Journal of Economic Behavior & Organization 142 (2017) 275–292

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Inter-industry labor flows

Frank M.H. Neffke[a,*], Anne Otto[b], Antje Weyh[b]

[a] *Harvard Kennedy School, Center for International Development, United States*
[b] *Institute for Employment Research (IAB), Germany*

ARTICLE INFO

Article history:
Received 26 January 2017
Received in revised form 12 June 2017
Accepted 5 July 2017
Available online 22 July 2017

Keywords:
Labor mobility
Relatedness
Skills
Regional growth
Germany
Human capital specificity

ABSTRACT

Using German social security data, we study inter-industry labor mobility to assess how industry-specific human capital is and to determine which industries have similar human capital requirements. We find that inter-industry labor flows are highly concentrated in just a handful of industry pairs. Consequently, labor flows connect industries in a sparse network. We interpret this network as an expression of industries similarities in human capital requirements, or skill relatedness. This skill-relatedness network is stable over time, similar for different types of workers and independent of whether workers switch jobs locally or over larger distances. Moreover, in an application to regional diversification and local industry growth, skill relatedness proves to be more predictive than colocation or value chain relations. To facilitate future research, we make detailed inter-industry relatedness matrices online available.

1. Introduction

Labor mobility plays an important role in economics. On the one hand, industry-specific technology or demand shocks often necessitate a transfer of productive capacity, and thus of workers, from shrinking to growing industries. On the other hand, labor mobility diffuses know-how across firms, industries and locations, and is therefore important in organizational learning (March 1991; Simon, 1991) and regional and national growth (Saxenian, 2007). Unsurprisingly, therefore, labor mobility has received much scholarly attention from both labor economists and innovation economists. However, one aspect of labor mobility has hitherto been relatively neglected, namely, the mobility patterns of workers across industry boundaries. As a consequence, the inter-industry structure of labor flows is still poorly understood. This is surprising, given that if inter-industry labor flows exhibit a high degree of structure, mobility of workers across industries will be constrained by this structure. Because any constraints to such mobility will limit both, the reallocation of labor, and the diffusion of knowledge, a deeper understanding of inter-industry labor flows may shed light on a wide range of economic phenomena, from individual careers to economic development, structural change and innovation.

In this paper, we contribute to our understanding of inter-industry labor flows, showing that they exhibit strong regularities. We summarize these regularities in a set of stylized facts that are organized around three related topics: (1) the expression of human capital specificities in the structure labor flows, (2) the use of labor-flow-networks as measures of inter-industry relatedness and (3) the way in which the constraints on inter-industry labor-flows these networks express affect diversification and labor reallocation in local economies. In particular, we ask a number of interrelated questions: Do labor flows concentrate in relatively few industry pairs? How stable is the network of inter-industry labor flows? Is this net-

* Corresponding author.
E-mail address: frank_neffke@hks.harvard.edu (F.M.H. Neffke).

http://dx.doi.org/10.1016/j.jebo.2017.07.003
0167-2681/

work general or specific to an occupation? Does the sparseness of the inter-industry labor-flow-network condition a region's growth path? And, finally, does this sparseness constrain a local economy's capacity to reallocate labor from contracting to expanding industries?

These questions complement a vast literature on general labor flows and job switching. For instance, labor economists have extensively studied job-switching rates (or, their complement, employment durations) and how they depend on business cycles, industry and worker characteristics. Recent work in this tradition finds that workers often change jobs across industries that belong to completely different sectors (Parrado et al., 2007; Kambourov and Manovskii, 2008). This finding may lead to the conclusion that human capital has no strong industry-specific component. However, because this research fails to take into consideration *which* industries exchange workers, it implicitly assumes that all industries are equidistant from one another in terms of human capital requirements. We show that a closer analysis of the network structure of inter-industry labor flows casts doubts on this conclusion. These analyses are collected in a first set of stylized facts that describe how much structure inter-industry labor flows exhibit.

A different group of scholars at the intersection of innovation economics and economic geography has studied the role of labor flows as conduits of knowledge diffusion, typically focusing on the mobility of highly skilled workers, such as inventors. However, although the spatial limits to mobility are central in much of this research (Breschi and Lissoni, 2005; Agrawal et al., 2006; Casper, 2007), also here, the question of whether there are *inter-industry* constraints to labor mobility has typically been neglected.

A second debate to which our work relates takes place in the literature on inter-industry relatedness. In spite of the relative neglect of inter-industry labor flows in labor and innovation economics, an increasing number of papers has turned to such flows as an expression of inter-industry relatedness. These papers assume that human capital is to some extent industry specific. Consequently, labor flows are constrained and will predominantly take place between industries with similar human-capital requirements. This has resulted in labor-flow-based *skill-relatedness*[1] measures (Neffke and Henning, 2013), which have been used in a variety of papers (e.g., Timmermans and Boschma, 2014; Boschma et al., 2014; Diodato and Weterings, 2015). In accordance with this literature, a second set of stylized facts analyzes inter-industry labor flows through the prism of skill relatedness. In particular, we are interested in four issues. First, how volatile are skill-relatedness structures? That is, do they change much from one year to the next or are they relatively stable? Second, how general are skill-relatedness measures? That is, do different types of workers exhibit different skill-relatedness patterns? Third, given that many workers tend to search for jobs in their own region, skill-relatedness measures may just reflect industrial colocation patterns. We therefore ask: do short-distance and long-distance flows differ in the skill-relatedness structure they exhibit? Fourth and finally, we ask: what is the predictive validity of skill-relatedness measures vis-à-vis alternative relatedness measures?

We derive stylized facts from Germany's social security records between 1999 and 2008, which cover over 80% of the working population. We find that, although workers often do switch industries, even at a very high level of aggregation (stylized fact 1), labor flows are highly structured (stylized fact 2). In particular, on average, related industries that together represent just 5% of total German employment absorb over 60% of an industry's total worker outflow. Moreover, the underlying network of labor flows is largely independent of a worker's occupation: workers in different occupations tend to make the same industry switches. This suggests that, independently of any occupational specificities, job switches are guided by a non-negligible industry-specific component in human capital. When we turn to labor flows as a measure of inter-industry relatedness, we find that the derived skill-relatedness index is remarkably stable (stylized fact 5), general across occupations and wage levels and similar in former East and West Germany (stylized fact 4). Furthermore, given that intra-regional flows follow a similar skill-relatedness structure as long-distance flows (stylized fact 3), skill-relatedness is not simply a reflection of the industrial composition of local economies. Moreover, in a direct comparison, our labor-flow-based measure outperforms commonly used alternative relatedness measures in predicting entry and growth rates of local industries (stylized fact 6). Finally, skill-related industries have uncorrelated growth patterns, suggesting that skill-relatedness should typically not impede the reallocation of labor from shrinking to growing industries (stylized fact 7).

Although we limit the analysis in the paper to skill-relatedness among the industries of the classification systems in use between 1999 and 2008, we have constructed skill-relatedness matrices for various industrial and occupational classification systems between 1975 and 2014. To facilitate future research, these matrices are available for online download.[2]

The paper is structured as follows. In Section 2, we discuss the literature on human capital specificities and job switches and the literature on inter-industry relatedness measures. Section 3 describes the data. In Section 4, we develop a number of statistical tools to analyze labor-flow networks and present the stylized facts uncovered with these tools. Section 5 discusses future research and concludes.

[1] Neffke and Henning (2013) argue that, although workers may select a new job for reasons other than skill relatedness (such as preferences and social networks), in the aggregate, inter-industry labor flows seem to predominantly express skill similarities. We will return to this issue later on.

[2] A link to these matrices, as well as detailed description of the procedure used to create these matrices, is provided on the first author's personal website.

2. Literature review

2.1. Human capital specificity and job switching

Human capital and skills are pivotal inputs in today's production processes, which is why a firm's workforce is regarded as an exceedingly important competitive asset (Porter, 1987; Grant, 1996; Grant and Spender, 1996). Moreover, today's workforces are highly specialized: individual workers often invest heavily in education and training to acquire specific skills that allow them to carry out specific tasks, running the gamut from engineering to financial management and from construction work to food preparation. Because workers specialize, their human capital is often held to be specific to the firm where they work (Becker, 1964), to an industry (Neal, 1995; Parent, 2000; Sullivan, 2010) and to occupations and tasks performed (Poletaev and Robinson, 2008; Gathmann and Schönberg, 2010). However, there is considerable debate over which of these dimensions of skill specificity dominate. For instance, Kambourov and Manovskii (2009) study the value of occupation tenure and industry tenure and report that only the former is rewarded with higher wages, casting doubt on the existence of industry specificities in human capital. In contrast, Sullivan (2010) shows that industry tenure sometimes does reap high rewards that are not explained by occupation tenure, but that these rewards *depend* on the occupation.

An alternative way of thinking about the skill content of human capital is described by Lazear (2009). Lazear argues that it is not skills that are specific to a firm, but rather the exact combination in which the firm uses them. Accordingly, human capital consists of a number of different skills, all of which are general in the sense that they are demanded by a wide range of firms. However, firms differ from one another in the weights they attach to each skill. Consequently, workers with a given skill profile are more productive in one firm than in another. Firm-specific skill weights yield many of the same predictions as firm-specific human capital, such as that workers will incur wage losses upon involuntary job separations, but they can also explain some empirical facts that are harder to square with a theory based on the existence of firm-specific human capital.[3]

We approach the question of whether human capital has an industry-specific component in a way that is closely aligned with Lazear's skill weights interpretation of human capital.[4] The starting point is that job-switching patterns contain valuable information on the nature of human capital. In Lazear's terminology, industries differ in how they weigh different skills. Consequently, when a worker switches jobs, she will render some human capital redundant, whenever the old and the new job require a different skill mix. To avoid such human capital depreciation, workers will predominantly switch to jobs that allow them to reuse as many of their skills as possible. This suggests that overlap in industries' human capital requirements, or, more accurately, an absence of such overlap, should constrain inter-industry labor movements. Therefore, inter-industry job switches contain information on which industries value similar skills and know-how, indirectly shedding light on the existence of human capital specificities.

Job switches have been studied in great detail in labor economics. Much of this literature's interest in labor flows is driven by their role in readjusting the allocation of labor across firms.[5] Research in this tradition has resulted in numerous papers summarized in various reviews (Davis and Haltiwanger, 1999; Farber, 1999; Davis et al., 2006). These papers mainly study the rate at which jobs are created and destroyed, the rate at which workers change jobs, and whether workers do so voluntarily or involuntarily. Other important questions are how job-switching rates develop over the business cycle and what this means for unemployment dynamics and labor-market institutions.

Although most of this work in labor economics has focused on the question of how often and why workers change *jobs*, more recently, scholars have turned to the issue of workers switching *industries*. These studies find that, in the United States, workers change 1-digit industries at relatively high rates of between 12% (Kambourov and Manovskii, 2008) and 20% (Parrado et al., 2007) a year. These findings may be interpreted as a sign that workers are not strongly constrained in their movements across industries, implying that human capital is not highly specific to an industry. But this conclusion rests on the assumption that the hierarchy of the industry classification system groups industries by their human capital requirements. However, a similar notion – that industry classification systems reflect similarities in use of strategic resources – has been heavily criticized in strategic management (Robins and Wiersema, 1995; Bryce and Winter, 2009; Neffke and Henning, 2013). In the empirical section, we will show that such assumptions are indeed not just problematic, but that evidence based on them is even misleading.

[3] For instance, in the latter kind of theory, firms should not be willing to pay for their employees' training in general skills, because the benefits of this training will be fully appropriated by the employee. Another implication of the skill weights approach is that workers will face lower wage drops when losing their job in thicker labor markets. If the wage losses were due to the fact that workers lost their firm-specific human capital, the thickness of the labor market should not matter.

[4] We than an anonymous reviewer for pointing out the connection between our analyses and Lazear's notion of skill weights.

[5] For instance, in an overview paper, Davis and Haltiwanger (1999, p. 2713) motivate the widespread interest in labor flows as follows: "[T]he extent to which the reallocation and matching process operates smoothly determines, in large measure, the difference between successful and unsuccessful economic performance."

2.2. Knowledge spillovers

Labor mobility has also been studied in a different context, namely, as a mechanism for knowledge diffusion. Accordingly, workers who switch jobs do not just reallocate labor, but often also carry with them valuable knowledge, expertise and networks they acquired at their previous employer (Almeida and Kogut, 1999; Cantner and Graf, 2006; Storz et al., 2015). Because most individuals change jobs within their regions, such knowledge spillovers tend to be spatially constrained. The spatially bounded nature of knowledge spillovers has received much attention (e.g., Agrawal et al., 2006). However, the scope for knowledge spillovers may also be limited by human capital specificities. In particular, economic activities that employ radically different skills and knowhow will have limited scope for exchanging workers, making it less likely that knowledge and technologies are shared.

2.3. Related diversification

Because inter-industry labor flows will mostly occur among related industries, inter-industry labor flows might be used to measure inter-industry relatedness. The concept of inter-industry relatedness has played an important role in the literature on firm diversification (Penrose, 1959; Teece et al., 1994) and, more recently, also in economic geography and the literature on geographical clusters of firms. For instance, Porter (1998) identifies local clusters of related economic activities as important sources of competitive advantage, with Silicon Valley as the archetypal example. Although originally stressing local value chains, the cluster concept has evolved to include more general linkages that "create externalities of various types" (Porter, 2003). To measure such linkages, scholars have used information on the co-location patterns of industries (Porter, 2003) and the co-occurrence of products in countries' export portfolios (Hidalgo et al., 2007). Although these and other relatedness measures have proven highly predictive of the growth of local industries (Boschma et al., 2013; Delgado et al., 2010; Delgado et al., 2014; Essletzbichler, 2013; Hausmann et al., 2014; Neffke et al., 2011; Rigby, 2015), recently, labor-flow based inter-industry relatedness measures have been gaining support. For instance, Greenstone et al. (2010) show that large-plant openings create spillovers to local firms, but in particular to firms in industries that are related to the new plant's industry as measured by labor flows. Similar labor-flow-based measures have been used in studies in economic geography (Dauth, 2010; Timmermans and Boschma, 2014), trade (Kaplan et al., 2011), strategic management (Neffke and Henning, 2013) and entrepreneurship research (Costa and Baptista, 2011). To our knowledge, however, hitherto there has not been any detailed investigation of the structure of inter-industry labor flows that would assess the legitimacy of such labor-flow based relatedness measures.

3. Data

Our data are constructed from Germany's social security records as compiled in the German Employee History (*Beschäftigtenhistorik*, EH) database.[6] The HES offers a large set of demographic and employment characteristics, such as a worker's daily wage,[7] occupation, work status (i.e., apprentice, part-time worker, full-time worker), gender and age. Furthermore, the industry[8] and location of each individual's work establishment are known.

We limit the analysis to full-time employees aged 18–65. Furthermore, we exclude apprentices and volunteers because they are still investing in education to acquire skills. Because upper limits to social security contributions result in right-censored wage information, we impute wages whenever they exceed the contribution limits following Gartner (2005). Due to changes in the industry-classification system (see for a more detailed description Appendix A in Supplementary material), we confine our analyses to the years 1999–2008. This results in a final data set with, on average, about 20 million workers a year.

3.1. Definition of labor flows

We use the EH data to construct inter-industry labor flows. Labor flows arise when workers switch establishments from one year to the next. Workers who enter or exit the social-security data in this period are ignored in these flows. As a consequence, these labor flows predominantly reflect job-to-job switches.[9] Moreover, establishment identifiers in the EH are not perfectly reliable. For instance, spin-offs, mergers, break-ups or mere recodings all would introduce new establishment identifiers that do not correspond to de-novo entries. Hethey and Schmieder (2010) find that for only 35% to 40% of all establishments with over three employees a new (or a disappearing) establishment identifier can be interpreted unambiguously as an entry (or as an exit) of an economic establishment. In the other cases, workers move in larger blocks from

---

6 Bender and Möller (2010) provide a detailed description of this database.

7 We deflate wages to 2005 EUR.

8 We drop workers employed through employment agencies, because we don't have information on the actual industry or region in which these individuals work.

9 Some job switches may feature short (unobserved) unemployment spells in between the two observed jobs. Although this eliminates some involuntary job switches that lead to prolonged unemployment spells, we have no reliable way of differentiating between voluntary and involuntary job switches.

one establishment to another. To avoid that such spurious identifier changes contaminate our labor-flow measurements, we remove 531,000 job switches (27.5%) from a total of 1.8 million yearly job switches.[10]

3.2. Labor-flow segments

In the empirical section of this paper, we decompose flows into different segments. We introduce three kinds of segmentations. The first is based on the geography of flows, the second captures workers' skills and the third distinguishes between an eastern and a western German labor market.

The first segmentation is motivated by the potential concern that the structure of inter-industry labor flows is governed by the availability of local jobs. In that case, inter-industry labor flows could simply be an expression of the co-location patterns of industries. We look into this by comparing job switches over short and long distances, where long distances refer to switches where the old and the new job are at least 100 kilometers apart.[11] Second, to proxy workers human capital *levels*, we segment the labor market into workers who earn below and those who earn above the median wage[12] in their industries. To explore whether the *type* of human capital matters, we also split workers into eight occupational groups that are associated with different broad sets of tasks. In particular, we distinguish among managers, sales-related employees, accountants, information technology (IT) workers, office clerks, cleaners, security personnel and other workers (see Appendix C in Supplementary material). These occupations were chosen because they are found in a wide range of industries so that we are not limiting the analysis to a small subset of the economy.[13] Furthermore, the latter two segmentations are based on a worker's wage and occupation at the origin of a job switch, not (necessarily) its destination. Third, we investigate whether there are regional differences in labor-flow patterns by splitting the German labor market along the former border between East and West Germany.[14]

4. Results

Below, we derive a number of stylized facts on inter-industry labor flows. We first describe the general structure of labor flows in terms of the number of job switches and the degree to which labor flows concentrate in relatively few industry pairs to shed light on the specificity of human capital. Next, we analyze these flows' structural characteristics by plotting the skill-relatedness network and comparing skill-relatedness matrices of different labor-market segments. Finally, we turn to the question of how skill relatedness affects local labor markets by estimating local industry-growth regressions and determining the extent to which mobility constraints across industries could hinder an efficient reallocation of workers from shrinking to growing industries.

4.1. Cross-industry labor-flow patterns

At their coarsest level of aggregation, industries in the European NACE Revision 1.1 classification are divided into sections. Sections consist of several sub-sections, which themselves are composed of 2-digit industries. These 2-digit industries are further subdivided into 3-, 4-, and 5-digit industries. Table 1 summarizes average yearly labor flows across industries at these different levels of aggregation. The first column of Table 1 reports figures for Germany as a whole. Of all workers who change jobs, 73.4% change industries at the 5-digit level, the most disaggregated level available. Most industry switching takes place across highly aggregated industries: 58.7% of 5-digit industry switchers switch industries at the section level, the most aggregated industry grouping in the NACE 1.1 classification. This echoes the findings in Kambourov and Manovskii (2008) on the US economy. These authors find that in the late 1990s, 10% of workers switch 1-digit industries, whereas only marginally more, 13%, switch industries at the 2- or even 3-digit level. Because demographics, sample restrictions and data cleaning procedures differ, it is hard to compare these numbers to ours in absolute terms. Moreover, the US and Germany use different industry classification systems. The 1-digit level of aggregation in the US corresponds relatively closely to the German section level (with 12 1-digit US industries compared to 15 industry sections in Germany). Similarly, the 2-digit US level corresponds more or less to the level of subsections in Germany (33 against 29), whereas at the 3-digit level there are more or less equal numbers of industries in both countries (212 against 219). Therefore, we can compare the ratio of 3-digit

[10] See Appendix B in Supplementary material for a more detailed description of the identification and elimination of spurious job switches.

[11] Distances between old and new jobs represent road distances between the centroids of the districts (*Kreise*) in which the corresponding establishments are registered. Given that only five percent of German employees commute over distances greater than 50 km (Winkelmann, 2010), we postulate that distances of over 100 km typically require a worker to relocate.

[12] We use a worker's wage to proxy worker quality, instead of for instance, his or her education and experience or, its complement, a Mincer residual, because wages will reflect both observed and unobserved quality characteristics.

[13] For instance, if we were to use occupations such as glass makers, we would not expect any labor flows to originate from banks, given that such occupations will simply not exist in that industry.

[14] To ensure that our results are not driven by the extensive outmigration from the East after Germany's reunification, we exclude flows between eastern and western Germany in these analyses.

Table 1
Cross-industry labor flows by labor-market segment.

LABOR-MARKET SEGMENT	WAGES			GERMANY		GEOGRAPHY		OCCUPATIONS						
	all	high	low	East	West	local	long-distance	managers	sales	accountants	office clerks	IT	cleaners	security
WORKERS (THOUSANDS)														
employment	19,897.1	9,947.6	9,914.8	3,890.2	15,958.8	n.a.	n.a.	435.9	1,440.2	293.1	2,600.8	390.5	220.2	254.2
job switchers	1,206.7	554.3	652.4	194.2	940.9	948.2	258.5	32.0	116.5	18.9	148.5	28.7	11.1	14.7
no industry switch	321.1	165.2	155.9	56.2	246.1	255.3	65.8	7.1	32.4	5.3	32.8	4.5	3.3	4.9
industry switch	885.7	389.1	496.5	138.0	694.8	692.9	192.8	24.9	84.1	13.7	115.7	24.2	7.8	9.8
different section	519.8	215.9	303.8	82.1	408.0	409.3	110.4	14.5	45.2	8.0	75.4	14.1	6.1	7.2
same section	365.9	173.2	192.7	56.0	286.8	283.6	82.3	10.3	38.9	5.7	40.3	10.0	1.7	2.6
same sub-section	301.6	145.4	156.2	50.1	231.1	229.8	71.7	8.6	36.2	5.2	34.5	9.4	1.6	2.5
same 2-digit industry	225.9	109.8	116.0	40.4	170.4	175.3	50.6	6.0	25.6	4.0	23.4	4.6	1.3	1.9
same 3-digit industry	117.3	58.8	58.5	20.5	88.6	90.4	26.9	2.8	11.4	3.0	10.8	2.1	0.3	0.6
same 4-digit industry	62.3	31.7	30.6	10.4	47.6	47.6	14.8	1.4	6.0	1.5	5.1	1.7	0.2	0.3
PERCENTAGES														
no industry switch	26.6%	29.8%	23.9%	28.9%	26.2%	26.9%	25.4%	22.2%	27.8%	27.9%	22.1%	15.7%	29.6%	33.3%
industry switch	73.4%	70.2%	76.1%	71.1%	73.8%	73.1%	74.6%	77.8%	72.2%	72.1%	77.9%	84.3%	70.4%	66.7%
different section	58.7%	55.5%	61.2%	59.5%	58.7%	59.1%	57.3%	58.4%	53.8%	58.4%	65.2%	58.5%	78.8%	73.9%
same section	41.3%	44.5%	38.8%	40.5%	41.3%	40.9%	42.7%	41.6%	46.2%	41.6%	34.8%	41.5%	21.2%	26.1%
same sub-section	34.1%	37.4%	31.4%	36.3%	33.3%	33.2%	37.2%	34.7%	43.1%	38.0%	29.8%	38.8%	20.2%	25.7%
same 2-digit industry	25.5%	28.2%	23.4%	29.3%	24.5%	25.3%	26.3%	23.9%	30.4%	29.3%	20.3%	19.0%	16.2%	19.1%
same 3-digit industry	13.2%	15.1%	11.8%	14.9%	12.7%	13.0%	14.0%	11.3%	13.5%	21.7%	9.3%	8.7%	4.4%	5.6%
same 4-digit industry	7.0%	8.1%	6.2%	7.5%	6.8%	6.9%	7.7%	5.7%	7.1%	10.9%	4.4%	6.8%	3.0%	3.3%

The table contains information about absolute and relative employment levels, jobs switching and cross-industry flows in each labor-market segment averaged over 1999–2007. In the section WAGES, column <all> refers to all workers, column <high> refers to workers with wages above their industry's median, column <low> to workers below this median. Long-distance flows are labor flows between establishments that are at least 100 km apart. The occupations are as defined in Table C1.

industry-switchers to 1-digit industry-switchers in the US to the ratio of 3-digit industry-switchers to section-switchers in Germany. These ratios are fairly similar at 1.3 in the US against 1.48 in Germany.[15]

Another way to evaluate the numbers in table is to compare them to a random benchmark. To do so, we simulate switches in which all workers who leave a certain industry choose new industries with probabilities equal to the industries' employment share in the overall economy. This exercise shows that workers tend to remain in their 5-digit industry 39 times more often than random. Moreover, switches to industries that are closer by in the classification system typically exceed this benchmark by more than switches between more aggregated sectors. For instance, in Germany as a whole, switches within 4-digit industries exceed the random benchmark by a factor 15, whereas switches to other sections happen at rates of 0.7 times the random benchmark.[16]

Distinguishing flows by their labor-market segment of origin, we find that workers in the high-income segment switch industries less often and undertake less drastic switches than low-wage workers do. For instance, 8.1% of high-wage workers who switch industries stay in their 4-digit industries (column 2), against only 6.2% for low-wage workers (column 3). This finding is in line with Parrado et al. (2007), who find that higher wages increase the likelihood that workers remain within their industry in the US. Workers in eastern Germany switch industries somewhat less than their colleagues in the west (columns 4 and 5). Columns 6 and 7 show that job switches over distances below 100 km ("local" labor flows) display patterns that are very similar to those that involve distances of over 100 km ("long-distance" labor flows). Differences by occupation are larger. Workers in lower-skilled occupations such as cleaning (78.8%) and security (73.9%) cross section borders much more often than workers in higher-skilled occupations, such as management (58.4%), accountancy (58.4%) and IT (58.5%), who tend to switch to industries that are classified more closely to their old jobs. This suggests that moving to a radically different industry is less attractive when human capital requirements are higher.[17] The fact that more skilled occupations exhibit lower levels of industry switching matches the fact that, in the US, higher levels of educational attainment are associated with lower likelihoods of switching industries (Kambourov and Manovskii, 2008). Overall, the following stylized fact emerges from Table 1:

Stylized fact 1: Workers frequently cross industry boundaries, even at the most aggregate level of the industry classification system. Moreover, such "distant" moves are most frequent for workers with low wages and in low-skill occupations.

4.2. Structuredness of inter-industry labor flows

The fact that workers tend to cross boundaries between highly aggregated sectors does not necessarily mean that labor flows are unstructured. Indeed, 3.3% of all possible pairs of 5-digit industries account for 80% of all job switches and in 56% of pairs we do not observe any job switches whatsoever in the nine-year period we study. How random then are labor flows? To answer this question, we determine for each industry which share of its labor outflows is absorbed by industries that represent a fraction q of overall employment in a segment. Let F_{ij}^s be the labor inflow into industry j that originates from industry i in segment s. Furthermore, let α_{ij}^s be:

$$\alpha_{ij}^s = \frac{F_{ij}^s}{N_j^s}$$

where N_j^s is the number of workers in industry j in segment s. α_{ij}^s can be thought of as the per-employee-rate at which industry j absorbs workers in segment s from industry i. Next, for each industry of origin, we sort destination industries in descending order of this ratio: $\alpha_{i1}^s \geq \alpha_{i2}^s \geq \ldots \geq \alpha_{in}^s$, where n is the number of industries in the economy. Throughout this paper, we will use a dot (".") to denote summation over an omitted dimension. Hence, $F_{i.}^s = \sum_j F_{ij}^s$ represents the total labor outflow in a segment s from industry i. Furthermore $N_.^s = \sum_i N_i^s$ represents the total employment in this segment. We define $STRUC_i^s(q)$ as:

$$STRUC_i^s(q) = \sum_{j=1}^{k} \frac{F_{ij}^s}{F_{i.}^s} \quad \text{with} \quad k = \text{argmin}_{k'}\left(q < \sum_{j=1}^{k'} \frac{N_j^s}{N_.^s}\right)$$

[15] The ratio of 2-digit to 1-digit switchers derived from the study by Kambourov and Manovskii (2008) of 1.3 compares to a 1.12 ratio of subsection-to-section switchers in Germany.

[16] An exception are cleaners, whose switching rates exceed the random benchmark when they change sections, but not when they stay within sections, subsections or detailed industries.

[17] Workers in security and cleaning jobs seem to switch industries less often than other workers. However, this is misleading because cleaners and security guards predominantly work in the cleaning and the security industry. As a consequence, most job switches occur within these two 5-digit industries. Once cleaners and security guards cross industry boundaries, their labor flows look much less structured than the ones in other occupations.

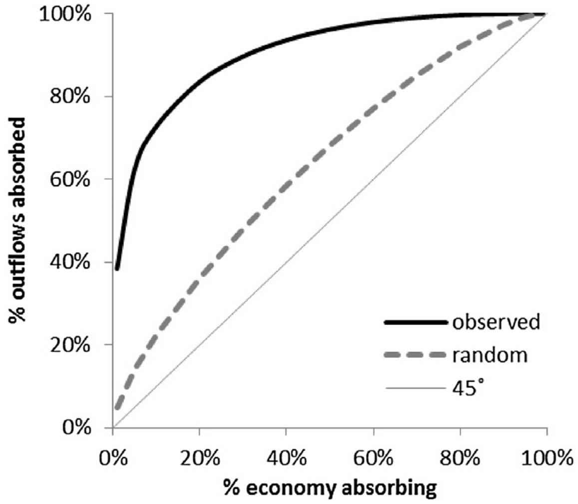

Fig. 1. Degree of structure in the German labor market.
The vertical axis shows the percentage of labor outflows that is absorbed by industries that together represent the share of total German employment depicted on the horizontal axis. Values include within-industry flows, are averaged across the years 1999–2007 and are calculated for the German labor market as a whole. The dotted line provides a random benchmark. The thin grey line represents the 45° line as a point of reference.

In other words, $STRUC_i^s(q)$ represents the maximum share of industry i's total labor outflow that is absorbed by a set of 5-digit industries that together represent at most q% of segment s's total employment.[18] To determine the rigidity of workers in an entire labor-market segment, we calculate the outflow-weighted average of $STRUC_i^s(q)$ across all industries of origin:

$$STRUC^s(q) = \sum_i \frac{F_{i.}^s}{F_{..}^s} STRUC_i^s(q)$$

where $F_{..}^s = \sum_i F_{i.}^s$ represents total inter-industry labor flows originating in labor-market segment s.

$STRUC^s(q)$ can be calculated both with and without within-industry labor flows, that is, with or without workers who change establishments, but not industries. Fig. 1 depicts the average $STRUC^G(q)$ over the period 1999–2007 against q, where G represents the "segment" that contains the entire German labor market. $STRUC^G(q)$ is highly concave, meaning that the bulk of labor outflows are absorbed by a small part of the economy. Because some of the curvature of $STRUC^G(q)$ would also have occurred, had flows simply followed employment, we create a benchmark based on simulated flows. In this simulated benchmark, job switchers in industry i randomly choose a destination industry j with a probability equal to j's employment share in the overall economy: $p_j^s = \frac{N_j^s}{N_.^s}$. The resulting null-model curve is depicted as a dashed line.

Table 2 compares the rigidity of workers in different labor-market segments by providing $STRUC$ (5%) values for each segment excluding (column 1) and including (column 2) within-5-digit-industry flows. The number in parentheses shows the corresponding random benchmark.

In general, labor flows are tightly structured, especially if we include within-industry flows.[19] For Germany as a whole, industries that represent just 5% of the economy absorb 62% of all workers that change jobs (51% if within-industry flows are excluded).[20] High-wage workers, with a $STRUC$ (5%) of 70% (60%), are much less flexible than low-wage workers, with a $STRUC$ (5%) of 59% (49%). Given that low- and high-wage groups have virtually identical null-model predictions, their estimates can be readily compared. For other labor-market segments, null-model predictions vary widely, which complicates comparisons across groups. However, because in all segments, $STRUC$ (5%) values are well below their null-model predictions, we arrive at the following stylized fact:

Stylized fact 2: Labor flows are channeled along tight paths: most workers are absorbed by sets of industries that represent only a small fraction of total employment in the economy.

---

[18] Where necessary, we interpolate $\sum_{j=1}^{k} \frac{F_{ij}^s}{F_{i.}^s}$ between k and $k+1$.

[19] The exception is occupational segments that are dominated by a single large industry (as in cleaning and security jobs), where most of the flows take place among firms within that 5-digit industry. Here including the within-industry flows substantially *decreases* the estimated rigidity or structuredness.

[20] Note that each industry will have a different set of industries that absorbs most of its worker outflow.

Table 2
FLEX(5%) estimates by labor-market segment.

segment	STRUC(5%) (inc. intra-ind flows)	STRUC(5%) (exc. intra-ind flows)
Germany		
all	62.1% (14.0%)	50.7% (14.0%)
high wage	70.4% (18.2%)	60.1% (18.3%)
low wage	59.4% (17.2%)	48.8% (17.3%)
East/West Germany		
East	69.2% (24.2%)	59.9% (24.3%)
West	62.1% (15.2%)	50.8% (15.2%)
Geography of flows		
local	62.1% (15.1%)	50.3% (15.1%)
long-distance	70.8% (24.0%)	63.1% (24.1%)
Occupations		
managers	75.6% (49.4%)	71.5% (49.6%)
sales	65.8% (23.2%)	59.2% (23.4%)
accountants	48.9% (31.9%)	52.4% (33.5%)
office clerks	61.9% (29.2%)	53.6% (29.3%)
IT	36.7% (25.1%)	36.4% (26.0%)
cleaners	42.9% (22.4%)	51.3% (26.0%)
security	47.1% (20.6%)	58.2% (22.1%)

STRUC(5%) equals the (out)flow-weighted maximum average share of (5-digit) industries' total labor outflow that is absorbed by a set of 5-digit destination industries that, taken together, represent 5% of a segment's total employment. All values represent averages across the period 1999–2007. The number provided in parentheses represents the simulated null-model *STRUC*(5%) value.

Stylized fact 2 is a first indication that human capital is highly specific to a worker's job: most potential industry destinations are all but completely rejected by workers. In the next section, we study to what extent this pickiness of workers in choosing a new industry can be interpreted as an expression of industry-specific human capital.

4.3. The skill-relatedness structure of labor-flow matrices

So far, we have documented patterns in raw labor flows. However, the size of labor flows will depend on the sizes and flow rates (i.e., the fraction of employees switching jobs) of the industries involved. To isolate the structure underlying inter-industry labor flows, we calculate the ratio between the observed volume of labor flows, and the one that would be expected from industries flow rates. If workers switched industries with probabilities proportional to the total outflow of the industry of origin, $F^s_{i.}$, and the total inflow into the destination industry, $F^s_{.j}$, the expected labor flow between i and j is given by $\hat{F}^s_{ij} = \frac{F^s_{i.}F^s_{.j}}{F^s_{..}}$ and the ratio of observed to expected flows by[21]:

$$R^s_{ij} = \frac{F^s_{ij}F^s_{..}}{F^s_{i.}F^s_{.j}} \tag{1}$$

Values for R^s_{ij} from 1 to ∞ indicate that labor flows are in excess of the random benchmark. Values between 0 and 1 indicate that labor flows are below this benchmark. Because the distribution of R^s_{ij} is strongly right-skew, we transform R^s_{ij} as follows:

$$\bar{R}^s_{ij} = \frac{R^s_{ij} - 1}{R^s_{ij} + 1}$$

which maps R^s_{ij} values between 0 and 1 onto the interval $[-1, 0]$ and values from 1 to ∞ onto the interval [0,1]. As a result, $\bar{R}^s_{ij}$ is symmetric with respect to zero.[22]

[21] Alternatively, we can derive a baseline expectation using the size of industries. That is, if we assume that workers switch from one industry to another industry with probabilities that are proportional to these industries' sizes, then the expected labor flow from industry i to j is $N^s_{.}\frac{F^s_{..}}{N^s_{.}}\frac{N^s_{i}}{N^s_{.}}\frac{N^s_{j}}{N^s_{.}-N^s_{i}}$. Relatedness is now calculated as: $R^s_{ij} = \frac{F^s_{ij}(N^s_{.}-N^s_{i})}{F^s_{..}N^s_{i}N^s_{j}}$ Eq. (1) yields R_{ij} values that have slightly higher year-on-year correlations, suggesting a somewhat higher consistency. In practice, however, both baselines give virtually the same results, with rank correlations of around 0.98. An intermediate solution is provided in Neffke and Henning (2013), who use the prediction from a regression of the size of the flow between two industries on the industries' sizes, wage levels and growth rates as a benchmark. Also such parametric methods yield skill-relatedness matrices that are all but indistinguishable from the ones we use here.

[22] In particular, observed flows that exceed expected flows by a factor a translate into $\bar{R}^s_{ij} = \frac{a-1}{a+1}$, whereas the opposite – expected flows exceeding observed flows by a factor a – yields $\bar{R}_{ij} = \frac{1/a-1}{1/a+1} = \frac{1-a}{1+a} = -\frac{a-1}{a+1}$.

In principle, flow patterns may reflect other factors than just an overlap in industries skill requirements. In fact, Campbell et al. (2012) discuss a number of constraints beyond pure human capital specificities that emerge on the demand as well as the supply side of labor mobility due to information asymmetries, job satisfaction and complementarities between human capital and other productive resources.[23] Following March and Simon (1958), Neffke and Henning (2013) argue that job-switching is affected by the desirability and (potentially subjective) availability of jobs. For instance, workers may find certain jobs desirable because they are well-aligned with their personal values. In this case, inter-industry labor flows will reflect industries value-, not skill-compatibility. Furthermore, because the perceived availability of job openings may depend on a worker's social network, inter-industry labor flows may in part reflect the industrial structure of social networks. Neffke and Henning list three reasons why they believe that inter-industry labor flows reflect predominantly an overlap in industries' skill requirements. First, even though value-compatibility and social networks may narrow down the set of jobs from which workers choose, ultimately "having the right skills would appear to be a sine qua non for gaining employment in another industry" (Neffke and Henning, 2013, p. 303). Second, social networks (and especially the parts thereof that help finding jobs) will to a considerable extent consist of former colleagues, business associates and classmates, i.e., of workers with similar skills. Therefore, the jobs found through social networks are more likely to be skill-related than not. Third, Neffke and Henning show that their skill-relatedness index correlates strongly with a direct measure of similarity in skill-use derived from a skills and tasks survey. Acknowledging that inter-industry labor flows may reflect more than just similarities in human capital requirements, we therefore follow Neffke and Henning (2013), and interpret these inter-industry labor flows as an expression of skill similarities. Therefore, we refer – in part as a shorthand notation – to $\bar{R}_{ij}^{s}$ as the *skill relatedness* of i to j in segment s and call i and j *skill related* if $\bar{R}_{ij}^{s} > 0$.

Fig. 2 shows two visualizations of the $\bar{R}$-matrix for the German labor market as a whole. Fig. 2a shows a heat map for skill-relatedness estimates among all possible 5-digit industry pairs, with rows and columns sorted by an average-linkages hierarchical clustering algorithm. The dark squares along the figure's diagonal indicate that the matrix exhibits a fair degree of clustering, i.e., certain sets of industries are densely connected among each other. However, there are also links across these clusters, as evidenced by various dark off-diagonal areas.

To get an impression of which industries are connected, Fig. 2b shows the network spanned by just the top 651[24] values in the $\bar{R}$-matrix, instead of depicting the entire skill-relatedness matrix. In this figure, nodes represent 3-digit industries (colored according to the NACE sections to which they belong). The size of a node represents the corresponding industry's average employment in the period we study. The layout of the network is based on an algorithm that aims at grouping closely related industries together such that nodes that cluster in the graph generally correspond to sets of skill-related industries.

Industries tend to cluster by section. For instance, we find a cluster of dark-blue textiles and leather industries (center-left), a cluster of hotels and transport-related services (green, upper-right), and a large metals-and-electronics cluster (blue, top-left). At the same time, there are large labor flows between industries with different colors, i.e., of different industry sections. However, these links across sections often connect industries that are intuitively related. For instance, high-technology manufacturing and service industries are connected in a cluster that links the *computer* and *telecommunications equipment* industries of the blue manufacturing section with the *software consulting* and *data processing* industries of the orange business-services section. Similarly, we find strong links between the manufacturing industries of *printing* and *publishing* and creative services like *radio & TV* and *advertising*.

4.4. Comparing skill relatedness across labor-market segments

Does the network in Fig. 2 depict a general structure or does skill relatedness differ by labor-market segment? To answer this question, we need to compare different skill-relatedness matrices to one another. We do this by first stacking all columns of a skill-relatedness matrix into one long vector. Next, we calculate the correlations among such vectors for different segments.

The estimated correlations are surprisingly low, typically between 0.3 and 0.5. However, even if we compare the skill-relatedness matrix for the same segment in two consecutive years, correlations rarely exceed the 0.5 mark. Although skill relatedness may change in the long run with shifts in technologies, it is implausible that it changes much on such a short horizon. This suggests that skill-relatedness matrices are estimated with a substantial amount of noise. Indeed, if we assume that skill relatedness does not change from one year to the next, a year-on-year correlation of 0.5 suggests that 75% $\left(1-0.5^2\right)$ of the variance in the estimated skill relatedness can be attributed to random noise. As a consequence, correlations with skill relatedness will exhibit a strong attenuation bias, i.e., they will be biased towards zero.

To resolve this, let skill-relatedness estimates be composed of two components: the (unobserved) actual skill relatedness, and measurement error. Formally, we write the stacked vector of skill-relatedness estimates for segment s in year t, $\hat{r}_{st}$, as the sum of a time-invariant, real skill-relatedness vector, r_s, and a year-specific measurement error component, ε_{st}:

$$\hat{r}_{st} = r_s + \epsilon_{st} \tag{2}$$

[23] Campbell et al. (2012) explore this issue in the context of the competitive advantage firms may acquire from their employees.

[24] We display three times as many links as nodes, which, as a rule of thumb, yields networks that are not overly cluttered. For the same reason, we aggregate data to the 3-digit level.

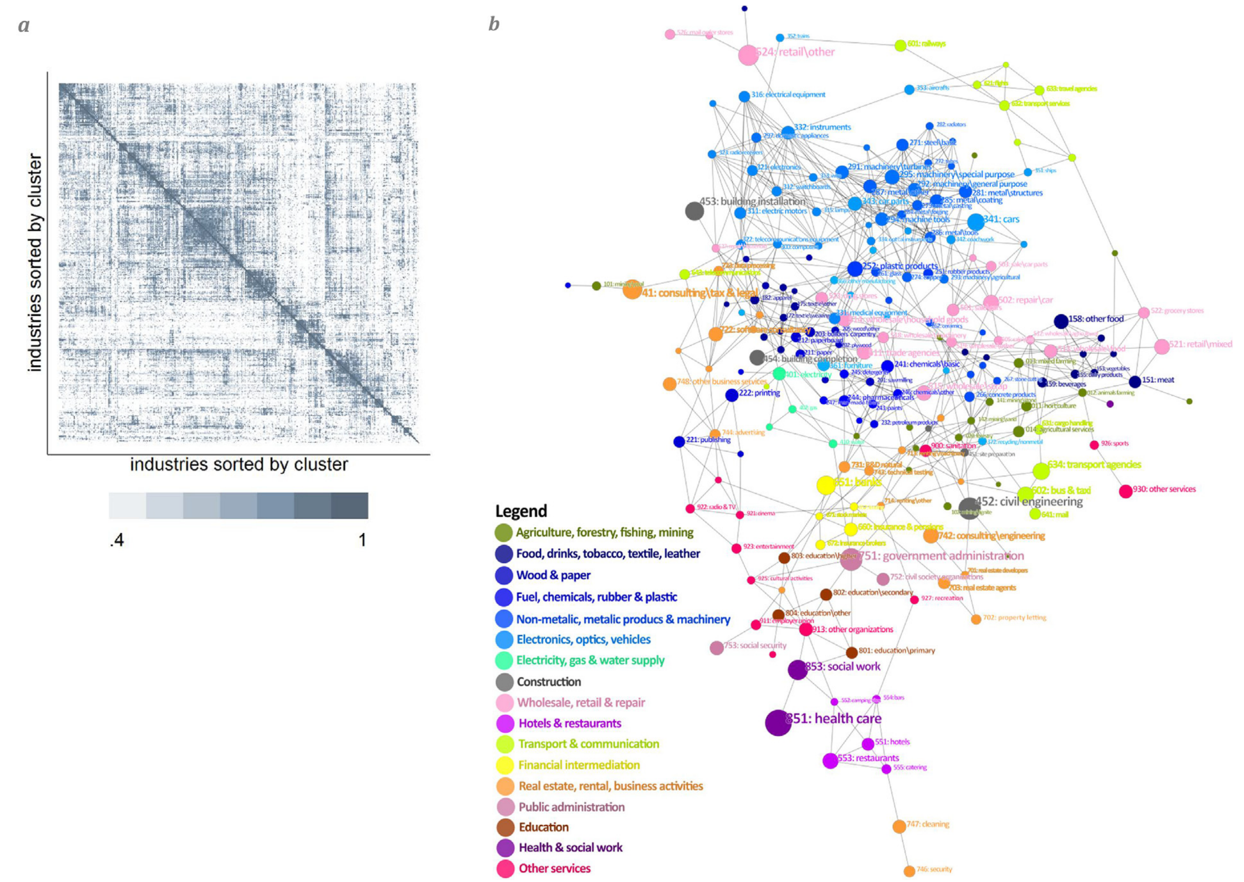

Fig. 2. (a) Calustering of inter-industry linkages in Germany (1999–2008). The shading in the figure represents the average symmetrized skill-relatedness matrix for Germany at the 5-digit level using all yearly flows between 1999 and 2008. Rows and columns are sorted according to an average-linkages hierarchical clustering algorithm (b) Network depiction of inter-industry linkages in Germany (1999–2008). The network depicts the strongest 651 links among 3-digit industries in the symmetrized average skill-relatedness matrix for Germany as a whole, using all yearly flows between 1999 and 2008. The layout is based on the organic layout procedure in the Cytoscape software, manually adjusted to increase the clarity of the graph by minimizing edge crossings. Labels are omitted for small industries. Color codes represent subsections in the NACE classification.

If we assume that ε_{st} is distributed identically and independently, Eq. (2) implies that the bias in the correlation between two observed skill-relatedness vectors, $\hat{r}_{st}$ and $\hat{r}_{s't}$, can be reduced, albeit – given that we only have nine yearly estimates – not eliminated, by averaging the skill-relatedness estimates of different years. Correlations between time-averaged skill-relatedness vectors, therefore, represent, lower bounds of the actual correlations of skill relatedness.

An alternative approach is to tackle the measurement error directly. This is possible, because our yearly $\bar{R}^s$ matrices, in principle, represent nine independent estimates of skill relatedness.[25] Consequently, we can use the bias-correction method introduced by Spearman (1904):

$$Corr[r_s, r_{s'}] = \frac{Corr[\hat{r}_{st}, \hat{r}_{s't}]}{\sqrt{Corr[\hat{r}_{st}, \hat{r}_{st+1}]}\sqrt{Corr[\hat{r}_{s't}, \hat{r}_{s't+1}]}}. \tag{3}$$

That is, the true correlation between skill-relatedness vectors for segments s and s' can be estimated by dividing the correlation between observed skill-relatedness vectors by the square root of the correlations between two consecutive skill-relatedness measurements in each segment. Detailed derivations are provided in Appendix D in Supplementary material.[26]

[25] Because we observe each worker only once a year, a worker who moves out of an industry in year t cannot undertake that same move again in year $t+1$. Consequently, moves within a given industry pair in consecutive years are necessarily composed of disjoint worker sets. The assumption that yearly relatedness estimates are uncorrelated is still quite strong. For instance, workers may imitate coworkers' previous moves and shocks that make an industry switch attractive in one year may be autocorrelated. However, in that case, the error term should exhibit positive autocorrelation, which would inflate the denominator in bias-correction Eq. (2) and imply that, if at all, our estimates would *underestimate* true cross-segment correlations.

[26] Appendix D in Supplementary material also provides outcomes using a third bias-correction method. This approach combines the two methods described above and yields very similar results as the bias-correction method of Eq. (3).

Table 3
Correlations of skill-relatedness estimates of local versus long-distance flows.

	local	long-distance
own correlation	0.56	0.51
local	n.a./1.00	
long-distance	0.87/0.81	n.a./1.00

The table presents correlations between the skill-relatedness ($\bar{R}_{ij}$) matrices for the combination of labor-market segments given in the rows and columns. The first row of the table provides the average correlation between skill-relatedness matrices for two consecutive years of the labor-market segment in the columns. In the remaining cells, the top row represents the bias-corrected correlation using Eq. (3) and the bottom row the correlation between relatedness matrices averaged over all available years. Industries have been aggregated to the 3-digit level. Local flows are flows taking place over a road-distance of below 100 km, flows beyond this distance are labelled long-distance.

Table 4
Correlations of skill-relatedness estimates by wage levels and region.

	high East	low East	high West	low West
own correlation	0.45	0.43	0.56	0.51
high East	n.a./1.00			
low East	0.94/0.80	n.a./1.00		
high West	0.79/0.72	0.74/0.69	n.a./1.00	
low West	0.75/0.67	0.79/0.73	0.93/0.84	n.a./1.00

The table presents correlations between the skill-relatedness ($\bar{R}_{ij}$) matrices for the combination of labor-market segments given in the rows and columns. The first row of the table provides the average correlation between skill-relatedness matrices for two consecutive years of the labor-market segment in the columns. In the remaining cells, the top row represents the bias-corrected correlation using Eq. (3) and the bottom row the correlation between relatedness matrices averaged over all available years. Industries have been aggregated to the 3-digit level. High (low) East (West) represents high-wage (low-wage) workers in eastern (western) Germany.

In the analyses below, we aggregate industries to the 3-digit level to strike a balance between using relatively homogeneous industries and maintaining a ratio of labor flows to industry pairs large enough to calculate precise skill-relatedness estimates. Tables 3–5 summarize, for each of the three segmentations described in Section 3, the correlations among skill-relatedness matrices of different labor market segments. The first row in these tables reports the correlation for a single segment's skill relatedness in two consecutive years.[27] The higher this year-on-year own-correlation, the less noisy estimates are. The remaining rows describe correlations between the skill-relatedness measurements in a given pair of segments.[28] The first value in a cell reports the average bias-corrected correlations using Eq. (3). To calculate the second value in these cells, we first average skill relatedness for the segments across years and then calculate the correlation of these averages. This provides a lower bound on the true correlation of two relatedness types.

One potential concern is that skill-relatedness does not measure similarities in skill requirements, but simply reflects industries co-location patterns. If this were the case, we would expect that labor flows within a region would be structurally different from those that cross into other regions. However, Table 3 shows that, at a bias-corrected correlation of 0.87, local and long-distance labor flows exhibit very similar skill-relatedness estimates. This suggests that the structure of the skill-relatedness network is not driven by industry co-location patterns.

Stylized fact 3: Given that local and long-distance flows exhibit very similar skill-relatedness structures, skill-relatedness estimates are not mainly driven by industry co-location patterns.

Table 4 shows that, in eastern as well as in western Germany, workers with different wage levels have almost identical skill-relatedness matrices: all bias-corrected correlations are well above 0.9 and even without bias correction, we observe correlations of at least 0.8. Although differences between eastern and western Germany are slightly larger, with bias-corrected correlation estimates typically above 0.75, also these differences are relatively small.

Table 5 compares skill relatedness in different occupational segments. With the exception of cleaning and security personnel, for whom bias-corrected correlations with the other occupational groupings rarely exceed the 0.5 mark, the different occupational groupings exhibit very similar skill-relatedness matrices.[29] Management occupations display skill-relatedness matrices that are almost identical to those of sales people (bias-corrected estimate: 0.92), accountants (0.86), office clerks (0.91) and, to a somewhat lesser extent, IT specialists (0.83). Indeed, even without correcting for measurement error, all

[27] We first calculate correlations for all eight pairs of consecutive years and then take the average.

[28] Because we can only estimate the relatedness among industries with at least some inflow or outflow in each of the two labor-market segments, the number of observations varies across the tables' cells.

[29] Apparently, cleaners and security guards' job transitions are quite distinct from those of the other groups. Interestingly however, with a bias-corrected correlation of 0.79, cleaners and security guards display very similar skill-relatedness matrices to one another. A closer inspection of their relatedness matrices suggests that even in these jobs, where skills are arguably not very industry specific, workers do not switch industries randomly. For instance, in both groups, we find strong connections among various construction industries.

Table 5
Correlations of skill-relatedness estimates for different occupations.

	managers	sales	accountants	office clerks	IT	security	cleaners
own correlation	0.42	0.43	0.29	0.44	0.33	0.34	0.27
managers	n.a./1.00						
sales	0.92/0.77	n.a./1.00					
accountants	0.86/0.69	0.80/0.64	n.a./1.00				
office clerks	0.91/0.78	0.89/0.76	0.78/0.62	n.a./1.00			
IT	0.83/0.69	0.78/0.63	0.89/0.68	0.76/0.60	n.a./1.00		
security	0.44/0.39	0.38/0.33	0.53/0.43	0.39/0.35	0.44/0.37	n.a./1.00	
cleaners	0.50/0.39	0.41/0.35	0.60/0.45	0.47/0.36	0.57/0.38	0.79/0.62	n.a./1.00

The table presents correlations between the skill-relatedness ($\bar{R}_{ij}$) matrices for the combination of labor-market segments given in the rows and columns. The first row of the table provides the average correlation between skill-relatedness matrices for two consecutive years of the labor-market segment in the columns. In the remaining cells, the top row represents the bias-corrected correlation using Eq. (3) and the bottom row the correlation between relatedness matrices averaged over all available years. Industries have been aggregated to the 3-digit level. Row and column labels refer to broad occupational groupings.

Table 6
Changes in skill relatedness between 1999 and 2008.

segment	stability
Germany	
all wages	0.93
high wages	0.93
low wages	0.93
East/West Germany	
East	0.90
West	0.93
Geography of flows	
local	0.93
long-distance	0.95
Occupations	
managers	0.94
sales	0.97
accountants	0.92
office clerks	0.95
IT	0.93
cleaners	0.91
security	0.93

Stability is the attenuation-bias-corrected correlation between skill relatedness in year 1999 and 2007 of a given labor-market segment. Measurement error variances are assessed by the correlations between skill-relatedness estimates for 1999/00 and 2000/01 and for 2006/07 and 2007/08 of the corresponding labor-market segment.

correlations in the first five occupational segments exceed 0.6. This stability across occupations is remarkable. Even though managers and IT specialists will carry out different tasks, they tend to switch jobs among the same industries. This contrasts sharply to the findings of Kambourov and Manovskii (2009) and Poletaev and Robinson (2008), who find that industry-specific work experience is scarcely important after taking occupational work experience into account. Contrary to this, the fact that inter-industry flow patterns are similar for different occupations suggests that human capital *does* have an industry-specific component that is independent of a worker's occupation.

Taking together the results in Tables 4 and 5, we arrive at the following stylized fact:

Stylized fact 4: Workers with different levels and types of skills have similar skill-relatedness networks, i.e., the same industries are connected by labor flows, regardless of the skills of workers involved.

Using the bias-correction method of Eq. (3), we can also assess the extent to which skill relatedness changes over time. The bias-corrected correlation between skill relatedness in the first and last year of our data set is 0.9 or higher in all segments (Table 6).[30] This shows that, if at all, skill relatedness changes very slowly over time.

Stylized fact 5: There is no indication of rapid change in skill relatedness.

[30] The measurement-error correction is in this case based on Eq. (3) and uses the year-on-year same-segment correlations between 1999/00 and 2000/01 and between 2006/07 and 2007/08 in the denominator.

4.5. Skill relatedness and the growth of local industries

In this subsection, we test the predictive validity of skill-relatedness in comparison to alternative relatedness measures. As explained in Section 2, a good context for such a test is regional economic development. In particular, we assess to what extent the presence of related industries predicts the growth and entry of local industries in German planning regions (*Raumordnungsregionen*). We define related industries on the basis of three different relatedness measures: skill relatedness, a co-location-based measure as in Porter (2003) and Hidalgo et al. (2007), and an input-output-based measure. All three measures are based on data for the period before 2003 and used to predict growth patterns in the period 2003–2008.

As a measure of skill relatedness, we take the average skill-relatedness for Germany as a whole across 1999/00, 2000/01 and 2001/02. Next, we symmetrize the resulting matrix and rescale it to the interval [0,1]:

$$SR_{ij} = \frac{1}{2}(\frac{\bar{R}_{ij}+\bar{R}_{ji}}{2}+1)$$

Co-location-based relatedness is calculated as the correlation between the regional employment vectors of two industries:

$$CL_{ij} = \frac{1+Corr\left(E_{ir},E_{jr}\right)}{2} \tag{4}$$

where E_{ir} represents the employment in industry i and region r in the year 2002. The transformation in (4) maps the correlation onto the interval [0,1]. CL_{ij} increases as the geographical distribution of employment in industries i and j becomes more similar.

For input-output relatedness, we use the German 2-digit input-output matrix of the year 2003 as provided by EUROSTAT.[31] The input-output matrix records for each pair of industries (i, j) the value of industry i's outputs purchased by industry j, V_{ij}. To establish the strength of input-output linkages between two industries, we express V_{ij} once as a percentage of all intermediate inputs into industry j and once as a percentage of all intermediate outputs of industry i. We do the same for the reverse value-flow, V_{ji}. Input-output relatedness is now defined as the average of these four figures:

$$IO_{ij} = \frac{1}{4}\left(\frac{V_{ij}}{V_{i.}}+\frac{V_{ij}}{V_{.j}}+\frac{V_{ji}}{V_{j.}}+\frac{V_{ji}}{V_{.i}}\right)$$

where a dot (".") once more denotes summation over the omitted dimension. For all relatedness measures, we define industries to be unrelated to themselves: $SR_{ii} = CL_{ii} = IO_{ii} \equiv 0$.

We use these relatedness measures to quantify how well a local industry's related industries are represented in the region. To do so, we calculate for each industry i the weighted average employment of all other industries, $j \neq i$, in the region, using the relatedness between i and j as weights:

$$E_{irt}^{REL} = \sum_{j \neq i} \frac{REL_{ij}}{\Sigma_{m \neq i} REL_{im}} E_{jrt}$$

where E_{jrt} is industry j's employment in region r and year t and REL is either SR, CL or IO. Next, we use these variables to predict annualized growth rates of existing local industries and the entry of new industries in a region. These regressions are limited to local industries in the traded, non-resource-based, private sector.[32] Moreover, for the growth regressions, we only use industries that exist at the start of the period, estimating (cross-sectional) regression equations of the following type:

$$\log\left(\frac{E_{irt}}{E_{irt+\tau}}\right)^{\frac{1}{\tau}} = \gamma\log(E_{irt}) + \beta_{SR}\log\left(E_{irt}^{SR}\right) + \beta_{CL}\log\left(E_{irt}^{CL}\right) + \beta_{IO}\log\left(E_{irt}^{IO}\right) + \eta_i + \rho_r + \varepsilon_{irt} \tag{5}$$

where γ estimates potential mean-reversion effects and η_i and ρ_r represent industry- and region-specific effects respectively. Furthermore, t is the base year of the growth rate and τ represents the time horizon over which growth is measured such that the dependent variable reflects annualized growth.

To analyze the entry of new local industries, we run Linear Probability Models (LPMs).[33] That is, our dependent variable is a dummy variable $ENTRY_{irt,\tau}$ that has a value of 1 if an industry i that did not yet exist in year t enters region r within the next τ years:

$$ENTRY_{irt,\tau} = \beta_{SR}\log\left(E_{irt}^{SR}\right) + \beta_{CL}\log\left(E_{irt}^{CL}\right) + \beta_{IO}\log\left(E_{irt}^{IO}\right) + \eta_i + \rho_r + \varepsilon_{irt} \tag{6}$$

Like Eq. (5), Eq. (6) describes a cross-sectional analysis with industry and region fixed effects.

Table 7a
Local industry growth and entry regressions (3-digit industries).

	(1) growth 2003–2004	(2) growth 2003–2008	(3) entry 2003–2004	(4) entry 2003–2008
log(Eirt)	−0.035***	−0.030***		
	(0.0036)	(0.0015)		
log(Eirt_SR)	0.171***	0.158***	0.074*	−0.061
	(0.0416)	(0.0169)	(0.0399)	(0.0705)
log(Eirt_CL)	0.126	0.082	−0.152	1.448*
	(0.1911)	(0.0854)	(0.5325)	(0.8588)
log(Eirt_IO)	0.004	−0.016	−0.003	−0.077
	(0.0234)	(0.0100)	(0.0264)	(0.0538)
R-squared	0.054	0.175	0.226	0.253
N	12,408	12,168	2046	2046

*** $p<0.01$, ** $p<0.05$, * $p<0.10$. Robust standard errors in parentheses. All models include region and industry fixed effects. Regional units are defined as Germany's planning regions and industries are aggregated at the 3-digit level, excluding industries in agriculture, fishing and mining (NACE 01-14), non-traded industries (NACE 40–59) and public sector industries (NACE 75–99). Columns (1) and (2) report OLS regressions of the logarithm of annualized growth rates over a one- and over a five-year period. Columns (3) and (4) report the outcomes of OLS regressions with as a dependent variable a dummy variable that evaluates whether or not a new local industry enters a region within a one- or a five-year period.

Table 7b
Local industry growth and entry regressions (4-digit industries).

	(1) growth 2003–2004	(2) growth 2003–2008	(3) entry 2003–2004	(4) entry 2003–2008
log(Eirt)	−0.043***	−0.033***		
	(0.0028)	(0.0011)		
log(Eirt_SR)	0.200***	0.146***	0.051***	0.104***
	(0.0333)	(0.0120)	(0.0140)	(0.0242)
log(Eirt_CL)	−0.041	0.035	0.067	−0.195
	(0.2279)	(0.0826)	(0.2409)	(0.3628)
log(Eirt_IO)	0.042**	0.005	0.015	−0.019
	(0.0192)	(0.0077)	(0.0127)	(0.0191)
R-squared	0.082	0.201	0.153	0.207
N	21,855	21,050	9309	9309

Robust standard errors in parentheses. All models include region and industry fixed effects. Regional units are defined as Germany's planning regions and industries are aggregated at the 4-digit level, excluding industries in agriculture, fishing and mining (NACE 01-14), non-traded industries (NACE 40–59) and public sector industries (NACE 75–99). Columns (1) and (2) report OLS regressions of the logarithm of annualized growth rates over a one- and over a five-year period. Columns (3) and (4) report the outcomes of OLS regressions with as a dependent variable a dummy variable that evaluates whether or not a new local industry enters a region within a one- or a five-year period.

Tables 7a and 7b report estimates for (5) and (6) with base years $t = 2003$ and growth and entry defined over a one-year and over a five-year period.[34] Table 7a aggregates industries to the 3-digit level, whereas Table 7b presents estimates for 4-digit industries. In all but one model, only skill-related employment has a statistically significant (positive) coefficient at the 5% level. The estimated coefficients suggest that a 10% rise in skill-weighted average employment is associated with a between 1.5% and 2% increase in growth rate and between 0.5 ($0.051 * \ln(1.1)$) and 1.0 ($0.104 * \ln(1.1)$) percentage points higher entry rates. This shows that industries tend to enter and grow faster in regions with large amounts of skill-related employment. In contrast, controlling for skill-related employment, employment that is related according to co-location or input linkages typically does not show any association with growth and entry rates.

This is not to say that co-location and input-output relatedness indicators have no merit. For one, the log $\left(E_{irt}^{REL}\right)$ terms are highly collinear.[35] When running separate regressions for each relatedness type, all exhibit some explanatory power. For another, if we measure E_{irt}^{REL} not as a relatedness-weighted average of local employment, but rather as the sum of local employment in related industries (where industries are considered related if the relatedness index exceeds a certain threshold), co-location-based relatedness (but not input-output based relatedness) gains importance. However, also in these

[31] To be precise, we use the domestic section of the product-based symmetric input-output table, ignoring international trade and sales to end-consumers.
[32] That is, we exclude industries in the 2-digit NACE classes 01-14 (agriculture, fishing and mining), 40–59 (non-traded services), and 75–99 (public sector and miscellaneous industries).
[33] The use of LPMs allows controlling for the same full set of region and industry fixed effects as used in the growth regressions. In contrast, nonlinear models such as probit or logit models cannot accommodate such a large number of regressors.
[34] Results are qualitatively similar for other base years and time windows (results available on request).
[35] This is not due to high correlations among inter-industry relatedness measures. Indeed, at the 2-digit level, the correlation is 0.50 between SR_{ij} and IO_{ij}, 0.36 between SR_{ij} and CL_{ij} 0.36, and 0.40 between IO_{ij} and CL_{ij}. At the 4-digit level, these correlations drop to 0.29, 0.31 and 0.19, respectively. However, at the industry-region level, correlations among log $\left(E_{irt}^{SR}\right)$, log $\left(E_{irt}^{IO}\right)$ and log $\left(E_{irt}^{CL}\right)$ run from a maximum of 0.98 at the 2-digit level to 0.94 at the 4-digit level. Indeed, at the 2-digit level, where correlations are highest and the number of observations is lowest, multicollinearity issues become unsurmountable.

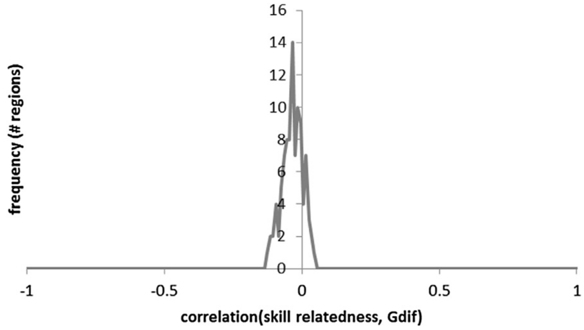

Fig. 3. Regional labor reallocation: correlation($\bar{R}_{ij}$,$Gdif_{ij}$) by region.
Histogram of region-specific rank correlations between skill relatedness and the absolute difference in local growth rates within an industry pair.

regressions, skill-related employment tends to remain highly significant.[36] Overall, we therefore conclude that the predictive validity of skill-relatedness vis-à-vis co-location-based and input-output relatedness is strong:

Stylized fact 6: Skill relatedness is a stronger predictor of industries' regional growth and entry rates than input-output and co-location-based relatedness.

4.6. Skill relatedness and reallocation frictions

The documented skill-relatedness patterns can be thought of as expressions of constraints to labor mobility. Such constraints may hinder the efficient reallocation of workers from declining to expanding industries, which may affect a region's resilience to adverse shocks. Indeed, just as a lack of *geographical* labor mobility makes it harder for economies to deal with asymmetric regional shocks,[37] the constrained mobility across *industries* raises similar concerns. However, the degree to which such mobility constraints cause reallocation frictions depends on whether or not related industries experience correlated shocks. That is, only if skill-related industries have strongly correlated growth patterns will it be hard to reallocate workers in skill-preserving ways.

To assess to what extent this is the case, we need to explore whether the differences in growth rates of two industries are associated with their skill relatedness. To do so, we first calculate skill relatedness among 3-digit industries, using only flows in the period 1999/00 to 2002/03. For the remaining years, 2004–2008, we calculate the absolute difference in growth rates for each pair of industries:

$$Gdif_{ij} = \left|\frac{E_{i,2008}}{E_{i,2004}} - \frac{E_{j,2008}}{E_{j,2004}}\right|$$

When industries i and j grow at exactly the same rate, $Gdif_{ij}$ is zero, whereas $Gdif_{ij}$ increases as the growth rates of i and j diverge. The Spearman rank correlation between $\bar{R}_{ij}$ and $Gdif_{ij}$ of -0.17 is negative yet small. Accordingly, skill-related industries tend to have only weakly similar growth rates, which should limit the problems of reallocating labor from shrinking to growing industries. Indeed, in Appendix E in Supplementary material, we show that if workers in industries with labor redundancies are reallocated to the most skill-related industries with labor shortages, this reallocation can be achieved in skill-preserving ways.

In spite of this finding, reallocation bottlenecks may still exist in individual regions. Therefore, we repeat these analyses, but define $Gdif_{ij}$ in terms of industries' growth rates within a planning region. Fig. 3 shows the histogram for the Spearman rank correlations between these regional growth differentials and skill relatedness.

At the regional level, reallocation problems are even less pronounced than for Germany as a whole. For some regions rank correlations are even positive and the lowest observed rank correlation in a region is -0.126. These findings suggest the following stylized fact:

Stylized fact 7: Skill-related industries typically do not exhibit highly similar growth rates. Therefore, in spite of inter-industry labor flows being highly constrained, labor surpluses of shrinking industries can typically be absorbed by growing industries in a skill-preserving way.

[36] Results are available upon request.
[37] See for instance, Siebert (1997) and Bentivogli and Pagano (1999) for a discussion of limited labor mobility among European regions in the context of Europe's Economic and Monetary Union as an optimal currency area.

5. Conclusions and future research

Our analyses of inter-industry labor flows in the German economy have yielded the following seven stylized facts. First, workers often switch jobs between industries that belong to different (highly aggregated) sections of the industry classification system (Stylized Fact 1). Although this fact may be taken as an indication that workers can change industries relatively freely and *a fortiori* that human capital cannot be particularly industry specific, a closer inspection reveals that industry switches are far from random. On the contrary, most labor flows take place within a narrow set of industry pairs (Stylized Fact 2): on average, 62% of job switchers move to industries that represent just 5% of total employment in Germany. Furthermore, after controlling for the overall flow rates of industries, the underlying structure of inter-industry labor flows hardly changes over time (Stylized Fact 5) and does not just reflect industries' co-location patterns (Stylized Fact 3). Moreover, job switchers with different wages and occupations follow remarkably similar flow-patterns (Stylized Fact 4).

These findings suggest that inter-industry labor flows reveal how industries are connected to one another in terms of their human-capital requirements. In line with recent literature, we therefore call labor flows between two industries that exceed a well-defined random benchmark *skill related*. Following the literature on related diversification in economic geography, we test whether this labor-flow-based skill-relatedness measure predicts local industries entry and growth rates. These analyses show that the skill-relatedness index yields better growth predictions than inter-industry relatedness measures based on co-location patterns or input-output relations (Stylized Fact 6). However, because skill-related industries don't exhibit strongly correlated growth patterns, the skill-relatedness structure governing inter-industry labor flows doesn't present major obstacles in the reallocation of workers from shrinking to growing industries (Stylized Fact 7).

Overall, we believe that these findings show that inter-industry labor flows are at present understudied and merit further scholarly attention. For instance, skill-relatedness matrices can be further exploited in research where inter-industry linkages plays a role, ranging from labor economics (Poletaev and Robinson, 2008; Gathmann and Schönberg, 2010) and economic geography (Porter, 2003; Neffke et al., 2011; Boschma et al., 2013) to development economics (Hidalgo et al., 2007) and strategic management (e.g., Teece et al., 1994; Farjoun, 1994; Bryce and Winter, 2009; Lien and Klein, 2009). Moreover, skill-relatedness matrices may find various policy applications, from employment and retraining programs to cluster policy. Furthermore, we have shown that analyzing inter-industry flows offers new ways to study labor markets and the flexibility of a labor force. Finally, although skill-relatedness does not change much in the relatively short period we study, if technological progress affects industries' skill requirements we expect more drastic changes over longer time horizons. Labor-flow matrices may, therefore, offer new ways to analyze technological change. We hope that making the tools and skill-relatedness matrices in this paper available online[38] will facilitate progress on these and other topics.

Acknowledgments

We thank David Autor, Alex Coad, Koen Frenken, Andres Gomez, Ricardo Hausmann, César Hidalgo, Paul Novosad, Stuart Russell, Muhammed Yildirim and the participants of the Workshop on Industry relatedness and Regional Qualitative Change in Aalborg for valuable comments. Frank Neffke is grateful to the MasterCard Center for Inclusive Growth & Financial Inclusion for their financial support.

Appendix Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2017.07.003.

References

Agrawal, A., Cockburn, I., McHale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. J. Econ. Geogr. 6 (5), 571–591.

Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks. Manage. Sci. 45 (7), 905–917.

Becker, Gary S., 1964. Human Capital Chicago. University of Chicago Press.

Bender, S., Möller, J., 2010. Data from the federal employment agency. In: German Data Forum, Rat für Sozial- und Wirtschaftsdaten (Eds.), Building on Progress. Expanding the Research Infrastructure for the Social, Economic, and Behavioral Sciences, Vol. 2. Budrich UniPress, Opladen, pp. 943–958.

Bentivogli, C., Pagano, P., 1999. Regional disparities and labour mobility: the Euro-11 versus the USA. Labour 13 (3), 737–760.

Boschma, R., Minondo, A., Navarro, M., 2013. The emergence of new industries at the regional level in Spain. A proximity approach based on product-relatedness. Econ. Geogr. 89 (1), 29–51.

Boschma, R., Eriksson, R., Lindgren, U., 2014. Labour market externalities and regional growth in Sweden: the importance of labour mobility between skill-related industries. Reg. Stud. 48 (10), 1669–1690.

Breschi, S., Lissoni, F., 2005. Cross-firm inven–tors and social networks: localised knowledge spillovers revisited. Annales d'Economie et de Statistique 79–80, 189–209.

Bryce, D.J., Winter, S.G., 2009. A general inter-industry relatedness index. Manage. Sci. 55, 1570–1585.

Campbell, B.A., Coff, R., Kryscynski, D., 2012. Rethinking sustained competitive advantage from human capital. Acad. Manage. Rev. 37 (3), 376–395.

[38] A link to skill-relatedness estimates for all different classification systems in use in Germany between 1975 and 2014 can be found on the first author's website.

Cantner, U., Graf, H., 2006. The network of innovators in Jena: an application of social network analysis. Res. Policy 35 (4), 463–480.
Casper, S., 2007. How do technology clusters emerge and become suitable? Social network formation and inter-firm mobility within the San Diego biotechnology cluster. Res. Policy 16, 438–455.
Costa, C., Baptista, R., 2011. Agglomeration versus organizational reproduction: the molds cluster in Portugal. Front. Entrepreneurship Res. 31 (4).
Dauth, W., 2010. The Mysteries of the Trade: Employment Effects of Urban Interindustry Spillovers. IAB Discussion Paper 15/2010.
Davis, S.J., Haltiwanger, J., 1999. Gross job flows. Handb. Labor Econ. 3, 2711–2805.
Davis, S.J., Faberman, R.J., Haltiwanger, J., 2006. The flow approach to labor markets: new data sources and micro–macro links. J. Econ. Perspect. 20 (3), 3–26.
Delgado, M., Porter, M.E., Stern, S., 2010. Clusters and entrepreneurship. J. Econ. Geogr. 10 (4).
Delgado, M., Porter, M.E., Stern, S., 2014. Cluster, convergence, and economic performance. Res. Policy 20 (4), 1785–1799.
Diodato, D., Weterings, A., 2015. The resilience of regional labour markets to economic shocks: exploring the role of interactions among firms and workers. J. Econ. Geogr. 15 (4), 723–742.
Essletzbichler, J., 2013. Relatedness, industrial branching and technological cohesion in US metropolitan areas. Reg. Stud., http://dx.doi.org/10.1080/00343404.2013.806793.
Farber, H.S., 1999. Mobility and stability: the dynamics of job change in labor markets. Handb. Labor Econ. 3, 2439–2483.
Farjoun, M., 1994. Beyond industry boundaries: human expertise, diversification and resource-related industry groups. Organiz. Sci. 5 (2), 185–199.
Gartner, H., 2005. The imputation of wages above the contribution limit with the German IAB employment sample. In: FDZ-Methodenreport (02/2005), Nürnberg.
Gathmann, C., Schönberg, U., 2010. How general is human capital? A task-based approach. J. Labor Econ. 28, 1–49.
Grant, R.M., Spender, J.-C., 1996. Knowledge and the firm: overview. Strateg. Manage. J. 17, 5–9.
Grant, R.M., 1996. Toward a knowledge-based theory of the firm. Strateg. Manage. J. 17, 109–122.
Greenstone, M., Hornbeck, R., Moretti, E., 2010. Identifying agglomeration spillovers: evidence from winners and losers of large plant openings. J. Political Econ. 118 (3), 536–598.
Hausmann, R., Hidalgo, C., Stock, D.P., Yildirim, M.A., 2014. Implied Comparative Advantage. Mimeo.
Hethey, T., Schmieder, J.F., Using worker flows in the analysis of establishment turnover. Evidence from German Administrative Data. (FDZ Methodenreport, 06/2010), Nürnberg.
Hidalgo, C.A., Klinger, B., Barabási, A.-L., Hausmann, R., 2007. The product space conditions the development of nations. Science 317, 482–487.
Kambourov, G., Manovskii, I., 2008. Rising occupational and industry mobility in the United States: 1968–97. Int. Econ. Rev. 49 (1), 41–79.
Kambourov, G., Manovskii, I., 2009. Occupational specificity of human capital. Int. Econ. Rev. 50 (1), 63–115.
Kaplan, D.S., Lederman, D., Robertson, R., 2011. Employment and Wage Responses to Trade Shocks: Evidence from Mexico During the 2008–09 US Recession , mimeo.
Lazear, E.P., 2009. Firm-specific human capital: a skill-weights approach. J. Political Econ. 117 (5), 914–940.
Lien, L.B., Klein, P.G., 2009. Using competition to measure relatedness. J. Manage. 35 (4), 1078–1107.
March, J.G., Simon, H.A., 1958. Organizations. Wiley, New York.
March, J.G., 1991. Exploration and exploitation in organizational learning. Organiz. Sci. 2 (1), 71–87.
Neal, D., 1995. Industry-specific human capital: evidence from displaced workers. J. Labor Econ. 13, 653–677.
Neffke, F.M.H., Henning, M., 2013. Skill-relatedness and firm diversification. Strateg. Manage. J. 34 (3), 297–316.
Neffke, F.M.H., Henning, M., Boschma, R.A., 2011. How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Econ. Geogr. 87 (3), 237–265.
Parent, D., 2000. Industry-specific capital and the wage profile: evidence from the national longitudinal survey of youth and the panel of income dynamics. J. Labor Econ. 18, 306–323.
Parrado, E., Caner, A., Wolff, E.N., 2007. Occupational and industrial mobility in the United States. Labour Econ. 14 (3), 435–455.
Penrose, E., 1959. The Theory of the Growth of the Firm. John Wiley & Sons, New York.
Poletaev, M., Robinson, C., 2008. Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys. J. Labour Econ. 26 (3), 387–420.
Porter, M.E., 1987. From competitive advantage to corporate strategy? Harv. Bus. Rev. 65 (3), 43–59.
Porter, M.E., 1998. Clusters and the new economics of competition. Harv. Bus. Rev. 76 (November–December), 77–90.
Porter, M.E., 2003. The economic performance of regions. Reg. Stud. 37 (6–7), 549–578.
Rigby, D.L., 2015. Technological relatedness and knowledge space: entry and exit of U.S. cities from patent classes. Reg. Stud. 49 (11), 1922–1937.
Robins, J., Wiersema, M.F., 1995. A resource-based approach to the multibusiness firm: empirical analysis of portfolio interrelationships and corporate financial performance. Strateg. Manage. J. 16 (4), 277–299.
Saxenian, A., 2007. The New Argonauts: Regional Advantage in a Global Economy. Harvard University Press.
Siebert, H., 1997. Labor market rigidities: at the root of unemployment in Europe. J. Econ. Perspect. 11 (3), 37–54.
Simon, H., 1991. Bounded rationality and organizational learning. Organiz. Sci. 2, 125–134.
Spearman, C., 1904. The proof and measurement of association between two things. Am. J. Psychol. 15 (1), 72–101.
Storz, C., Riboldazzi, F., Moritz, J., 2015. Mobility and innovation: a cross-country comparison in the video games industry. Res. Policy 44 (2015), 121–137.
Sullivan, P., 2010. Empirical evidence on occupation and industry specific human capital. Labour Econ. 17 (3), 567–580.
Teece, D.J., Rumelt, R., Dosi, G., Winter, S., 1994. Understanding corporate coherence: theory and evidence. J. Econ. Behav. Organiz. 23, 1–30.
Timmermans, B., Boschma, R., 2014. The effect of intra- and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. J. Econ. Geogr. 14 (2), 289-311.
Winkelmann, U., 2010. Manche pendeln weit –Berufspendler im bundesländervergleich statistisches. Monatsheft Baden-Württemberg 4, 40–43.

Further reading

Spearman, C. (1910). Correlation calculated from faulty data, *Br. J. Psychol.* **3** (3), 1910, 271–295.