and routine/nonroutine[9], improve insights into labor trends but still obfuscate distinctive occupation skill requirements and the role of skill specialization. More detailed calculations of skill similarity better predict worker transition rates between occupations (Fig. 1), and so, we encode skill similarity into the edge weights of occupation networks describing each US city's labor market (Fig. 2b).

This data-driven approach connects workers' skills and employment history to their position in their local labor market while identifying opportunities for career mobility. In particular, our analysis of resumes suggests that workers who increase their skill specialization relative to their local labor market earn higher wages (Fig. 2d and Table 2). Correspondingly, a cross-sectional analysis of annual salaries for individual jobs within each worker's resume shows that lower employment-weighted embeddedness corresponds to higher earnings (Supplementary Table 26). Thus, modeling city's as occupation networks reveals the specialized occupations relative to the local labor market and the earnings implications for skill specialization.

Highly embedded subsets of a city's occupation network represent parts of the local labor market that may be more saturated since, despite having potentially different job titles, there are many workers leveraging similar skills. This is reflected in the negative relationship between embeddedness and wages (Fig. 2d and Table 2). Thus, future work might investigate how the would-be embeddedness of currently absent occupations relates to the emergence of new occupations in a given city. However, employment-weighted embeddedness is not easily separable into employer and worker dynamics during the job-matching process. Still, these results may highlight a trade-off between the presence of specialized high-paying work (that is, occupations with low embeddedness) and creating an economically resilient labor market where workers possess the skills required by many occupations[13]. The specificity of these insights related to embeddedness would be missed with traditional labor models that fail to consider occupations' skill requirements.

Combining occupations' skill requirements with employment connects microscopic labor requirements to the macroscopic trends shaping urban, regional and national economies in the United States. While often described at the national scale[8,48], recent work[19,44] highlights the spatiality of job polarization as divisions

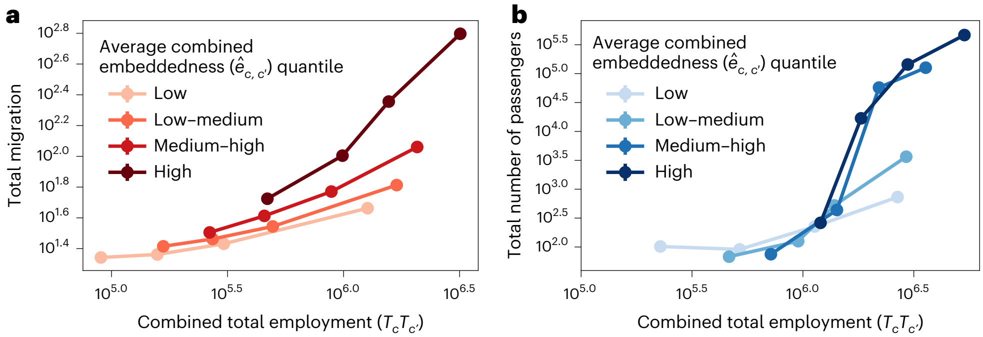

Fig. 3 | Average combined embeddedness moderates the relationship between city size and intercity mobility. a,**b**, City pairs with greater combined embeddedness have greater migration of people (**a**) and more enplaned passengers (**b**) moving between them. Colors represent embeddedness quantiles, and points of the same color represent combined employment quantiles within the embeddedness quantile. Points are plotted according to the average combined employment and average total migration (**a**) or enplaned passengers (**b**) for city pairs. The 99% confidence intervals are provided but are small. See Supplementary Fig. 10 for underlying scatter plots.

between large and small cities[44,50,51]. Since cities divide workers with different skills[41] and employment opportunities influence urbanization[30,31,38], bottlenecks to career mobility may hinder spatial mobility as well. To this end, traditional models for inter-city mobility are improved when considering combined occupation embeddedness (Fig. 3 and Table 3). Comparing the labor market in one city to that in another—beyond simply comparing their size (that is, through the gravity or radiation models)—reveals additional information about intercity mobility. Our study is a step in this direction by demonstrating how workers' positions in city's occupation networks relate to mobility. Future work might use the concept of embeddedness and skill similarity to better understand modern urbanization trends; for example, to better understand why modern urbanization is increasingly dominated by only high-skill workers[44].

We interpret our analysis of embeddedness, relocation and wages as revealing the role of skill specialization in labor dynamics when skill demands are in equilibrium. That is, O*NET's occupation skill profiles do not change substantially between versions (that is, each occupation is updated every 5 years) and an occupation's important skills in O*NET probably result from a balance between what workers can do and what employers need workers to do. Similarly, employment statistics from the BLS OEWS are not easily separated into the dynamics of employers and the dynamics of job seekers. Further, embeddedness may play a different role during out-of-equilibrium labor shocks. For example, a recent study[13] found that cities' with densely connected occupation networks experienced lower unemployment following the Great Recession. Future work may better describe the trade-off between embeddedness during normal economic periods compared to economic shocks. Another limitation is that O*NET does not include regional variation in occupations' skill requirements. Although local employment share does not explain employment-weighted embeddedness (Supplementary Section 3), we cannot rule out that an occupation's skill requirements may vary regionally. Further, occupation classifications may change over longer periods of time. Alternative data sources, including job postings, may offer new insights into regional skill variability within occupations and new understanding into how changing skill demands within an occupation lead to the birth of new occupations.

Methods

O*NET skills data

We quantify occupations' skill requirements using the O*NET skills database. These occupation skill profiles include over 700 different job titles from the SOC system and details the importance of 120–230 examples of workplace knowledge, abilities and skills (henceforth, 'skills') in each year from 2012 to 2018. We use $\mathrm{onet}_{j,s} \in [0,1]$ to denote the real-valued importance of skill $s \in S$ to occupation $j \in J$ such that $\mathrm{onet}_{j,s} = 1$ identifies an essential skill and $\mathrm{onet}_{j,s} = 0$ indicates an irrelevant skill. Visualizations of the O*NET data for each year from 2012 to 2018 are provided in Supplementary Section 1. O*NET skills data have been used to explore labor dynamics in US cities and to explain nationwide labor trends[13,19,41,52,53].

Resume data

Resumes from BGT, FF and Revelio are probably not representative of the US economy on whole, but they do allow us to track individuals throughout their careers. Further, the three resume datasets differ in sampling but are not disjointed; therefore, we analyze each resume dataset separately in this study. Running our analysis on nationally representative CPS data and with resume data with a greater sample of high earning occupations (Supplementary Section 9) strengthens the robustness of our results if consistent across all data sets.

Migration and enplaned passengers

Our study used inter-city migration data from the US Census Bureau and the number of enplaned passengers flying between city pairs according to the US Bureau of Transportation Statistics. Our comprehensive data includes 262,754 directed migration flows and 44,749 directed flows of enplaned passengers between pairs of US cities during 2013 through 2018. These two data sources represent different types of spatial mobility: a migration is a permanent relocation, while flights include temporary relocations (for example, business travel). While employment opportunities are known to shape migration, they are a smaller factor in flight data. However, a large part of flight travel is work related. For example, 12% of enplaned passengers fly business class according to the US Department of Transportation. There were over 463 million business trips inside the USA in 2018. According to the Global Business Travel Association, 1.3 million business trips are taken in the United States every day. There is a small caveat that not all business class flight tickets are necessarily for business-related reasons and business-related trips need not be entirely dedicated to business, but social travel can also approximate the social connections that create career opportunities[54,55]. Still, these statistics suggest that flight mobility between cities may too be linked to the similarities in local economies, albeit to a lesser extent than intercity migration. Although the bulk of inter-city flights, even business class flights, are not related to job seeking, they are related to economic activity more generally.

Tracking career mobility through the community population survey

The US Census Bureau and BLS produce monthly CPS data through a continuous survey process that produces representative samples of the US population. Providing high-resolution labor statistics is one of the primary goals of CPS, and, in particular, CPS records changes in occupations of survey participants over a 1.5 year period for which that participant is an active contributor to the survey. However, some studies have noted limits to CPS's effectiveness at predicting aggregate labor trends since, for example, it misses immigration[56]. For our purpose, we were interested only in participants who reported one occupation when they were first surveyed and reported working a different occupation when they were surveyed one year later. There are several methods for joining different time periods of the CPS data[57], so we employed strict merging criteria including participant identification, gender, sex, state of residency and age to verify the validity of our occupation transitions. The result was a dataset of 36,536 occupation transitions for individual US workers in between 2012 and 2018.

Mapping occupations from FF resumes

We used a publicly available contextual embedding created using a DistilRoBERTa model[58] pre-trained on paraphrasing data[59] to map each FF job title to the SOC occupation used by the US Department of Labor with the largest cosine similarity in the contextual embedding. We used a word embedding model instead of simpler methods (for example, fuzzy string matching) because the model includes semantic information that string matching misses (for example, 'civil engineer' and 'mechanical engineer' both contain 'engineer' but are very different occupations in actuality) and because the model is trained on a diverse set of data (for example, all of English language Wikipedia), which we expect is semantically diverse enough to capture the diversity among US occupations.

Aggregations of workplace skills

While the present study argues for the value of granular workplace skills for studying worker mobility, several existing studies instead utilize broad aggregations of skills (for example, cognitive versus physical skills) to study labor. These simplifications are necessary for tractable analytical work, but this has the potential to stymie empirical forecasting of worker flows between occupations. Therefore, our analysis of worker flows (Fig. 1) compares a model using the granular O*NET taxonomy of skills to alternative baseline models using expertly derived aggregations of skills. We measure an occupations reliance on an aggregated skill category C according to

$$\mathrm{reliance}_{j,C} = \sum_{s\in C} \mathrm{onet}_{j,s}. \tag{6}$$

Definitions for routine or nonroutine and cognitive or physical O*NET variables are provided in Supplementary Section 2. Similar to our granular skills model discussed in the main text, we use Jaccard similarity to measure the skill similarity between occupation pairs based on occupations' reliance on the aggregated skill categories.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

This study uses several publicly available datasets from federal government sources. The O*NET database https://www.dol.gov/agencies/eta/onet and Occupation Employment and Wage Statistics https://www.bls.gov/oes/ are available for download from the US BLS website. The Current Population Survey and intercity migration data are available for download from the US Census Bureau website https://www.census.gov/topics/population/migration/data/tables/cps.html. Domestic Enplaned Passengers data are available for download through the US Bureau of Transportation Statistics https://www.bts.gov/browse-statistical-products-and-data/bts-publications. This study also uses three proprietary third-party resume datasets provided by Burning Glass Technologies Inc., FutureFit AI and Revelio Labs. We cannot make this proprietary data immediately available because of risks to individuals' privacy. Access to these data is available through the third-party data vendors.

Code availability

Code used to produce figures and tables in this article is available at https://bit.ly/3BiZoOT.

References

1. Frey, C. B. & Osborne, M. A. The future of employment: how susceptible are jobs to computerisation? *Technol. Forecast. Soc. Change* https://doi.org/10.1016/j.techfore.2016.08.019 (2016).
2. Benzell, S. G., Collis, A. & Nicolaides, C. Boosting business value by reducing COVID-19 transmission risk. *MIT Sloan Manag. Rev.* **62**, 1–6 (2020).
3. del Rio-Chanona, R. M., Mealy, P., Pichler, A., Lafond, F. & Farmer, J. D. Supply and demand shocks in the COVID-19 pandemic: an industry and occupation perspective. *Oxf. Rev. Econ. Policy* **36**, S94–S137 (2020).
4. Mealy, P. & Teytelboym, A. Economic complexity and the green economy. *Res. Policy* **51**, 103948 (2022).
5. Delgado, M., Porter, M. E. & Stern, S. Defining clusters of related industries. *J. Econ Geogr.* **16**, 1–38 (2016).
6. Jun, B., Glaeser, E. & Hidalgo, C. The role of industry, occupation, and location-specific knowledge in the survival of new firms. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1808.01237 (2018).
7. Autor, D. H. & Dorn, D. The growth of low-skill service jobs and the polarization of the US labor market. *Am. Econ. Rev.* **103**, 1553–1597 (2013).
8. Alichi, A., Kantenga, K. & Sole, J. *Income Polarization in the United States* (International Monetary Fund, 2016).
9. Acemoglu, D. & Autor, D. Skills, tasks and technologies: implications for employment and earnings. *Handbook Labor Econ.* **4**, 1043–1171 (2011).
10. Rosen, S. Specialization and human capital. *J. Labor Econ.* **1**, 43–49 (1983).
11. Fairchild, M. *in Frank and Lillian Gilbreth: Critical Evaluations in Business and Management* **2**, 102 (Routledge, 2003).
12. Frank, M. R. et al. Toward understanding the impact of artificial intelligence on labor. *Proc. Natl Acad. Sci USA* https://doi.org/10.1073/pnas.1900949116 (2019).
13. Moro, E. et al. Universal resilience patterns in labor markets. *Nat. Commun.* **12**, 1–8 (2021).
14. Roy, A. D. Some thoughts on the distribution of earnings. *Oxf. Econ. Papers* **3**, 135–146 (1951).
15. Heckman, J. J. & Honore, B. E. The empirical content of the roy model. *Econometrica* **58**, 1121–1149 (1990).
16. Dawson, N., Williams, M.-A. & Rizoiu, M.-A. Skill-driven recommendations for job transition pathways. *PLoS ONE* **16**, e0254722 (2021).
17. Goos, M., Rademakers, E., Salomons, A. & Willekens, B. Markets for jobs and their task overlap. *Labour Econ.* **61**, 101750 (2019).
18. Robinson, C. Occupational mobility, occupation distance, and specific human capital. *J. Hum. Resour.* **53**, 513–551 (2018).
19. Frank, M. R., Sun, L., Cebrian, M., Youn, H. & Rahwan, I. Small cities face greater impact from automation. *J. R. Soc. Interface* **15**, 20170946 (2018).

20. Farinha, T., Balland, P.-A., Morrison, A. & Boschma, R. What drives the geography of jobs in the US? Unpacking relatedness. *Ind. Innov.* **26**, 988–1022 (2019).
21. Wheaton, W. C. & Lewis, M. J. Urban wages and labor market agglomeration. *J. Urban Econ.* **51**, 542–562 (2002).
22. Angel, D. P. High-technology agglomeration and the labor market: the case of Silicon valley. *Environ. Plan.* **23**, 1501–1516 (1991).
23. Ellison, G. & Glaeser, E. L. The geographic concentration of industry: does natural advantage explain agglomeration? *Am. Econ. Rev.* **89**, 311–316 (1999).
24. Glaeser, E. L. & Gottlieb, J. D. The wealth of cities: agglomeration economies and spatial equilibrium in the United States. *J. Econ. Lit.* **47**, 983–1028 (2009).
25. Bettencourt, L., Samaniego, H. & Youn, H. Professional diversity and the productivity of cities. *Sci. Rep.* **4**, 5393 (2014).
26. Youn, H. et al. Scaling and universality in urban economic diversification. *J. R. Soc. Interface* **13**, 20150937 (2016).
27. del Rio-Chanona, R. M., Mealy, P., Beguerisse-Díaz, M., Lafond, F. & Farmer, J. D. Occupational mobility and automation: a data-driven network model. *J. R. Soc. Interface* **18**, 20200898 (2021).
28. Shutters, S. T. et al. Urban occupational structures as information networks: the effect on network density of increasing number of occupations. *PLoS ONE* **13**, e0196915 (2018).
29. Tian, L. *Division of Labor and Productivity Advantage of Cities: Theory and Evidence from Brazil* (CEPR, 2021).
30. Greenwood, M. J. & Hunt, G. L. Migration and interregional employment redistribution in the united states. *Am. Econ. Rev.* **74**, 957–969 (1984).
31. Rabe, B. & Taylor, M. P. Differences in opportunities? Wage, employment and house-price effects on migration. *Oxf. Bull. Econ. Stat.* **74**, 831–855 (2012).
32. Schubert, G., Stansbury, A. and Taska, B. Employer concentration and outside options. Preprint at *SSRN* https://doi.org/10.2139/ssrn.3599454 (2021).
33. Bijwaard, G. E., Schluter, C. & Wahba, J. The impact of labor market dynamics on the return migration of immigrants. *Rev. Econ. Stat.* **96**, 483–494 (2014).
34. Stark, O. & Bloom, D. E. The new economics of labor migration. *Am. Econ. Rev.* **75**, 173–178 (1985).
35. Beine, M., Bertoli, S. & Moraga, J. F.-H. A practitioner's guide to gravity models of international migration. *World Econ.* **39**, 496–512 (2016).
36. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
37. Petrongolo, B. & Pissarides, C. A. Looking into the black box: a survey of the matching function. *J. Econ. Lit.* **39**, 390–431 (2001).
38. Westerlund, O. Employment opportunities, wages and international migration in Sweden 1970–1989. *J. Reg. Sci.* **37**, 55–73 (1997).
39. Gathmann, C. & Schönberg, U. How general is human capital? A task-based approach. *J. Labor Econ.* **28**, 1–49 (2010).
40. Mealy, P., del Rio-Chanona, R. M. & Farmer, J. D. What you do at work matters: new lenses on labour. Preprint at *SSRN* https://doi.org/10.2139/ssrn.3143064 (2018).
41. Alabdulkareem, A. et al. Unpacking the polarization of workplace skills. *Sci. Adv.* **4**, eaao6030 (2018).
42. Becker, G. S. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (Univ. Chicago Press, 2009).
43. Ellison, G., Glaeser, E. L. & Kerr, W. R. What causes industry agglomeration? Evidence from coagglomeration patterns. *Am. Econ. Rev.* **100**, 1195–1213 (2010).
44. Autor, D. *Work of the Past, Work of the Future* (National Bureau of Economic Research, 2019).
45. Backhaus, A., Martinez-Zarzoso, I. & Muris, C. Do climate variations explain bilateral migration? A gravity model analysis. *IZA J. Dev. Migr.* **4**, 3 (2015).
46. Garcia, A. J., Pindolia, D. K., Lopiano, K. K. & Tatem, A. J. Modeling internal migration flows in sub-saharan africa using census microdata. *Migr. Stud.* **3**, 89–110 (2015).
47. Czaika, M. & Parsons, C. R. The gravity of high-skilled migration policies. *Demography* **54**, 603–630 (2017).
48. Autor, D. The polarization of job opportunities in the us labor market: implications for employment and earnings. *Community Invest.* **23**, 11–41 (2010).
49. Nedelkoska, L. et al. *Skill Mismatch and Skill Transferability: Review of Concepts and Measurements* (Papers in Evolutionary Economic Geography, 2019).
50. Feldman, M., Guy, F. & Iammarino, S. Regional income disparities, monopoly and finance. *Cambridge J. Reg. Econ. Soc.* **12**, 25–49 (2020).
51. Hong, I., Frank, M. R., Rahwan, I., Jung, W.-S. & Youn, H. The universal pathway to innovative urban economies. *Sci. Adv.* **6**, eaba4934 (2020).
52. MacCrory, F., Westerman, G., Alhammadi, Y. & Brynjolfsson, E. Racing with and against the machine: changes in occupational skill composition in an era of rapid technological advance. In *Proceedings of the International Conference on Information Systems* (2014).
53. Handel, M. J. The O*NET content model: strengths and limitations. *J. Labour Mark. Res.* **49**, 157–176 (2016).
54. Granovetter, M. S. The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973).
55. Gee, L. K., Jones, J. J., Fariss, C. J., Burke, M. & Fowler, J. H. The paradox of weak ties in 55 countries. *J. Econ. Behav. Organ.* **133**, 362–372 (2017).
56. Kambourov, G. & Manovskii, I. A cautionary note on using (March) Current Population Survey and Panel Study of Income Dynamics data to study worker mobility. *Macroecon. Dyn.* **17**, 172–194 (2013).
57. Madrian, B. C. & Lefgren, L. J. *A Note on Longitudinally Matching Current Population Survey (CPS) Respondents* (National Bureau of Economic Research, 1999).
58. Sanh, V., Debut, L., Chaumond, J. & Wolf, T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. Preprint at *arXiv* https://doi.org/10.48550/arXiv.1910.01108 (2019).
59. Reimers, N. & Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing* 11 (Association for Computational Linguistics, 2019).
60. Eucalyp. *flaticon* (2019); https://www.flaticon.com/authors/eucalyp

Acknowledgements

This research is supported by the University of Pittsburgh Pitt Momentum Fund and the Center for Research Computing. E.M. acknowledges support by Ministerio de Ciencia e Innovación, Agencia Española de Investigación (MCIN/AEI/10.13039/501100011033) grant PID2019-106811GB-C32 and the National Science Foundation grant no. 2218748. The authors thank E. Brynjolfsson for his insightful comments.

Author contributions

M.R.F. performed calculations and produced figures. M.R.F. and E.M. secured funding for this project. T.S. prepared the resume data from FutureFit AI. B.T. secured resume data from Burning Glass Technologies. All authors designed the research and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44284-023-00009-1.

Correspondence and requests for materials should be addressed to Morgan R. Frank.

Peer review information *Nature Cities* thanks Teresa Farinha, Marian-Andrei Rizoiu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Morgan Frank

Last updated by author(s): Nov 15, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	*Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*
Data analysis	Code has been made available at https://bit.ly/3BiZoOT

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

This study uses several publicly-available data sets from federal government sources.
The O*NET database \url{https://www.dol.gov/agencies/eta/onet} and Occupation Employment and Wage Statistics \url{https://www.bls.gov/oes/} are available for download from the US Bureau of Labor Statistics website.

The Current Population Survey and intercity migration data are available for download from the US Census Bureau website \url{https://www.census.gov/topics/population/migration/data/tables/cps.html}.
Domestic Enplaned Passengers data are available for download through the US Bureau of Transportation Statistics \url{https://www.bts.gov/browse-statistical-products-and-data/bts-publications}.
This study also uses three proprietary third-party resume data sets provided by Burning Glass Technologies Inc., FutureFit AI, and Revelio Labs.
We cannot make this proprietary data immediately available because of risks to individuals' privacy.
Access to these data are available through the third-party data vendors.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This quantitative study uses labor and skills data from the US Bureau of Labor Statistics to model US urban labor markets as networks
Research sample	Over 100 million resumes provided by Burning Glass Technologies, FutureFit Ai, and Revelio Labs. Data was collected by the companies. Representativeness of the data is discussed in the Supplementary Material.
Sampling strategy	Resume data is collected by Burning Glass Technologies, FutureFit AI, and Revelio Labs. Data was collected by the companies. Datasets were selected in the study because they are the most comprehensive resume studies available.
Data collection	Resume data is collected by Burning Glass Technologies, FutureFit AI, and Revelio Labs. Researchers have no role in the collection of those resumes databases
Timing	Data covers 2012 through 2018
Data exclusions	No data exclusions.
Non-participation	There are no "participants" in the study. The resume data was collected by the companies
Randomization	The resume data was collected by the companies. Includes all resumes provided by users. There are no "participants" in the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging