Review of Accounting Studies
https://doi.org/10.1007/s11142-024-09821-z

Labor market peer firms: understanding firms' labor market linkages through employees' internet "also viewed" firms

Nan Li[1]

Accepted: 26 January 2024

Abstract
This paper studies the grouping of firms based on their labor-market connections, a significant departure from the traditional approach of grouping based on product-market connections. It also proposes a measure of labor market peers by analyzing the "also viewed" companies on two major online labor market platforms, LinkedIn and Glassdoor. Using the labor market peer measure, I examine whether firms that hire employees with similar skills and that are presumably exposed to the same labor-related risks and shocks exhibit a strong comovement of stock returns and accounting-based performance variables. I find that labor market peers overlap but differ from traditional product-market-based industry groupings, have significant incremental power to explain stock return and accounting-based performance measure comovements, and outperform traditional industry groupings in explaining return and wage comovements when a base firm shares more labor skills with its peers. Overall, the study highlights that labor market peers capture fundamental linkages between firms that are challenging to identify using traditional industry measures.

Keywords Peer firms · Labor market · Industry classification · Online search · LinkedIn · Glassdoor · Benchmarking

JEL Classification D83 · G0 · J01 · M2

> *Flying on the Delta Shuttle with Bill Gates 12 years ago, I asked, " What Microsoft competitor worries you most?"*
> *"Goldman Sachs." I gave Gates a startled look. Was Microsoft about to try the investment banking business? " Software," he said, " is an IQ business. Microsoft must win the IQ war, or we won't have a future. I don't worry about Lotus or IBM, because the smartest guys would rather come to work for Microsoft. Our*

✉ Nan Li
nanjulie.li@utoronto.ca

[1] University of Toronto, 27 King's College Circle, Toronto, ON, Canada

Published online: 28 March 2024

competitors for IQ are investment banks such as Goldman Sachs and Morgan Stanley."[1]

1 Introduction

Understanding fundamental linkages between firms is central to economic and financial analyses. Most of this effort has focused on defining industries for benchmarking and on identifying groups of peers. Standard industry groupings are primarily based on product-market relations and measure output market linkages but not necessarily input market linkages. Labor inputs play an increasingly important role in driving firm performance and economic growth (Becker 1962). Although such firms as Microsoft and Goldman Sachs hire similar types of labor and face common input risks, their linkages may be overlooked because they are not in the same product market. This paper proposes a new approach for identifying labor market peer firms by analyzing "also viewed" firms on online labor markets Glassdoor and LinkedIn. The results suggest that labor market peers capture economic linkages that are not captured by traditional industry groupings and have incremental power in explaining stock returns and accounting-based performance variable comovements. This suggests that labor market connections are economically significant in transmitting risks and shocks to firm performance and stock returns. Additionally, labor market peers are found to help explain the selection of peers for CEO compensation benchmarking, highlighting the importance of considering labor market connections in economic and financial analyses.

Why do labor market peers offer incremental power over standard industry groupings in identifying firms' economic linkages? First, these peers represent common human-capital risk and can help identify demand and supply shocks to a firm's talent pool that standard industry groupings cannot.[2] Limits on labor can raise wages and impede a firm's production of goods. For example, technology shocks to new firms likely increase the demand for certain types of employees, leading to a labor demand shock (Bresnahan et al. 2002). Therefore shareholders require a risk premium for the threat of losing key talent (Eisfeldt and Papanikolaou 2013). Changes in immigration policy (Bernstein et al. 2022) and increased occupational licensing requirements (Barrios 2022) can also reduce the supply of labor and increase wage premiums. Examining labor market peers can help identify these common shocks, even when the affected firms do not operate in the same product market.

Second, labor market peers can help capture shared production risks and shocks, as firms that hire similar types of employees may have similar production processes and inputs. For example, chemical manufacturers and oil-and-gas companies both use chemical processes to transform raw materials into products for intermediate or final consumption. These firms are exposed to common input price volatilities and techno-

[1] Talent Wars, Forbes, 2005.

[2] Human capital is a risk factor that impacts stock returns. Early research introduces the idea of nonmarketable assets (human capital) to the CAPM framework (Mayers 1973; Jaganathan and Wang 1996). Recent studies show how labor market frictions affect asset prices (Belo et al. 2014; Donangelo 2014; Kuehn et al. 2017).

logical advancements that may affect production costs. They may also face common regulatory and capital market risks, due to their environmental impacts. Identifying this type of commonality can be challenging because these companies operate in different product markets.[3] However, overlaps in their labor markets can illuminate commonalities because both types of firms may hire chemical and mechanical engineers to develop new products, material scientists to optimize production, and financial analysts to help hedge the price fluctuations of raw materials.

Third, labor market peers can also capture output market shocks, as firms in the same product market may hire similar kinds of workers. Thus it is natural for a firm's labor market and product market peers to overlap. I don't focus on labor market peers' ability to capture output market shocks and use traditional industry peers to control for product market shocks.

To examine the incremental explanatory power of labor market peers over traditional industry groupings in identifying firms' economic linkages, I analyze the comovement of stock returns and earnings-based measures. Firms in the same product market are exposed to similar output market shocks that affect their revenues, while firms using the same inputs, including labor, experience similar input market shocks that affect expenses. These shocks, which can include exogenous one-time shocks, changes in risk factors, or a combination of these, can significantly impact firms' stock prices. When firms are exposed to common shocks, their stock prices and earnings are likely to comove. In a stylized model, I show that, when labor market peers hire similar types of employees and are subject to similar labor shocks, their earnings are more likely to comove. The strength of this comovement increases as the peers share more types of labor. While it is difficult to directly identify or observe shocks, I use tests of stock return and earnings comovement to infer the channels through which shocks affect firm value. The incremental explanatory power of labor market peers over standard industry groupings suggests that the labor markets capture important input shocks.

Identifying connections between firms in the labor market can be difficult, due to the limited information that firms disclose about their labor. Data on employment histories from sources such as the Census are not publicly available, and this limits the types of questions that can be answered using these data. However, the emergence of online labor markets has provided a new opportunity to study labor market activities (Horton and Tambe 2015). These sites offer a wealth of data that can be used to better understand the linkages between firms in the labor market.

This paper proposes a novel approach to identify labor market peers by examining "also viewed" firms on LinkedIn and Glassdoor. LinkedIn is the world's largest professional network, with more than 700 million members and 50 million companies worldwide as of 2020. Its members create online profiles and build professional networks on this platform. They use LinkedIn to research companies and explore career opportunities. They can also access job postings, information about corporate culture, and insights about firms' business prospects. At the time of data collection, for each firm on LinkedIn, the site shows six other firms that members "also viewed."

[3] The North American Industry Classification System (NAICS) is a widely used industry classification system that considers a firm's production processes. However, it uses similar data sources as the SIC and thus resembles the SIC. For example, even though chemical manufacturing and oil-and-gas extraction involve similar production processes, they are still assigned to different NAICS industries.

For example, members who "viewed" Tesla "also viewed" BMW, Ford, Solar City, Space Exploration Technologies, Google, and Apple on LinkedIn. While the Standard Industrial Classification (SIC) code would identify BMW and Ford as Tesla's peers, it would not consider Google and Apple, even though they are important peers for Tesla, due to their hiring of similar types of labor. The "also viewed" firms are based on LinkedIn's item-to-item collaborative filtering platform, where a variety of factors are considered and the top-N peer firms are presented.

To further validate the concept of identifying labor market peers based on employee "also viewed" patterns, I apply this approach to Glassdoor.com, an employee review and recruitment website. For each company, Glassdoor lists the top 12 firms "job seekers also viewed." For example, those interested in Amazon "also viewed" Walmart; tech companies such as Apple, Facebook, Google, IBM; the financial services company American Express; and the delivery service company UPS.

I classify the "also viewed" firms on LinkedIn and Glassdoor as the base firm's labor market peers. Considering that the display of "also viewed" companies aims to enhance user engagement and recommend content, and given the career-oriented nature of both LinkedIn and Glassdoor platforms, I expect the "also viewed" firms to reflect firms' labor market connections. Further details on the "also viewed" algorithm and measurement errors are discussed in Section 3.1.

Labor market peers identified from LinkedIn and Glassdoor both capture a firm's connections in the labor market. Approximately 72% of the base firms in my sample have overlapping LinkedIn and Glassdoor peers. Among these base firms, on average, 49% of a base firm's LinkedIn labor market peers are also identified as peers by Glassdoor. In addition to analyzing the LinkedIn and Glassdoor peers, I form a more comprehensive set of labor market peers based on the union of the LinkedIn and Glassdoor peers and a core set of peers based on the intersection of the LinkedIn and Glassdoor peers. This approach allows me to cross-validate the results obtained from the two platforms and provide a more comprehensive understanding of firms' labor market connections.

In this paper, I name the labor market peer measure the LMP and host the data for download on an external website. The LMP measure can identify commonalities between firms that traditional industry groupings cannot. For the union of the LinkedIn and Glassdoor labor market peers, 53% of the peers for S&P 1500 firms come from firms in different six-digit GICS industries and 70% come from firms in different four-digit SIC industries. The intersection sample exhibits less but still significant differences: 29% of S&P 1500 firms' labor market peers are from firms in different six-digit GICS industries, and 54% are from firms in different four-digit SIC industries. This illustrates that labor market peers capture commonalities beyond the traditional output market.

I find that the LMP measure has economically significant and meaningful incremental power in explaining stock return comovements over standard industry groupings (SIC, NAICS, GICS) and text-based product market industry groupings (TNIC, Hoberg and Phillips 2010, 2016). While traditional industry groupings can explain an average of 18.1% of the cross-sectional variation in returns, labor market peers on average increases this explanatory power to 20.8%, a 14.9% increase. Evidence from accounting-based comovement tests suggests that the LMP measure also has incremental power over standard industry groupings in explaining contemporaneous correlations in accounting-based performance measures, including valuation multiples, profitability, and expense ratios.

I also investigate whether there is a relationship between the strength of the LMP measure and the degree of a base firm's labor market connections. I sort the base firms into terciles based on their percentage of shared labor skills with their peers. My analysis shows that the LMP measure offers a higher explanatory power for stock returns when the base firm shares more labor skills with its peers. On average, it explains 22.5% of the cross-sectional variation in returns for the group of firms that share the highest proportion of labor skills, while it only explains 4.7% for the group that shares the least. My analysis also shows that the LMP measure has higher explanatory power for employee wages when the base firm has greater labor market connections with peers. On average, it explains 57.9% of the cross-sectional variation in wages for the group of firms that share the most labor skills, while it only explains 22.3% for the group that shares the fewest skills. Moreover, the measure significantly outperforms traditional industry classifications only when firms share a higher proportion of labor skills, suggesting that labor market peers do better than traditional industry measures at capturing firms' labor market connections.

Next, to demonstrate the practical implications of the LMP measure, beyond identifying labor market links that drive stock return comovements, I examine, in online appendix, whether the measure helps explain the selection of compensation benchmarking peers. I find that a firm's labor market peers are more likely to be chosen as compensation peers, even when controlling for such factors as industry, size, talent flow, peer pay level, and other factors that have been shown to influence the selection of compensation peers. CEO pay at labor market peers also helps explain the median CEO pay of a firm's compensation benchmarking peers and, as a result, the pay of the base firm's CEO. This suggests that the LMP measure can provide valuable insights into factors that influence compensation benchmarking.

Although traditional industry classifications, such as SIC or GICS, are convenient for research, they have limitations. First, traditional industry classifications impose transitivity, requiring that, if firms B and C are in firm A's industry, then B and C are also in the same industry. Thus traditional industry classifications tend to mix closely and remotely related firms. The LMP measure relaxes the transitivity requirement and likely identifies firms' most relevant peers and crucial linkages from a potentially noisy large sample of industry peers. Second, traditional industry groupings are relatively static. Labor market peers, however, are the outcome of real-time employee searches and are more likely to reflect firms' changing economic conditions. Third, no standard industry grouping easily incorporates employees' perspectives, though employees are industry and corporate insiders.

In sum, this study presents initial evidence that labor market peers, identified through "also viewed" firms on Glassdoor and LinkedIn, capture important labor market connections not captured by traditional industry groupings.

My approach also has limitations. First, the algorithm used to identify labor market peers is proprietary to the online platforms. Second, the peers identified through this approach may be influenced by LinkedIn and Glassdoor user demographics and firm activities on the platforms. As such, the LMP measure should be used cautiously for policy making. Third, labor market peers based on LinkedIn and Glassdoor "also viewed" firms likely capture a firm's most relevant peers but may miss other less related firms. Despite these limitations, the concept of extracting labor market connections between firms using employees' online browsing histories can be extended to various online labor markets, offering an alternative to traditional industry peers.

This paper makes several contributions. First, the primary contribution is to the peer selection and industry classification literature (see Section 2). To my knowledge, this paper is among the first to group firms based on labor market similarities and to provide a measure of labor market peers. The economic significance of the input market linkages to firms is not well understood, due to lack of good measure of these linkages. However, the issue is important, given that labor is a critical and scarce resource for many modern firms, especially as the world shifts to a talent economy (Martin 2014). I show that these relations transmit shocks to stock returns and firm performance and represent an important human capital risk. The LMP measure has potential applications in diverse areas, such as benchmarking in performance evaluation and human capital management.

Second, this paper contributes to the budding area of accounting, finance, and labor. Recent studies examine the impact of accounting and financial policies, such as financial distress (Brown and Matsa 2016) and financial misreporting (Choi and Gipper 2019; Teoh et al. 2023), on the labor market. It is challenging to use traditional industry peers to measure the scope of impact, as product and labor markets do not always overlap. This study should help future studies that examine specific labor-market impacts of corporate events and policies.

Third, this paper contributes to the growing body of literature that underscores the latent intelligence found in internet users' online activities, an area that has garnered considerable attention from academics and practitioners. Studies have used aggregate search trends and browsing histories to predict house prices (Wu and Brynjolfsson 2009), stock prices (Da et al. 2011), stock comovement (Lee et al. 2015; Leung et al. 2017), and employee turnover (Dehaan et al. 2023). This paper shows how the online activities of a specific group-employees-reveal firms' fundamental connections. It also contributes to the growing body of research on how information technology reshapes a firm's information environment (Blankespoor et al. 2013; Miller and Skinner 2015; Lee et al. 2015; Jung et al. 2018; Teoh 2018; Chiu et al. 2023).

Section 2 discusses related literature. Section 3 describes the data and provides summary statistics for the LMP measure. Section 4 presents the main results. Section 5 performs additional analyses and robustness tests. Section 6 concludes.

2 Related literature on industry and peer firm groupings

Practitioners and academic researchers identify industry classifications and peer firms for many purposes, including equity valuation (Brown and Ball 1967; King 1966; Fama and French 1997; Bhojraj et al. 2003; De Franco et al. 2015), executive compensation (Albuquerque 2009), corporate financial policy (Rauh and Sufi 2012; Leary and

Roberts 2014), information transfers (Foster 1981; Ramnath 2002; Cohen and Frazzini 2008), and risk management. Industry groupings are also essential for understanding vertical and horizontal integration in industrial organizations (Fan and Lang 2000; Hoberg and Phillips 2010; Lee et al. 2019).

Three industry classifications-the SIC, NAICS, and GICS-are widely used by academics and practitioners in accounting, finance, and economics. The Standard Industry Classification (SIC) System was developed by the Central Statistical Board in the 1930s and last updated in 1987. It uses a demand-based conceptual framework, where establishments are grouped into industries based on the similarity of their product markets. The North American Industry Classification System (NAICS) replaced the SIC method in 1997 and uses a supply-based framework, where establishments are grouped into industries according to similarities in production processes. Although NAICS considers firms' input commonalities, NAICS and SIC use similar data sources, and research shows that they have similar power to explain firms' stock return comovement and financial multiples (Bhojraj et al. 2003). The Global Industry Classification Standard (GICS), developed by Standard & Poor's and MSCI in 1999, offers a market-oriented industry classification. Companies are classified based on their principal businesses and by the market perceptions revealed by investment research reports. Bhojraj et al. (2003) compare the different industry classifications and find that GICS outperforms SIC, NAICS, and the Fama and French 49-industry groupings (Fama and French 1997) in explaining contemporaneous correlations in stock returns.

Text Network Industry Classification (TNIC) is a recent industry classification based on firm pairwise similarity scores from textual analysis of 10K product descriptions. TNIC is updated annually and offers more research flexibility than traditional industry classifications. Research shows that it improves upon SIC and NAICS codes in explaining firm characteristics, such as profitability (Hoberg and Phillips 2010, 2016). However, TNIC is still primarily based on product market similarities. I compare the incremental power of the LMP measure with the four existing industry classifications.

This paper also relates to recent papers that identify peer firms through internet search patterns. Lee et al. (2015) identify peer firms using internet traffic on the EDGAR website, while Leung et al. (2017) identify peers using the online searches of individuals who visit Yahoo! Finance. These measures are based on investor perceptions and most likely work through a mix of fundamental and investor-sentiment channels. In contrast, labor market peers based on LinkedIn and Glassdoor data are less subject to investor sentiment and can capture new dimensions of production functions in the labor market.

Lastly, this paper complements two recent studies that also examine firms' labor market connections. All three papers construct novel measures of labor market peers, and all show that firms' labor market connections differ from their product market connections. However, the three methods are distinct. Liu and Wu (2022) construct their peers based on a database of job postings and link firms based on the job postings' Occupational Information Network (O*NET) codes. Their approach likely offers a comprehensive set of labor market peers but may classify remotely related or unrelated jobs as related because O*NET codes may classify different jobs as related if they have the same job titles.[4] Bae et al. (2022) build their measure based on data from the Bureau of Labor Statistics' Occupational Employment Statistics (OES) survey and O*NET's classification of occupational knowledge. This approach, while likely also offers a thorough analysis of labor market connections, but is constrained to an industry-level measure due to the nature of the OES data.

In contrast, my approach is based on employees' "also viewed" patterns when they are researching companies on online labor market platforms. These patterns likely not only reflect labor demand information, such as job postings, but may also reflect information that affects labor supply and may incorporate employees' private information (Hales et al. 2018; Huang et al. 2020; Li et al. 2019). These sources of information are considered by employees but are likely unobservable to outsiders or researchers grouping firms. Thus Liu and Wu (2022) and Bae et al. (2022) likely capture a more comprehensive set of labor market peers, while the LMP measure in this paper likely captures a firm's most relevant labor market peers.

3 Data and measures

3.1 Data from online labor market "also viewed" firms

I assume that people primarily use online labor market platforms for career-related purposes. When company A is listed as company B's "also viewed" firm on LinkedIn or Glassdoor, it is likely that both companies attract a similar pool of talents. As such, I classify these "also viewed" firms as labor market peers.[5] To construct the LMP measure, I use data from LinkedIn and Glassdoor, respectively. To ensure robustness and gain a deeper understanding of the measure, I create two additional samples: one based on the union of the LinkedIn and Glassdoor labor market peers, which likely captures a broader set of peers, and another based on the intersection of the LinkedIn and Glassdoor peers, which likely captures a core set of peers. I describe the data below.

3.1.1 LinkedIn company page "people also viewed" data

LinkedIn is the world's largest online professional networking and recruiting site. Members, recruiters, and companies comprise of its users. There were more than 700 million members and 50 million companies listed on the site as of 2020. Members use

[4] For example, robotics engineers may require different skills across industries: robotics engineers in the automotive industry design and program physical robots, while those in the financial services industry tend to implement digital robotic processes to eliminate repetitive tasks. Using O*NET codes, these varied roles of robotics engineers could be grouped together under the same job title, despite the differences in their job functions (Ying 2019).

[5] Job search is a process that involves activities from the initial thoughts about quitting, to collecting information and researching new opportunities, to actually applying for a job (Mobley 1977). Thus I do not require viewing job posting to classify career interest. Job seekers could obtain valuable career information from LinkedIn or Glassdoor, and research in labor economics and management shows that new information about compensation, current job quality, and outside job prospects causes employees to reevaluate their positions, and can lead to job search (Stigler 1962; Rogerson et al. 2005)

LinkedIn to build their professional profiles and networks, allowing them to interact with other users and discover new career opportunities. Companies use their LinkedIn pages for recruiting and branding. LinkedIn company homepages have subpages such as "jobs," "life," "people," and "insights," which provide members not only access to job postings but also information about a company's culture, professional development opportunities, employee distribution, and headcount growth based on LinkedIn data. Members can also learn whether any alumni or anyone in their network works at a company and connect with them for information or referrals. Recruiters on LinkedIn, for their part, can use filters to find candidates who have viewed their company page, indicating that recruiters consider these members to be potential employees.

LinkedIn features "People Also Viewed" companies, based on its item-to-item collaborative filtering program known as "browsemaps." This system uses members' browsing histories to build a latent graph of co-occurrences of entities, showcasing relationships between pairs of companies based on the aggregated behavior of many LinkedIn users. Browsemap powers many navigational aids on LinkedIn and contributes a significant portion of traffic and engagement on the platform. The system uses techniques such as damping entities that are overly popular (e.g., preventing Google from being overly correlated throughout the ecosystem) and incorporating a form of hysteresis, where newer views are weighted more heavily than older ones. Figure 1 shows the interface of Dow's web page on LinkedIn including its "People Also Viewed" firms.[6]

The "People Also Viewed" feature on LinkedIn is likely determined by various factors, including counting the co-occurrences of views for different companies (Wu et al. 2014; LinkedIn 2022). LinkedIn may have considered such collaborative filtering features as co-follow or co-employed-at as well as content-based features, such as industry or location.[7] Generally, I would expect labor market-related factors to reveal firms' labor market connections, while nonlabor market-related factors to add noise and influence the outcome of the browsemap.

Recall that the aim of the "People Also Viewed" browsemap is to serve as a navigational aid, enhance user engagement, and recommend relevant content. Given LinkedIn's career focus, labor market peer firms are considered suitable candidates to fulfill this purpose. Nevertheless, it is crucial to acknowledge the possibility of measurement error and selection bias when interpreting the results obtained from the browsemap. I will further discuss these aspects in Sections 3.1.3 and 3.1.4.

[6] LinkedIn's current "People Also Viewed" browsemap displays 10 firms on a company homepage. Between the time the data was collected and January 2023, LinkedIn may have experimented with presenting different number of firms on the browsemaps.

[7] The collaborative filtering features involve co-occurrence browsemaps created from activities like "view," "follow," and "employer-at" on LinkedIn. For example, if users who follow Company A also follow Company B, they may be displayed in each other's "People Also Viewed" list. The "employed-at" activities capture LinkedIn members' job transitions, showing relationships like "People who worked at Company C also worked at Company Y." Content-based features from the company entities consider similarities in industry, location, website content, and other relevant aspects (Wu et al. 2014; LinkedIn 2022). Wu et al. (2014) discusses two company browsemaps which are different from "People Also Viewed": "Similar Companies" browsemap is no longer displayed on LinkedIn as of July 2023, and "People Also Follow" is an individualized browsemap that does not serve the purpose of this study.

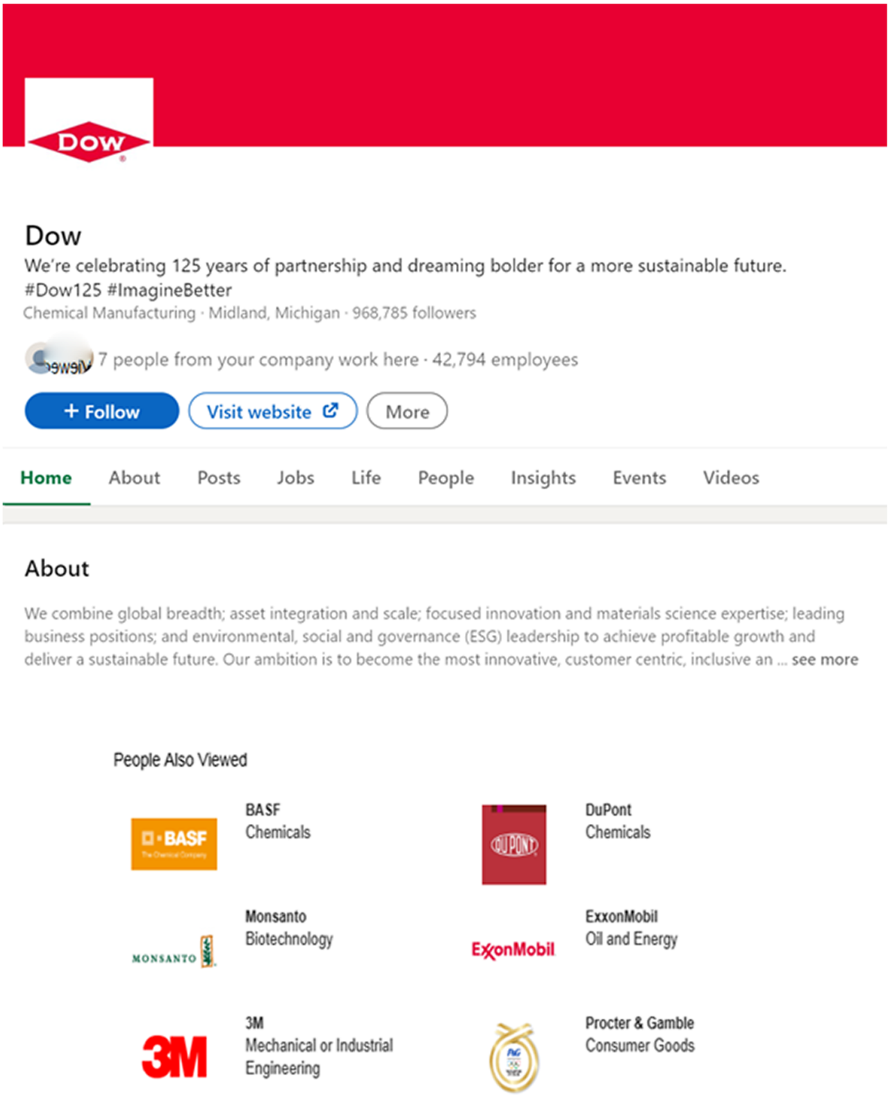

Fig. 1 Dow's LinkedIn homepage and "people also viewed" peers

LinkedIn's internal applications demonstrate that LinkedIn engineers use the "People Also Viewed" company browsemap to identify labor market peers and assist recruiters and hiring managers in finding similarly qualified candidates in peer firms. One such application is the "Similar Profiles" feature, which is a recommender system that pools similar candidates from "also viewed" companies based on the current company of members viewed by recruiters and hiring managers. An A/B test showed that using the company-view browsemap alone increased profile reviews by more than 30% (Wu et al. 2014). Another application, "Search by Ideal Candidates"

(Ha-Thuc et al. 2016), constructs a company browsemap using co-viewing relationships to generate a set of companies that are likely to have candidates similar to the ideal candidates. These applications demonstrate that LinkedIn leverages the company-view browsemap, based on the idea that "People Also Viewed" companies hire individuals with similar skills.

3.1.2 Glassdoor company page "job seekers also viewed" data

Glassdoor is a leading review and recruitment website, founded in 2007. Current and former employees can anonymously write reviews about corporate culture, career opportunities, and salaries on the website. This information is not easily available elsewhere, making Glassdoor useful for both current employees and job seekers. Glassdoor also hosts job postings, and employees can apply for jobs on Glassdoor. As of 2021, Glassdoor had 67 million unique monthly visitors to its website and mobile app and hosted company homepages and job openings.[8]

Glassdoor lists "Job Seekers Also Viewed" firms on its company homepages. According to inquiry with Glassdoor, this feature was based on the analysis of job-clicking by users during my sample period. Two employers are considered related in the labor market if the same user clicks on jobs for both. The top 12 firms of a given firm's "Job Seekers Also Viewed" firms are listed on the Glassdoor company homepage (Chen-Zion 2015). Over time, the feature expanded to include co-views of other employer pages on Glassdoor, such as employer review pages and salary pages, as of July 2023. I classify these firms as Glassdoor labor market peers. Figure 2 shows the interface of Amazon's web page on Glassdoor and the 12 "Job Seekers Also Viewed" firms. Karabarbounis and Pinto (2018) found that the wage distribution from Glassdoor reviews resembles that based on U.S. Census data and industries are widely represented on Glassdoor, although certain industries are more represented than others. Further discussion of Glassdoor's contents, user demographics, and user incentives can be found in the work of Hales et al. (2018); Huang et al. (2020); Marinescu et al. (2021); Dehaan et al. (2023).

3.1.3 Measurement error

The assumption of constructing peer firms based on LinkedIn and Glassdoor's "also viewed" data is that members primarily use online labor market platforms for career-related purposes and the platforms optimize their algorithms to serve this purpose effectively. Thus measurement noise may arise if members use LinkedIn and Glassdoor for other purposes or if nonlabor market factors are used in the construction of "also viewed" firms.

The first type of noise is generated when LinkedIn users perform noncareer-related activities on the platform. According to LinkedIn, the second-largest and

[8] See LinkedIn (https://news.linkedin.com/about-us#Statistics) and the Pew Research Center (https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/) for more details about LinkedIn statistics. Glassdoor statistics are from https://www.glassdoor.com/employers/resources/hr-and-recruiting-stats/.

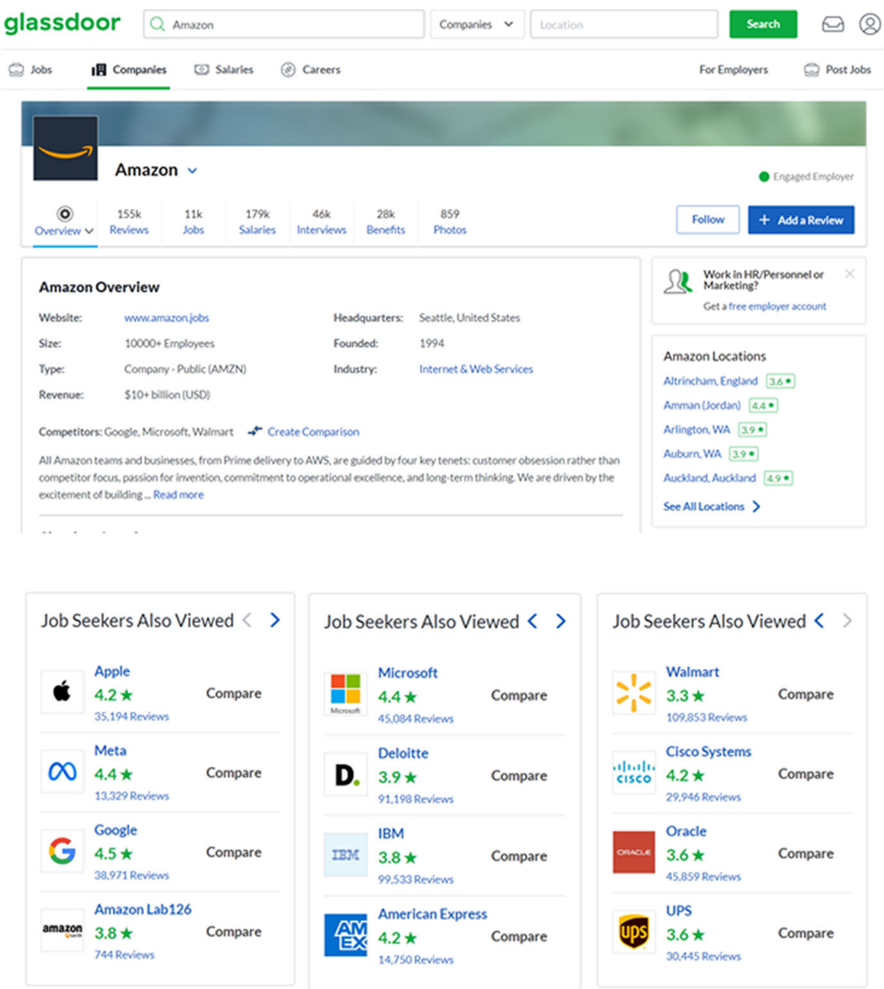

Fig. 2 Amazon's Glassdoor homepage and "job seekers also viewed" peers

fastest-growing activity on its site is B2B marketing. As a result, LinkedIn users may view company profiles for lead generation or to contact people who work for their customers or suppliers.[9]

The second type of noncareer-related activity is using LinkedIn or Glassdoor for leisure purposes. For example, LinkedIn members might connect with people they meet locally, view those people's employers out of curiosity, and use LinkedIn for communication. I address the measurement error issues in Sections 5.1.1 and 5.1.2.

The third source of noise arises from nonlabor market factors incorporated in the "also viewed" firms algorithm. For instance, the algorithm may consider standard

[9] https://business.linkedin.com/marketing-solutions. Accessed on Jan 2023.

industry classification similarity as a factor. It is essential to acknowledge that there are natural overlaps between a firm's product-market and labor-market peers, and including this factor could enhance the algorithm's performance. However, the downside is that nonlabor-related product market peers may introduce noise to the algorithm. To address the potential influence of product market peers, I employ standard industry measures as controls for the empirical tests.

3.1.4 Selection bias

As labor market peers reflect the aggregate behavior of LinkedIn or Glassdoor users, user demographics, firm representation, and how users and firms use the platform would affect their composition.

First, this composition likely depends on who uses LinkedIn and Glassdoor. Notably, 51% of LinkedIn members have a college degree, as compared to 35% based on the most recent U.S. Census Bureau's Current Population Survey (CPS).[10] These numbers indicate that the employees recruited on LinkedIn are likely the essential labor for the firm and the pool of talent that is in scarce supply. As such, the labor market peers identified through LinkedIn and Glassdoor should be interpreted as a firm's critical peers, and it is unclear whether the results would differ for the types of employees who do not use LinkedIn or Glassdoor.

Second, the composition of labor market peers may also depend on firm representation and how firms choose to use the sites. For example, certain industries are more represented on the sites than others (Karabarbounis and Pinto 2018), and this may affect relative web traffic. In addition, firm size may also affect the composition of labor market peers. Big firms likely attract more web traffic, are better known, and may have more accurate labor market peers. Small ones have less web traffic but may still attract interested users because they are likely less well known and have fewer alternative information channels. Sections 5.1.3 and 5.1.4 discusses several robustness tests that further examine this issue.

3.2 Sample selection

The base sample of firms is the S&P 1500 universe as of January 2013. Table 1 panel A describes sample selection. In July 2013, I searched LinkedIn for each firm in the S&P 1500 and found LinkedIn homepages for 1,464 unique firms, of which 1,456 firms had the "People Also Viewed" company browsemap. I collected the "People Also Viewed" firms for each company and manually matched them with firms in the CRSP/Compustat database. This resulted in 5,419 unique peer firms, including 3,366 U.S. public firms. Private and foreign peer firms were excluded due to data availability. I further restricted my sample to firms that could be matched with Compustat and were listed on the NYSE, NASDAQ, or AMEX trading with ordinary common shares (CRSP share code of 10 or 11). This resulted in 1,325 base firms from the S&P 1500 universe,

[10] Census Bureau statistics are from Educational Attainment in the United States: 2018 (https://www.census.gov/data/tables/2018/demo/education-attainment/cps-detailed-tables.html).

Table 1 Summary statistics of labor market peers

Panel A: Coverage of S&P 1500 firms

Representation of S&P 1500 universe		**S&P 1500 firms**	**S&P 500 firms**
		(1)	**(2)**
(1)	Number of firms with a LinkedIn company homepage	1464	500
(2)	Number of firms with "people also viewed" feature on LinkedIn	1456	500
(3)	At least one LinkedIn labor market peer with common shares and listed on NYSE, NASDAQ or AMEX	1325	449
(4)	Number of firms with a Glassdoor company homepage	1309	442
(5)	At least one labor market peer with common shares Listed on NYSE, NASDAQ or AMEX	1250	425

Panel B: Composition of S&P 1500 firms' labor market peers

Number of public labor market peers	**S&P 1500 firms LinkedIn**	**S&P 500 firms LinkedIn**	**S&P 1500 firms Glassdoor**	**S&P 500 firms Glassdoor**
	(1)	(2)	(3)	(4)
1	87	12	42	4
2	145	41	80	6
3	258	50	89	18
4	273	94	122	21
5	290	106	143	33
6	272	146	174	44
7-12	NA	NA	600	299
Total	1325	449	1250	425

This table provides summary statistics for S&P 1500 firms' LinkedIn and Glassdoor labor-market peer firms. Panel A reports the coverage of S&P 1500 firms on LinkedIn and Glassdoor. The first row reports the number of base firms with a LinkedIn company homepage. The second row removes firms without a "People Also Viewed" browsemap on its LinkedIn company homepage. The third row restricts the sample to firms that could be matched with Compustat and with at least one public labor market peer listed on NYSE, NASDAQ, or AMEX with a CRSP share code of 10 or 11. The third and fourth rows repeat steps in the second and third rows and report the number of firms for the Glassdoor sample. Columns (1) and (2) report the number of firms within the S&P 1500 and S&P 500 samples, respectively. Panel B reports the number of base firms with a corresponding number of public labor market peers. For the LinkedIn sample, each firm has at most six labor market peers, which are a mix of private and public peers. Columns (1) and (2) report the number of base firms with a corresponding number of public LinkedIn labor market peers for the S&P 1500 and S&P 500 samples, respectively. For example, in the S&P 1500 firms, 87 base firms have only one public labor market peer, 145 have two, and 258 have three. For the Glassdoor sample, each firm has at most 12 labor market peers. Columns (3) and (4) report the number of base firms with a corresponding number of public Glassdoor labor market peers for the S&P 1500 and S&P 500 samples, respectively including 449 S&P 500 firms.[11] I apply a similar process to collect the "Job Seekers Also Viewed" firms on Glassdoor for each company in February 2017, which further reduces my sample to 1,250 firms. Table 1 panel B provides the count of public labor market peers. Ninety-three percent (82%) of the LinkedIn S&P 1500 sample has at least two (three) public labor market peers. I keep base firms with at least two public peers, but the results are robust to requiring three public peers or to removing this restriction. I obtain stock return data from CRSP, financial information and executive compensation data from Compustat and ExecuComp. TNIC industry data are obtained from the Hoberg and Phillips data library. Firm location data are from Glassdoor. Median employee salary is computed based on Glassdoor salary data or from the Equilar database based on firm proxy statements. Compensation benchmarking peers are obtained from ISS Incentive Lab. Customer-supplier relationships are based on the Compustat Segments Customer File. The number of observations differs across tests, due to data availability.

3.3 Summary statistics

3.3.1 Comparing labor market peers identified from LinkedIn and Glassdoor

The LinkedIn algorithm is mainly based on "also viewed" company pages, while the Glassdoor algorithm is primarily based on "also viewed" job postings. Despite this difference, I expect an overlap in labor market peers identified from LinkedIn and Glassdoor because most activities on the sites are career-related. Indeed 72% (897/1250) of the base firm's LinkedIn labor market peers share at least one firm with the Glassdoor labor market peers. Among these firms, 49% of LinkedIn peers are also captured by the Glassdoor peers. This significant overlap validates the LMP measure and suggests that employee "also viewed" patterns on the two sites reveal a firm's essential labor market peers. Despite this commonality, LinkedIn and Glassdoor peers also differ. I further discuss this difference in Appendix B.

3.3.2 Comparison with product market peers

Table 2 shows the extent to which a firm's labor market peers differ from its product market peers. The differences are measured by the proportion of labor market peers that belong to different industry classification (two-digit GICS, six-digit GICS, two-digit SIC, four-digit SIC, three-digit NAICS, and six-digit NAICS) compared to the base firm. The results show that labor market peers overlap with but differ from product market peers.

[11] Labor market peers may not frequently change (Wu et al. 2014). For example, the "People Also Viewed" firms for Google were Amazon, Microsoft, Apple, Facebook, IBM, and Hewlett-Packard in my sample. As of November 25, 2020, Netflix and LinkedIn had replaced IBM and Hewlett-Packard, with the other four peers unchanged. As of May 25, 2021, five of Google's six labor market peers in my sample are still listed as "People Also Viewed" firms, except for Hewlett-Packard. Bae et al. (2022) also find their industry-level labor connections exhibit a degree of stability. Still, further study is needed to analyze the composition of LinkedIn and Glassdoor peers over time.

Table 2 Difference between labor market peers and standard industry classifications

Panel A: LinkedIn labor market peers

GICS2 groupings	Number of firms	Same GICS2	Same GICS6	Same SIC2	Same SIC4	Same NAICS3	Same NAICS6
Energy	85	0.04	0.10	0.33	0.43	0.34	0.60
Health care	147	0.09	0.28	0.30	0.53	0.43	0.57
Financial	176	0.11	0.26	0.23	0.34	0.20	0.37
Information technology	236	0.18	0.42	0.40	0.65	0.50	0.68
Consumer discretionary	211	0.20	0.38	0.47	0.59	0.48	0.68
Consumer staples	72	0.25	0.47	0.46	0.84	0.51	0.85
Utilities	58	0.25	0.65	0.23	0.69	0.25	0.70
Materials	88	0.26	0.37	0.47	0.77	0.45	0.81
Industrials	191	0.31	0.47	0.50	0.71	0.53	0.75
Communication services	44	0.42	0.65	0.36	0.61	0.47	0.67
Real estate	17	0.52	0.54	0.62	0.78	0.65	0.81
S&P1500	1325	0.20	0.38	0.39	0.60	0.43	0.65
S&P500	449	0.20	0.40	0.40	0.60	0.43	0.64

Panel B: Glassdoor labor market peers

GICS2 groupings	Number of firms	Same GICS2	Same GICS6	Same SIC2	Same SIC4	Same NAICS3	Same NAICS6
Energy	80	0.24	0.34	0.51	0.60	0.55	0.71
Health care	132	0.33	0.54	0.57	0.75	0.62	0.77
Financial	160	0.20	0.45	0.43	0.54	0.35	0.56
Information technology	225	0.32	0.69	0.60	0.79	0.64	0.80
Consumer discretionary	201	0.38	0.62	0.70	0.77	0.71	0.82
Consumer staples	73	0.52	0.70	0.64	0.93	0.73	0.93
Utilities	55	0.74	0.88	0.72	0.89	0.74	0.93
Materials	84	0.69	0.74	0.72	0.90	0.72	0.92
Industrials	183	0.53	0.73	0.74	0.87	0.79	0.89
Communication services	44	0.76	0.85	0.65	0.79	0.77	0.85
Real estate	13	0.80	0.85	0.85	0.92	0.87	0.96
S&P1500	1250	0.42	0.64	0.63	0.77	0.68	0.80
S&P500	425	0.43	0.65	0.64	0.78	0.65	0.80

Panel C: Union labor market peers

GICS2 groupings	Number of firms	Same GICS2	Same GICS6	Same SIC2	Same SIC4	Same NAICS3	Same NAICS6
Energy	68	0.18	0.25	0.47	0.56	0.49	0.67
Health Care	105	0.26	0.44	0.46	0.67	0.53	0.69
Financial	124	0.17	0.39	0.36	0.46	0.31	0.48

Table 2 continued

Information technology	166	0.25	0.59	0.53	0.73	0.59	0.75
Consumer discretionary	153	0.29	0.53	0.63	0.71	0.63	0.77
Consumer staples	48	0.42	0.63	0.58	0.90	0.63	0.91
Utilities	41	0.54	0.80	0.51	0.83	0.54	0.81
Materials	45	0.54	0.58	0.57	0.84	0.58	0.87
Industrials	114	0.42	0.61	0.65	0.80	0.68	0.82
Communication services	24	0.61	0.74	0.51	0.73	0.64	0.77
Real estate	9	0.81	0.83	0.86	0.91	0.85	0.95
S&P1500	897	0.31	0.53	0.53	0.70	0.56	0.73
S&P500	368	0.35	0.57	0.57	0.72	0.58	0.75

Panel D: Intersection labor market peers

GICS2 groupings	Number of firms	Same GICS2	Same GICS6	Same SIC2	Same SIC4	Same NAICS3	Same NAICS6
Energy	68	0.04	0.11	0.29	0.41	0.34	0.55
Health Care	105	0.06	0.17	0.26	0.48	0.40	0.52
Financial	124	0.04	0.19	0.20	0.30	0.17	0.35
Information technology	166	0.12	0.35	0.36	0.62	0.50	0.66
Consumer discretionary	153	0.11	0.29	0.39	0.53	0.39	0.59
Consumer staples	48	0.20	0.41	0.32	0.82	0.43	0.81
Utilities	41	0.18	0.59	0.19	0.68	0.19	0.70
Materials	45	0.27	0.31	0.32	0.60	0.34	0.90
Industrials	114	0.18	0.29	0.37	0.61	0.42	0.66
Communication services	24	0.39	0.59	0.22	0.46	0.41	0.57
Real estate	9	0.53	0.53	0.67	0.78	0.67	0.89
S&P1500	897	0.13	0.29	0.31	0.54	0.37	0.59
S&P500	368	0.15	0.30	0.34	0.53	0.38	0.59

This table summarizes the difference between labor market peers and other standard industry classification schemes by each GICS2 industry sector. Panels A to B report the fraction of LinkedIn and Glassdoor labor market peers with different industry classifications (for GICS2, GICS6, SIC2, SIC4, NAICS3, and NAICS6) from the base firm's, respectively. Panel C reports the fraction for union labor market peers, which is the union of the LinkedIn and Glassdoor. Panel D reports the fraction for intersection labor market peers, which is the intersection of the LinkedIn and Glassdoor. In Panels C and D, I require the base firms to have overlapping LinkedIn and Glassdoor labor market peers for comparison purposes

For LinkedIn labor market peers (Panel A), 18% have a different two-digit GICS classification from the base firm, 39% have a different two-digit SIC, and 60% have a different four-digit SIC. Meanwhile, the Glassdoor peers (Panel B) have a greater difference, with 42% having a different two-digit GICS, 63% having a different two-digit SIC, and 77% having a different four-digit SIC. These variations between the LinkedIn and Glassdoor peers can be attributed to platform-specific differences, such as the maximum number of “also viewed” firms displayed and other factors discussed by Appendix B.[12]

For the union and intersection peers, I require the base firm to have at least one overlapping LinkedIn and Glassdoor labor market peer for the purpose of comparison. The union peers (Panel C) tend to have differences between those of the LinkedIn and Glassdoor peers, while the intersection peers (Panel D), which are likely the core set of peers, has slightly more overlap with product market peers.

The degree of difference between the firm’s input and output markets varies by industry, with some industries having less difference, such as energy, healthcare, financial, and information technology, and others having greater difference, such as materials, consumer staples, and industrials. An example is the Dow Chemical, which belongs to the materials industry. Its LinkedIn labor market peers span five different two-digit GICS industries: materials (BASF and Dupont), industrials (3M), energy (Exxon Mobil), healthcare (Monsanto), and consumer staples (Procter & Gamble).

3.3.3 Labor skill similarity

One way to measure the relatedness of firms in the labor market is analyzing the proportion of skills that are shared among them. I use self-reported skills listed in the Skills & Expertise section of LinkedIn member profiles. These skills are publicly available to LinkedIn members and are endorsed by their network, which serves as a credibility indicator. LinkedIn also provides an overview of the five most common skills of a company’s current employees on each company’s LinkedIn homepage. Additionally, LinkedIn uses a proprietary algorithm to calculate related skills for each of the top skills, for example, the top skills and expertise of Google are Google Adwords, Python, machine learning, AdSense, and Google technologies. The related skills for machine learning are feature selection, text mining, pattern recognition, etc.[13]

I measure the skill similarity $s_{i,j}$ between firm i and firm j as the proportion of skills that are shared by both, divided by the number of total unique skills of both. This measure is defined as:

$$s_{i,j} = \frac{|C_i \cap C_j|}{|C_i \cup C_j|}, \tag{1}$$

where C_i is the skill set of firm i and C_j is the skill set of firm j. For example, Google has 88 skills, and Facebook has 76 skills. They have 32 skills in common. Their skill

[12] LinkedIn displays six “also viewed” firms, while Glassdoor shows twelve. If both platforms rank the same pool of labor market peers in the same manner, the design of LinkedIn would limit the display to the top six, whereas Glassdoor would show the top twelve firms. Based on the intersection sample statistics from Table 2 Panel D, it is evident that the top related firms are more likely to have a higher degree of overlap with standard industry classifications. This suggests that interface design could be one of the factors contributing to the difference in overlap between LinkedIn and Glassdoor peers with standard industry classifications.

[13] LinkedIn discontinued displaying firm-level skills after August 2013 but continues to display individual members’ skills. LinkedIn now hosts a database of skills for certain job titles and produces annual reports on top skills for emerging jobs (https://linkedin.github.io/future-of-skills, https://business.linkedin.com/talent-solutions/emerging-jobs-report#all, accessed on August 2022). These resources can provide valuable insights for policymakers, as demonstrated by LinkedIn’s partnership with the World Bank, which uses LinkedIn skills data to inform policies related to workforce development and education (Zhu et al. 2018).

Table 3 Summary statistics of skill similarity ratio

GICS2 groupings	Number of firms (1)	LinkedIn labor market peers (2)	Other labor market related firms (3)	(2)-(3) (4)
Energy	82	0.19	0.09	0.10***
Utilities	61	0.16	0.10	0.06***
Real Estate	15	0.15	0.06	0.09***
Financial	172	0.14	0.10	0.04***
Consumer discretionary	209	0.12	0.06	0.06***
Health care	147	0.10	0.05	0.05***
Materials	87	0.10	0.05	0.05***
Consumer staples	72	0.10	0.06	0.04***
Industrials	187	0.08	0.04	0.04***
Communication services	44	0.07	0.05	0.02***
Information technology	236	0.07	0.05	0.02***
S&P1500	1311	0.11	0.06	0.05***

This table provides summary statistics of the average skill similarity scores between the base firm and its LinkedIn labor market peers by each GICS2 industry sector. Skill similarity $s_{i,j}$ is defined in Eq. 1, and is measured as the fraction of skills shared by firm i and firm j divided by the number of total unique skills of firm i and firm j. Column (1) reports the number of firms in each GICS2 sector. Column (2) reports the average skill similarity between the base firm and its labor market peers. Column (3) reports the average skill similarity between the base firm and any firm in my sample that shares at least one skill with the base firm. Column (4) reports the difference between Columns (2) and (3). *** denotes significance at the 1% level (two-tailed t-test)

similarity would be $s_{Google,Facebook} = \frac{32}{88+76-32} = 0.24$. The top and the related skills of Google and Facebook can be seen in Appendix D.

Table 3 summarizes the labor market similarity between the base firm and its LinkedIn labor market peers by measuring the proportion of shared skills. The table presents data on the number of firms and average skill similarity in each two-digit GICS industry. The skill similarity between the base firm and its labor market peers is compared to the average skill similarity between the base firm and other firms that share at least one skill in my sample. The expectation is that, if labor market peers capture firms' critical labor market peers, the average skill similarity between the base firm and its peers will be higher than the benchmark. On average, S&P 1500 firms share 11% of skills with their labor market peers, as compared to 6% with the benchmark groups. Firms share significantly more skills with their labor market peers than with the benchmark in every GICS2 industry. Firms in the information technology industry have the lowest commonality with peers (7%), while energy, utility, and financial firms share 19%, 15%, and 14% of skills with their labor market peers respectively.[14] Generally, the more skills the base and peer firms share, the more related they are in

[14] Skill similarity should be interpreted with caution because skill definition may vary across industries. For example, IT firms have a very specific delineation of skills; each programming language, such as C, JAVA, and Perl, is listed separately. Financial industries have broad definitions of skills, such as corporate finance and mortgage lending.

the labor market. This result supports the conclusion that labor market peers capture labor market relatedness.[15]

4 Labor market peers and comovements

There is a large economics and finance literature on the impact of labor market shocks or risks on employee wages, firm earnings, and stock returns. Studies have found that labor market shocks or risks can affect not only employee wages but also firm earnings and stock returns (Mayers 1973; Jaganathan and Wang 1996; Eisfeldt and Papanikolaou 2013; Donangelo 2014). I develop a stylized model in Appendix A and illustrate how firms' earnings comove when they hire similar types of scarce labor, and how that comovement strengthens when they share more types of scarce labor. As a result, firms' stock returns will also comove when they are subject to common labor market shocks or risks.

In this section, I control for output market shocks and risks using existing industry groupings and show that labor market peers have significant incremental explanatory power over traditional industry groupings in understanding the cross-sectional variation of returns, valuation multiples, profitability, expense ratios, and wages. This supports the idea that the labor market captures firms' fundamental connections beyond the output market.

Moreover, the LMP measure explains more of the cross-sectional variation of returns and wages when firms share more skills and outperforms alternative industry peers when firms share more skills with their peers, supporting the idea that it captures firms' labor market similarities.

4.1 Return comovements of labor market peers

An important aspect of firms' economic linkages is the degree of contemporaneous correlation in their stock returns. In this analysis, I investigate to what extent labor market peers outperform standard industry classifications in explaining stock return comovements.

My first test compares labor market peers with existing industry peers to explain the base firm's stock return variation. I estimate the cross-sectional regression specification following the method used by Bhojraj et al. (2003) and Lee et al. (2015), for every month from 2014 to 2019. The regression specification is as follows:

$$R_{i,t} = \alpha_t + \beta_t \, R_{peer,t} + \epsilon_{i,t}, \tag{2}$$

where $R_{i,t}$ is the monthly stock return for each base firm i drawn from CRSP monthly files and where $R_{peer,t}$ is the equally weighted average monthly portfolio return based on the base firm's LMP, GICS, SIC, NAICS, or TNIC peers, excluding the base firm. I

[15] Similarities in skills can serve as a contributing factor in constructing the "also viewed" algorithm. In such scenarios, the selection of "also viewed" firms might be engineered to possess a greater probability of sharing a broader range of labor skills.

estimate Eq. 2 across firms for every month from 2014 to 2019 and obtain an average R^2 based on the 72 regressions.

I construct peer portfolio returns using the same number of firms as the number of labor market peers. Since the LMP measure likely picks a firm's closest labor market peers, I select a firm's closest product-market or industry peers for the benchmark peer portfolio construction. Different methods are used to select the closest product-market or industry peers, such as using the Hoberg-Phillips score data for the TNIC peer construction and selecting peer firms closest in size and in the same industry as the base firm for the GICS, SIC, and NAICS industry peers. For example, to construct the TNIC peer portfolio for a base firm with six labor market peers, I would select six TNIC peers with the highest Hoberg-Phillips score data. To construct the GICS peer portfolio for a base firm with six labor market peers, I would select the six firms closest in size and in the same most granular GICS industry classification as the base firm. If the number of firms in the most granular industry specification is less than the number of labor market peers, I would move up to the next level of industry specification.[16]

Four sets of labor market peers are analyzed in this section: the LinkedIn peers, the Glassdoor peers, the union of the LinkedIn and Glassdoor peers, and the intersection of the LinkedIn and Glassdoor peers. For a given base firm, the union set has the most labor market peers, while the intersection set has fewest. The union set likely captures a broader set of labor market peers, and the intersection set likely captures a core set of labor market peers. I don't have an expectation as to which set would perform best but do expect my results to be robust and consistent across the four sets if both the LinkedIn and Glassdoor "also viewed" algorithms capture a similar construct. To facilitate comparison, all regressions are conducted using the same underlying set of base firms. This reduces the number of firms in my sample because only 72% of the base firms' LinkedIn and Glassdoor labor market peers intersect. In online appendix, I drop the intersection sample and show that the results are robust to using a larger sample without requiring overlap between Glassdoor and LinkedIn labor market peers.

Suppose that the LMP measure captures firms' connections that transmit common shocks beyond the output market. In that case, it should exhibit incremental power over standard industry classifications in explaining the base firm's stock-return variation. I test the regression specification

$$R_{i,t} = \alpha_t + \beta_{LMP,t}\, R_{LMP,t} + \sum \beta_{peer,t}\, R_{peer,t} + \varepsilon_{i,t}, \tag{3}$$

where $R_{LMP,t}$ is the equally weighted average monthly portfolio return based on the LMP measure and $R_{peer,t}$ is the equally weighted average monthly portfolio return based on the same number of closest GICS, SIC, NAICS, or TNIC peers.

[16] Specifically, when forming the GICS peers for a base firm with six labor market peers, I first select six firms closest in size with the same eight-digit GICS code as the base firm. If there are less than six firms with the same eight-digit GICS as the base firm, I select six firms closest in size with the same six-digit GICS as the base firm. If there are still less than six firms with the same six-digit GICS as the base firm, I select six firms closest in size with the same two-digit GICS as the base firm. For SIC industry classification, I start from four-digit SIC and then move to three-digit SIC and two-digit SIC if necessary. For NAICS industry classification, I start from six-digit NAICS and then move to three-digit NAICS if necessary.

Columns (1) to (5) of Panels A to D of Table 4 report estimation results for Eq. 2 for the LinkedIn, Glassdoor, union, and intersection labor market peers, respectively. The average R^2 is the main measure for explanatory power. The LinkedIn labor market peers explain an average of 14.8% of the cross-sectional variation in realized returns, significantly outperforming output-market-based or investment-oriented industry groupings, that is, GICS (12.8%), SIC (10.6%), NAICS (10.7%), and TNIC (11.3%). The Glassdoor labor market peers explain an average of 11.9% of the cross-sectional variation in realized returns, significantly underperforming benchmark peers based on GICS but on par with peers based on SIC, NAICS, and TNIC. The union labor market peers explain an average of 14.6% of the cross-sectional variation

Table 4 Return comovement tests based on labor market peers and on alternative industry groupings

Panel A: LinkedIn labor market peers

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_LN	0.579***						0.323***
	(33.344)						(22.350)
GICS_LN		0.461***				0.244***	0.186***
		(28.093)				(19.660)	(15.581)
SIC_LN			0.388***			0.084***	0.060***
			(24.317)			(6.655)	(4.947)
NAICS_LN				0.398***		0.104***	0.068***
				(26.445)		(8.702)	(5.695)
TNIC_LN					0.407***	0.237***	0.159***
					(27.294)	(24.029)	(17.739)
Observations	39,600	39,600	39,600	39,600	39,600	39,600	39,600
Avg. R-squared	0.148	0.128	0.106	0.107	0.113	0.185	0.216
Number of groups	72	72	72	72	72	72	72

Panel B: Glassdoor labor market peers

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_GD	0.702***						0.324***
	(31.160)						(21.749)
GICS_GD		0.523***				0.278***	0.235***
		(29.066)				(19.730)	(17.081)
SIC_GD			0.427***			0.092***	0.078***
			(25.524)			(6.541)	(5.509)
NAICS_GD				0.429***		0.095***	0.067***
				(24.876)		(6.659)	(4.947)
TNIC_GD					0.461***	0.250***	0.192***
					(27.821)	(22.403)	(18.626)

Table 4 continued

Observations	39,600	39,600	39,600	39,600	39,600	39,600	39,600
Avg. R-squared	0.119	0.140	0.113	0.112	0.122	0.191	0.209
Number of groups	72	72	72	72	72	72	72

Panel C: Union labor market peers

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_UNI	0.797***						0.398***
	(35.937)						(23.630)
GICS_UNI		0.560***				0.297***	0.234***
		(30.466)				(19.413)	(15.703)
SIC_UNI			0.451***			0.091***	0.073***
			(26.782)			(6.282)	(5.128)
NAICS_UNI				0.449***		0.100***	0.062***
				(25.731)		(6.961)	(4.506)
TNIC_UNI					0.500***	0.257***	0.177***
					(29.937)	(21.048)	(16.513)
Observations	39,600	39,600	39,600	39,600	39,600	39,600	39,600
Avg. R-squared	0.146	0.147	0.116	0.115	0.127	0.194	0.217
Number of groups	72	72	72	72	72	72	72

Panel D: Intersection labor market peers

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_INT	0.422***						0.264***
	(25.500)						(21.097)
GICS_INT		0.295***				0.161***	0.127***
		(19.904)				(15.477)	(12.607)
SIC_INT			0.263***			0.057***	0.041***
			(19.190)			(4.840)	(3.622)
NAICS_INT				0.278***		0.107***	0.085***
				(20.914)		(8.816)	(7.261)
TNIC_INT					0.263***	0.180***	0.128***
					(26.199)	(23.301)	(17.852)

Table 4 continued

Observations	39,600	39,600	39,600	39,600	39,600	39,600	39,600
Avg. R-squared	0.114	0.088	0.078	0.082	0.078	0.153	0.189
Number of groups	72	72	72	72	72	72	72

This table reports the average of 72 monthly cross-sectional regressions of return comovement tests $R_{i,t} = \alpha_t + \beta_t R_{peer,t} + \epsilon_{i,t}$, that are based on the labor market peer and alternative industry groupings. The dependent variable is the base firm's stock-level monthly returns. In columns (1) to (5), the independent variable is the equally weighted contemporaneous average return of the LinkedIn labor market peers (LMP_LN), GICS, SIC, NAICS, or TNIC peers formed with the top-N number of industry peers closest in size with the base firm, where N is the number labor market peer firms. Column (6) includes peer returns from all alternative industry groupings. Column (7) adds portfolio returns from the LinkedIn peers. Panels A to D report results for the LinkedIn, Glassdoor, union, and intersection labor market peers, respectively. The time-series average of monthly cross-sectional regression coefficients and average R-squared values are reported. The t-statistics of coefficient estimates are reported in parentheses, where *, **, and *** denote significance at the 10%, 5%, and 1% significance levels, respectively. The sample includes S&P 1500 firms, whose LinkedIn and Glassdoor labor market peers overlap and have at least two publicly traded labor market peers between 2014 and 2019

in realized returns, insignificantly different from the performance of GICS peers and outperforming other peers used in this paper. The intersection labor market peers explain an average of 11.4% of the cross-sectional variation in realized returns, outperforming peers based on GICS, SIC, NAICS, and TNIC. Two-tailed t-tests are performed to test the difference significance. Note that the performance of the same benchmark industry grouping differs from Panel A to D; this is because I require the same number of firms as the corresponding labor market peer to construct a benchmark industry portfolio, and thus the number of firms used to construct a benchmark portfolio likely differs across panels.

Next I examine whether labor market peers capture firms' fundamental connections beyond traditional output-market-based peers. Columns (6) to (7) of Panels A to D of Table 4 show the incremental power of labor market peers. In Panel A column (6), GICS, SIC, NAICS, and TNIC together explain an average of 18.5% of cross-sectional variation in realized returns. Adding the LinkedIn labor market peer portfolio in column (7) increases the overall explanatory power to 21.6%. Similarly, adding the Glassdoor portfolio increases the overall explanatory power from 19.1% to 20.9%, the union portfolio increases the overall explanatory power from 19.4% to 21.7%, and the intersection portfolio increases the overall explanatory power from 15.3% to 18.9%. Two-tailed t-tests, conducted on the time-series of monthly differences from 2014 to 2019, indicate that the explanatory power of column (7) is significantly greater than that of column (6) for all four portfolios at the 1% significance level. On average, labor market peers achieve an improvement of 14.9% in average R^2 values over the aggregate explanatory power of GICS, SIC, NAICS, and TNIC for the S&P 1500 sample.[17] The results suggest that the labor market peer measure captures firms' important fundamental connections that are not identified by traditional industry peers.

In Section 5.2, I show that my results are robust to alternative methods of selecting closest product market peers. This includes selecting peers that are in the same industry

[17] Specifically, (21.6% + 20.9% + 21.7% + 18.9%)/(18.5% + 19.1% + 19.4% + 15.3%) - 1=14.9%.

and have the closest performance or firms that are closest in both size and performance to the base firm. I also show that my findings are robust when using benchmark portfolios constructed with more peers, such as those based on industry and size quartile-matched portfolios or portfolios that include all firms in the same industry.

4.2 Accounting ratios and valuation multiples

Another measure of firms' economic connections is the degree of contemporaneous correlation in their accounting-based performance metrics. Brown and Ball (1967) show that industry earnings explain a substantial amount of an individual firm's earnings. I use the same sample and methodology as in the stock return analysis, and the LMP measure should have incremental explanatory power over traditional industry peers in explaining profitability ratios and valuation multiples if it measures input market shocks that affect firm value. I use annual data from Compustat on a range of expense ratios, valuation multiples, and a number of other financial ratios, including research and development expenses scaled by net sales (*rdpersales*); selling, general and administrative expenses scaled by net sales (*sgapersales*); the price-to-book ratio (*pb*); the enterprise value-to-sales ratio (*evs*); the price-to-earnings ratio (*pe*); returns on net operating assets (*rnoa*); the return on equity (*roe*); the inverse of assets turnover (*at*); the profit margin (*pm*); leverage (*leverage*); and one-year-ahead realized sales growth (*salesgrowth*). The methods used to calculate these ratios are detailed in Appendix C.

For each of these variables, I start by running the cross-sectional regression including traditional industry groupings:

$$Ratio_{i,t} = \alpha_t + \sum \beta_{peer,t}\ Ratio_{peer,t} + \epsilon_{i,t} \tag{4}$$

and then add the accounting-based performance ratio of my measure:

$$Ratio_{i,t} = \alpha_t + \beta_{LMP,t}\ Ratio_{LMP,t} + \sum \beta_{peer,t}\ Ratio_{peer,t} + \epsilon_{i,t} \tag{5}$$

where $Ratio_{i,t}$ is the variable of interest for each base firm i and $Ratio_{LMP,t}$ is the equally weighted average of the same variable for the closest labor market peers. The independent variables are the same number of equally weighted ratios of the closest GICS, SIC, NAICS, and TNIC peers. I estimate these regressions for every calendar year from 2014 to 2019. I use yearly ratios because different industries are more comparable at the yearly level, due to seasonalities.[18]

Table 5 presents the results. Panels A to D report results for the LinkedIn, Glassdoor, union, and intersection labor market peers respectively. Column (1) shows the number

[18] Following Bhojraj et al. (2003), I drop observations with missing total assets, long-term debt, net income before extraordinary items, or operating income after depreciation. I drop observations with negative common or total equity, keep share prices greater than \$3 at the end of the fiscal year, and keep net sales exceeding \$100 million. I also require that net income before extraordinary items be positive in computing *pe*. Finally, to mitigate the effect of outliers, I truncate observations at the 1st and 99th percentiles for each variable. These restrictions result in fewer annual observations per year. The actual observations used in various tests vary depending on data availability.

Table 5 Incremental R^2 values: accounting-based performance ratios

Panel A: LinkedIn labor market peers

	Num. Obs	**GICS+SIC NAICS+TNIC**	**GICS+SIC NAICS+TNIC +LMP_LN**	
		Avg. R-squared	**Avg. R-squared**	**(3)-(2)**
	(1)	**(2)**	**(3)**	**(4)**
Expense ratios				
rdpersales	509	0.827	0.841	0.014***
sgapersales	509	0.723	0.746	0.023***
Valuation multiples				
pb	508	0.105	0.109	0.004*
evs	508	0.734	0.746	0.012***
pe	384	0.061	0.075	0.046
Financial statement ratios				
rnoa	370	0.149	0.164	0.015**
roe	509	0.110	0.117	0.007
at	509	0.892	0.901	0.009***
pm	509	0.664	0.691	0.027***
leverage	509	0.080	0.081	0.001*
salesgrowth	354	0.222	0.245	0.023

Panel B: Glassdoor labor market peers

	Num. Obs	**GICS+SIC NAICS+TNIC**	**GICS+SIC NAICS+TNIC +LMP_GD**	
		Avg. R-squared	**Avg. R-squared**	**(3)-(2)**
	(1)	**(2)**	**(3)**	**(4)**
Expense ratios				
rdpersales	509	0.822	0.830	0.008***
sgapersales	509	0.725	0.731	0.006***
Valuation multiples				
pb	508	0.088	0.099	0.011*
evs	508	0.734	0.735	0.001**
pe	384	0.046	0.067	0.021*
Financial statement ratios				
rnoa	370	0.167	0.180	0.012*
roe	509	0.084	0.092	0.009
at	509	0.895	0.902	0.007***
pm	509	0.659	0.676	0.017***

Table 5 continued

leverage	509	0.065	0.067	0.002
salesgrowth	354	0.219	0.231	0.012*

Panel C: Union labor market peers

	Num. Obs	**GICS+SIC NAICS+TNIC**	**GICS+SIC NAICS+TNIC +LMP_UNI**	
		Avg. R-squared	**Avg. R-squared**	**(3)-(2)**
	(1)	**(2)**	**(3)**	**(4)**
Expense ratios				
rdpersales	509	0.822	0.831	0.009***
sgapersales	509	0.724	0.731	0.007***
Valuation multiples				
pb	508	0.100	0.112	0.012
evs	508	0.730	0.734	0.004*
pe	384	0.048	0.065	0.017
Financial statement ratios				
rnoa	370	0.146	0.161	0.015**
roe	509	0.088	0.103	0.015
at	509	0.893	0.900	0.008***
pm	509	0.650	0.677	0.027***
leverage	509	0.075	0.077	0.002*
salesgrowth	354	0.216	0.232	0.017*

Panel D: Intersection labor market peers

	Num. Obs	**GICS+SIC NAICS+TNIC**	**GICS+SIC NAICS+TNIC +LMP_INT**	
		Avg. R-squared	**Avg. R-squared**	**(3)-(2)**
	(1)	**(2)**	**(3)**	**(4)**
Expense ratios				
rdpersales	509	0.816	0.839	0.023***
sgapersales	509	0.706	0.739	0.034***
Valuation multiples				
pb	508	0.076	0.087	0.011
evs	508	0.722	0.733	0.011***
pe	384	0.079	0.085	0.006

Table 5 continued

Financial statement ratios				
rnoa	370	0.161	0.192	0.031***
roe	509	0.088	0.094	0.006**
at	509	0.898	0.906	0.008***
pm	509	0.639	0.672	0.033***
leverage	509	0.082	0.101	0.019
salesgrowth	354	0.213	0.238	0.025*

This table reports the average R^2 values from yearly cross-sectional regressions for accounting-based performance measure comovement tests based on the LMP measure and on alternative industry groupings from 2014 to 2019. Column (1) shows the number of firms involved in each regression. Column (2) reports the average R^2 from a yearly cross-sectional regression of the form $Ratio_{i,t} = \alpha_t + \beta_{1,t} Ratio_{GICS,t} + \beta_{2,t} Ratio_{SIC,t} + \beta_{3,t} Ratio_{NAICS,t} + \beta_{4,t} Ratio_{TNIC,t} + \epsilon_{i,t}$. The dependent variable is the financial ratio of the base firm, and the independent variables are the average financial ratios of the top-N firms in each industry grouping. N is the number of firms in the corresponding labor market peers. The closest peers were selected based on size and industry similarity (for GICS, SIC, and NAICS) and Hoberg-Phillips score (for TNIC). Each row shows a different financial ratio defined in Appendix C. Panel A column (3) adds the average financial ratio of the LinkedIn labor market peers to the regression form in column (2) and reports average R^2 values from yearly cross-sectional regressions. Column (4) reports the difference in average R^2 values between the regressions in columns (3) and (2) and reports significance based on two-tailed t-tests. Panels B to D report results based on the Glassdoor, union, and intersection labor market peers respectively. The sample includes S&P 1500 firms whose LinkedIn and Glassdoor peers overlap, and each have at least two publicly traded firms. I drop observations with missing total assets, long-term debt, net income before extraordinary items, operating income after depreciation, negative common or total equity, share prices smaller than $3 at the end of the fiscal year, and net sales smaller than $100 million. Significance levels are indicated by *, **, and *** for 10%, 5%, and 1% respectively

of firms for each estimation. Columns (2) and (3) report the average value of R^2 on yearly cross-sectional regressions for Eqs. 4 and 5, respectively. Column (4) reports two-tailed t-tests for the R^2 difference between columns (2) and (3) based on the time-series of R^2 from 2014 to 2019.

Table 5 shows that labor market peers exhibit significant incremental explanatory power over the aggregate power of GICS, SIC, NAICS, and TNIC in explaining cross-sectional variation in accounting-based performance metrics, such as expense ratios and valuation multiples. For example, adding LinkedIn peer portfolio helps explain a significantly greater proportion of the cross-sectional variation for expense ratios (*rdpersales*, *sgapersales*), valuation multiples (*pb*, *evs*), and financial statement ratios (*rnoa*, *at*, *pm*, *leverage*). The increases for LinkedIn labor market peers are statistically significant at the 1% level for *rdpersales, sgapersales, evs, at, pm*, at the 5% level for *rnoa*, and at the 10% level for *pb, leverage*. The results for the Glassdoor, union, and intersection peers in Panels B to D are also similar, further supporting the robustness of the findings.

In sum, the LMP measure increases the explanatory power of existing industry-peer portfolios in explaining cross-sectional accounting-based performance ratios. If GICS, SIC, NAICS, and TNIC peer portfolios together capture common output market shocks, the results in Tables 4 and 5 illustrate that the LMP measure captures important common input-market shocks that are not reflected by existing industry measures.

4.3 Cross-sectional evidence

4.3.1 Return comovements by skill similarity

In this section, I provide cross-sectional evidence to examine whether the LMP measure captures labor market relatedness. If it does capture firms' labor market similarities, I expect labor market peers to explain more cross-sectional variation in returns when firms share more skills with their peers and are thus exposed to more labor shocks and risks, as illustrated in the stylized model in Appendix A. Additionally, I expect labor market peers to outperform other industry measures when firms are more closely connected with their peers in the labor market.

Specifically, I sort the sample into terciles based on the proportion of shared skills between the base firm and its LinkedIn labor market peers. I use skill similarity (defined in Eq. 1), as a proxy for firms' labor market shocks and risks. Because the skill similarity is based on LinkedIn data, I examine the performance of only LinkedIn labor market peers and the corresponding GICS, SIC, NAICS, and TNIC peers in this section. The bottom tercile shares an average of 2.5% skills with their LinkedIn peers, the median shares an average of 10.4%, and the top tercile shares an average of 23.2%.

Columns (1) to (3) of Panel A of Table 6 report the results of estimating Eq. 2 every month from 2014 to 2019, using LinkedIn peer portfolios for each skill similarity group. As firms share more skills with their peers, the average R^2 increases from 4.7% for the bottom tercile, to 11.6% for the median tercile, and then to 22.5% for the

Table 6 Return comovements tests by skill similarity

Panel A: LinkedIn labor market peers			
	(1)	**(2)**	**(3)**
LMP_LN	0.31***	0.49***	0.73***
	(16.30)	(23.98)	(33.49)
Observations	17,685	17,635	17,652
Avg. R-squared	0.047	0.116	0.225
Number of groups	72	72	72
Skill similarity	Low	Medium	High
Panel B: GICS			
	(1)	**(2)**	**(3)**
GICS_LN	0.27***	0.38***	0.57***
	(15.95)	(20.70)	(25.26)
Observations	17,685	17,635	17,652
Avg. R-squared	0.052	0.099	0.182
Number of groups	72	72	72
Skill similarity	Low	Medium	High

Table 6 continued

Panel C: SIC

	(1)	(2)	(3)
SIC_LN	0.22***	0.31***	0.50***
	(15.62)	(17.50)	(23.50)
Observations	17,685	17,635	17,652
Avg. R-squared	0.039	0.082	0.159
Number of groups	72	72	72
Skill similarity	Low	Medium	High

Panel D: NAICS

	(1)	(2)	(3)
NAICS_LN	0.22***	0.35***	0.51***
	(15.69)	(22.55)	(21.42)
Observations	17,685	17,635	17,652
Avg. R-squared	0.039	0.094	0.162
Number of groups	72	72	72
Skill similarity	Low	Medium	High

Panel E: TNIC

	(1)	(2)	(3)
TNIC_LN	0.22***	0.36***	0.53***
	(16.22)	(22.02)	(23.55)
Observations	17,685	17,635	17,652
Avg. R-squared	0.041	0.089	0.170
Number of groups	72	72	72
Skill similarity	Low	Medium	High

This table reports the average monthly cross-sectional regressions of return comovement tests, $R_{i,t} = \alpha_t + \beta_t R_{peer,t} + \epsilon_{i,t}$, for each skill similarity group. I sort the sample into terciles based on the proportion of shared labor skills between the base firm and its labor market peers. The dependent variables are the base firm's monthly stock returns. From Panels A to D, the independent variable is the equally weighted contemporaneous average return of LinkedIn labor market peers, closest GICS peers, SIC peers, NAICS peers, and TNIC peers with the same number of LinkedIn labor market peers. Firms in column (1) are in the bottom tercile and share the least proportion of labor skills with their labor market peers. Firms in Column (2) are in the middle tercile. And firms in Column (3) are in the top tercile. The time-series average of monthly cross-sectional regression coefficients and the R-squared values are both reported. The t-statistics of coefficient estimates are reported in parentheses, where *, ** and *** denote significance at the 10% , 5% , and 1% significance levels, respectively. The sample includes the S&P 1500 sample that has at least two publicly traded LinkedIn labor market peer firms between 2014 and 2019 top tercile. Two-tailed t-tests indicate that these differences are significant at the 1% level.

Next I examine the relative performance of the LMP measure, compared with traditional industry groupings, for each skill similarity group. Product- and labor-market peers may overlap, so they may capture common risks and shocks. But when a firm shares more skills with its peers, it is likely exposed to more labor market shocks and risks, and thus peer measures that capture more common labor market shocks and risks will outperform other industry measures in explaining return variation of the base firm.

In Table 6, for each tercile, I estimate Eq. 2 separately for GICS, SIC, NAICS, and TNIC portfolios formed by the same number of closest industry peers. Panels B to E show the results of the comparison. The average R^2 values for each tercile are also presented in Fig. 3. When firms share the least amount of labor skills, the explanatory power of the LMP measure is not significantly different from other industry portfolios. However, as firms share more labor skills, the measure significantly outperforms other industry measures. The sharpest contrast comes from comparing the average R^2 for the high skill-similarity group. The LMP measure explains 22.5% of the cross-sectional

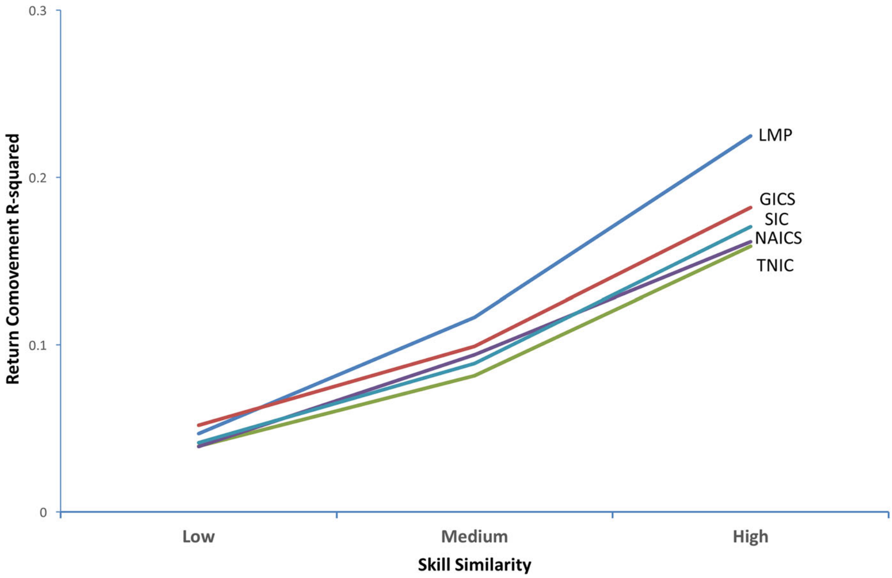

Fig. 3 Comparison of return comovement R^2 values among the LMP measure and alternative industry groupings by skill similarity.This figure reports the average R^2 values from monthly cross-sectional regressions of the form $R_{i,t} = \alpha_t + \beta_t R_{Ind,t} + \epsilon_{i,t}$, for different industry-peer and skill-similarity groups from 2014 to 2019. The dependent variables are the base firm's monthly stock returns. The independent variable is the equally weighted contemporaneous average return of LinkedIn labor market peer, closest GICS, SIC, NAICS, and TNIC peer firms. I sort the sample into terciles based on the average skill similarity between the base firm and its labor market peers. Skill similarity $s_{i,j}$ is defined in Eq. 1 and is measured as the number of skills shared by firm i and firm j divided by the number of total skills of firm i and firm j. High skill similarity implies more shared labor skills between the base firm and its peers. The horizontal line is the skill-similarity group, and the vertical line is the average R^2 values. The sample includes S&P 1500 firms with at least two publicly traded LinkedIn labor market peer firms

returns, while GICS, SIC, NAICS, and TNIC explain 18.2%, 15.9%, 16.1%, 17.0% of the cross-sectional returns respectively. The R^2 differences between the LMP measure and these alternative industry measures are significant at the 1% level based on two-tailed t-tests.[19]

4.3.2 Salary comovements by skill similarity

One direct test of firms' labor market similarities is the similarities of employee salaries. Firms that share similar labor markets should provide comparable salaries. If the LMP measure captures firms' labor market peers, I expect it to explain more of the base firm's salary when that firm and its labor market peers share more skills. I also expect the measure to outperform other industry measures at explaining employee salaries when the base firm and its labor market peers share more skills. I use median employee salary as a proxy for firm salary. For the year 2014 to 2016, I use Glassdoor salary review data to construct annual median employee salary at the firm level. Since 2017, SEC has required firms to disclose the ratio of the compensation of its chief executive officer (CEO) to the median compensation of its employees. For the years 2017 to 2019, I obtain the pay ratio and CEO compensation data from the Equilar database to compute the median employee salary.[20]

Columns (1) to (3) of Panel A of Table 7 report the results of regressing annual employee median salary on the average median salary of LinkedIn labor market peers every year from 2014 to 2019 for each skill similarity group. The decrease in the number of observations was significant because these tests require all firms to have available median employee salary data. From columns (1) to (3), the average R^2 increases from 22.3% to 40.9% and then to 57.9% as firms start sharing more labor skills with their peers. Two-tailed t-tests show that the differences are significant at the 1% level.

Next I compare the performance of the LMP measure with traditional industry groupings for each skill similarity group. Similar to the return comovement tests by skill similarity group, when a firm shares more labor market skills with its peers, I expect peer measures that capture more common labor market shocks and risks to outperform other industry measures at explaining salary variation of the base firm.

In Table 7, for each tercile, I estimate the explanatory power separately for GICS, SIC, NAICS, and TNIC portfolios formed by closest peers as the same number of LinkedIn peers. Panels B to D show the results for comparison. Figure 4 presents the average R^2 values for each regression and tercile. When firms share the fewest skills, the explanatory power of the LMP measure is not significantly different from other industry portfolios. It significantly outperforms other industry measures for the median skill-similarity group. For the high skill-similarity group, it explains 57.9% of the cross-sectional returns, compared to 44.7%, 40.8%, 45.3%, 47.8% for GICS, SIC, NAICS, and TNIC portfolios respectively. The R^2 differences are significant at

[19] Note that the average R^2 produced by output-market-based industry portfolios exhibits a similar pattern as LMP. This is likely due to the overlap among different industry classifications.

[20] This method may have limitations, as firms hire employees with different occupations and each occupation may have a different labor market. Thus results in this section should be viewed as initial evidence, and future research might examine more granular labor markets or use more detailed salary data.

Table 7 Salary comovements test based on the LMP measure by skill similarity

Panel A: LinkedIn labor market peers

	(1)	**(2)**	**(3)**
Salary_LMP_LN	0.54***	0.75***	0.89***
	(10.29)	(35.06)	(39.46)
Observations	1,232	1,296	1,293
Avg. R-squared	0.223	0.409	0.579
Number of groups	6	6	6
Skill similarity	Low	Medium	High

Panel B: GICS

	(1)	**(2)**	**(3)**
Salary_GICS_LN	0.65***	0.69***	0.75***
	(14.99)	(64.86)	(12.95)
Observations	1,232	1,296	1,293
Avg. R-squared	0.288	0.300	0.447
Number of groups	6	6	6
Skill similarity	Low	Medium	High

Panel C: SIC

	(1)	**(2)**	**(3)**
Salary_SIC_LN	0.61***	0.63***	0.71***
	(8.29)	(21.20)	(22.09)
Observations	1,232	1,296	1,293
Avg. R-squared	0.234	0.300	0.408
Number of groups	6	6	6
Skill similarity	Low	Medium	High

Panel D: NAICS

	(1)	**(2)**	**(3)**
Salary_NAICS_LN	0.56***	0.57***	0.75***
	(7.19)	(52.49)	(21.96)
Observations	1,232	1,296	1,293
Avg. R-squared	0.213	0.275	0.453
Number of groups	6	6	6
Skill similarity	Low	Medium	High

Table 7 continued

Panel E: TNIC

	(1)	**(2)**	**(3)**
Salary_TNIC_LN	0.61***	0.69***	0.74***
	(11.84)	(33.78)	(15.44)
Observations	1,232	1,296	1,293
Avg. R-squared	0.329	0.396	0.478
Number of groups	6	6	6
Skill similarity	Low	Medium	High

This table reports the average monthly cross-sectional regressions of salary comovement tests, $Salary_{i,t} = \alpha_t + \beta_t\ Salary_{peer,t} + \epsilon_{i,t}$, for each skill similarity group. I sort the sample into terciles based on the proportion of shared labor skills between the base firm and its LinkedIn labor market peers. The dependent variables are the base firm's annual median employee salary. From Panels A to D, the independent variable is the equally weighted contemporaneous average salary of LinkedIn labor market peers, GICS peers, SIC peers, NAICS peers, and TNIC peers closet in size, and with the same number of firms as LinkedIn labor market peers. Firms in column (1) are in the bottom tercile and share the fewest skills with their labor market peers. Firms in column (2) are in the middle tercile. And firms in Column (3) are in the top tercile. The time-series average of annual cross-sectional regression coefficients and the R-squared values are both reported. The t-statistics of coefficient estimates are reported in parentheses, where *, ** and *** denote significance at the 10%, 5%, and 1% significance levels, respectively. The sample includes the S&P 1500 sample that has at least two publicly traded LinkedIn labor market peer firms between 2014 and 2019

the 1% level between the LMP measure and GICS, SIC, NAICS peers and at the 5% level for the TNIC peers.

Overall the results show that the LMP measure explains a higher proportion of return and salary comovements when firms are closely connected in the labor market. The measure outperforms traditional measures when firms share more labor skills. These findings allay concerns that labor market peers capture closely related peers that have no connection in the labor market (e.g., an alternative explanation is that labor market peers reflect general interest among LinkedIn and Glassdoor users but have no relationship to a firm's labor market). Under this alternative explanation, it is unlikely to see a systematic relationship between skill similarity and the performance of the LMP measure.

5 Additional analyses and robustness tests

5.1 Investigating measurement error and selection bias

5.1.1 Measurement error from customer and supplier relationships

As discussed in Section 3.1.3, some users use LinkedIn for lead generation, and thus labor market peers may reflect firms' customer and supplier relationships. It is unlikely that Glassdoor labor market peers are formed through lead generation because Glassdoor reviews are anonymous and users cannot contact each other. However, for the

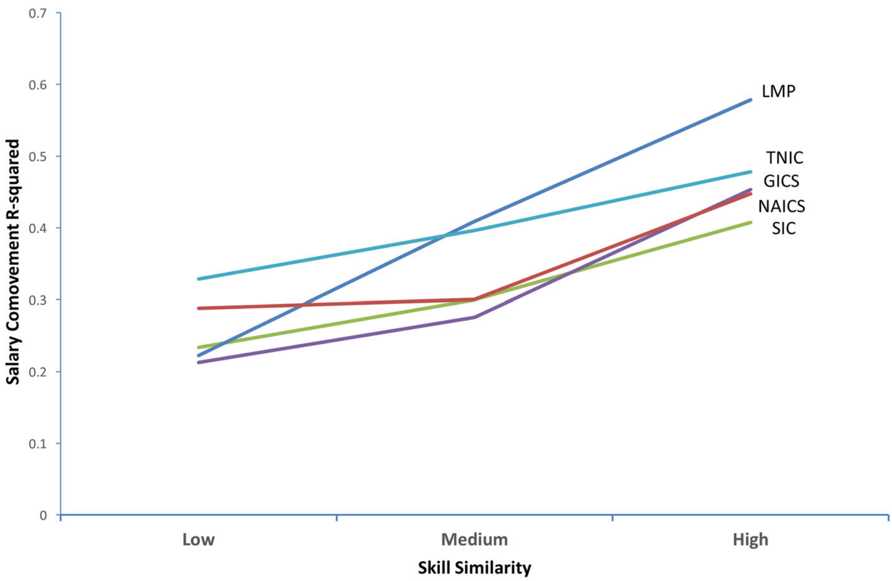

Fig. 4 Comparison of salary comovement R^2 values among the LMP measure and alternative industry groupings by skill similarity. This figure reports the average R^2 values from monthly cross-sectional regressions of the form $Salary_{i,t} = \alpha_t + \beta_t Salary_{Ind,t} + \epsilon_{i,t}$, for different industry-peer and skill-similarity groups from 2014 to 2019. The dependent variables are the base firm's annual median employee salary. The independent variable is the equally weighted contemporaneous average salary of LinkedIn labor market peer, closest GICS, SIC, NAICS, and TNIC peer firms. I sort the sample into terciles based on the average skill similarity between the base firm and its labor market peers. Skill similarity $s_{i,j}$ is defined in Eq. 1, and is measured as the number of skills shared by firm i and firm j divided by the number of total skills of firm i and firm j. High skill similarity implies more shared labor skills between the base firm and its peers. The horizontal line is the skill-similarity group, and the vertical line is the average R^2 values. The sample includes S&P 1500 firms with at least two publicly traded LinkedIn labor market peer firms

completeness of the analysis, I will examine the four sets of labor market peers as outlined in previous sections.

According to SFAS No.131, firms must disclose the existence of and sales to customers representing more than 10% of total firm revenues (principal customers). Therefore I classify two firms as having a customer-supplier relationship if one firm is listed as the other firm's principal customer. Information on customer-supplier relationships is based on the Compustat Segments Customer File (Cen et al. 2018). I exclude these customer or supplier firms from the labor market peer sample and replicate columns (6) and (7) of Table 4. The significance level of two-tailed t-test of the R^2 difference is also reported. Results are presented in Table 8 Panel A and are qualitatively unchanged.

5.1.2 Measurement error from geographic proximity

Labor market peers may be local firms because geographic distance is a natural component in career considerations (Marinescu and Rathelot 2018) and firms tend to cluster to attract talent (Pouder and St. John 1996). I do not consider these firms as measurement error, but using LinkedIn as a communication channel with local friends and viewing their employer with zero career-related interest may cause measurement error. Although it is difficult to tease out this type of noise from local firms, I form a portfolio of geographically close firms ($R_{geopeer,t}$) for each base firm to control for local firms and local shocks these firms represent. I then examine the incremental return explanatory power of labor market peers over local peers and traditional industry peers. Specifically, I use location data from Glassdoor to identify the metropolitan statistical areas from which employees submit reviews (Dehaan et al. 2023). A firm is considered a local peer if it shares at least one metropolitan statistical area as a work location with the base firm between 2014 to 2019. Results are presented in Table 8 Panel B and show that labor market peers still have significant incremental explanatory power after controlling for local peer returns.

5.1.3 Selection bias based on firm size

As discussed in Section 3.1.4, the amount of web traffic and informativeness of LinkedIn or Glassdoor web content may affect the composition of labor market peers. These features likely relate to firm size, and I divide the sample into two groups by base firm size to address this issue. Big firms may attract more users on LinkedIn and Glassdoor, providing a stronger signal to construct labor market peers. However, small

Table 8 Robustness tests: comparison of R^2 values

Panel A: Excluding customer and supplier firms			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.187	0.218***	0.193	0.210***
Union		**Intersection**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.196	0.218***	0.154	0.190***
Panel B: Control for local peer firms			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.171	0.202***	0.177	0.195***
Union		**Intersection**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.179	0.202***	0.146	0.180***
Panel C: Small firms			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.164	0.191***	0.172	0.190***

Table 8 continued

Union		Intersection	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.177	0.198***	0.133	0.164***
Panel D: Large firms			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.246	0.289***	0.248	0.274***
Union		**Intersection**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.246	0.278***	0.223	0.270***
Panel E: Closest industry peers matched by performance and size			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.188	0.221***	0.197	0.217***
Union		**Intersection**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.202	0.226***	0.152	0.192***
Panel F: Closest industry peers matched by performance			
LinkedIn		**Glassdoor**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.172	0.205***	0.173	0.194***
Union		**Intersection**	
Standard industries	**Standard industries + LMP**	**Standard industries**	**Standard industries + LMP**
0.174	0.202***	0.165	0.192***

Panel G: Size quartile matched industry groups				
Standard industries	**(1)+LMP_LN**	**(1)+LMP_GD**	**(1)+LMP_UNI**	**(1)+LMP_INT**
(1)	**(2)**	**(3)**	**(4)**	**(5)**
0.199	0.226***	0.214***	0.221***	0.220***
Panel H: Granular industry classification				
Standard industries	**(1)+LMP_LN**	**(1)+LMP_GD**	**(1)+LMP_UNI**	**(1)+LMP_INT**
(1)	**(2)**	**(3)**	**(4)**	**(5)**
0.174	0.211***	0.197***	0.207***	0.200***

Table 8 continued

Panel I: R^2 by base firm's industry

	Base	**(1) + LMP_LN**	**Base**	**(3) + LMP_GD**	**Base**	**(5) + LMP_UNI**	**Base**	**(7) + LMP_INT**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Energy	0.284	0.329***	0.283	0.326***	0.273	0.316***	0.268	0.331***
Health Care	0.153	0.192***	0.147	0.177***	0.154	0.187***	0.162	0.192***
Financial	0.186	0.218***	0.192	0.212***	0.190	0.213***	0.152	0.188***
Information Technology	0.123	0.158***	0.122	0.145***	0.121	0.150***	0.111	0.137***
Consumer Discretionary	0.122	0.153***	0.127	0.147***	0.127	0.152***	0.109	0.140***
Consumer Staples	0.235	0.292***	0.227	0.279***	0.222	0.278***	0.238	0.292***
Materials	0.324	0.414***	0.312	0.369***	0.304	0.370***	0.302	0.393***
Industrials	0.165	0.197***	0.163	0.190***	0.159	0.186***	0.145	0.181***
Communication Services	0.464	0.559***	0.439	0.534***	0.422	0.531***	0.458	0.541***
Utilities	0.189	0.241***	0.208	0.253***	0.195	0.246***	0.179	0.219***

This table tests the robustness of the main results in Table 4 columns (6) to (7). Panel A reports results for excluding the base firm's customer or supplier firm from the labor market peers. The closest industry peers are adjusted accordingly if the number of corresponding labor market peer firms change. In the interest of space, I only report the average R^2 values before and after adding the labor market peer portfolio and the significance of the two-tailed t-test of the difference between the two R^2. Panel B reports results controlling for the average return of local peers. Local peers are defined as firms that share at least one work location with the base firm. I use metropolitan statistical areas (MSAs) from which Glassdoor employee reviews are submitted to identify firm locations. In Panels C and D, I sort the sample into two groups based on base firm size (market cap at the beginning of the year) and present results for small and large firms respectively. Panels E to F report results using different ways of classifying closest peers within traditional industry classification. Panel E selects industry peers that are in the same size quartile as the base firm and are closest in performance, as proxied by ROA. Panel F selects industry peers that are closest in performance, as proxied by ROA. Panel G reports the results for size-industry-matched peers following Albuquerque (2009). Panel H reports results based all firms within the same most granular level of traditional industry peers: eight-digit GICS, four-digit SIC, six-digit NAICS, and TNIC industry peers. The first column reports the average R^2 values from monthly cross-sectional regressions of the following form:

$$Ratio_{i,t} = \alpha_t + \beta_{1,t}\ Ratio_{GICS,t} + \beta_{2,t}\ Ratio_{SIC,t} + \beta_{3,t}\ Ratio_{NAICS,t} + \beta_{4,t}\ Ratio_{TNIC,t} + \epsilon_{i,t}.$$

The next column reports the average R^2 values when adding the LinkedIn, Glassdoor, union and intersection labor market peers, respectively. Panel I reports results by the two-digit GICS classification of base firms in the S&P 1500 universe (excluding real estate due to data availability). T-tests for the significance of the differences in average R^2s values between the two regressions are reported. *, ** and *** denote significance at the 10%, 5%, and 1% significance levels, respectively. The sample includes S&P 1500 firms with LinkedIn and Glassdoor company homepages that have at least two publicly traded labor market peers between 2014 and 2019 firms can still attract users because there may be less alternative information available from other sources, making it unclear which type of firms better captures labor market peers. Panels C and D of Table 8 replicate columns (6) and (7) of Table 4 for small and large firms respectively. The result suggests that labor market peers capture labor market relatedness, regardless of a firm's size. That said, the effect of website content and usage on the LMP measure is an endogeneity threat that I cannot eliminate.

5.1.4 Selection bias by industry

Lastly, some industries are better represented on LinkedIn and Glassdoor than others, and employees in some industries may use these platforms more frequently, leading to selection bias. Although I cannot eliminate this selection bias, I examine whether the results are industry-specific. I classify each base firm in the S&P 1500 sample by its two-digit GICS code and rerun the same specification in Table 4 columns (6) to (7) for each two-digit GICS industry sector (excluding real estate, due to lack of observations). Table 8 Panel I reports the results. Labor market peers significantly increase the explanatory power of stock returns for all two-digit GICS industries, with an increase in R^2 ranging from 2.2% for industrial firms to 7.4% for communication services firms.

5.2 Alternative ways of forming benchmarking industry peers

This section examines the robustness of the main results in Table 4 Panels A to D columns (6) to (7) to alternative ways of forming industry peers as benchmarks for LinkedIn, Glassdoor, union, and intersection labor market peers.

5.2.1 Top-N peer firms with the closest size and/or performance

In this section, the robustness of the main results is tested by using two alternative methods to select a base firm's top-N closest industry (GICS, SIC, or NAICS) peers to form benchmark industry peers. The first alternative approach is to select top-N firms that are closest in performance and in the same industry as the base firm (Kothari et al. 2005). I use beginning-of-year ROA as a proxy for performance. The second alternative approach is to classify top-N peers as firms that are closest in size and performance. Specifically, I choose firms that are in the same industry and size quartile and closest in performance as the base firm. I select the top-N TNIC peers with the highest Hoberg-Phillips score data, as in Section 4.1. Both of these alternative methods are used to replicate Table 4 columns (6) to (7) for the corresponding industry peers, and the results are presented in Table 8 Panels E and F. The R^2 values increase for all estimations, and the significance of the R^2 differences remains at the 1% level.

5.2.2 Comparison against industry-size peers

In this section, I relax the restriction on the number of firms to form an alternative industry peers. This is done to address concerns that averaging a smaller number of peers may bias the results in favor of or against alternative industry groupings. While averaging a smaller number of firms may perform better if the peers are more relevant, this approach may perform worse at cancelling out the idiosyncratic component in stock returns. This section aims to explore this issue further.

I first apply the approach based on the findings of Albuquerque (2009), who show that industry-size-matched peer groups perform better at capturing peers affected by similar shocks than industry-matched peer groups. To implement this approach, firms are sorted into size quartiles within a two-digit SIC industry and matched with firms in the same industry-size peer group (excluding the firm itself) to compute an equally weighted portfolio return. This method is used to generate size-adjusted industry peers for six-digit GICS industry, two-digit SIC industry, and three-digit NAICS. I then rerun Table 4 columns (6) to (7) for the corresponding size-adjusted industry peers. Table 8 panel G presents the results. The R^2 differences remain significant and large.

5.2.3 Comparison against benchmark industry peers formed by all industry firms

In this section, the incremental performance of labor market peers is compared to benchmark peers formed by all firms within an industry at the industry's most granular level. To do this, the base firm is matched with its eight-digit GICS, four-digit SIC, and six-digit NAICS industry peers and an equally weighted portfolio return is computed. I rerun Table 4 columns (6) to (7) for corresponding peers. The results are presented in Table 8 Panel H, and the R^2 differences remain significant and are of similar magnitude to the main results.

5.3 The effects of online labor markets on labor market connections

In this section, I explore an alternative explanation that the "also viewed" feature on LinkedIn and Glassdoor enables job seekers to discover companies previously unknown to them, leading to new labor market connections and return comovements, rather than simply reflecting a company's existing connections in the labor market.

To test this alternative explanation, I examine the performance of labor market peers prior to the introduction of "also viewed" functions on online labor market platforms. As both Glassdoor and LinkedIn were founded in 2007, "also viewed" functions likely became available online between 2007 and 2013. I analyze labor market peer performance during sample periods between 2004 to 2006, 2007 to 2009, 2010 to 2012, and 2013 to 2015, when the influence of the online labor market on the real labor market was likely limited. Under this alternative explanation, labor market peers should not comove before the introduction of "also viewed" firms feature on LinkedIn and Glassdoor.

The results, shown in Table 9, indicate that labor market peers still demonstrate significant added value over traditional industry classifications in explaining stock returns, suggesting that the introduction of "also viewed" algorithms is not an impetus for the labor market peer performance. However, this does not rule out the possibility

Table 9 Robustness tests: different sample periods

Panel A: LinkedIn labor market peers								
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Year	**2004-2006**		**2007-2009**		**2010-2012**		**2013-2015**	
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_LN		0.264***		0.267***		0.294***		0.287***
		(17.164)		(23.200)		(17.052)		(20.060)
GICS_LN	0.203***	0.157***	0.180***	0.138***	0.183***	0.136***	0.180***	0.138***
	(16.379)	(12.615)	(16.261)	(12.872)	(15.717)	(12.419)	(12.634)	(10.066)
SIC_LN	0.049***	0.030**	0.100***	0.083***	0.076***	0.059***	0.078***	0.053***
	(4.046)	(2.560)	(9.382)	(8.263)	(6.458)	(4.966)	(5.839)	(3.795)
NAICS_LN	0.113***	0.090***	0.062***	0.038***	0.084***	0.054***	0.106***	0.080***
	(9.111)	(7.983)	(5.640)	(3.589)	(7.873)	(4.978)	(6.929)	(4.715)
TNIC_LN	0.238***	0.175***	0.227***	0.161***	0.228***	0.158***	0.214***	0.150***
	(18.519)	(13.802)	(17.942)	(15.046)	(20.594)	(15.719)	(13.696)	(12.599)
Observations	29,194	29,194	32,285	32,285	33,329	33,329	31,596	31,596
Avg. R-squared	0.133	0.155	0.128	0.152	0.117	0.146	0.142	0.168
Number of groups	36	36	36	36	36	36	36	36
Panel B: Glassdoor labor market peers								
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Year	**2004-2006**		**2007-2009**		**2010-2012**		**2013-2015**	
VARIABLES	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**	**ret**
LMP_GD		0.268***		0.238***		0.279***		0.274***
		(15.708)		(15.661)		(17.727)		(15.335)

Table 9 continued

GICS_GD	0.213***	0.178***	0.214***	0.185***	0.204***	0.175***	0.204***	0.176***
	(15.178)	(12.505)	(17.589)	(15.661)	(16.611)	(14.798)	(11.728)	(10.430)
SIC_GD	0.079***	0.067***	0.090***	0.079***	0.065***	0.050***	0.087***	0.074***
	(5.855)	(5.336)	(7.438)	(6.667)	(5.317)	(4.014)	(5.876)	(5.050)
NAICS_GD	0.096***	0.078***	0.060***	0.044***	0.097***	0.083***	0.106***	0.092***
	(7.358)	(6.291)	(4.685)	(3.496)	(8.871)	(7.890)	(6.581)	(6.100)
TNIC_GD	0.251***	0.208***	0.253***	0.216***	0.249***	0.212***	0.221***	0.183***
	(16.658)	(14.563)	(17.619)	(15.288)	(19.292)	(16.784)	(15.618)	(14.924)
Observations	29,386	29,386	32,584	32,584	33,438	33,438	31,627	31,627
R-squared	0.131	0.145	0.134	0.147	0.119	0.135	0.143	0.157
Number of groups	36	36	36	36	36	36	36	36

This table reports the average of several monthly cross-sectional regressions of return comovement tests based on labor market peers and alternative industry groupings for different sample periods. The sample period is from 2004 to 2006 in columns (1) and (2), from 2007 to 209 in columns (3) and (4), from 2010 to 2012 in columns (5) and (6), and from 2013 to 2015 in columns (7) and (8). The dependent variables are the base firm's monthly stock returns. In columns (1), (3), (5), (7), the independent variable is the equally weighted contemporaneous average return of LMP, GICS, SIC, NAICS, and TNIC peers formed with the top-N number of industry peers closest in size with the base firm, where N is the number labor market peer firms. Columns (2), (4), (6), (8) include peer portfolio returns from the LMP measure. The time-series average of monthly cross-sectional regression coefficients and average R-squared values are reported. Results for LinkedIn and Glassdoor sample are shown in Panels A and B respectively. The t-statistics of coefficient estimates are reported in parentheses, where *, **, and *** denote significance at the 10%, 5%, and 1% significance levels, respectively. The sample includes S&P 1500 firms, whose LinkedIn and Glassdoor company homepages have at least two publicly traded labor market peers that the "also viewed" feature could strengthen labor market connections and enhance comovements. The impact of the "also viewed" function on the real labor market remains an open area of investigation for further research.

5.4 Additional analyses

The online appendix discusses additional analyses and sensitivity tests. First, I find that labor market peers are more likely to be selected as compensation benchmarking peers and help explain CEO compensation. Second, I show that the results are robust to the use of a larger sample without requiring overlap between Glassdoor and LinkedIn labor market peers. Third, I find the results are robust to the use of a pooled regression design.

6 Conclusion

This study offers a new measure of labor market peers (LMP) by analyzing the "also viewed" companies on online labor market platforms such as Glassdoor and LinkedIn. The results suggest that labor market peers reveal important economic relationships that are not captured by traditional industry classifications and can provide valuable insights, particularly when firms are closely connected in the labor market.

I see this study as an initial step in the growing field of using online labor market data for economic and financial analysis. Future research can test the labor market peer construct using different methods, explore differences between online labor market platforms, and investigate how website content and usage impact the composition of labor market peers. Additionally, future research can examine the impact of financial events on labor market peers and the real effects of online labor market platforms on labor market and financial outcomes. Overall there is ample opportunity for further research on firms' financial performance and their labor market connections through the lens of online labor market platforms.

Appendix A: A stylized model of labor market peers and earning comovements

In this section, I propose a stylized model in which firms that hire common types of labor in limited supply are exposed to similar labor market shocks, which can affect wages and ultimately firm profits. The model also suggests that the strength of earnings comovements will increase as firms share more types of labor.

The stylized model represents two firms, 1 and 2. They share the same types of specific labor in the production process. Specific labor is in limited supply in the short run. To keep the focus on the dynamics of labor, capital is assumed away.[21] Outputs

[21] A simple extension of the model would be adding capital as an additional input. This does not affect the results of the stylized model.

of firm 1 and 2 are given by:

$$Y_1 = \sum_{s \in S_1} \beta_{1s} L_{1s}^{\alpha_s},$$
$$Y_2 = \sum_{s \in S_2} \beta_{2s} L_{2s}^{\alpha_s},$$

where S_1 and S_2 are the labor skills used by firm 1 and 2, respectively; L_{1s} and L_{2s} denote the mass of employed labor with skill s, $0 < \alpha_s < 1$ so that firms' production functions are decreasing returns to scale; and β_{1s} and β_{2s} are parameters denoting production technologies.

The firm's profit is $\pi_i = p_i Y_i - \sum_{s \in S_i} w_s L_{is}$, where $i = 1, 2$. Each firm chooses labor to maximize profit. Perfect competition in the labor market drives firms to equate the marginal profitability of employed skills to wages. Hence firms' profits are:

$$\pi_1^* = \sum_{s \in S_1} (\alpha_s p_1 \beta_{1s})^{\frac{1}{1-\alpha_s}} ((\alpha_s p_1)^{-1} - 1) w_s^{\frac{\alpha_s}{\alpha_s - 1}},$$

$$\pi_2^* = \sum_{s \in S_2} (\alpha_s p_2 \beta_{2s})^{\frac{1}{1-\alpha_s}} ((\alpha_s p_2)^{-1} - 1) w_s^{\frac{\alpha_s}{\alpha_s - 1}},$$

where p_1 and p_2 are product prices for firm 1 and 2, w_s is the wage for labor skill s, and demand for labor skill s are $L_{1s} = (\frac{w_s}{\alpha_s p_1 \beta_{1s}})^{\frac{1}{\alpha_s - 1}}$ and $L_{2s} = (\frac{w_s}{\alpha_s p_2 \beta_{2s}})^{\frac{1}{\alpha_s - 1}}$. I assume that there are barriers to entering the product market in the short run, and thus firms have nonzero profits.

The supply of labor with skill s is assumed to be $L^S = w_s^{b_s}$, where the supply increases with wages and with labor supply elasticity ($b_s > 0$). The supply of workers with skill s is more elastic when b_s is higher, suggesting that a small increase in wages leads to a large increase in labor supply. It takes time for workers to acquire new skills, and so I assume that labor supply is inelastic and that labor markets for different skills are segmented in the short run.

Labor markets are in equilibrium when the demand for skill s equals its supply, and $L_{1s} + L_{2s} = L^S \epsilon_s$, where ϵ_s is a demand or a supply $i.i.d.$ shock to the labor market. When the labor market clears ($L^S = L^D$), wages per unit of skill s are endogenously determined in equilibrium:

$$w_s = [\alpha_s^{\frac{1}{\alpha_s - 1}} ((p_1 \beta_{1s})^{\frac{1}{1-\alpha_s}} + (p_2 \beta_{2s})^{\frac{1}{1-\alpha_s}})^{-1} \epsilon_s]^{\frac{\alpha_s - 1}{1 + (1-\alpha_s) b_s}}.$$

Proposition 1: Firms' earnings comove when they hire similar types of scarce labor.

Proof: The profits (earnings) of firm 1 and 2 comove due to shocks to shared labor markets. The comovement is represented by the covariance of firms 1's and 2's earnings:

$$cov(\pi_1^*, \pi_2^*) = \sum_{s \in S_1 \cap S_2} \lambda_s var(w_s^{\frac{\alpha_s}{\alpha_s - 1}}) = \sum_{s \in S_1 \cap S_2} \lambda_s var[(\gamma_s \epsilon_s)^{\frac{\alpha_s}{1+(1-\alpha_s)b_s}}],$$

where $\lambda_s = (\alpha_s^2 p_1 p_2 \beta_{1s} \beta_{2s})^{\frac{1}{1-\alpha_s}} ((\alpha_s p_1)^{-1} - 1)((\alpha_s p_2)^{-1} - 1)$ and $\gamma_s = \alpha_s^{\frac{1}{\alpha_s - 1}} ((p_1 \beta_{1s})^{\frac{1}{1-\alpha_s}} + (p_2 \beta_{2s})^{\frac{1}{1-\alpha_s}})^{-1}$, which are both positive parameters, and where $S_1 \cap S_2$ is the labor skills that firm 1 and 2 share. Thus the covariance of earnings for firms 1 and 2 is the summation of the wage variances in shared labor markets. The more skills the two firms share, the stronger the covariance, and thus we have the following proposition.

Proposition 2: Earning comovements strengthen when firms share more labor skills.

Appendix B: LinkedIn and Glassdoor comparison

I compare the LinkedIn with the Glassdoor sample in this appendix. The table below shows the average firm characteristics of the LinkedIn and Glassdoor sample. Glassdoor peers tend to be larger and have more employees. LinkedIn peers tend to have a higher median employee salary of $67, 395, compared to $65, 365 of the Glassdoor peers (see Section 4.3.2 for the computation of median employee salary). To analyze the geographic relation between the base firm and its labor market peers, I use the location data from Glassdoor to track metropolitan statistical areas (MSAs) from which employee reviews are submitted. For each firm, I keep MSAs with more than 10 reviews and assume these MSAs to be the firms' work locations. On average, the base firm shares more MSAs with Glassdoor labor market peers (2.71) than with LinkedIn labor market peers (2.00), indicating that Glassdoor better captures the geographic proximity between firms.

A few factors might explain the differences in the composition of LinkedIn and Glassdoor labor market peers. First, the also-viewed firms on LinkedIn may reflect more of an employee network because an important engagement for LinkedIn members is through the employee network. Research shows that people with more connections are more likely to be associated with higher paid jobs (Chetty et al. 2022), and referral networks help match high-ability workers to high-paying firms (Burks et al. 2015; Schmutte 2015; Dustmann et al. 2016). So the labor market peers identified through LinkedIn may better reflect the labor market for higher-paid workers.

Second, Glassdoor is more likely to capture geographic distance between firms. First, the Glassdoor algorithm for my sample period is based on also-viewed jobs, while the LinkedIn algorithm is likely based on also-viewed companies. Users may be more likely to view job postings from firms located close to them Marinescu and Rathelot (2018). Additionally, the difference in global presence between LinkedIn and Glassdoor could also influence the labor market peers' differences, as LinkedIn's larger global presence may lead to more diverse labor market peers.[22]

Third, firms likely post more jobs on Glassdoor than on LinkedIn, so larger firms may be more likely to be selected as the "job seekers also viewed" firms on Glassdoor. For example, Southwest only has 13 job postings, most of which are data analysts, on LinkedIn. In contrast, there are 86 Southwest jobs posted on Glassdoor on the same day (August 22, 2022), including jobs such as assistant admin, customer service agent, etc. Thus Glassdoor peers are more likely to capture large firms, as Glassdoor labor market peers are based on the browsing histories of a greater variety of employees. However, I hesitate to draw strong inferences for the LinkedIn and Glassdoor labor market peer differences, and further investigation is needed to fully understand the reasons for these differences and their implications.

Variables	LinkedIn sample	Glassdoor sample	Difference
Percentage of Shared labor market peers	49%	35%	14%
Median Employee Salary	67,395	65,365	2,030***
Median Num. of Shared MSAs	2.00	2.71	-0.70***
Num Employee	66.16	122.37	-56.21***
Assets	115,320	184,851	-69,531***
Market Value	59,165	84,910	-25,745***

This table compares the average of key firm characteristics for the LinkedIn and Glassdoor sample that share at least one labor market peer. The percentage of shared labor market peers for each sample (Percentage of Shared labor market peers), the median employee salary (Median Employee Salary), the median number of MSAs shared by a base firm and its labor market peers (Median Num. of Shared MSA), the number of employees (Num. Employee, in thousands), total assets (Assets, in millions), the market value of equity (Market Value, in millions), the differences between the LinkedIn and Glassdoor sample for each firm characteristic, and the t-statistics of the differences are reported

22 As of 2022, over 78% of LinkedIn members are from outside the United States. Although most recent data for Glassdoor global presence is not readily available, a Glassdoor report in 2015 showed that approximately 25 percent of Glassdoor traffic came from outside the United States. LinkedIn statistics are from https://kinsta.com/blog/linkedin-statistics/. Glassdoor statistics are from https://www.glassdoor.com/about-us/glassdoor-launches-germany/. Accessed on December 22, 2022.

Appendix C: Variable definitions

Variable names	Variable definitions
Expense Ratios	
rdpersales	R&D expense (xrd)/net sales (sale)
sgapersales	SG&A expense (xsga)/net sales (sale)
Employment Ratios	
empgrowth	(one year ahead number of employees (emp_{t+1})-emp_t)/emp_t
Valuation Multiples	
pb	market cap/total common equity (ceq)
evs	(market cap+long-term debt(dltt))/net sales(sale)
pe	market cap/net income before extraordinary items (ib)
Financial Statement Ratios	
rnoa	net operating income after depreciation (oiadp)/ (property, plant, and equipment(ppent) + current assets (act) - current liabilities(lct))
roe	net income before extraordinary items(ib)/total common equity (ceq)
at	total assets (at) / net sales (sale)
pm	net operating income after depreciation (oiadp) / net sales (sale)
leverage	long term debt (dltt)/total stockholder's equity (seq)
salesgrowth	(one year ahead realized sales ($sale_{t+1}$)-$sale_t$)/$sale_t$

This table provides detailed definitions for the accounting-based performance variables used in the paper, including the corresponding Compustat item names

Appendix D: Example of labor market skill sets

Panel A: Google's LinkedIn-identified labor skill set

Top Skills & Expertise	**Related Skills**
Google Adwords	MSN AdCenter, Yahoo Search Marketing, Search Advertising Organic Search, Adsense, Paid Search Strategy, Google Merchant Center Google Ad Planner, Conversion Optimization, Keyword Research, Paid Search Campaigns, Marin Software, Google Website Optimizer, Kenshoo, Google Adwords Professional, Adgooroo Landing Page Optimization, Google Webmaster Tools, Search Analysis
Python	NumPy, Django, SciPy, SQLAlchemy, PyQt, Matplotlib, wxPython, Celery, NLTK, WSGI, Flask, CherryPy, Web2py, TurboGears, Pygame, Pylons, PyGTK, SCons, Zope, PyUnit
Machine Learning	Feature Selection, Semi-supervised Learning, Classifiers, Dimensionality Reduction, Graphical Models, Reinforcement Learning, Unsupervised Learning, Text Classification, Pattern Recognition, Recommender Systems, Natural Language Processing, Text Mining, Object Detection, Collaborative Filtering, SVM, Statistical Machine Translation, Mahout, Bayesian networks, NLTK, Natural Language

Adsense	DoubleClick for Publishers, Google Base, ADX, URL, DFP, Google Website Optimizer, Clickbank, Sitemaps, OpenX, Link Popularity Keyword Density Link Exchange, Google Ad Planner, Search Engine Submission, Off Page, Yahoo Site Explorer, Off-Page, Google Products XML, Word Tracker, Website Monetization
Google Technologies	Google Local, Google Products Local Search Optimization, Google Merchant Center, Yelp, Google Search, Google Search Appliance, Google Trends, Adgooroo, Google Base, URL, Off-Page, Del.icio.us Keyword Density, Kenshoo, Google Insights Online Video Marketing Market Samurai, Conversion Tracking, Backlinks
Panel B: Facebook's LinkedIn-identified labor skill set	
Top Skills & Expertise	**Related Skills**
Hive	Sqoop, Oozie, Flume, Amazon Elastic MapReduce, Mahout, HBase, Avro, Cascading, CDH, Apache Pig, Cascalog, MapReduce, Voldemort, Google Ad Planner, Conversion Optimization, Keyword Research, Paid Search Campaigns, Marin Software, Google Website Optimizer, Kenshoo, Google Adwords Professional, Adgooroo Landing Page Optimization, Google Webmaster Tools, Search Analysis
Facebook API	LinkedIn API, YouTube API, FQL, Google API, Paypal Integration, OAuth, OpenSocial, Custom Facebook Pages, Social Engine, eBay API, Google Checkout, OpenSceneGraph, Papervision3D, Twilio, Away3D, Authorize.net, Kohana, Chrome Extensions, Social Graph, Socket.io
Machine Learning	Same as those of Google
MapReduce	Oozie, Sqoop, Flume, BigTable, Mahout, HBase, Amazon Elastic MapReduce, Cascading, Hive, Avro, Voldemort, Katta, Cascalog, Collaborative Filtering, Apache Pig, CDH, Relevance, Nutch, Stream processing, Recommender Systems
Hadoop	HBase, Oozie, Sqoop, Flume, Mahout, MapReduce, Hive, Cascading, Amazon Elastic MapReduce, Cascalog, Nutch, Voldemort, Apache Pig, Katta, Avro, Cassandra, Greenplum, Vertica, Collaborative Filtering, CDH

This table provides the proxy for Google's and Facebook's labor market skill sets based on user reported skills on LinkedIn in Panels A and B respectively. For each firm, I include the top five skills and expertise listed on its LinkedIn company homepage and the related skills for each of the top five skills. Google has 88 skills, and Facebook has 76. They have 32 skills in common, so their skill similarity score is $s_{Google,Facebook} = \frac{32}{88+76-32} = 0.24$

Acknowledgements I thank Patricia Dechow, the editor, and two anonymous referees for their very helpful feedback. This paper is based on my Ph.D. dissertation at University of Chicago in 2014. I am very grateful to members of my dissertation committee for their invaluable help and support: Douglas Skinner (Chair), Ray Ball, Gary Becker, Pingyang Gao, and Michael Minnis. I also appreciate comments from and discussions with Jim Baer, Philip Berger, Andrew Chamberlain, Hans Christensen, Rui Cui, Eugene Fama, Pengjie Gao, Joseph Gerakos, João Granja, Zhiguo He, Harrison Hong, Danqi Hu, Yunzhi Hu, Erik Hurst, Filipe Lacerda, Christian Leuz, Amber Zichao Li, Meng Li, Hai Lu, Mark Maffett, Sara Moreira, Valeri Nikolaev, Dan Nguyen, Elizabeth Pontikes, Haresh Sapra, Jee-Eun Shin, Kelly Shue, Abbie Smith, Barbara Su, Prasanna Tambe (discussant), Fangfang Tan, Regina Wittenberg-Moerman, Franco Wong, Wei Wu, Danxia Xie, Qiping Xu, Anastasia Zakolyukina, Sarah Zechman, Mingyue Zhang, Daniel Zhao, Wuyang Zhao, Frank Shuang Zhou, and seminar participants at Chinese University of Hong Kong, Glassdoor.com, Hong Kong University, Nanyang Technological University, National University of Singapore, NBER Summer Institute (Economics of IT and Digitization), New York University, Peking University, Singapore Management University, UCLA, University of Chicago, University of Pittsburg, University of Toronto, and Zhejiang University. I thank Chuan Du for excellent research assistance. I thank people at LinkedIn and Glassdoor for helpful discussions. I am grateful to Glassdoor.com for providing employee salary review data. I thank Ling Cen for sharing customer supplier relationship data. I acknowledge the generous financial support provided by the Accounting Research Center and PhD program at Booth School of Business, University of Chicago and University of Toronto. The author is supported in part by funding from the Social Sciences and Humanities Research Council of Canada. A prior version of this paper was entitled "Labor Market Peer Firms".

References

Albuquerque, Ana. 2009. Peer firms in relative performance evaluation. *Journal of Accounting and Economics* 48 (1): 69–89.

Bae, Joon Woo, Turan G. Bali, Ali Sharifkhani, and Xiaofei Zhao. 2022. Labor market networks and asset returns. https://ssrn.com/abstract=3951333. Accessed 17 Jan 2024.

Barrios, John M. 2022. Occupational licensing and accountant quality: Evidence from the 150-hour rule. *Journal of Accounting Research* 60 (1): 3–43.

Becker, Gary S. 1962. Investment in human capital: A theoretical analysis. *The Journal of Political Economy* 70 (5): 9–49.

Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch. 2014. Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy* 122 (1): 129–177.

Bernstein, Shai, Rebecca Diamond, Abhisit Jiranaphawiboon, Timothy McQuade, and Beatriz Pousada. 2022. The contribution of high-skilled immigrants to innovation in the united states. https://ssrn.com/abstract=4312044. Accessed 22 Dec 2023.

Bhojraj, Sanjeev, Charles MC. Lee, and Derek K. Oler. 2003. What's my line? a comparison of industry classification schemes for capital market research. *Journal of Accounting Research* 41 (5): 745–774.

Blankespoor, Elizabeth, Gregory S. Miller, and Hal D. White. 2013. The role of dissemination in market liquidity: Evidence from firms' use of twitter. *The Accounting Review*.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics* 117 (1): 339–376.

Brown, Philip, and Ray Ball. 1967. Some preliminary findings on the association between the earnings of a firm, its industry, and the economy. *Journal of Accounting Research*, 55–77.

Brown, Jennifer, and David A. Matsa. 2016. Boarding a sinking ship? an investigation of job applications to distressed firms. *The Journal of Finance* 71 (2): 507–550.

Burks, Stephen V., Bo. Cowgill, Mitchell Hoffman, and Michael Housman. 2015. The value of hiring through employee referrals. *The Quarterly Journal of Economics* 130 (2): 805–839.

Cen, Ling, Yu. Feng Chen, and Hou, and Gordon D. Richardson. 2018. Strategic disclosures of litigation loss contingencies when customer-supplier relationships are at risk. *The Accounting Review* 93 (2): 137–159.

Chen-Zion, Ayal. 2015. Who competes for job seekers? visaulizing the labor market with glassdoor data. *Glassdoor Economic Research Blog*. https://www.glassdoor.com/research/employer-networks/. Accessed 22 Dec 2023.

Chetty, Raj, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. 2022. Social capital i: measurement and associations with economic mobility. *Nature* 608 (7921): 108–121.

Chiu, Peng-Chia., Siew Hong Teoh, Yinglei Zhang, and Xuan Huang. 2023. *Using google searches of firm products to detect revenue management*, 101457. Accounting: Organizations and Society.

Choi, Jung Ho, and Brandon Gipper. 2019. Fraudulent financial reporting and the consequences for employees. https://ssrn.com/abstract=3346759. Accessed 22 Aug 2022.

Cohen, Lauren, and Andrea Frazzini. 2008. Economic links and predictable returns. *The Journal of Finance* 63 (4): 1977–2011.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. In search of attention. *The Journal of Finance* 66 (5): 1461–1499.

Dehaan, Ed., Nan Li, and Frank S. Zhou. 2023. Financial reporting and employee job search. *Journal of Accounting Research* 61 (2): 571–617.

Donangelo, Andres. 2014. Labor mobility: Implications for asset pricing. *The Journal of Finance* 69 (3): 1321–1346.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. Referral-based job search networks. *The Review of Economic Studies* 83 (2): 514–546.

Eisfeldt, Andrea L., and Dimitris Papanikolaou. 2013. Organization capital and the cross-section of expected returns. *The Journal of Finance* 68 (4): 1365–1406.

Fama, Eugene F., and Kenneth R. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 153–194.

Fan, Joseph PH., and Larry HP. Lang. 2000. The measurement of relatedness: An application to corporate diversification. *The Journal of Business* 73 (4): 629–660.

Foster, George. 1981. Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics* 3 (3): 201–232.

Franco, De., Ole-Kristian Hope. Gus, and Stephannie Larocque. 2015. Analysts' choice of peer companies. *Review of Accounting Studies* 20 (1): 82–109.

Hales, Jeffrey, James R. Moon Jr, and Laura A. Swenson. 2018. A new era of voluntary disclosure? empirical evidence on how employee postings on social media relate to future corporate disclosures. ss Accounting. *Organizations and Society* 68: 88–108.

Ha-Thuc, Viet, Ye Xu, Satya Pradeep Kanduri, Xianren Wu, Vijay Dialani, Yan Yan, Abhishek Gupta, and Shakti Sinha. 2016. Search by ideal candidates: Next generation of talent search at linkedin. In *Proceedings of the 25th International Conference Companion on World Wide Web*, 195–198.

Hoberg, Gerard, and Gordon Phillips. 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies* 23 (10): 3773–3811.

Hoberg, Gerard, and Gordon Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124 (5): 1423–1465.

Horton, John J., and Prasanna Tambe. 2015. Labor economists get their microscope: big. *Big Data* 3 (3): 130–137.

Huang, Kelly, Meng Li, and Stanimir Markov. 2020. What do employees know? evidence from a social media platform. *The Accounting Review* 95 (2): 199–226.

Jagannathan, Ravi, and Zhenyu Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51: 3–54.

Jung, Michael J., James P. Naughton, Ahmed Tahoun, and Clare Wang. 2018. Do firms strategically disseminate? evidence from corporate use of social media. *The Accounting Review* 93 (4): 225–252.

Karabarbounis, Marios, and Santiago Pinto. 2018. What can we learn from online wage postings? evidence from glassdoor. *Economic Quarterly* 4Q: 173–189.

King, Benjamin F. 1966. Market and industry factors in stock price behavior. *The Journal of Business* 39 (1): 139–190.

Kothari, Sagar P., Andrew J. Leone, and Charles E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Kuehn, Lars-Alexander., Mikhail Simutin, and Jessie Jiaxu Wang. 2017. A labor capital asset pricing model. *The Journal of Finance* 72 (5): 2131–2178.

Leary, Mark T., and Michael R. Roberts. 2014. Do peer firms affect corporate financial policy? *The Journal of Finance* 69 (1): 139–178.

Lee, Charles MC., Paul Ma, and Charles CY. Wang. 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116 (2): 410–431.

Lee, Lian Fen, Amy P. Hutton, and Susan Shu. 2015. The role of social media in the capital market: evidence from consumer product recalls. *Journal of Accounting Research* 53 (2): 367–404.

Lee, Charles MC., Stephen Teng Sun, Rongfei Wang, and Ran Zhang. 2019. Technological links and predictable returns. *Journal of Financial Economics* 132 (3): 76–96.

Leung, Alvin Chung, Ashish Agarwal Man, Prabhudev Konana, and Alok Kumar. 2017. Network analysis of search dynamics: The case of stock habitats. *Management Science* 63 (8): 2667–2687.

Li, Nan, M. H. Wong, and Frank Zhou. 2019. Analysts' underreaction to employee business outlooks. https://ssrn.com/abstract=3409397. Accessed 22 Dec 2023.

LinkedIn. 2022. People also viewed feature - overview. https://www.linkedin.com/help/linkedin/answer/a541888/people-also-viewed-feature-overview?lang=en. Accessed 22 Dec 2023.

Liu, Yukun, and Xi Wu. 2022. Labor links, comovement and predictable returns. https://ssrn.com/abstract=3175958. Accessed 22 Dec 2023.

Marinescu, Ioana, Andrew Chamberlain, Morgan Smart, and Nadav Klein. 2021. Incentives can reduce bias in online employer reviews. *Journal of Experimental Psychology: Applied*.

Marinescu, Ioana, and Roland Rathelot. 2018. Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics* 10 (3): 42–70.

Martin, Roger L. 2014. The rise (and likely fall) of the talent economy. *Harvard Business Review* 92 (10): 16.

Mayers, David. 1973. Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset. *The Journal of Business* 46 (2): 258–267.

Miller, Gregory S., and Douglas J. Skinner. 2015. The evolving disclosure landscape: How changes in technology, the media, and capital markets are affecting disclosure. *Journal of Accounting Research* 53 (2): 221–239.

Mobley, William H. 1977. Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology* 62 (2): 237–240.

Pouder, Richard, H.. St.. Caron, and John. 1996. Hot spots and blind spots: Geographical clusters of firms and innovation. *Academy of management review* 21 (4): 1192–1225.

Ramnath, Sundaresh. 2002. Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research* 40 (5): 1351–1376.

Rauh, Joshua D., and Amir Sufi. 2012. Explaining corporate capital structure: Product markets, leases, and asset similarity. *Review of Finance* 16 (1): 115–155.

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. Search-theoretic models of the labor market: A survey. *Journal of Economic Literature* 43 (4): 959–988.

Schmutte, Ian M. 2015. Job referral networks and the determination of earnings in local labor markets. *Journal of Labor Economics* 33 (1): 1–32.

Stigler, George J. 1962. Information in the labor market. *Journal of Political Economy* 70 (5, Part 2): 94–105.

Teoh, Siew Hong., Yuqing Zhou, and Christos Makridis. 2023. Understanding reputation damage to firm culture: Insights from employee perceptions post financial misconduct. https://ssrn.com/abstract=3467787. Accessed 22 Dec 2023.

Teoh, Siew Hong. 2018. The promise and challenges of new datasets for accounting research. *Accounting, Organizations and Society* 68: 109–117.

Wu, Lili, Sam Shah, Sean Choi, Mitul Tiwari, and Christian Posse. 2014. The browsemaps: Collaborative filtering at linkedin. In *RSWeb@ RecSys*.

Wu, Lynn, and Erik Brynjolfsson. 2009. The future of prediction: How google searches foreshadow housing prices and sales. *Proc. 30TH International Conference Information Systems (Association for Information Systems, Atlanta)*, 923–936.

Ying, J. 2019. How we mapped the “skills genome” of emerging jobs. https://engineering.linkedin.com/blog/2019/how-we-mapped-the-skills-genome-of-emerging-jobs. Accessed 22 Aug 2022.

Zhu, Tingting Juni, Alan Fritzler, and Jan Alexander Kazimierz Orlowski. 2018. World bank group-linkedin data insights: Jobs, skills and migration trends methodology and validation results. https://documents1.worldbank.org/curated/en/827991542143093021/pdf/World-Bank-Group-LinkedIn-Data-Insights-Jobs-Skills-and-Migration-Trends-Methodology-and-Validation-Results.pdf. Accessed 22 Aug 2022.