On-the-job training and labor market competition

Abigail Adams-Prassl
Thomas Le Barbanchon
Alberto Marcato [*]

August 2022
Preliminary version
Please do not circulate

Abstract

In this paper, we provide a new set of stylised facts on firm provision of on-the-job training and labor market competition. We exploit job ad texts taking a supervised machine learning approach to identify training offers. We find that around 20% of US job posts offer on-the-job training, with an upward trend over the last decade. Training offers are positively correlated with local labor market concentration, a finding that is robust to an instrumental variables strategy based on the local differential exposure to national firm-level trends. Moving from the first to the third quartile of labor concentration increases training by almost 5%. We interpret our results through the lens of a directed search model where training acts to reduce the time to fill a vacancy and training has a greater expected benefit to the employer in less competitive labor markets given the lower separation rates.

Keywords: training, monopsony, job vacancies

JEL Codes: J24, J31, J42, J63, M53

---

[*]Adams: Oxford University. Le Barbanchon: Bocconi University. Marcato: Bocconi University. For helpful comments, we would like to thank Tito Boeri, Peter Kuhn, as well as participants at Bocconi, and the IZA-IDSC conference on matching workers and jobs online. We gratefully acknowledge financial support from the ERC StG 758190 ”ESEARCH”. All errors are our own. Adams is also affiliated with CEPR, Le Barbanchon with CEPR, IGIER, IZA, J-PAL.

1 Introduction

On-the-job training is crucial in many industrialized countries. In light of an aging population and rapid technological changes, promoting lifelong training is a key challenge for policymakers to maintain an efficient and productive workforce (U.S. Council of Economic Advisers, 2018; EU Council, 2019). Yet, the government has limited influence on the on-the-job training provision, relying primarily on private employers (Carnevale et al., 2015). Therefore, it is essential to explore the determinants that stimulate employer-provided training, especially in the aftermath of the Covid-19 pandemic, which caused the displacement of a profuse number of workers, who will need to adjust and find a job in less familiar occupations (World Economic Forum, 2020; OECD, 2021). Despite its evident relevance, little is known about the mechanisms that drive employer-provided training.

In an influential paper, Acemoglu and Pischke (1998) suggests that labor market power can stimulate employers to provide training. By exerting monopsony power, i.e., by setting wages below the workers' marginal productivity, employers have an incentive to increase labor productivity. However, this relationship between labor market power and training relies on the strong assumption that the labor supply elasticity does not change after training. As pointed out by Manning (2003), it is entirely plausible for trained workers to have better outside options than untrained ones, leading to a more extensive bargaining power from the trained workers and thus a lower incentive for monopsonistic employers to provide on-the-job training. By virtue of recent evidence documenting strong labor market power, the question of whether monopsony power can affect on-the-job training provision experience a recent revival (U.S. Council of Economic Advisers, 2016). Yet, the question is still unsettled, and the empirical evidence is scant. A key reason for this is the lack of high-quality data on training provision and of an identification strategy to deal with the measurement of labor market competition.

This paper studies the effect of labor market competition on employer-provided training and sheds light on the mechanisms behind this effect. We exploit the job vacancy data from Burning Glass Technologies (BGT) database to construct training and labor market competition measures. Relying on a Bartik-style instrument, we find that labor market competition favors on-the-job training provisions and is associated with reducing the probability of a vacancy posting wage information.

Traditionally, training information is collected using questionnaire-based surveys. Still, these have some crucial drawbacks as they are costly, have time lags, lack extensive coverage at the geographical level, prevent conducting across labor market analysis, and require the truthful participation of workers or firms. In contrast, nearly every firm has online job vacancies, which has been shown to contain useful information (Adams-Prassl et al., 2020; Ash et al., 2020). Therefore, this paper provides a new measure of employer-provided training based on job vacancies in the U.S. We use a sample of manually tagged vacancies to train a supervised machine learning (ML) model that can capture content on the employer-training provision. Comparing our ML indicator with survey-based data shows that our training indicator leads to plausible results. We find that around 20 percent of job vacancies offer on-the-job training, with an upward trend over the last decade.

Consistent with classical and recent literature, we quantify labor market competition with a measure of employer concentration.[1] We follow Azar et al. (2020b) in measuring employer concentration with a Herfindahl-Hirschman Index (HHI) measured at a combination between a metropolitan statistical area (MSA), 6-digit Standard Occupation Classification code (SOC), and a year. We further show that employer concentration is correlated with lower job transitions, which is consistent with a theory of oligopsonistic labor markets in which employers compete for workers a la Cournot. We use a Bartik-style instrumental variable approach to deal with the endogeneity of the labor market concentration measure. We instrument labor market concentration changes in a local labor market with the predicted variations in the number of vacancies constructed from the variations at the national level of firms posting behavior, excluding that specific local labor market (Schubert et al., 2020). Under the key assumption that a firm's decision to post vacancy at the country level is not affected by local labor market specificities, this approach provides shocks independent of market unobservable characteristics that can affect on-the-job training decisions.

Next, we investigate how labor market competition affects what an employer posts in a job vacancy. A related concern is that labor market power could affect employers' decision to post wage information. Policymakers have called for employers to disclose wage information on job listings, as the lack of pay transparency has been a source of gender wage discrimination.[2] Therefore, understanding what influence the decision of an employer to

[1]For the classical motivation of employer concentration as a source of monopsony power, as well as for the other potential sources of monopsony power, see (Robinson, 1969; Boal and Ransom, 1997; Manning, 2003); among the most recent papers modeling the relationship between employer concentration and monopsony power, see (Berger et al., 2022; Jarosch et al., 2019).

[2]Starting from 2021, many jurisdictions are introducing regulations to require wage disclosure in job postings. For example, from January 1, 2023, the State of Washington is the third U.S. jurisdiction (and counting) to introduce this regulation (Washington State, 2022), following Colorado and New York City. It is

disclose wage information on job listings is of great importance. We showed that, indeed higher level of concentration leads to lower wage posting. We also provide evidence that employer concentration reduces the requirements of education and experience.

The main findings of the paper can be summarized as follows. First, we find a positive and statistically significant effect of employer concentration on on-the-job training offers. Going from the 25 percentile of the HHI distribution to the 75 percentile is associated with an increase of around 5 percent. Then, we find a negative and statistically significant effect on the probability of posting wages. The same interquartile movement is related to a decrease in the probability of a vacancy posting wage information of 10 percent. We further observed a reduction in the education demanded by around 8 percent and experience by about 3 percent.

Related Literature: This paper builds and extends on different strands of the literature. First, we contribute multiple ways to the scarce literature on the effect of labor market competition on employer-provided training. This paper is the first to develop a machine learning model to extract training information from job vacancy text. Consequently, developing a cost-effective way to measure employer-provided training, which can be quickly updated and covers the near-universe of job vacancies – thus, not having the usual disadvantages of being survey-based. Contrary to most, this paper is one of the few to examine the role of employer concentration and to address the endogeneity of employer concentration through an instrumental variable approach, as well as the first to investigate the U.S. market. Harhoff and Kane (1997), Brunello and Gambarotto (2007), and Muehlemann and Wolter (2011) observed a positive correlation between the number of firms in a market and employer-provided training provisions in Germany, the U.K., and Switzerland, respectively. However, neither of those papers addresses the endogeneity problem.[3] Second, we contribute to the flourishing empirical literature that analyzes the effect of employer concentration on wages (Martins, 2018; Abel et al., 2018; Rinz, 2022; Lipsius, 2018; Qiu and Sojourner, 2019; Azar et al., 2022; Benmelech et al., 2020; Azar et al., 2020a,b; Arnold, 2020; Schubert et al., 2020; Marinescu et al., 2021; Bassanini et al., 2021; Popp, 2021), as well as a

worth noting that these legislatures do not affect our results, as they are all after our analysis period, which ends in 2019. A similar policy has also been recently discussed in the European Parliament (E.U. Parliament, 2021).

[3]Other related papers are Starr (2019), which observes that workers in U.S. States with more restricted use of non-competing agreements receive less training. Rzepka and Tamm (2016) observed that employer concentration is correlated with larger employer-provided training in Germany, yet they do not address the endogeneity issue besides the use of fixed effects. Finally, Marcato (2021) showed that employer concentration positively affects employers' investment in training in Italy, using a similar instrumental variable approach. Other related papers are (Mohrenweiser et al., 2019; Méndez, 2019; Arellano-Bover, 2020; Bratti et al., 2021).

more theoretical literature connecting employer concentration to wage markdown or labor share (Berger et al., 2022; Jarosch et al., 2019; Azkarate-Askasua and Zerecero, 2020; Hershbein et al., 2022). This paper adds to this existing literature by focusing on an additional effect of employer concentration, namely employer-provided training. The paper is also close to the literature on using text analysis to extract information from job vacancy text (Deming and Kahn, 2018; Adams-Prassl et al., 2020; Ash et al., 2020) and the literature on vacancy data and wages (Brenčič, 2012; Kuhn and Shen, 2013; Marinescu and Wolthoff, 2020; Le Barbanchon et al., 2021).

The paper proceeds as follows. We describe the data in Section 2 and provide some descriptive statistics in Section 3. We illustrate the conceptual framework in Section 4. We introduce our empirical strategy in Section 5 and describe the results in Section 6. We explore additional effects in Section 7 and we conclude in Section 8.

2 Data

To make progress, we turn to information on training contained in job vacancies. Our primary data source is a database of online job ads provided by Burning Glass Technologies (BGT). By scraping more than 40,000 online job boards and company websites, BGT provides job postings data covering the near-universe of occupation, industries, and geographic areas in the U.S. Comparing BGT data to the Job Opening and Labor Force Turnover Survey (JOLTS), Hershbein and Kahn (2018), Deming and Kahn (2018) and Berkes et al. (2018) observed that BGT provides a good coverage of the job openings in the U.S.[4]

Besides the vast coverage and the rich information, a clear advantage of the BGT data is that it provides the text of each job post. This allows us to predict through a machine learning method whether a vacancy offers on-the-job training and to investigate how on-the-job training is impacted by the level of competition faced by employer in her local labor market.

[4] Recently, the same data source has been widely used in academic research (Blair and Deming, 2020; Forsythe et al., 2020; Schubert et al., 2020; Azar et al., 2020b; Burke et al., 2020; Modestino et al., 2020; Clemens et al., 2021; Kuhn et al., 2018).

2.1 Measuring training in job vacancies

Our goal is to take the information in the job vacancies text and predict the probability that a given vacancy offers on-the-job training. To this end, we take a supervised machine learning approach that relies on manual annotations similar in spirit to Adams-Prassl et al. (2020). The method proceeds as follows:

1. Manually tag a set of job vacancies if they are explicitly offering on-the-job training
2. Define the vocabulary and represents each job vacancy in a matrix format
3. Train a machine learning model to classify training offers based on the vacancy text;
4. Apply the machine learning model to all job vacancies for use in the subsequent analysis.

Compared to a word search approach, the machine learning approach allows us to predict better the actual disclosure of training offerings in the vacancy text. Indeed, searching for keywords like ”training” or ”paid training” on the one hand will neglect all those vacancies that will offer training using different expressions such as ”we will train you” or ”we will provide you paid in-class courses to acquire the required skills”. On the other hand, the word search approach will incorrectly label as offering training vacancies, which, for example, requires the newly hired to train customers. Figure F1 compares the percentage of false positives and negatives of the machine learning and word search approaches. Appendix A reports a few examples of false positives and negatives from the word search approach.

2.1.1 Manual tagging

This section describes our manual tagging procedure. We provide a detailed guideline to twelve research assistants, who inspect the text of around 6,000 job vacancies. They manually annotate if a job vacancy offered training and additional label when training was not an employee benefit but a task the employee has to perform; further information is in Appendix A.

We do not distinguish between general and specific training, but we consider training as any program that will help new hires acquire new competencies or skills. As the differentiation between the two types is difficult to correctly implement, especially given that the information on the training content is scarcely reported in the job vacancies text. Some examples are: ”Paid training”, ”new employee training”, ”tuition reimbursement”, ”continuous professional development”, ”practice-paid continuing education”; for more information see Appendix A.

Before getting to the modeling, we build a vocabulary from the set of annotated vacancies. Next, we tune the vocabulary by filtering out stop words (e.g., and, the, it). Stop words have limited lexical content, and their presence adds a lot of noise and little signal to help us identify when a job vacancy is offering training. We finally break down the vocabulary into countable features. Each vacancy text is represented as a frequency distribution over a vocabulary of words, bigrams, and trigrams (two and three words phrases). The vocabulary is filtered to the 5000 most frequent phrases.

2.1.2 Machine learning model

This section describes how we train the model to predict whether a vacancy offers training. We use a logistic prediction model using as regressors the aforementioned matrix of 5,000 phrases. To select the hyper-parameters, we use five-fold cross-validation in the training set.

Table 1 reports the relevant test-set evaluation metrics. Accuracy is the proportion of out-of-sample observation for which the machine-predicted model correctly predict the true label. Recall is the proportion of correct predicted training within the set of vacancies actually offering training (TP/(TP+FN)). Precision instead is the percentage of correct training predictions relative to total number of training predictions (TP/(TP+FP)). Therefore, while precision penalizes false positives, recall penalizes false negatives. F1 is the weighted harmonic mean of precision and recall, thus it penalizes equally false negatives and positives. Finally, another standard metric in binary classification is the AUC-ROC (Area Under the Receiver Operating Characteristics), which takes values between 0 and 1 and tells how much the model is capable of distinguishing between classes. It can be interpreted as the probability that a randomly sampled vacancy offering training is ranked more highly by predicted probability than a randomly selected vacancy not offering training.[5] Performance are quite good with test-set accuracy = 0.80, AUC-ROC = 0.86, and F1 = 0.71.[6] The

[5] In Appendix, Figure F6 plots the ROC curve which display the percentage of true positives predicted by the model along the predicted probabilities of offering training.

[6] The performance is similar to that of other recent economic papers using similar machine learning mod-

model is somewhat conservative in identifying training offers (recall = 0.67), but with good precision (0.76). Figure F1 shows the comparison in term of false positives and false negatives using our machine learning model instead of a more straightforward word search approach. Comparing the metrics, the word search approach largely underperforms the machine learning approach with a F1 score of 0.616, with a good recall of 0.806 but a poor precision of around 0.498.

In addition to the classification metrics, Figure F2 shows how the logistic regression is well-calibrated. Specifically, the figure divided the the out-of-sample test set into bins according to their predicted probability to offering training and it compares their average predicted probability of offering training to the average true probability in each bin. Finally, to provide a qualitative assessment, we report the most predictive phrases in the word cloud in Figure F3.

2.1.3 On-the-job training in US online ads

This section describes new summary statistics of on-the-job training posted in US ads. It also compares with survey results on actual training.

2.1.4 Random sample of job vacancies

Even limiting the analysis from 2013 to 2019 and to only those vacancies with non-missing occupation, employer name, and metropolitan statistical area (MSA) code the number of vacancies exceeds the 100 millions. Therefore, to maintain the tractability of the analysis we randomly select 10% of the employers posting at least a vacancy in 2019. This yields a sample of $12,634,777$ unique job vacancies. Figure F4 compares the number of posted vacancy in the full sample and the 10% random sample. We can notice how the two series are well correlated, however, we can also notice that in 2018 there is a significant jump in the number of vacancies collected by BGT, this is due to an improvement in the collecting of online vacancy. To check if the estimates are not driven by the difference in the collection procedure, we have performed the analysis also only on those vacancy posted before 2018.

els, such as Ash et al. (2021) with an AUC-ROC of 0.78; Kleinberg et al. (2018) of 0.71; and Mullainathan and Obermeyer (2019) of 0.73.

2.1.5 Comparison with other measures of on-the-job training

We compare the job-ad measure of on-the-job training to two other measures. First, we use BLS data on the typical on-the-job training needed to attain competency in each occupation.[7] The BLS distinguishes six on-the-job training categories based on the occupational description from O*NET: none, short-term/moderate-term/long-term training, residency and apprenticeship.[8] We pool the five categories with some training together, and we aggregate the BLS measures at the SOC 2-digit level by taking the simple average of the corresponding training dummy. Similarly, we compute the average probability of on-the-job training offer from our ML model at the same occupational level. Figure 1b shows that the BLS and job-ad measures are positively correlated. The fitted line accounts for the occupational number of vacancies posted within the year 2019. This is illustrated by the marker size in the right-hand panel (b). Furthermore, Figure 1c replicates the analysis at a finer level of the SOC occupation classification (5-digit). The binscatter plot confirms the positive relation between the job-ad measure and the BLS index.

Second, we use the Survey of Income and Program Participation (SIPP) of 2008. This is the main survey recording employees' answers to on-the-job training related questions. The SIPP describes if each worker has received in the previous year some kind of training to improve her skills and if this training was paid by her employer. Using these two questions we construct the average percentage of workers who have received on-the-job training per US state and compare this estimate with those obtain on BGT data.[9] Figure 1d shows the positive relation between the SIPP and job-ad measure.

[7]The BLS data are available at https://www.bls.gov/emp/tables/education-and-training-by-occupation.htm

[8]Long-term on-the-job training takes more than 12 months; Moderate-term takes more than 1 month and up to 12 months; Short-term on-the-job training takes 1 month or less. See https://www.bls.gov/emp/documentation/education/tech.htm

[9]There are more recent SIPP surveys: 2014 and 2018. However, unfortunately, these surveys have changed the relevant questions regarding on-the-job training and restricted the sample of respondents. Specifically, the training related questions are asked only to those workers who are not graduated and are below 200% of the poverty line. We have further restrict the SIPP sample to only those workers with an age between 15 and 65 years old that had a paid job during the reference period.

2.2 Labor market competition

A flowering recent literature, both empirical (Azar et al., 2020a,b; Benmelech et al., 2020; Rinz, 2022; Schubert et al., 2020; Marinescu et al., 2021; Azar et al., 2022; Hershbein et al., 2022; Arnold, 2020) and theoretical (Jarosch et al., 2019; Berger et al., 2022; Azkarate-Askasua and Zerecero, 2020) has demonstrated a negative relationship between wages and local labor market concentration. Building on this research, we adopt local labor concentration as our primary measure of labor market competition. The rationale is that an increase in concentration reduces labor supply elasticity, possibly due to a reduction in outside options or worker and job heterogeneity.[10]

The main point is that due to a reduced labor supply, an employer can offer lower wages without fearing the possibility that the worker will change jobs. However, as a corollary, sometimes ignored, lower labor supply works both ways, i.e., an employer in a concentrated market will find it more difficult to hire new workers as they have a lower labor supply elasticity.[11] This decreasing labor supply elasticity in a highly concentrated market leads employers to reduce their labor demand to keep wages low. This idea is observed, for example, in Marinescu et al. (2021) and Berger et al. (2022). Consequently, an employer in concentrated markets would be more willing to offer training when this allows for increasing their labor supply without raising wages.[12] We further detail this mechanism in Section 4. In subsection 2.2.1 we define our measure of local labor market concentration, while in section 2.2.2 we compare this measure with other potential proxies of labor market competition, such as unemployment, job creation, and job destruction.

2.2.1 Local labor market concentration measure

To measure the effect of local labor market concentration on the employers' on-the-job training and wage posting decisions, we first have to define the relevant labor market.

[10]See (Manning, 2003, 2011; Robinson, 1969) for more in-depth information and a review of the literature on monopsony power and its potential sources.

[11]For clarity, although the employer can offer lower wages to incumbent workers, to increase her workforce and hire new workers the employer will have to offer a larger wages. If the employer cannot perfectly wage-discriminate across her workers, a marginal increase in the wages will lead to an increase in all the incumbent wages.

[12]Two recent surveys provide anecdotal evidence on the potential effect of training as an attracting device (Monster, 2021; Gallup, 2021). Monster (2021) found that among workers who recently quit a job, 45% would have remained if they were offered more training, while Gallup (2021) documented that 48% of American workers would switch jobs if the new job provided skill training opportunities.

Most literature represents a local labor market as a combination of a geographical area, a time, and an industry or occupation; see for example (Schubert et al., 2020; Azar et al., 2020b; Marinescu et al., 2021; Hershbein et al., 2022; Rinz, 2022). Following this literature, we define a local labor market as a combination of a year, a SOC 6-digit occupation, and a U.S metropolitan statistical area (MSA).

We use the Burning Glass Technologies (BGT) database of online vacancy postings to measure employer concentration. We calculate the Herfindahl-Hirschman Index (HHI) of the share of vacancies posted by each employer at the local labor market level.

Specifically, we compute the vacancy HHI as the sum of the squared vacancy employer shares for each local labor market, which is defined as the combination of 6-digit SOC occupation, metropolitan statistical area, and year.

$$HHI_{o,l,t} = \sum_{i}^{N} \left(\frac{v_{i,o,l,t}}{V_{o,l,t}} \right)^2 \tag{1}$$

where $v_{i,o,l,t}$ is the number of vacancies posted by employer i in the local labor market defined by the combination of occupation o, metropolitan statistical area l, and year t; while $V_{o,l,t}$ is the total number of vacancies posted in that local labor market.

2.2.2 Comparison with other measures of labor market competition

We compare our employment HHI measure with other potential proxies of labor market competition: unemployment rate, job creation rate, and job destruction rate.

We use the local unemployment rate from the Local Area Unemployment Statistics (LAUS) program provided by the U.S Bureau of Labor Statistics, measured at the metropolitan statistical area (MSA) and year. The job creation and destruction rates are obtained from the Business Dynamics Statistics (BDS) provided by the U.S. Census and are measured at the combination between a sector (2-digit NAICS code), MSA, and year. The job creation and destruction rates are constructed considering the increase and decrease (respectively) in the share of employment for each establishment in the segment previously defined. For more information on these measures see Appendix B.

To compare these measures with our HHI measure, which is defined at the combination between MSA, occupation (6dig SOC), and year; we take the weighted average of our HHI at the same level of the unemployment rate (year $\times$ MSA) and of the job creation and destruction rate (year $\times$ sector $\times$ MSA). Figures 3, 4, and 5 shows that these alternative measures are strictly correlated with the HHI, even when controlling for fixed effects, such as year, MSA, and industry.

Despite the strong correlation, we prefer our HHI measure for two main reasons. First, by virtue of what discuss in Section 2.2 and Section 4, we consider HHI as a better measure for labor market competition. Second, because HHI can be constructed at the finest level of aggregation, which by virtue of the recent literature that shows how labor markets are very local, we consider a better local labor market definition (Manning and Petrongolo, 2017; Marinescu and Rathelot, 2018; Kaplan and Schulhofer-Wohl, 2017; Le Barbanchon et al., 2021). However, we can see in Appendix B that the results using these alternative measures are in line with our main results.

3 Descriptive statistics

Focusing on our representative sample consisting on 10% of the employers posting a vacancy in 2019, Figure 2 shows the share of vacancies offering on-the-job training at the monthly and quarterly level. The share of vacancies is slightly increasing over time floating between 25% and 30%. Figure 6 plots the average training predicted probability and the HHI across Metropolitan Statistical Areas (MSAs) in 2019. The training map is divided by quartiles: $[0, 0.24)$, (0.24-0.33], (0.33-0.37], and (0.37-0.65]. While for the HHI map, we group the MSA according to the DOJ/FTC (2010) classification, which defines a market with an HHI of 0-0.15 as low concentrated, 0.15-0.25 moderately concentrated, 0.25-0.35 highly concentrated, and $\geq$ 0.35 very highly concentrated.

Figure 7 shows the logarithmic distribution of the HHI at the local labor market and at vacancy level. While the average job ads is posted in a lowly concentrated labor market, the average local labor market is moderately concentrated, with a mean above log(HHI) = 7, equivalent to an HHI of 0.11 or an Inverse Herfindahl-Hirschman Index (IHHI) of 9.2.[13] The Inverse Herfindahl-Hirschman Index (IHHI) can be interpreted as the number of

[13]Note that as a standard procedure we have taken the log of the HHI multiplied by 10 000, this is to avoid

equal-sized firms that will induce the same observed HHI.[14]

Figure 8 displays the 2019 distribution of the average number of vacancies, the average training probability, and the HHI across 2-digit SOC occupations.

Figure 9 show the average predicted training probability according to the level of education and experience demanded in the vacancy posted. One can see that experience and training are negatively correlated. Vacancies demanding many years of experience are unlikely to offer training. About education, it also appears that employers offering training do not request graduate applicants.

Finally, Figure 10 provides (left) the average training probability and (right) the share of vacancies providing some information on the wage offered across the HHI quartiles. The figure supports our idea that employers strategically exploit the labor market conditions. Specifically, an employer located in a highly concentrated labor market, given the difficulty to poach workers, will try to attract new workers from other occupations, highlighting the willingness to train them in the new occupation. On the other hand, if an employer is in a lowly concentrated labor market, she can poach workers from her competitors. Thus, she will try to attract these new experienced workers by posting an attractive salary.

4 Conceptual framework

Although the main focus of our paper is empirical, to guide the discussion we construct a stylized framework combining a oligopsony model results to a direct search model.

There are a number of models of the labor market in which labor market concentration matters for wages. Specifically, according to the traditional monopsony theory (Robinson, 1969; Manning, 2003) labor market concentration can generate an upward-sloping labor supply curve to individual firms, making for them marginally more expensive to hire workers. Consequently, the most productive firms in highly concentrated market will reduce their optimal employment causing misallocation inefficiencies.[15]

having negative numbers.

[14]For example, an IHHI of 10 implies that the market has the same HHI that a market consisting of 10 firms with the same number of employees would have.

[15]For a theoretical description see Robinson (1969) and Manning (2003), or more recently Berger et al. (2022). For empirical results on this idea, see Marinescu et al. (2021). An alternative approach, similar in

We include this result of an upward labor supply curve in a direct search model framework similar to Shimer (2005); Faberman and Menzio (2018); Marinescu and Wolthoff (2020), and discuss its predictions.

The model distinguishes firms and workers into two groups, representing the labor market's different occupations. We further allow the possibility for the firms to pay a fixed cost to rule out the difference in productivity that workers have in various occupations. The idea is to address the opportunity for firms to train workers and make them equally productive with already trained workers. Our model shows that firms will opt for training to reduce queuing time to attract valuable applicants.

4.1 Setting

We consider a static economy composed by two agents, workers and firms. Both agents are divided into two different segments, characterizing a horizontal heterogeneity in productivity, underlining two different occupations. The rationale behind is that a welder can not be able to do the job of a lawyer and vice-versa.[16]

result, is the one in Jarosch et al. (2019), where labor market concentration reduces the number of feasible outside options for workers, as the average worker in a given labor market has few different firms as possible alternative employers. Analogously, labor market concentration reduces the number of suitable candidates for firms. Considering that workers require considerable skills to perform a specific job profitably and that this skill can be properly acquired only through on-the-job training and experience. A dominant firm can only obtain suitable candidates from other firms in its market by poaching its employees. However, as the number of firms in a market decreases, the number of poachable workers also decreases. The same idea of an upward labor supply curve can also be derived from a model à la Burdett and Mortensen (1998), where firms with a larger labor force have to offer higher wages. Therefore, at a high concentration level, i.e., few large firms dominate the market, an employer has to offer marginally higher wages to attract more workers. Although an additional worker will be beneficial, the firms find it more profitable to renounce hiring new workers as this would create upward pressure on the wages, reducing their profits consequently. Note that is the concentration level and not exclusively the number of firms the key to the mechanism. The employment size of a firm depends not only on the wage offered by the same firms but also on the distribution of wages offered by the other competing firms. Thus, for a given firm with a fixed employment share, being in a market with several small or few large firms has a big difference. As in the former case, the distribution of the wages by the competitors will be lower than in the latter case due to the strict link between wages and size.

[16]Clearly, the model can be extended to account for a broader set of occupations, each of them with a different degree of substitutability. However, this extensions would not change our empirical predictions. Additionally, it can be extended to assume that workers within an occupation are divided into M different types, $m = 1, \ldots, M$. Similarly, in each segment there are N different types, $n = 1, \ldots, N$. These different types depict workers and firms different productivity, a worker $m = 2$ will be more productive than a worker $m = 1$, analogously for the firms.

4.2 Direct search and matching

Firms compete for workers by posting wages that are type-independent $\{w_j\}$. After observing the wages, workers decide to apply to the firm, maximizing their expected payoff based on both the probability of being hired and the wage. Following a standard assumption in the literature of direct search, we assume that workers can apply to only one firm, and firms hire only the best applicant. The assumption of a single application captures the opportunity cost for applying to multiple jobs. However, extending the model and allowing workers to apply to a finite number of firms in the same period will not substantially change the empirical predictions.

We further assume that posted wages are binding, i.e., firms cannot decide to bargain the wage after the match. Moreover, firms can choose to do not to post wages. In this case, for simplicity, we assume that workers correctly predict the average expected wages. However, they are also risk-averse; thus, given the same expected wage, they will prefer to apply to a firm posting wage to a non-posting wage firm.

Finally, we assume that the measure of potential candidates for a firm is proportional to the extent of concentration in that market. The idea behind this assumption is that at a higher concentration, there is a lower labor supply faced by the firm, i.e., a firm finds it more challenging to attract new candidates.[17] A straightforward way to assume it is to consider a search model with an exogenous separation rate and without recall as in Jarosch et al. (2019). In this case, the number of potential candidates for a firm j will be $\delta(1-s_j)$, where δ is the exogenous separation rate and s_j is the employment share of firm j in that market. Therefore, the average number of potential candidates in the market will be:[18]

$$\sum_j s_j \delta(1-s_j) = \delta(1-HHI)$$

[17]In this, we differentiate by the standard literature of direct search, which generally uses measures of market tightness such as the unemployment rate. However, as observed recently by Faberman et al. (2017) the majority of job transitions are job-to-job transitions; thus, concentration could be a better-suited proxy for the potential number of workers, including the possibility for employers to poach workers from her competitors.

[18]Alternatively, an analogous result arises with a Cournot-style oligopsonistic model of labor quantity competition based on differentiated firm-specific amenities in a nested framework that separate within-market and between-market labor supply behavior (Berger et al., 2022).

4.3 Production and Payoffs

A match between a worker (i) and a firm (j) produces $y_{i,j}$, which depends on the quality of the match (i, j). The worker's payoff is the wage w_j, while the firms keeps the remaining output $y_{i,j} - w_j$. Unmatched workers and firms get a null payoff. Therefore, the expected payoff of a worker (i) depends both on the probability to be hired and her productivity.

4.4 Symmetric equilibrium and Queue length

As standard in the literature, we consider the symmetric equilibrium where workers and firms of the same type behave identically. This assumption implies that the expected number of applicants of type (i) at a firm (j) follows a Poisson distribution with endogenous mean $q_{i,j}$ which is known as the "queue length", see Shimer (2005).

4.5 Benchmark: no training allowed

Assuming an horizontal differentiation, i.e. matches between workers and firms of the same types are always more productive: $y_{1,1} > y_{2,1}$ and $y_{2,2} < y_{2,1}$; the number of applicants that a firm received will depend on the number of workers in her segment and the difference in productivity between working in their own occupation or working in the other one. Intuitively, firms in markets in which they expect few applicants will offer longer queue. Therefore, in more concentrated markets where the number of potential applicants is limited, firms will wait longer to obtain a match. However, if the gap in productivity is relatively small $y_{2,1} - y_{2,2} = \varepsilon$, firms in high concentrated markets will attract also workers from the other market, which in turn reduces their queue time.

Moreover, as posting wages is binding, firms with a short queue, i.e with a few expected number of applicants, will be less willing to post wages. As by posting wages they will preclude themselves to only one time of applicant. If they post a wage low in order to attract also the less productive workers, they will disincentive the matched workers to apply. On the other hand, if the post a wage tailored for the matched workers, they will be bound to pay the same wages also to the less suited matches. Consequently, in high concentrated markets where employers will attract also workers from other segments, we can expect to observe less wage posting.

4.6 Extension: Allowing training

Assume now that firms could pay a fixed cost for training their workers. After training, the workers are equally productive in both occupations. As before, two forces are at play: labor market tightness (concentration) and difference in productivity. There will be no training if the differences in productivity and concentration are slight across the two markets. However, suppose the productivity in the more concentrated market is remarkably higher. In that case, offering training will drastically increase the number of applicants for the firms in concentrated markets, which will reduce their lost profits due to the time for finding a suitable candidate. More prominent is the gap in productivity or concentration across the two markets; more significant will be the incentive to train.

Empirical predictions:

1. At higher employer concentration there will be more on-the-job training offers.
2. Higher concentration lowers the probability a vacancy posts wage information.

5 Empirical strategy

To estimate the effect of employer concentration on training and wage posting probability, we rely on the following specification:

$$Y_{ijolt} = \gamma \log(HHI_{olt}) + \delta X_{ijolt} + \text{fixed effects} + \varepsilon_{ijolt} \tag{2}$$

where subscript i, j, o, l, and t denote respectively vacancy i, employer j, occupation o, MSA l, and year t. $\log(HHI_{olt})$ denotes the log. of the Herfindahl-Hirschman Index (HHI) as defined in Equation 1; Y_{ijolt} is the outcome variable, which in the main analysis is either (i) the predicted probability that vacancy i offers on-the-job training, or (ii) a binary variable on whether the vacancy has posted wage information. X_{ijolt} denotes a set of controls at the firm, vacancy, or market level. We saturate the model with the inclusion of different combinations of fixed effects, the preferred specification includes year, MSA, 6-digit SOC occupation code, 2-digit NAICS industry code, and employer fixed. Yet, the results are robust to different combination of fixed effects. We estimate equation 2 with standard errors clustered at the local labor market level, combination between a year, MSA, and a 6-digit SOC occupation.

5.1 Instrumental variable approach

The use of the concentration measure likely raises endogeneity concerns. An increase in concentration could be driven by the expansion of more productive and larger firms, leading to higher on-the-job training offers or wage information postings. One could think that larger firms have more resources or dedicated human resources departments, which might affect the provision of training or the quality of the job ads text, for example, by including more information. On the other hand, an increase in concentration could also be driven by worsening the business conditions in that local labor market, leading to a reduction of training provisions. Therefore, although the issues on the endogeneity of the concentration measure, the direction of the bias is ambiguous. To address the endogeneity issue, we follow the strategy of Schubert et al. (2020) and use an instrumental variable approach based on the granular instrumental variable approach of Gabaix and Koijen (2020) and a ”double shift-share Bartik” approach (Chodorow-Reich and Wieland, 2020).

Following the Bartik strategy, we can decompose the variation in the vacancy concentration for an occupation o, location l, and year t as a function of the previous vacancy share of each firm j ($s_{j,o,l,t-1}$) and its growth rate $g_{j,o,l,t}$ with respect to the market vacancy growth rate $g_{o,l,t}$. Formally,

$$\Delta HHI_{i,l,t} = \sum_j s^2_{j,o,l,t} - \sum_j s^2_{j,o,l,t-1} = \sum_j s^2_{j,o,l,t-1} \left(\frac{(1+g_{j,o,l,t})^2}{(1+g_{o,l,t})^2} - 1 \right) \quad (3)$$

where $s^2_{j,o,l,t}$ is the employment squared share of employer j in the occupation o, metropolitan statistical area l and year t; $\tilde{g}_{jot}$ is the national vacancy growth of firm j in occupation o and year t leaving out metropolitan statistical area l.

Following Schubert et al. (2020), we instrument the vacancy growth for each firm j in occupation o and location l with the national vacancy growth of that firm for that occupation in the other locations. Formally, the instrument is as follows:

$$\log(HHI^{\text{instr.}}_{o,l,t}) = \log\left[\sum_j s^2_{j,o,l,t-1}\left(\frac{(1+\tilde{g}_{j,(-l),o,t})^2}{(1+\tilde{g}_{o,l,t})^2} - 1\right)\right] \tag{4}$$

where $\tilde{g}_{olt}$ is the vacancy growth for MSA l, occupation o, and year t predicted by the predicted growth in vacancy in occupation i for each employer j

$$\tilde{g}_{olt} = \sum_j \sigma_{jolt}\tilde{g}_{j,o,(-l),t}$$

To avoid that relatively small employers drive the instrument, we construct $\tilde{g}_{j,o,(-l),t}$ considering only "large employers", i.e., those employers with vacancies in occupation o in at least five metropolitan statistical areas in that year t, as in Schubert et al. (2020). This restriction implies that the sum of the vacancies shares computed in the instrument construction does not sum up to 1. To address this issue, we add an "exposure control", defined as the sum of these large employers' squared vacancy shares. Given our econometric model, firms' vacancy growth should affect training only through vacancy concentration, which is a quadratic term. However, firms' vacancy growth could linearly affect local labor market features, such as labor demand and training decisions. Therefore, following the literature, we include two additional controls: (i) the actual vacancy growth rate in occupation o, location l, at year t ($g_{o,l,t}$) and (ii) the predicted vacancy growth rate $\tilde{g}_{o,l,t}$. Conceptually, this should capture the potential direct linear effects of firms' vacancy growth on training decisions.

6 Main Results

This section tests our main empirical predictions that local labor market concentration (1) increases the on-the-job training provision and (2) decreases the share of vacancies posting wage information.

Table 4 reports the estimates of labor market concentration on the predicted probability for a vacancy to offer training, as specified in Equation 2. Columns 1, 2, and 3 show the basic OLS estimates considering different specifications of fixed effects: Column 1 considers year, MSA, and occupation fixed effects; Column 2 adds sector (2digit NAICS) fixed effect, and Column 3 adds employer fixed effects. Columns 4, 5, and 6 adopt the instrumental variable approach described in Section 5 and use the same fixed effect specifications of Columns 1, 2, and 3, respectively. On-the-job training is positively and significantly correlated with labor market concentration across all six specifications. Moreover, the IV estimates are notably larger in magnitude, suggesting that some combination of omitted variable bias or measurement error biases the coefficients toward zero in the basic OLS regressions.[19] Additionally, it is worth noting as the results are robust to the inclusion of employer fixed effects, hinting that employers consider the labor market conditions in their hiring decisions.

To give a sense of the average results of labor market concentration on training offers, consider Table 4, Column (6), which is the IV specification including year, MSA, occupation, industry, and employer fixed effects. This specification implies that an interquartile range change in the HHI vacancy distribution, which consists in a vacancy posted in the 25th percentile of the HHI distribution (0.011) to the 75th percentile (0.078), increases the probability that that vacancy offers training by 1.4 percentage points, consisting of almost 5 percent increase of the likelihood that an employer provides on-the-job training with respect to the mean.[20]

Tables 5 shows the analogous impact of local labor market concentration on the probability that an employer posts wage information. in both the OLS and the IV specifications an increase in HHI reduces the probability of wage information posting. All coefficients are also all statistically significant at the 1 percent level. Considering our preferred specification (Column, 6), it documents that an interquartile range increase in the HHI decreases the probability that a vacancy provides wage information by 1.3 percentage points, consisting of a 10 percent decrease with respect to the mean.

[19]The first stage regressions are reported in Table T1.

[20]The percentage points are computed as: $[log(p75) - log(p25)] * \hat{\gamma} * 100$. The percent effect ($\beta_{\%}$) with respect to the mean is computed as follows:

$$\beta_{\%} = (\hat{\gamma} * log(\delta + 1))\frac{100}{\text{MDV}}$$

$$\text{where} \qquad \delta = \frac{p75(HHI) - p25(HHI)}{p25(HHI)}$$

where $\hat{\gamma}$ is the estimated coefficient, p75(HHI) and p25(HHI) are taken from Table 2, and MDV is the mean of the dependent variable, i.e. predicted training probability, taken from Table 4.

7 Additional effects of employer concentration

In this section, we investigate other potential effects that employer concentration may have. First, we provide evidence that employer concentration decreases the years of experience and education required. Second, we show that vacancies in highly concentrated markets have more words, but these words have fewer syllables. Moreover, those vacancies demanded more intellect personality traits than vacancies in low concentrated markets.

7.1 Experience and Education demanded

At high concentration levels, given the limited presence of suitable job seekers in that market and the difficulty of attracting workers from other markets, employers may find it more challenging to find job candidates. For this reason, they could reduce the requirements requested to fill the vacancy. Thus, for example, an employer could reduce the years of experience or education generally required to attract even those workers who work in other occupations or who can be taught the necessary missing skills through on-the-job training. This effect is indeed what we observed.

Using the same empirical strategy described in Section 5, in Table 6, we show how the years of experience and education demanded in the vacancies decrease significantly as the level of concentration increases. In particular, the first two columns show how, if the concentration increases by an interquartile range, the years of experience demanded decreases between 0.05 (OLS) and 0.04 (IV) index points, which consists of a reduction of around 4% and 3% with respect to the mean, respectively. Similarly, with the same increase in HHI, our education index variable loses 0.04 (OLS) and 0.1 (IV) index points, representing a reduction of 3% and 8%, respectively. Finally, the last two columns show how the same increase in concentration reduces the probability that the job advertisement requires at least a Bachelor's degree of about 1.2 and 2.5 percentage points, consisting of a reduction of 5% and 9%, respectively.[21]

[21] The experience variable is approximated at the year, 0 is no experience demanded, 1 is less or equal than 1 year of experience, 2 more than 1 but less or equal than 2 years of experience, and so on. The education variable takes values: 0 if no qualification required, 1 if High School diploma, 2 if Associate's degree, 3 if Bachelor's degree, 4 if Master's degree, and 5 if PhD. Graduate takes value 1 if education takes value greater or equal than 3.

7.2 Job text complexity and type of skill demanded

Does the concentration level in a local labor market affect the amount of information and the complexity of the job ad text? To answer this question, we consider four different variables (1) the log of the number of words in the job ad text, (2) the average number of syllables per word, and if the job ad text has some words linked to specific personality traits. In this regard, we distinguished between (3) intellect traits and (4) non-intellect traits. The intellect traits are identified by keywords such as intellectual, complex, creative, imaginative, and innovative. On the other hand, the non-intellectual traits concern more traits like conscientiousness, agreeableness, and surgency; determined by words such as talkative, assertiveness, cooperative, kind, neat, systematic, practical, sympathetic.

Table 7 and Table 8 provides the OLS and IV estimates for each dependent variable, respectively, following the empirical strategy described in section 5. As concentration increases, the number of words in the job ad's text increases while the number of syllables per word decreases. The results are significant and robust in both specifications (OLS and IV), except for the number of words variable that loses significance in the IV model. These results suggest that as concentration increases, employers increase the information provided to candidates and tend to use simpler words underlined by fewer syllables. Assuming that workers from other occupations require more information because they are less aware of that specific occupation's job features and tasks. These outcomes seem to support the hypothesis that employers are trying to attract workers from different occupations when they are located in highly concentrated markets.

Finally, in the same Tables 7 and 8, we can observe that employers in highly concentrated markets demand more intellectual traits. The effect is significant and robust to both specifications. One rationale could be that intellect traits can be more helpful in enabling workers to learn skills faster and better. If the employers plan to train her new workers, she would like workers who learn new competencies more quickly and thus have more intellect traits.

8 Conclusion

We have examined the effects of labor market competition on employer-provided training decisions and on their decision on posting wages. Taking advantage of the information disclosed by employers in the job vacancy text, we develop a ML measure of the probability that a vacancy is offering on-the-job training. We investigate whether the different level of labor market competition, measured as vacancy concentration, affects the employer's decision to provide training and disclose wage information. The empirical evidence documents how employers increase their training offers at a lower level of labor competition but decrease the probability of revealing wage information. We further observe that employers have lower requirements regarding experience and education at a lower level of labor competition.

The paper adds to the literature by developing a new way to measure employer-provided training and contributes to the literature by examining the effects of labor market competition. Our measure of employer-provided training, contrary to the survey-based measures, can be quickly updated and used in the universe of online job postings. Our findings are consistent with the classical monopsony theory's upward slope labor supply curve to wages. Despite employers could still extract significant wage markdowns from workers, they reduce their labor demand to keep wages low. This upward labor supply curve to wage encourages employers to increase their labor supply by providing training or reducing their education or experience requirements. Overall, this paper has clear implications for policymakers, showing that recruitment behavior differs in monopsonistic markets. Consequently, labor market competition should be considered when designing policies to mitigate anticompetitive or antidiscrimination practices, as well as labor market policies aimed at bridging the skill gap of displaced workers.

References

Abel, W., S. Tenreyro, and G. Thwaites (2018): “Monopsony in the UK,” *Working Paper*.

Acemoglu, D. and J.-S. Pischke (1998): “Why do firms train? Theory and evidence,” *The Quarterly journal of economics*, 113, 79–119.

Adams-Prassl, A., M. Balgova, and M. Qian (2020): “Flexible work arrangements in low wage jobs: Evidence from job vacancy data,” *Working Paper*.

Arellano-Bover, J. (2020): “The Effect of Labor Market Conditions at Entry on Workers’ Long-Term Skills,” *Review of Economics and Statistics*, 1–45.

Arnold, D. (2020): “Mergers and acquisitions, local labor market concentration, and worker outcomes,” *Working paper*.

Ash, E., S. Galletta, and T. Giommoni (2021): “A Machine Learning Approach to Analyze and Support Anti-Corruption Policy,” *Working Paper*.

Ash, E., J. Jacobs, W. B. MacLeod, S. Naidu, and D. Stammbach (2020): “Unsupervised extraction of workplace rights and duties from collective bargaining agreements,” *Working Paper*.

Azar, J., S. Berry, and I. E. Marinescu (2022): “Estimating labor market power,” *NBER, Working paper*, no. w30365.

Azar, J., I. Marinescu, and M. Steinbaum (2020a): “Labor market concentration,” *Journal of Human Resources*, 1218–9914R1.

Azar, J., I. Marinescu, M. Steinbaum, and B. Taska (2020b): “Concentration in US labor markets: Evidence from online vacancy data,” *Labour Economics*, 66, 101886.

Azkarate-Askasua, M. and M. Zerecero (2020): “The Aggregate Effects of Labor Market Concentration,” *Working Paper*.

Bassanini, A., C. Batut, and E. Caroli (2021): “Labor Market Concentration and Stayers’ Wages: Evidence from France,” *Working Paper*.

Benmelech, E., N. K. Bergman, and H. Kim (2020): “Strong employers and weak employees: How does employer concentration affect wages?” *Journal of Human Resources*, 0119–10007R1.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2022): “Labor market power,” *American Economic Review*, 112, 1147–93.

Berkes, E., P. Mohnen, and B. Taska (2018): “The Consequences of Initial Skill Mismatch for College Graduates: Evidence from Online Job Postings,” *Working Paper*.

Blair, P. Q. and D. J. Deming (2020): “Structural Increases in Demand for Skill after the Great Recession,” *AEA Papers and Proceedings*, 110, 362–65.

Boal, W. M. and M. R. Ransom (1997): "Monopsony in the labor market," *Journal of economic literature*, 35, 86–112.

Bratti, M., M. Conti, and G. Sulis (2021): "Employment protection and firm-provided training in dual labour markets," *Labour Economics*, 69, 101972.

Brenčič, V. (2012): "Wage posting: evidence from job ads," *Canadian Journal of Economics/Revue canadienne d'économique*, 45, 1529–1559.

Brunello, G. and F. Gambarotto (2007): "Do spatial agglomeration and local labor market competition affect employer-provided training? Evidence from the UK," *Regional Science and Urban Economics*, 37, 1–21.

Burdett, K. and D. T. Mortensen (1998): "Wage differentials, employer size, and unemployment," *International Economic Review*, 257–273.

Burke, M. A., A. Sasser, S. Sadighi, R. B. Sederberg, and B. Taska (2020): "No longer qualified? Changes in the supply and demand for skills within occupations," *Working Paper*.

Carnevale, A. P., J. Strohl, and A. Gulish (2015): "College is Just the Beginning: Employers' Role in the $ 1.1 Trillion Postsecondary Education and Training System," *Center on Education and the Workforce McCourt School of Public Policy*.

Chodorow-Reich, G. and J. Wieland (2020): "Secular labor reallocation and business cycles," *Journal of Political Economy*, 128, 2245–2287.

Clemens, J., L. B. Kahn, and J. Meer (2021): "Dropouts need not apply? The minimum wage and skill upgrading," *Journal of Labor Economics*, 39, S107–S149.

Deming, D. and L. B. Kahn (2018): "Skill requirements across firms and labor markets: Evidence from job postings for professionals," *Journal of Labor Economics*, 36, S337–S369.

EU Council (2019): "Council conclusions on the implementation of the Council Recommendation on Upskilling Pathways: New Opportunities for Adults," *Official Journal of the European Union*, C 189, 23—27.

E.U. Parliament (2021): "Proposal for a directive of the European Parliament and of the Council to strengthen the application of the principle of equal pay for equal work or work of equal value between men and women through pay transparency and enforcement mechanisms," *Commission Staff Working Document*.

Faberman, R. J. and G. Menzio (2018): "Evidence on the Relationship between Recruiting and the Starting Wage," *Labour Economics*, 50, 67–79.

Faberman, R. J., A. I. Mueller, A. Şahin, and G. Topa (2017): "Job search behavior among the employed and non-employed," *NBER, Working paper*, no. w23731.

Forsythe, E., L. B. Kahn, F. Lange, and D. Wiczer (2020): “Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims,” *Journal of public economics*, 189, 104238.

Gabaix, X. and R. S. Koijen (2020): “Granular instrumental variables,” *NBER, Working paper, no. w28204*.

Gallup (2021): “The American Upskilling Study: Empowering Workers for the Jobs of Tomorrow,” Tech. rep.

Harhoff, D. and T. J. Kane (1997): “Is the German apprenticeship system a panacea for the US labor market?” *Journal of population economics*, 10, 171–196.

Hershbein, B. and L. B. Kahn (2018): “Do recessions accelerate routine-biased technological change? Evidence from vacancy postings,” *American Economic Review*, 108, 1737–72.

Hershbein, B., C. Macaluso, and C. Yeh (2022): “Monopsony in the US Labor Market,” *American Economic Review*, 112, 2099–2138.

Jarosch, G., J. S. Nimczik, and I. Sorkin (2019): “Granular search, market structure, and wages,” *NBER, Working Paper*, no. w26239.

Kaplan, G. and S. Schulhofer-Wohl (2017): “Understanding the long-run decline in interstate migration,” *International Economic Review*, 58, 57–94.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2018): “Human decisions and machine predictions,” *The quarterly journal of economics*, 133, 237–293.

Kuhn, P., P. Luck, and H. Mansour (2018): “Offshoring and Skills Demand,” *Working Paper*.

Kuhn, P. and K. Shen (2013): “Gender discrimination in job ads: Evidence from china,” *The Quarterly Journal of Economics*, 128, 287–336.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2021): “Gender differences in job search: Trading off commute against wage,” *The Quarterly Journal of Economics*, 136, 381–426.

Lipsius, B. (2018): “Labor market concentration does not explain the falling labor share,” *Working Paper*.

Manning, A. (2003): *Monopsony in motion: Imperfect competition in labor markets*, Princeton University Press.

——— (2011): “Imperfect competition in the labor market,” *Handbook of labor economics*, 4, 973–1041.

Manning, A. and B. Petrongolo (2017): “How local are labor markets? Evidence from a spatial job search model,” *American Economic Review*, 107, 2877–2907.

Marcato, A. (2021): “Lights and Shadows of Employer Concentration: On-the-Job Training and Wages,” *Working Paper*.

Marinescu, I., I. Ouss, and L.-D. Pape (2021): "Wages, hires, and labor market concentration," *Journal of Economic Behavior & Organization*, 184, 506–605.

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

Marinescu, I. and R. Wolthoff (2020): "Opening the black box of the matching function: The power of words," *Journal of Labor Economics*, 38, 535–568.

Martins, P. S. (2018): "Making their own weather? Estimating employer labour-market power and its wage effects," *Working Paper*.

Méndez, F. (2019): "Training opportunities in monopsonistic labour markets," *Applied Economics*, 51, 4757–4768.

Modestino, A. S., D. Shoag, and J. Ballance (2020): "Upskilling: Do employers demand greater skill when workers are plentiful?" *Review of Economics and Statistics*, 102, 793–805.

Mohrenweiser, J., T. Zwick, and U. Backes-Gellner (2019): "Poaching and firm-sponsored training," *British Journal of Industrial Relations*, 57, 143–181.

Monster (2021): "Fall 2021, Hiring Report," Tech. rep.

Muehlemann, S. and S. C. Wolter (2011): "Firm-sponsored training and poaching externalities in regional labor markets," *Regional Science and Urban Economics*, 41, 560–570.

Mullainathan, S. and Z. Obermeyer (2019): *A machine learning approach to low-value health care: wasted tests, missed heart attacks and mis-predictions*, National Bureau of Economic Research.

OECD (2021): *Training in Enterprises*, OECD Publishing, Paris.

Popp, M. (2021): "Minimum wages in concentrated labor markets," *Working paper*.

Qiu, Y. and A. Sojourner (2019): "Labor-market concentration and labor compensation," *Working Paper*.

Rinz, K. (2022): "Labor market concentration, earnings, and inequality," *Journal of Human Resources*, 57, S251–S283.

Robinson, J. (1969): *The economics of imperfect competition*, Springer.

Rzepka, S. and M. Tamm (2016): "Local employer competition and training of workers," *Applied Economics*, 48, 3307–3321.

Schubert, G., A. Stansbury, and B. Taska (2020): "Employer Concentration and Outside Options," *Working paper*.

Shimer, R. (2005): "The assignment of workers to jobs in an economy with coordination frictions," *Journal of political Economy*, 113, 996–1025.

Starr, E. (2019): "Consider this: Training, wages, and the enforceability of covenants not to compete," *ILR Review*, 72, 783–817.

U.S. COUNCIL OF ECONOMIC ADVISERS (2016): "Labor Market Monopsony: Trends, Consequences, and Policy Responses," Tech. rep.

——— (2018): "Addressing America's Reskilling Challenge," Tech. rep.

U.S. DEPARTMENT OF JUSTICE/FEDERAL TRADE COMMISSION (DOJ/FTC) (2010): "Horizontal merger guidelines," *Report, Federal Trade Commission, Washington, DC*.

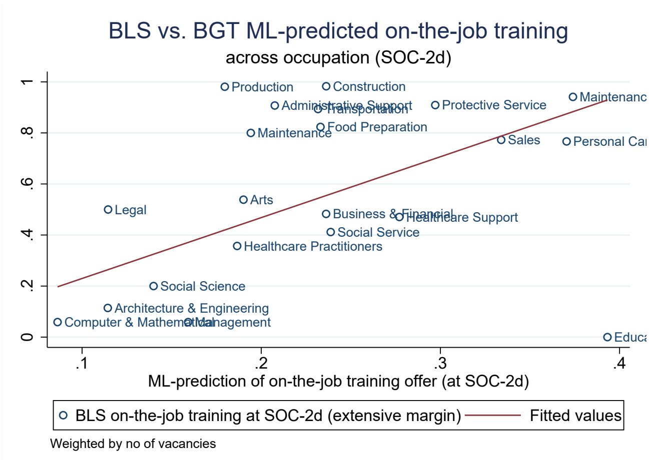
WASHINGTON STATE (2022): "Wage and salary information – applicants for employment," *Engrossed substitute Senate Bill 5761, amendment March 31, 2022*.

WORLD ECONOMIC FORUM (2020): "The Future of Jobs Report," Tech. rep.

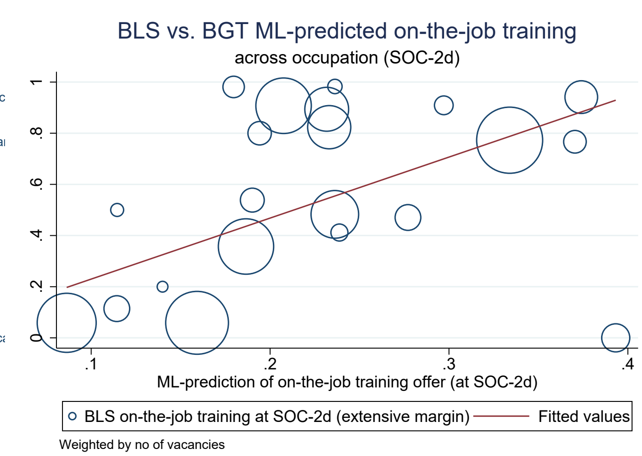
Main Figures

Figure 1: Comparison btw BLS, SIPP and BGT measures of on-the-training

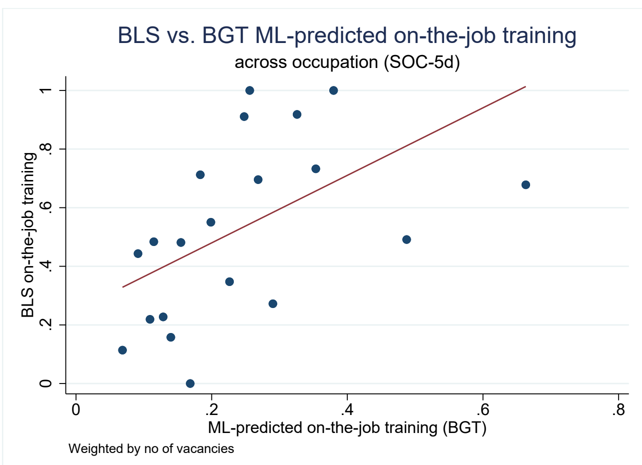
(a) 2-digit SOC



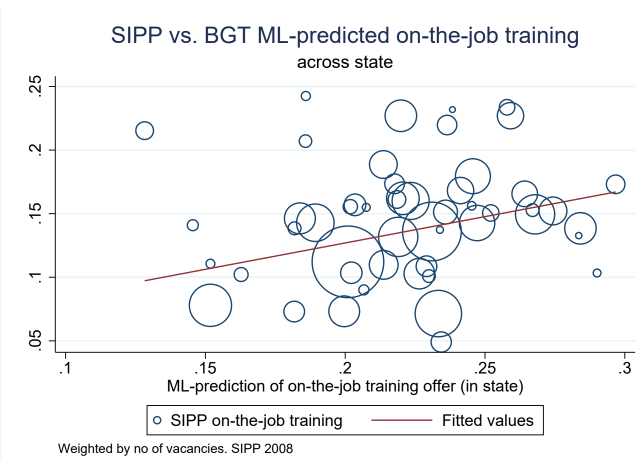
(b) 2-digit SOC



(c) 5-digit SOC

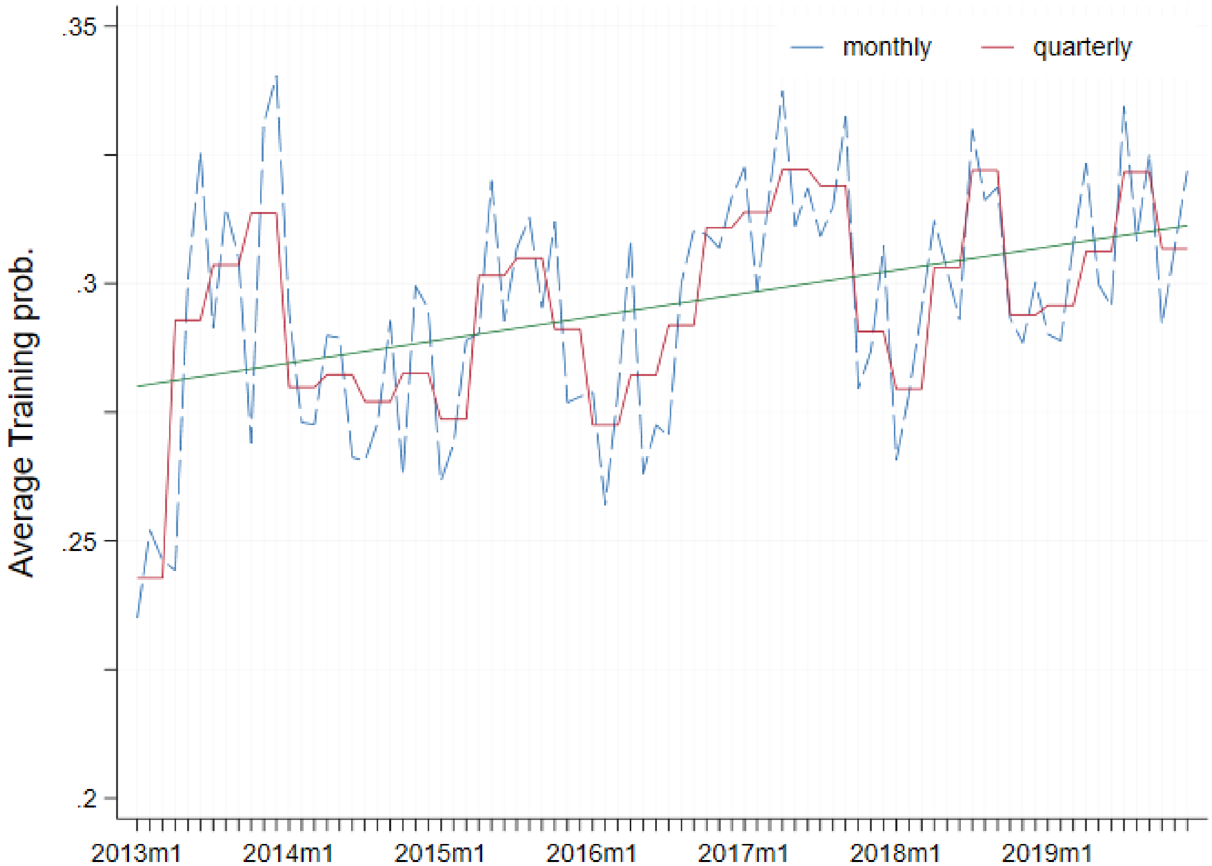


(d) US states



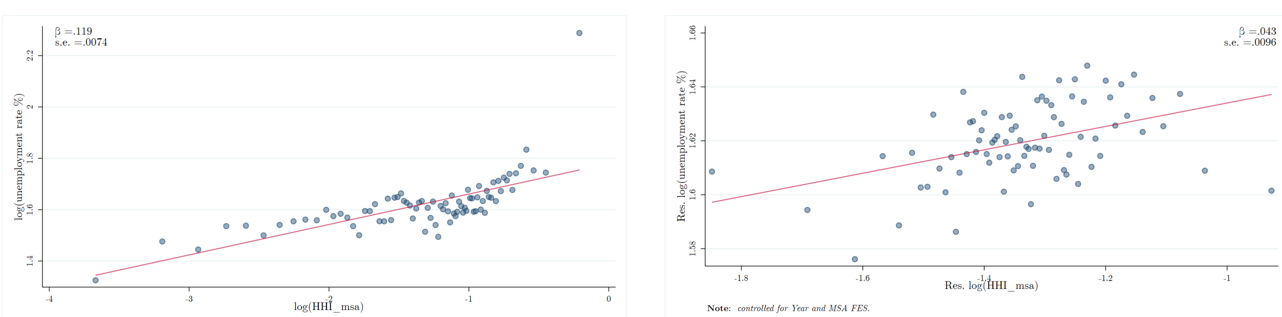
Note: This Figure compares various on-the-job training measures. Our new measure obtained from BGT job ads is on the X axis, while Panels 1b, 1c, have the BLS measure on the y-axis. We look at the correlation across SOC occupations at the 2-digit level in Panels 1b and at the 5-digit level in Panel 1c. Occupations are weighted by their number of job ads posted in 2019.

Figure 2: On-the-training offer (BGT)



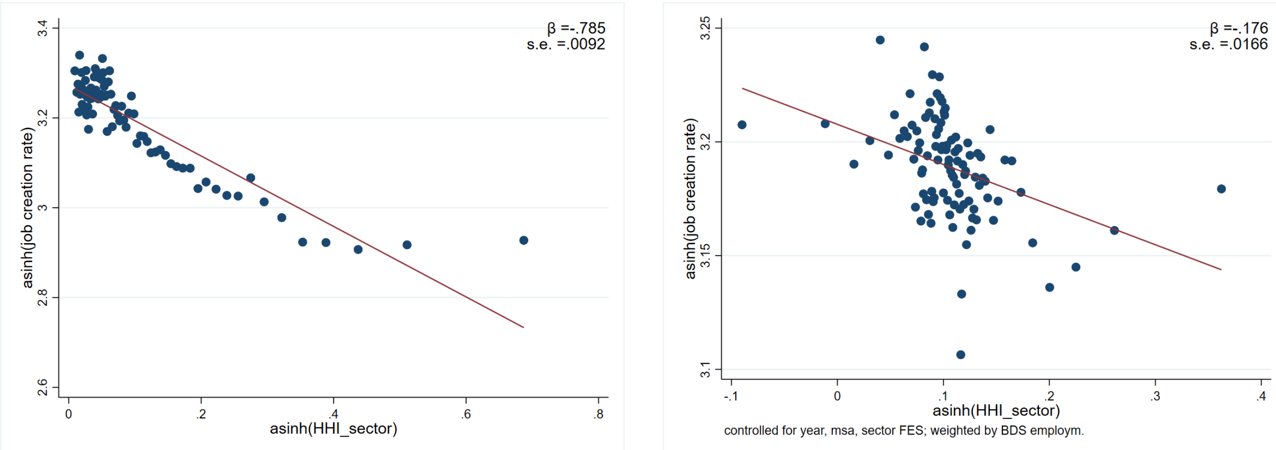
Note: This Figure plots both the monthly and quarterly time series evolution of the share of vacancies offering on-the-job training in our random sample consisting of all the vacancies posted by 10% of the employers posting vacancies in 2019.

Figure 3: Unemployment rate: Binned scatter plots



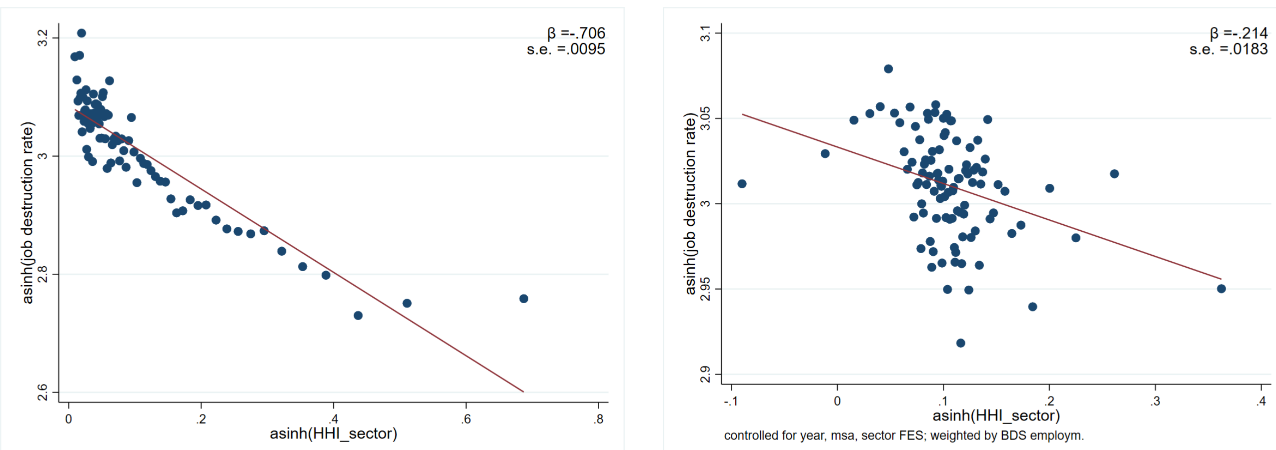
Note: Binned scatter plots between the LAUS unemployemnt rate and log HHI_MSA, for the years 2013-2019. An observation is a combination between a year, and MSA.

Figure 4: Job creation rate: Binned scatter plots



Note: Binned scatter plots between the BDS job creation rate and log HHI_sector, for the years 2013-2019. An observation is a combination between a year, MSA, and sector (NAICS-2d). Asinh stands for "Inverse Hyperbolic Sine Transformation" function. Both plots are weighted by the employment size of the market (BDS data).

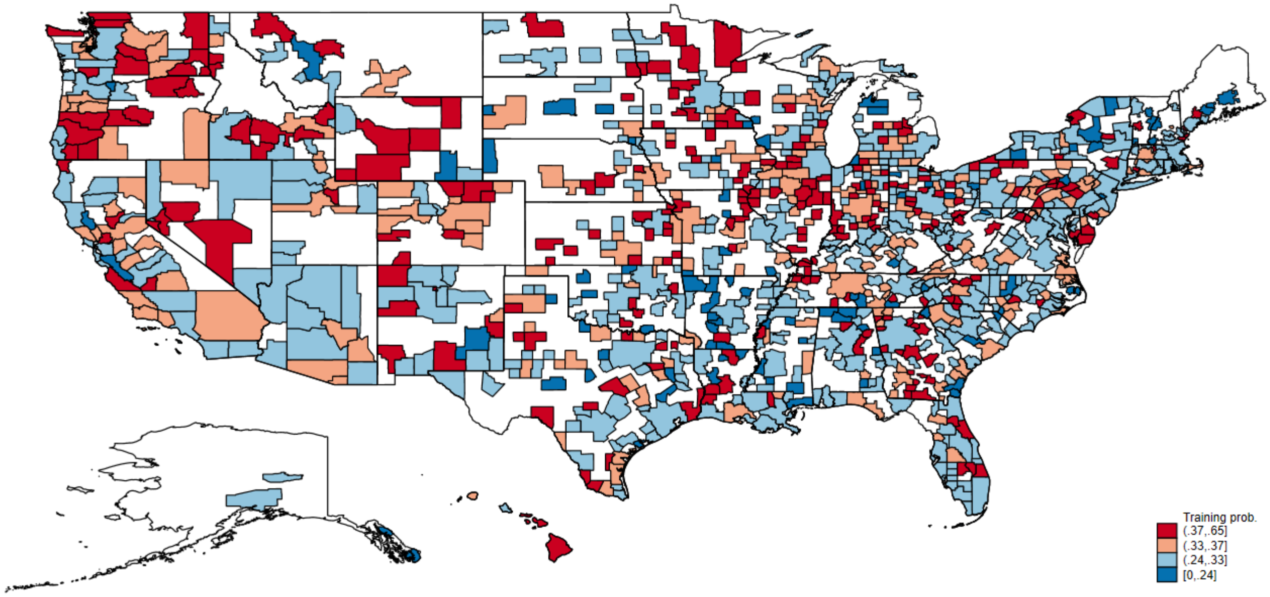
Figure 5: Job destruction rate: Binned scatter plots



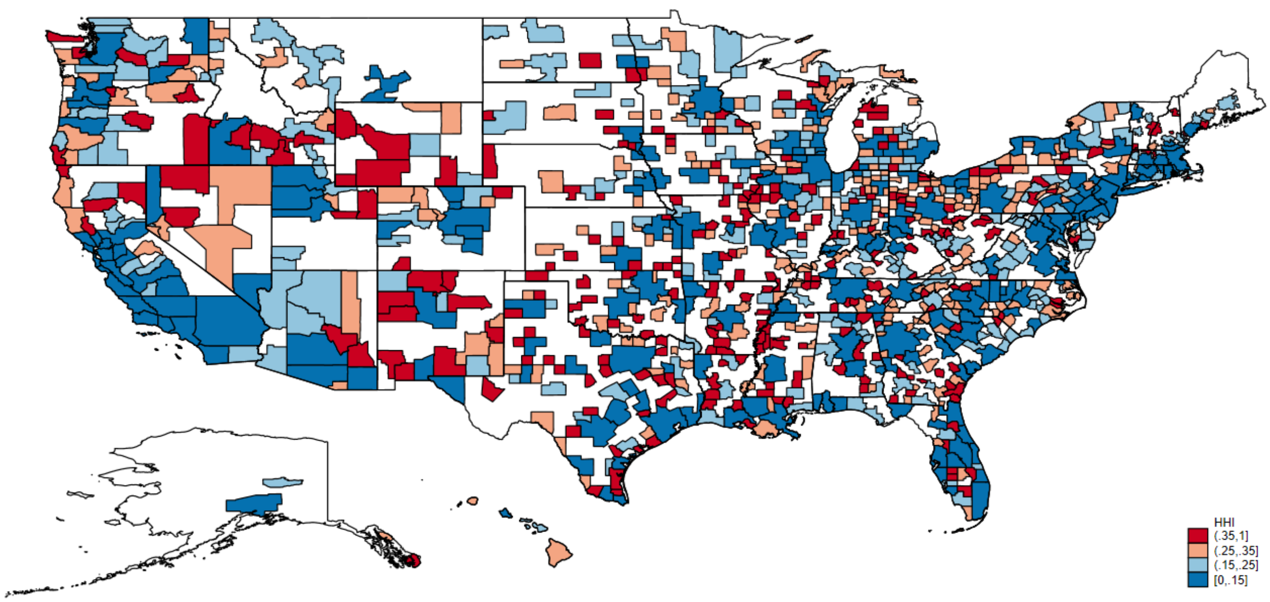
Note: Binned scatter plots between the BDS job destruction rate and log HHI_sector, for the years 2013-2019. An observation is a combination between a year, MSA, and sector (NAICS-2d). Asinh stands for "Inverse Hyperbolic Sine Transformation" function. Both plots are weighted by the employment size of the market (BDS data).

Figure 6: On-the-training offer and HHI concentration across MSA

(a) Average training predicted probability



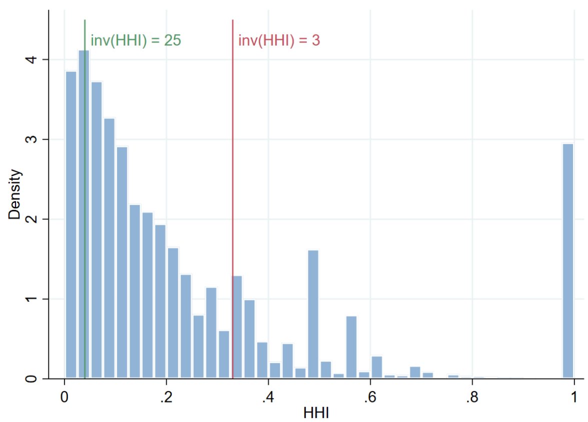
(b) Average HHI



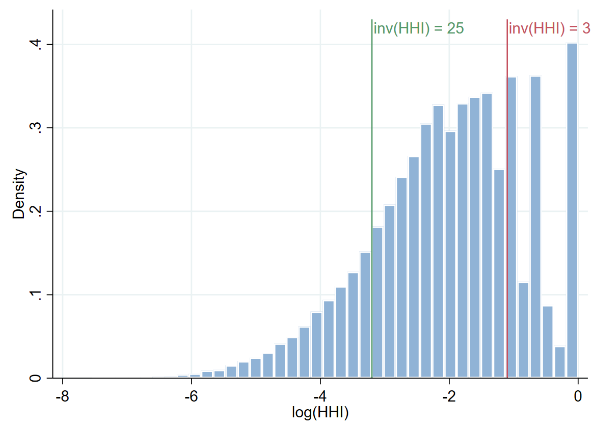
Note: These Figures plot the average training predicted probability (Panel 6a) and HHI (Panel 6b) across MSAs for the year 2019.

Figure 7: Employment concentration in the local labor markets

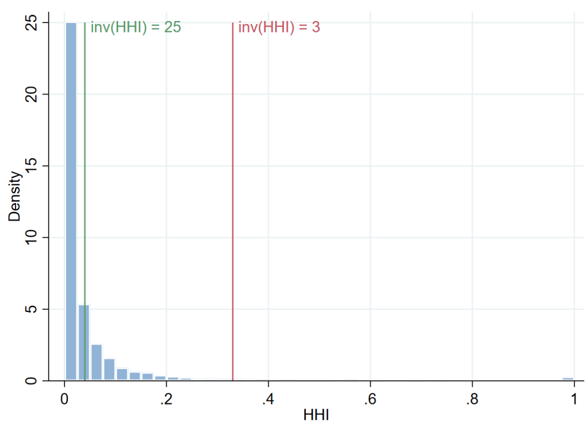
(a) *Average across Markets, level*



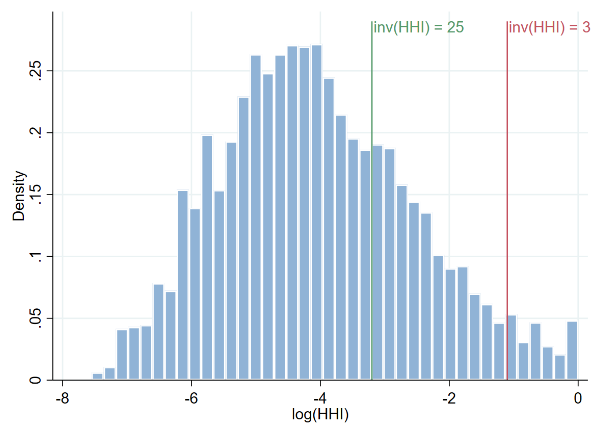
(b) *Average across Markets, logarithm*



(c) *Average across Job Ads, level*



(d) *Average across Job Ads, logarithm*

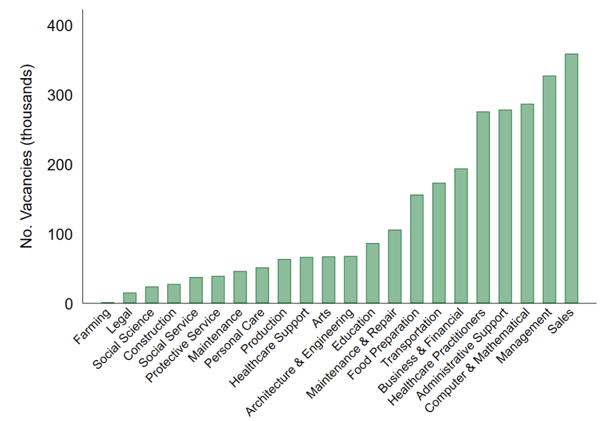


Sample: all vacancies posted in 2019, by a random sample of 10% of all the employers posting vacancy in 2019.

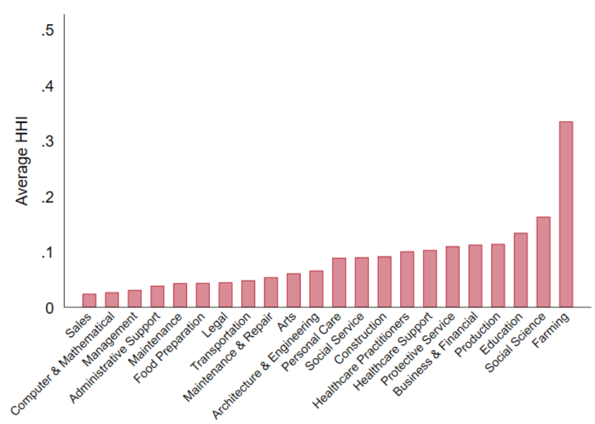
Note: The HHI is computed at the local labor market level, which is defined as a combination of MSA, 6-digit SOC code, and year. The two graphs in the top of the figure are calculated taking the average across local labor market. The two graphs in the bottom of the figure are calculated taking the average across job ads. The inv(HHI) defines the Inverse Herfindahl-Hirschman Index, which can be interpreted as the number of equally sized firms that will obtain the same HHI.

Figure 8: On-the-training offer and HHI concentration across 2-digit occupations

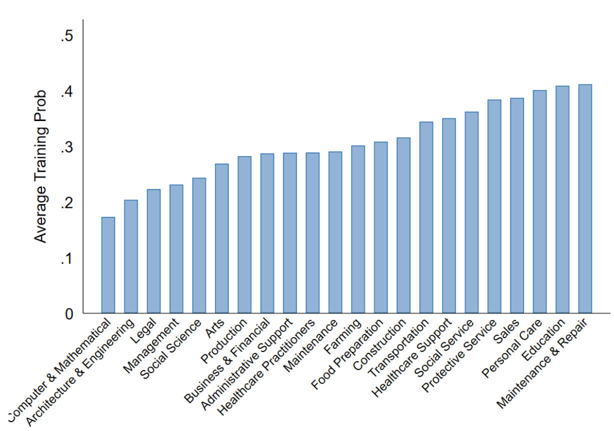
(a) Average number of vacancies



(b) Average HHI



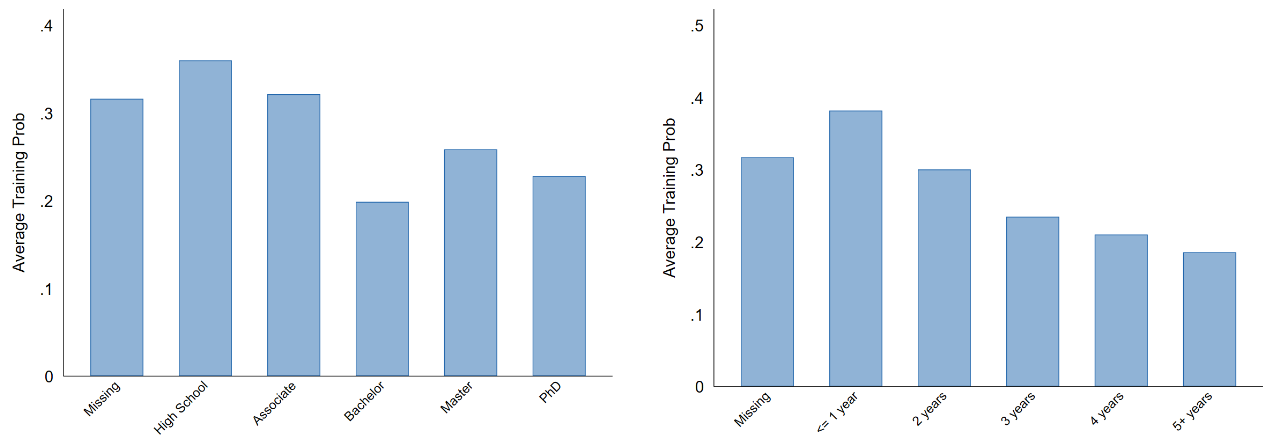
(c) Average training predicted probability



Note: These Figures plot the average number of vacancies, training predicted probability, and HHI across 2-digit SOC codes for the year 2019.

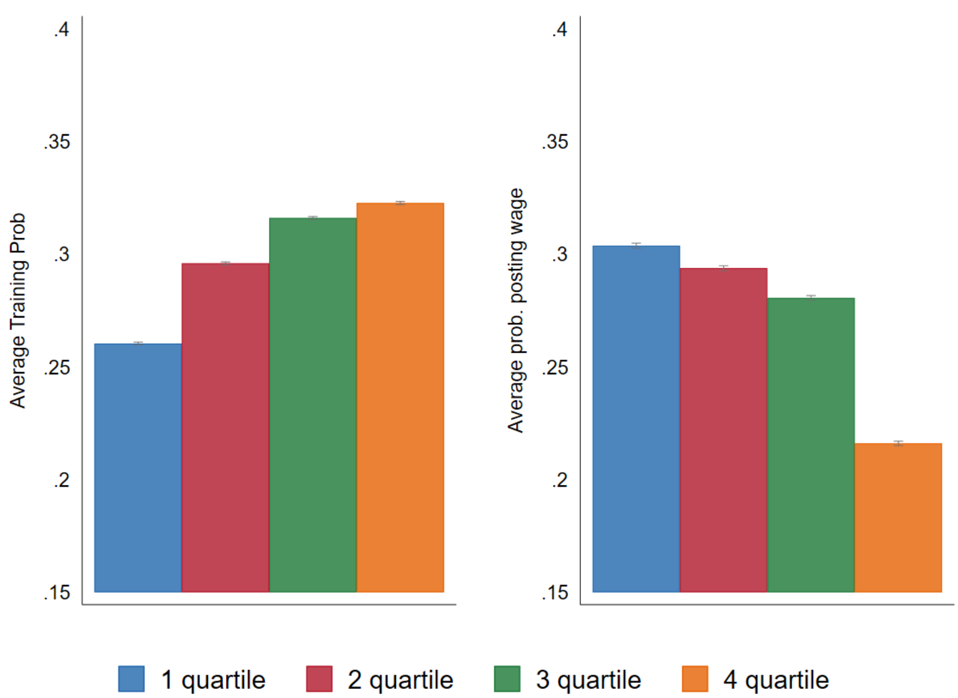
Figure 9: On-the-job training offer by education and experience requirements

(a) Education requirement (b) Experience requirement



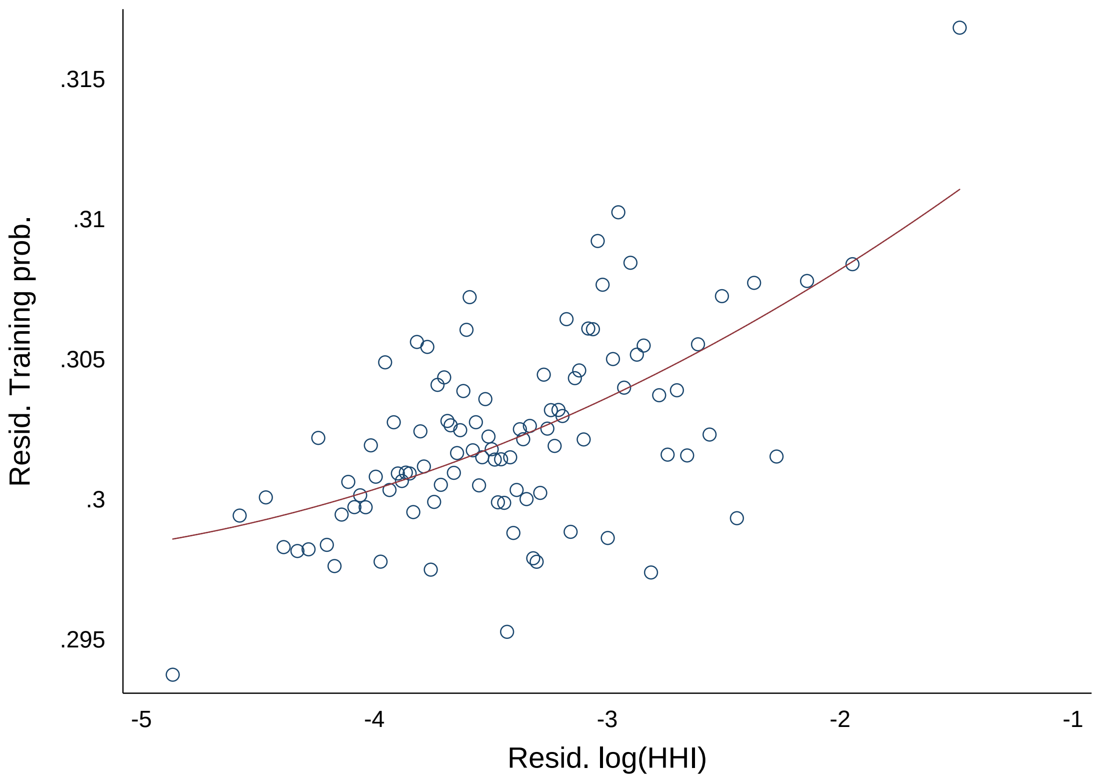
Note: Using BGT 2019 data on our 10% random sample of employers, the Figure plots the probability of offering training according to the level of education (Panel 9a) and experience demanded (Panel 9b) in the vacancy.

Figure 10: Average training offer and wage posted across HHI quartiles



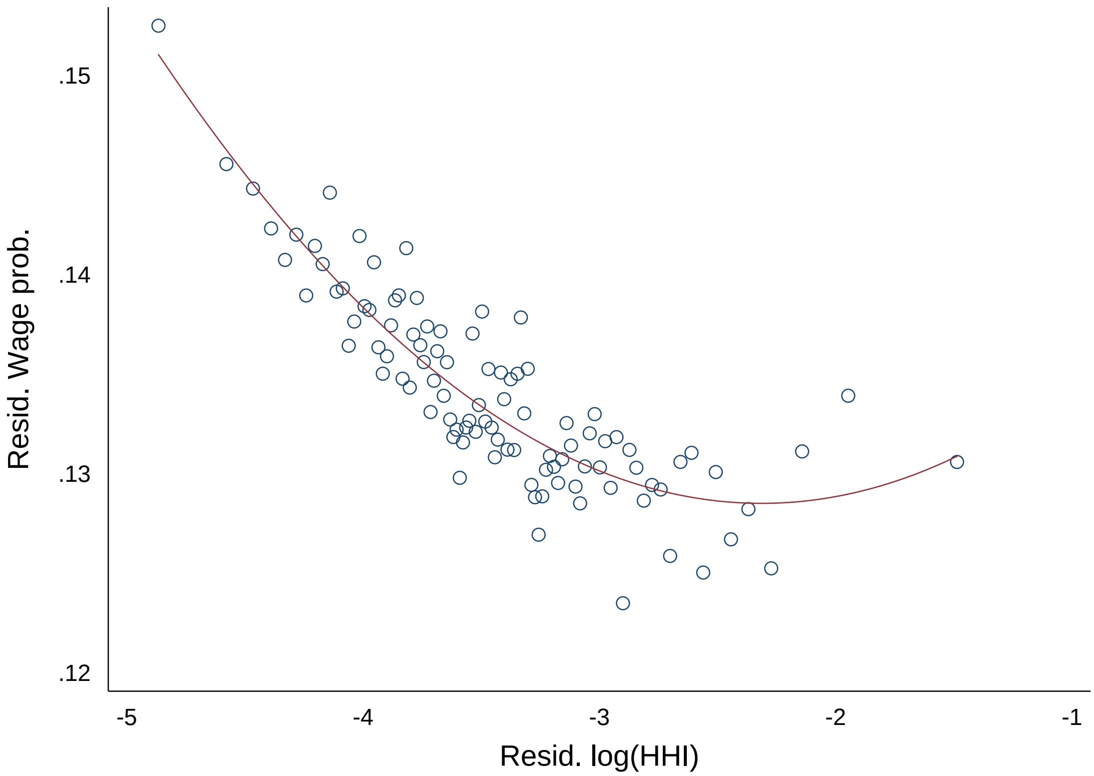
Note: Using BGT 2019 data on our 10% random sample of employers, the Figure (left) shows the average training probability across the HHI distribution quartiles (Q1=0.006, Q2=0.15, Q3=0.48) whereas the (right) Figure displays the share of vacancies posting the wage offered across the different HHI quartiles.

Figure 11: Binscatter plot, residualized regression of labor market concentration and training offer



Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Note: The residuals are computed using as regressors SOC 6-dig, MSA, year, sector, employer, education and experience level fixed effects.

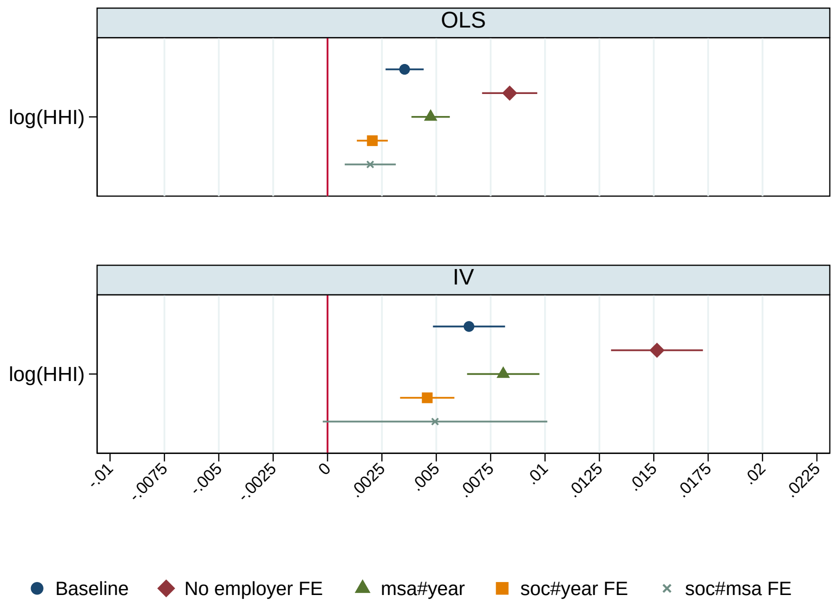
Figure 12: Binscatter plot, residualized regression of labor market concentration and wage posted



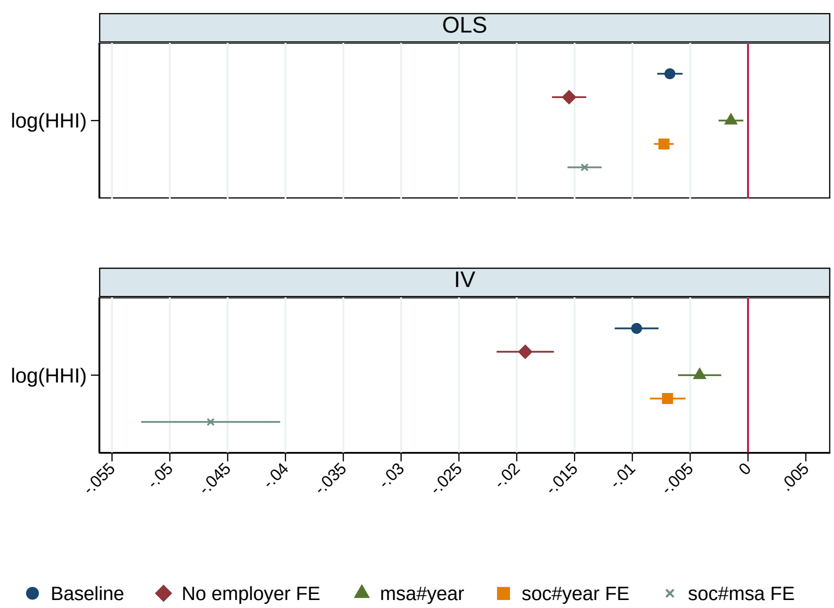
Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Note: The residuals are computed using as regressors SOC 6-dig, MSA, year, sector, employer, education and experience level fixed effects.

Figure 13: Coefficients of training and wage on HHI regressions: robustness checks

(a) Training probability



(b) Posted Wage probability



Note: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019. These Figures plot the coefficient of log. HHI on the probability on offering training (Panel 13a) and on disclosing wage information (Panel 13b). The blue circle shows the estimate when we use individual fixed effects for MSA, year, occupation, sector, and employer. The red diamond uses the same fixed effects of the blue circle one except that it removes the employer fixed effects. The green triangle differs from the baseline has instead of having individual fixed effects for MSA and year, it uses their combination. The yellow square considers the combination of occupation and year, while the teal cross the combination between MSA and occupation. All the regressions controlled for the level of experience and education demanded. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year.

Main Tables

Table 1: Out-of-Sample Metrics for Predicting Training offers

Accuracy	0.795	True Negatives (TN)	1072
F1	0.708	False Negatives (FN)	244
AUC-ROC	0.862	True Positives (TP)	487
Recall	0.666	False Positives (FP)	157
Precision	0.756		

Notes: The table reports the test-set evaluation metrics. Accuracy is the proportion of out-of-sample observation for which the machine-predicted model correctly predict the true label. Recall is the proportion of correct predicted training within the set of vacancies actually offering training (TP/(TP+FN)). Precision instead is the percentage of correct training predictions relative to the total number of training predictions (TP/(TP+FP)). F1 is the weighted harmonic mean of precision and recall. The AUC-ROC identifies the Area Under the Receiver Operating Characteristics.

Table 2: Summary Statistics

	N	mean	sd	25 pct.	median	75 pct.	95 pct.
Training prob.	12,633,515	0.292	0.268	0.083	0.183	0.439	0.873
Wage info.	12,634,279	0.146	0.353	0.000	0.000	0.000	1.000
HHI	12,634,279	0.082	0.150	0.011	0.029	0.078	0.360
log(HHI)	12,634,279	-3.511	1.426	-4.534	-3.551	-2.547	-1.022
HHI(market level)	773,677	0.266	0.269	0.078	0.167	0.344	1.000
log(HHI_market)	773,677	-1.849	1.115	-2.557	-1.792	-1.068	0.000
Education [0,5]	12,634,279	1.181	1.324	0.000	1.000	3.000	3.000
NA	12,634,279	0.440	0.496	0.000	0.000	1.000	1.000
HighSchool	12,634,279	0.248	0.432	0.000	0.000	0.000	1.000
Associate	12,634,279	0.046	0.210	0.000	0.000	0.000	0.000
Bachelor	12,634,279	0.232	0.422	0.000	0.000	0.000	1.000
Master	12,634,279	0.025	0.155	0.000	0.000	0.000	0.000
PhD	12,634,279	0.009	0.097	0.000	0.000	0.000	0.000
Experience [0,5]	12,634,279	1.374	1.798	0.000	0.000	2.000	5.000
NA	12,634,279	0.507	0.500	0.000	1.000	1.000	1.000
$\leq$ 1 year	12,634,279	0.157	0.364	0.000	0.000	0.000	1.000
2 years	12,634,279	0.097	0.296	0.000	0.000	0.000	1.000
3 years	12,634,279	0.074	0.261	0.000	0.000	0.000	1.000
4 years	12,634,279	0.025	0.156	0.000	0.000	0.000	0.000
$\geq$ 5 years	12,634,279	0.140	0.347	0.000	0.000	0.000	1.000
unemploy. rate	12,615,947	4.619	1.566	3.550	4.275	5.300	7.600
log. avg. OES Wages	8,407,096	3.216	0.542	2.769	3.157	3.625	4.143
log BGT wages	1,845,306	10.696	0.585	10.243	10.636	11.082	11.716
High skill occ.	12,634,279	0.531	0.499	0.000	1.000	1.000	1.000
No Words	12,633,515	338.293	212.301	200.000	310.000	436.000	694.000
Avg no Syllables	12,630,562	2.132	0.212	2.015	2.157	2.269	2.418

Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Note: Each observation consists in a vacancy. This table displays the summary statistics. The training probability defines the probability that a vacancy is offering training, which is measured using our machine learning algorithm. Wage Info determines whether the vacancy is posting any information regarding the wages proposed. The HHI (Herfindahl-Hirschman index) is the vacancy employer concentration in the local labor market where the vacancy is posted. Each market is defined as the combination of an occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. The unemployment rate is the BLS Local Area Unemployment Statistics (LAUS), measured at the same market level definition of the HHI. *log. avg. OES wage* is the logarithm of the average wage in the same local labor market, obtained from the BLS Occupational Employment and Wage Statistics (OEWS). The *log. avg. BGT wage* is instead the log. of the average between the min and max yearly wage displayed in the vacancy text. *High skill occ.* defines whether the vacancy is in high skilled occupations, which are those in the 1-3 digit SOC category. *No. Words* counts the number of words in the job ad text, while *Avg. no. Syllables* is the average number of syllables of all the words in the job ad text.

Table 3: Summary Statistics by Predicted Training

	No Training.		Training		
	mean	sd	mean	sd	diff.
Wage Info.	0.141	0.348	0.163	0.370	-0.022
HHI	0.079	0.146	0.093	0.166	-0.014
log(HHI)	-3.554	1.426	-3.353	1.415	-0.201
Education					
NA	0.430	0.495	0.477	0.499	-0.047
HighSchool	0.222	0.416	0.343	0.475	-0.121
Associate	0.047	0.211	0.043	0.203	0.004
Bachelor	0.262	0.440	0.120	0.324	0.143
Master	0.028	0.165	0.013	0.113	0.015
PhD	0.011	0.103	0.005	0.069	0.006
Experience [0,5]	1.497	1.881	0.929	1.366	0.568
Missing	0.499	0.500	0.535	0.499	-0.036
$\leq$ 1 year	0.132	0.339	0.247	0.431	-0.115
2 years	0.096	0.294	0.101	0.301	-0.005
3 years	0.082	0.274	0.045	0.208	0.036
4 years	0.028	0.166	0.013	0.114	0.015
$\geq$ 5 years	0.163	0.369	0.058	0.235	0.105
unemploy. rate	4.620	1.555	4.613	1.605	0.007
log. avg. OES Wage	3.259	0.548	3.067	0.492	0.192
log. avg. BGT Wage	10.701	0.592	10.678	0.563	0.024
High skill occ.	0.566	0.496	0.403	0.490	0.163
No. Words	322.462	209.167	395.821	213.663	-73.359
Avg. no. Syllables	2.154	0.215	2.052	0.177	0.102
Observations	9907234		2726281		12633515

Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.

Note: Each observation consists in a vacancy. This table displays the summary statistics. The training probability defines the probability that a vacancy is offering training, which is measured using our machine learning algorithm. Wage Info determines whether the vacancy is posting any information regarding the wages proposed. The HHI (Herfindahl-Hirschman index) is the vacancy employer concentration in the local labor market where the vacancy is posted. Each market is defined as the combination of an occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. The unemployment rate is the BLS Local Area Unemployment Statistics (LAUS), measured at the same market level definition of the HHI. *log. avg. OES wage* is the logarithm of the average wage in the same local labor market, obtained from the BLS Occupational Employment and Wage Statistics (OEWS). The *log. avg. BGT wage* is instead the log. of the average between the min and max yearly wage displayed in the vacancy text. *High skill occ.* defines whether the vacancy is in high skilled occupations, which are those in the 1-3 digit SOC category. *No. Words* counts the number of words in the job ad text, while *Avg. no. Syllables* is the average number of syllables of all the words in the job ad text.

Table 4: Estimates of labor market concentration on predicted training probability

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
log(HHI)	0.0101***	0.0101***	0.0042***	0.0140***	0.0145***	0.0072***
	(0.0007)	(0.0007)	(0.0005)	(0.0029)	(0.0032)	(0.0023)
Year FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
SOC_6d FE	✓	✓	✓	✓	✓	✓
NAICS_2d FE		✓	✓		✓	✓
Employer_FE			✓			✓
Controls-IV				✓	✓	✓
MDV	0.294	0.304	0.305	0.288	0.298	0.299
mean(HHI)	0.067	0.071	0.071	0.066	0.069	0.069
std(log(HHI))	1.357	1.341	1.340	1.340	1.326	1.325
R2	.205	.221	.472	.	.	.
F	.	.	.	1,258	1,238	1,370
no employers	121,382	94,893	62,390	93,924	72,733	45,937
N	12,092,827	10,687,657	10,655,154	7,266,614	6,428,006	6,401,210

Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019. Notes: Each observation consists in a vacancy. This table reports the TSLS and OLS regression outputs using as dependent variable *Training*, which defines the estimated probability that that vacancy is offering on-the-job training. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. The "Control-IV" identifies the exposure, vacancy growth, and predicted vacancy growth controls as described in Section 5. MDV reports the Mean of the Dependent Variable. F shows the Kleibergen-Paap F Statistics from the regression. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year. ***, **, and * indicate significance level at the 1%, 5%, and 10% level, respectively.

Table 5: Estimates of labor market concentration on wage information

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
log(HHI)	-0.0169***	-0.0154***	-0.0063***	-0.0222***	-0.0192***	-0.0066***
	(0.0008)	(0.0008)	(0.0006)	(0.0033)	(0.0033)	(0.0025)
Year FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
SOC_6d FE	✓	✓	✓	✓	✓	✓
NAICS_2d FE		✓	✓		✓	✓
Employer_FE			✓			✓
Controls-IV				✓	✓	✓
MDV	0.144	0.133	0.132	0.138	0.128	0.126
mean(HHI)	0.067	0.071	0.071	0.066	0.069	0.069
std(log(HHI))	1.357	1.341	1.340	1.340	1.326	1.325
R2	.124	.196	.418	.	.	.
F	.	.	.	1,259	1,238	1,370
no employers	121,382	94,893	62,390	93,924	72,733	45,937
N	12,093,571	10,688,366	10,655,863	7,267,157	6,428,522	6,401,726

Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Notes: Each observation consists in a vacancy. This table reports the TSLS and OLS regression outputs using as dependent variable *Wage*, which defines whether that vacancy is posting wage information. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. The "Control-IV" identifies the exposure, vacancy growth, and predicted vacancy growth controls as described in Section 5. MDV reports the Mean of the Dependent Variable. F shows the Kleibergen-Paap F Statistics from the regression. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year. ***, **, and * indicate significance level at the 1%, 5%, and 10% level, respectively.

Table 6: OLS estimates of labor market concentration on experience and education demanded

	Experience		Education		Graduate	
	OLS	IV	OLS	IV	OLS	IV
log(HHI)	-0.0242***	-0.0199*	-0.0192***	-0.0498***	-0.0064***	-0.0131***
	(0.0026)	(0.0111)	(0.0022)	(0.0093)	(0.0010)	(0.0042)
MDV	1.362	1.413	1.186	1.216	0.262	0.275
mean(HHI)	0.071	0.069	0.071	0.069	0.071	0.069
std(log(HHI))	1.340	1.325	1.340	1.325	1.340	1.325
R^2	0.371	.	0.458	.	0.458	.
F	.	1,370	.	1,370	.	1,370
no employers	62,390	45,937	62,390	45,937	62,390	45,937
N	10,655,863	6,401,726	10,655,863	6,401,726	10,655,863	6,401,726

Sample: all vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Note: Each observation consists in a vacancy. This table reports the OLS and IV regression outputs using as dependent variables: (1) *Experience*, which is an index variable that takes values 0 if no experience demanded (or missing), 1 if "$\leq$ 1 year", 2 if 2 years, 3 if 3 years, 4 if 4 years, 5 if 5 or more years). (2) *Education*, which is an index variable that takes values 0 if no education or missing, 1 if High School diploma, 2 if Associate's degree, 3 if Bachelor's degree, 4 if Master's degree, and 5 if PhD. (3) *Graduate* that takes value 1 if the vacancy demanded at least a Bachelor's degree, 0 otherwise; those vacancies with missing information on the education requirement are not included. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. All the regressions uses year, SOC 6-digit, MSA, NAICS 2-digit, and employer fixed effects. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year.

Table 7: OLS estimates of labor market concentration on the job ad text information

	log. no. words	Avg. Syllables	Intellect	Non-Intellect Personality
log(HHI)	0.0123***	-0.0024***	0.0024***	0.0031***
	(0.0019)	(0.0003)	(0.0009)	(0.0010)
MDV	5.628	2.129	0.291	0.413
mean(HHI)	0.071	0.071	0.071	0.071
std(log(HHI))	1.340	1.340	1.340	1.340
R^2	0.310	0.578	0.331	0.288
no employers	62,390	62,390	62,390	62,390
N	10,652,695	10,652,695	10,655,863	10,655,863

Sample: all the vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.
Note: Each observation consists in a vacancy. This table reports the OLS regression outputs using as dependent variables: (1) *No. Words*, which counts the number of words in the job ad text; (2) *Average Syllables*, which measures the average number of syllables each word in the job ad text has. (3) *Intellect* which identifies if a vacancy is requiring some intellectual skills; while *Non-intellect* if the vacancy is demanding a skill which is not directly related to intellect, but more on agreeableness, conscientiousness, and surgency. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. All the regressions uses year, SOC 6-digit, MSA, NAICS 2-digit, and employer fixed effects. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year.

Table 8: IV estimates of labor market concentration on text information

	log. no. words	Avg. Syllables	Intellect	Non-Intellect Personality
log(HHI)	0.0110	-0.0080***	0.0078**	0.0084**
	(0.0084)	(0.0012)	(0.0037)	(0.0041)
MDV	5.619	2.132	0.294	0.408
mean(HHI)	0.069	0.069	0.069	0.069
std(log(HHI))	1.325	1.325	1.325	1.325
F	1,369	1,369	1,370	1,370
no employers	45,937	45,937	45,937	45,937
N	6,400,211	6,400,211	6,401,726	6,401,726

Sample: all the vacancies posted between 2013-2019, by a random sample of 10% of all the employers posting vacancy 2019.

Note: Each observation consists in a vacancy. This table reports the 2SLS IV regression outputs using as dependent variables: (1) *No. Words*, which counts the number of words in the job ad text; (2) *Average Syllables*, which measures the average number of syllables each word in the job ad text has. (3) *Intellect* which identifies if a vacancy is requiring some intellectual skills; while *Non-intellect* if the vacancy is demanding a skill which is not directly related to intellect, but more on agreeableness, conscientiousness, and surgency. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. All the regressions uses year, SOC 6-digit, MSA, NAICS 2-digit, and employer fixed effects. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year.

Online Appendix

This is the online Appendix of Title...

A Tagging process

Figures A1 and A2 show the tagging instructions given to twelve research assistants in December 2020. We prepare a sample of vacancy posted in 2019. First, we take a random sample of 5,000 vacancies, denoted $\mathcal{S}_a$. The random sampling is stratified per month. Second, we select among the 2019 vacancies all vacancies whose text comprises the words ”train”, ”tuition reimbursement”, ”personal development plan”, ”career development”, ”professional development program”, ”education assistance”, ”continuing education”. From this subsample, we take a random sample of 5,000 vacancies, denoted $\mathcal{S}_b$, still stratifying per month. We append both samples to obtain the $\mathcal{S}$ sample to be tagged. From this sample of 10,000 vacancies, we ask each Research Assistant to tag 500 vacancies each.

Figure A1: Tagging instructions: page 1

TAGGING Instructions

Each of the following tags is a binary variable. Write 1 for TRUE and leave it blank for FALSE. Write U if you are unsure.
One job ad can have multiple flags.
For each of the TRUE or Unsure flags, report, in its respective "key phrase" column, the phrase(s)/group of words for which you made that decision.
If there are more phrases referring to a Tag, separate them with a semicolon ";".

TAGS list

- **Y (training)**
  - The job ad offers/sponsors some form of training or education program that will help new hires to acquire new competences or skills
    - *Examples*: "we'll train you"; "Paid training"; "New employee training"; "Training and room to advance"; "Practice-paid continuing education opportunities"; "* Paid Training & Continuous Professional Development"; "8-week comprehensive training program"; "No Experience Needed - Paid Training!"; "continuing education and advancements are among our top priorities"; …

- **G (general training)**
  - If the training/education offered has a general purpose, that can be recognized or transferred to other employers. The training is not limited to some employer's specific competences. In particular, flag those training offers for which a worker will receive some type of certification for the training received.
    - Ex: "Educational Assistance programs"; "Tuition Reimbursement"; "online training courses"; …

- **S (specific training)**
  - If the training is specific to that occupation/employer. The training is intended to provide the job applicants with specific skills/competences required to perform that job for that employer, but it's hardly exploitable by competing firms. For example, training individuals on firm products' features, or training program designed to introduce workers to the company organizational policies or guidelines.
    - Ex: "You'll be trained to educate clients on our products, services, and benefits"; "Brand training are provided before going into the field"; …

***Note:** General and specific training are training, but not all trainings are either generic or specific.*
When the training offer is clearly neither generic nor specific, flag exclusively the Y tag. If it is specific (general), flag both Y and S (G) tags. A single job ad can offers simultaneously general and specific training.

Note:

Figure A2: Tagging instructions: page 2

The rationale behind the following last tags is to enable the machine learning algorithm to distinguish job ads that offer "proper" training from those that instead refer to training as a task or requirement demanded to the job candidate.

- **T (task)**
  - Training is not a "benefit" but a task/mansion for the applicants. The job applicants will not receive training, but they are those that will train other employees or clients.
    - "Set up training and mentoring to grow team"; "Train and mentor new team members"; "provide education, training and support to patients and families"; "Conducts field training or retraining and instructs crew on new or revised job units"; …

- **R (requirement)**
  - The employer requires job applicants to possess some sort of training, or the possess of some sort of training is considered a plus. Do not consider demands for previous experience/skill/competence in this flag but consider only those ads where training is explicitly demanded.
    - "Experience in implementation, training and documentation preferred."; "Subspecialty training/certification is also highly desirable but not required"; "should have prior military or law enforcement experience, or comparable training or certification," …

- **D (disclaimer)**
  - The job ad refers to training, however this is not specific to this job ad, but it's included in a generic disclaimer.
    - "This policy applies to all terms and conditions of employment, including recruiting, hiring, placement, promotion, termination, layoff, recall, and transfers, leaves of absence, compensation and training."; "[We use this information for ..] manage workforce activities and personnel generally, including for recruitment, background screening, performance management, career development, payments administration, employee training, leaves and promotions"; ..

- **E (equality & diversity)**
  - The job ad has an explicit statement on equality and diversity in hiring
    - "We are an equal opportunity employer"; "We encourage applications from under-represented groups"…

Note:

Figure A3: Example of a False Negative

SWAT Inventory Specialist
Best Buy
Mobile, Alabama

What does a Best Buy SWAT Inventory Specialist do?

At Best Buy our mission is to leverage the unique talents and passions of our employees to inspire, delight, and enrich the lives our customers through technology and all its possibilities. If you have a passion and curiosity for what is possible and enjoy people, we invite you to join us on this mission.

A Best Buy SWAT Inventory Specialist ensures inventory integrity in the store through a variety of inventory adjustments and data collection tools. The SWAT Specialist consistently and accurately completes and communicates stock count. They identify, determine and communicate high shrink categories. After identifying the root cause of replenishment issues, they follow up with leadership until the problem is resolved.

Job responsibilities include:

*Executing the inventory integrity process from end to end
*completing inventory daily tasks as assigned
*communicating and coaching store employees and leadership on the importance of inventory integrity and any process gaps that were identified
*Other duties as assigned.

What are the Professional Requirements of a Best Buy SWAT Inventory Specialist?

Basic Qualifications
*Ability to work successfully as part of a team
*Ability to work a flexible schedule inclusive of holidays, nights and weekends
*Ability to lift or maneuver 50-100 pounds, with or without accommodations
Preferred Qualifications
*3 months experience in retail, customer service or related fields

Additional Job Information
What are my rewards and benefits?
Discover your career here! At Best Buy we offer much more than a paycheck. Surrounded by the latest and greatest technology, a team of amazing coworkers and a work environment where anything is possible, you'll find it easy to be your best when you work with us. We provide an exciting work environment with a community of techno learners where you can be yourself while investing in your career. Empowered with knowledge you will discover endless opportunities to grow. From deep employee discounts to tuition reimbursement, to health, wealth and wellness benefits, to learning and development programs, we believe the success of our company depends on the passion of employees for learning, technology and people.

Figure A4: Example of a False Positive

IT Systems Engineer - Federal (N)
Centurylink
WORKS FROM HOME - Colorado, United States
Information Technology

Job Summary:
CenturyLink (NYSE: CTL) is a global communications and IT services company focused on connecting its customers to the power of the digital world CenturyLink offers network and data systems management, big data analytics, managed security services, hosting, cloud, and IT consulting services The company provides broadband, voice, video, advanced data and managed network services over a robust 265,000-route-mile US fiber network and a 360,000-route-mile international transport network Visit CenturyLink for more information
As an integral member of the Tier II group, the IT Systems Engineer II provides advanced Tier II support of a nationwide diverse fiber optic OTN/DWDM transport backbone, as well as two (2) geographically, separated Network Operation Centers (NOC) The position serves as the NOC primary point of contact (POC) for problem escalations to include troubleshooting network devices, tools, and data services The Tier II Engineer will interact with Network Provisioners and/or equipment vendors to identify and develop the solution of escalated troubles The Tier II Engineer is expected to provide accurate documentation such as SOPs, Training Modules, Reports and Project Narratives. The position requires a high-level understanding of Layers 1/2 inter-dependences and Transport proficiency. The network is expected to sustain continual growth to include additional sites and customers across the life of the program and the ITSE responsibilities continue to grow as new customers are added to the network.
Job Description:
-Perform Transport, IP, and Network Management functions in engineering and operations environment and operates as technical lead on problem escalations Serves as the first level of escalation to Tier 1 NOC personnel on Transport related issues
-Serves as technical POC and Lead for Provisioning, Configurations, and TTU Interface with Network Provisioner and configuration managers for requirements, designs, and any other information needed to perform tasks
-Performs advanced diagnostics using software and hardware tools to determine network status and optimize network performance
-Assist provisioning, configurations, and TTU tasks by supporting new system/circuit activations or deactivations as driven by customer requirements
-The ITSE II works closely with Provisioner in implementing system requirements to provide connectivity for TDM (DS-1, DS-3), SONET (OC-12/48/192), OTN, IP, and DWDM paths
-Provide technical guidance, direction, and training to junior technicians with emphasis on Transport equipment
-Create technical operational documentation and SOP for Transport discipline supporting the NOC
-Create and maintain Circuit Layout Records (CLR), Network Diagrams, and other supporting documentation Review and provide feedback or updates to network as-built, Engineering Development Plans (EDP), and configuration management drawings
-On-call after hours/on-call support as required
-Junior Technical to Network Provisioner and ITSE III: receive technical guidance and mentorship
-85% Initial travel is required 13;
Qualifications:
-Bachelor's degree in an Information Technology field and 5 years experience required or seven years of applicable work experience
-The individual is required to keep up with the high demand for the position
-Firm knowledge of SONET, OTN, DWDM, and IP, as well as a high level of proficiency in troubleshooting service troubles in this arena, is required
-Experience with Ciena and Infinera Optical equipment and software is required Long haul optical equipment certifications and training are highly desired
-Demonstrates a high-level of proficiency in IP technical discipline Able to operate and accomplish the technical task with minimum guidance Proficient knowledge of common network architectures for routing, switching, and security technologies are required
-Familiar with the uses of Transport management tool and network management protocols, including but not limited to One-Control, Node Manager, Site Manager
-Security+ or equivalent IAT/IAM/IASAE Level 2 of DoD 85701 is required
-JNCIA or CCNA is preferred
-Firm knowledge of the encryption devices KG-175G/D, KIV-7M, KG-340
Other Requirements
-This position requires a Top Secret/SCI clearance
-Expected to work in a shift environment in support of 24x7 operations
-Up to 15% travel can be expected due to possible deployments to lab facilities and other NOC locations (85% initial on the job travel should be expected) 13;
Requisition:
This job may require successful completion of an online assessment A brief description of the assessments can be viewed on our website at findcenturylinkjobs/testguides/
EEO Statement:
We are committed to providing equal employment opportunities to all persons regardless of race, color, ancestry, citizenship, national origin, religion, veteran status, disability, genetic characteristic or information, age, gender, sexual orientation, gender identity, marital status, family status, pregnancy, or other legally protected status (collectively, protected statuses) We do not tolerate unlawful discrimination in any employment decisions, including recruiting, hiring, compensation, promotion, benefits, discipline, termination, job assignments or training.
Disclaimer:
The above job definition information has been designed to indicate the general nature and level of work performed by employees within this classification It is not designed to contain or be interpreted as a comprehensive inventory of all duties, responsibilities, and qualifications required of employees assigned to this job Job duties and responsibilities are subject to change based on changing business needs and conditions.

B Alternative approaches to measure labor market competition and comparison with employment concentration

In Section 2.2.2, we compare the Herfindhal-Hirschman index (HHI) with the unemployment rate from Local Area Unemployment Statistics (LAUS) from U.S. Bureau of Labor Statistics and the job creation and job destruction from Business Dynamics Statistics (BDS) from U.S Census.

The LAUS data combines three different sources: the Current Population Survey (CPS), the Current Employment Statistics (CES) survey, and state unemployment insurance (UI) systems, to create estimates that are adjusted to the statewide measures of unemployment.[22]

We use the LAUS annual information on the unemployment rate for each metropolitan statistical area (MSA), however LAUS does not provide additional information on the occupation or industry of these unemployed workers. Therefore, for comparability reason, we average our HHI measure to the MSA $\times$ year; as follows

$$HHI_{m,t} = \sum_{j} \mu_{j,m,t} HHI_{j,m,t}$$

were $\mu_{j,m,t}$ is the share of vacancies posted in MSA m by occupation j in year t, and $HHI_{j,m,t}$ is the aforementioned HHI measured at MSA$\times$SOC$\times$year.

The Business Dynamics Statistics (BDS) is a product of the U.S. Census Bureau, and it is compiled from the Longitudinal Business Database (LBD), a confidential database.

The Business Dynamics Statistics (BDS) tracks changes over time, providing annual measures of establishment openings and closings, firm startups and shutdowns, and job creation and destruction for each establishment. These measures are available at finest dimension as at a combination between an industrial sector, 2-digit NAICS, MSA, and year.[23]

We focus mainly on the job creation and job destruction rate measures. These measure start from job destruction and job creation, which are the number of jobs created or destructed in all the establishment in a specific segment (MSA$\times$NAICS2-dig) each year. These absolute

[22] More information and data are available at: https://www.bls.gov/lau/home.htm

[23] More information and access to the data at the following link: https://www.census.gov/programs-surveys/bds/data.html

measure are then divided by the average number of employees across establishment in that segment to construct the job creation and destruction rates.

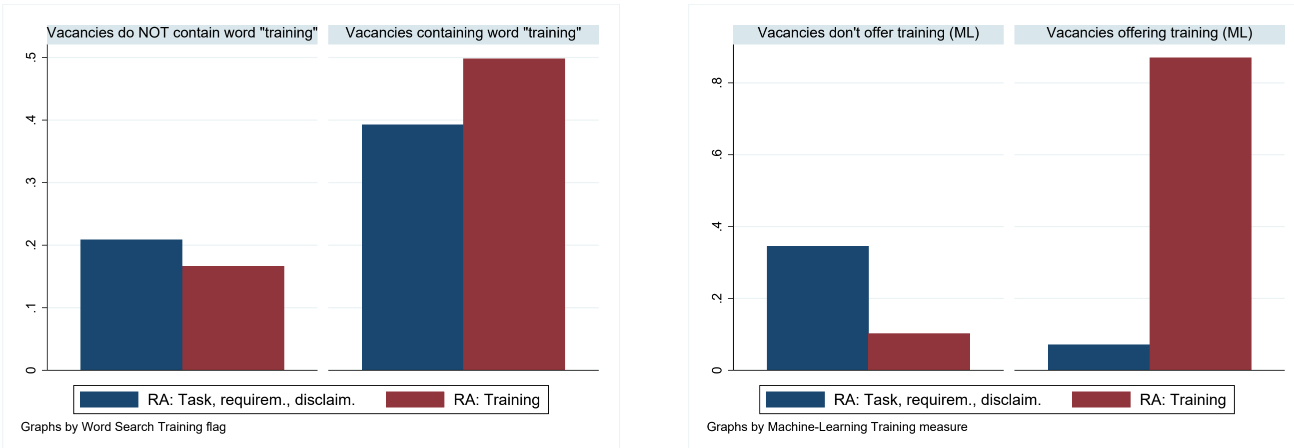
As for the LAUS unemployment measure, also for this measure there is an issue of comparability with our HHI measure, for this reason we average our HHI measure at the same level of the BDS measure, as follows:

$$HHI_{m,k,t} = \sum_{j} \mu_{jkmt} HHI_{jmt}$$

where μ_{jkmt} is the share of vacancies for occupation j in sector k, for the same MSA m and year t.

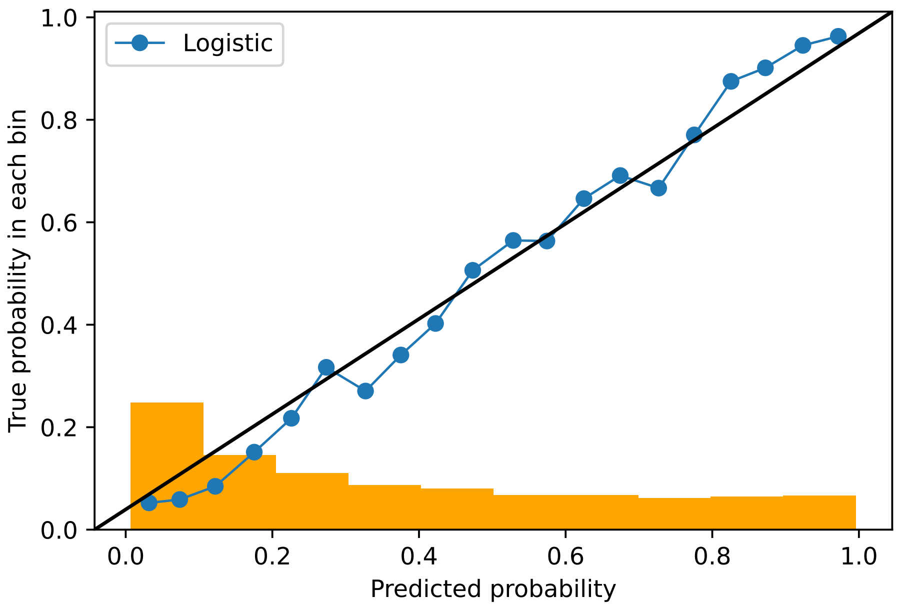
C Extra Figures

Figure F1: Comparison between Machine learning and Word search approaches



Notes: Figures compare the share of correct predictions using our machine learning algorithm (ML) and searching for the word "training" within a job vacancy text. Using the manual annotations, in blues are indicated those vacancies that requires training as a requirement or task, or mention training in a disclaimer. In red those vacancies that actually offer training.

Figure F2: Calibration Plot for Training classifier

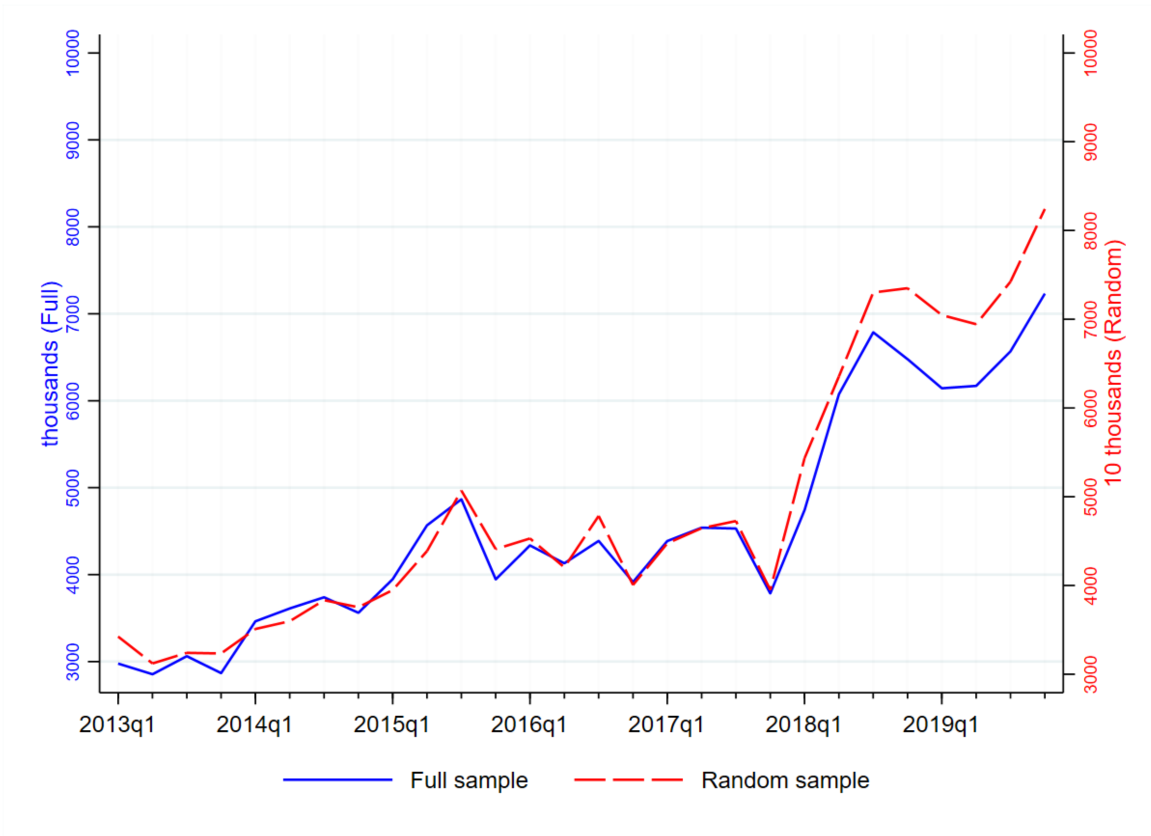


Notes: Calibration plot showing the average true training offers probability (on the y-axis) and the average predicted training probability (x-axis) for each of the 20 bins in which the (out-of-sample) test set is divided according to the predicted training probability. The orange histogram displays the density of the predicted training probability. The 45-degree line identifies perfect calibration.

Figure F3: Word Cloud: most predictive phrases for offering training

Notes: The Figure displays the most predictive n-grams to identifies vacancies offering on-the-job training.

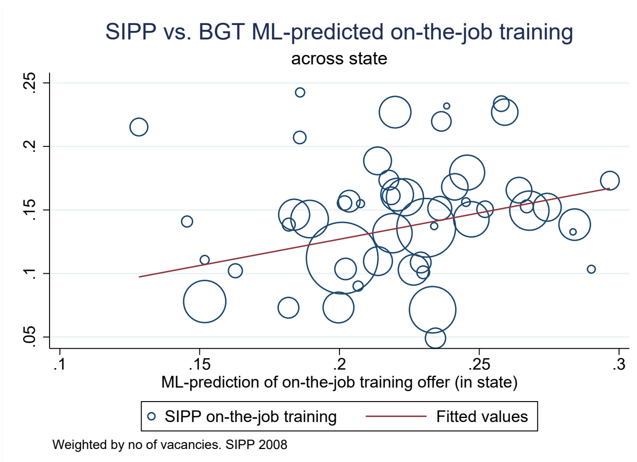
Figure F4: Number of job ads: Full sample vs Random sample



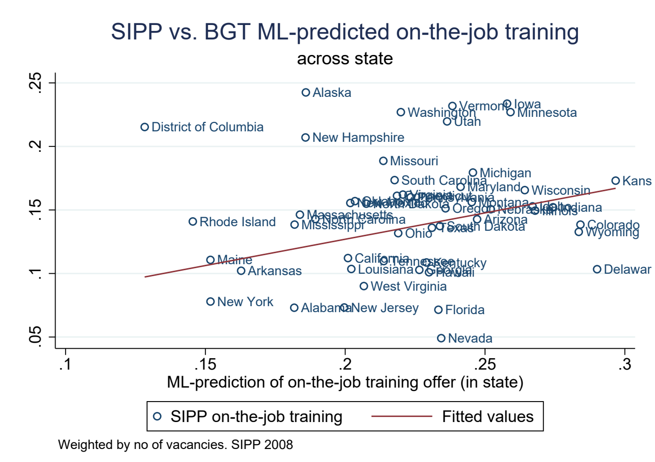
Note: This Figure plots time series evolution of the quarterly number of vacancies in the entire BGT data excluding those vacancies with no employer name and our random sample consisting of all the vacancies posted by 10% of the employers posting vacancies in 2019. The quarterly number of vacancies are expressed in thousands of units.

Figure F5: Comparison btw BLS, SIPP and BGT measures of on-the-training

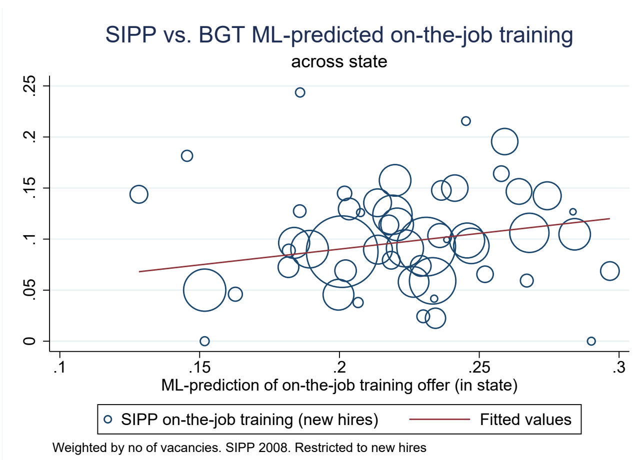
(a) US states



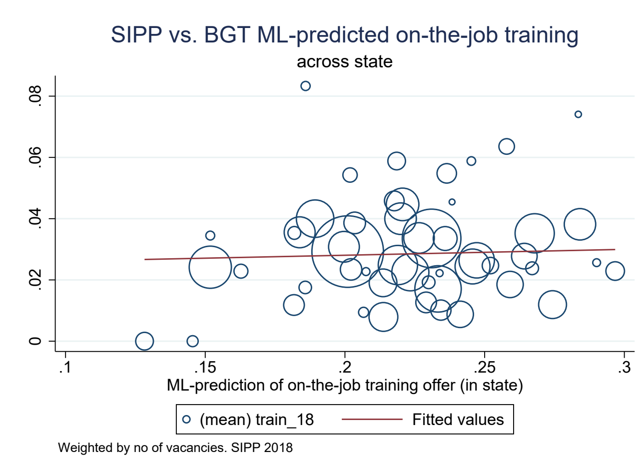
(b) US states



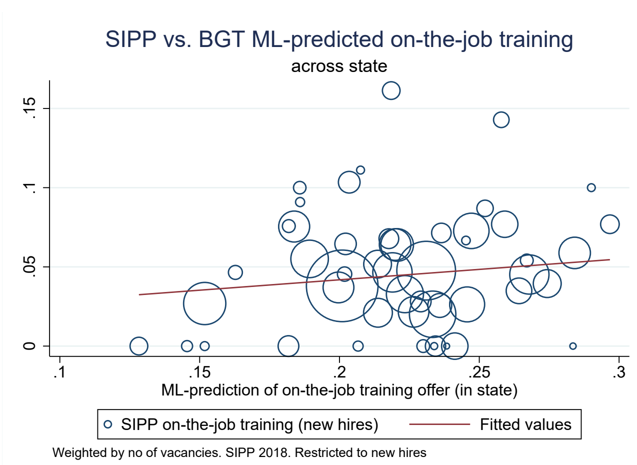
(c) US states



(d) US states

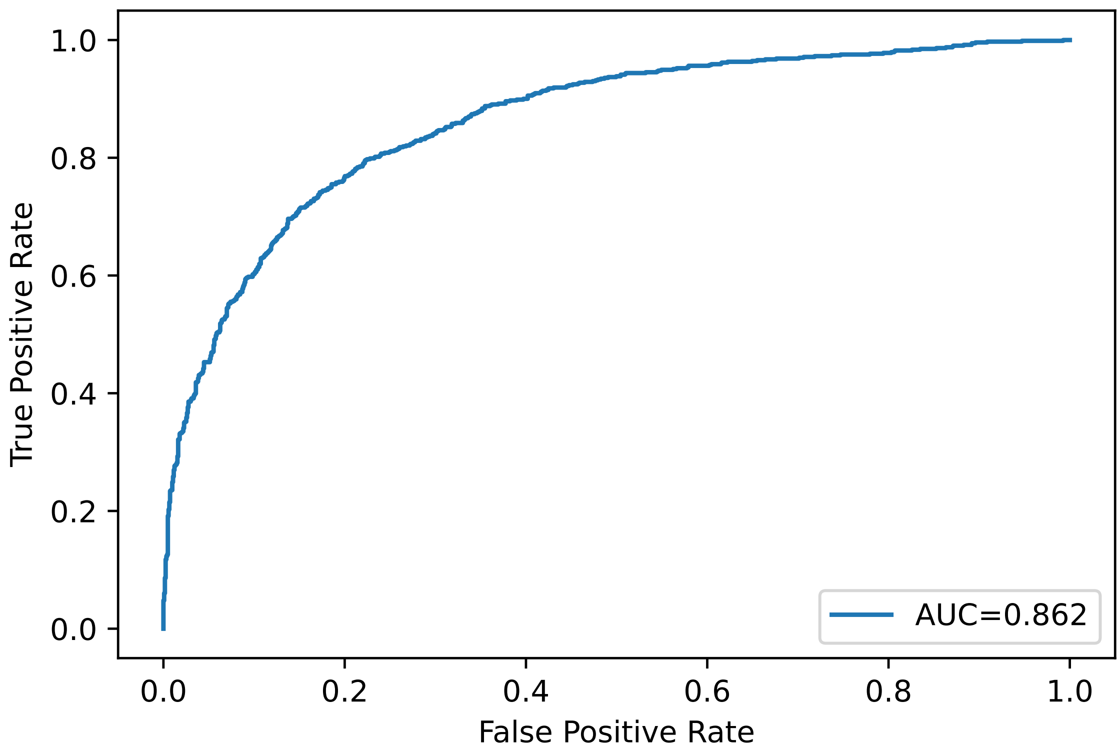


(e) US states



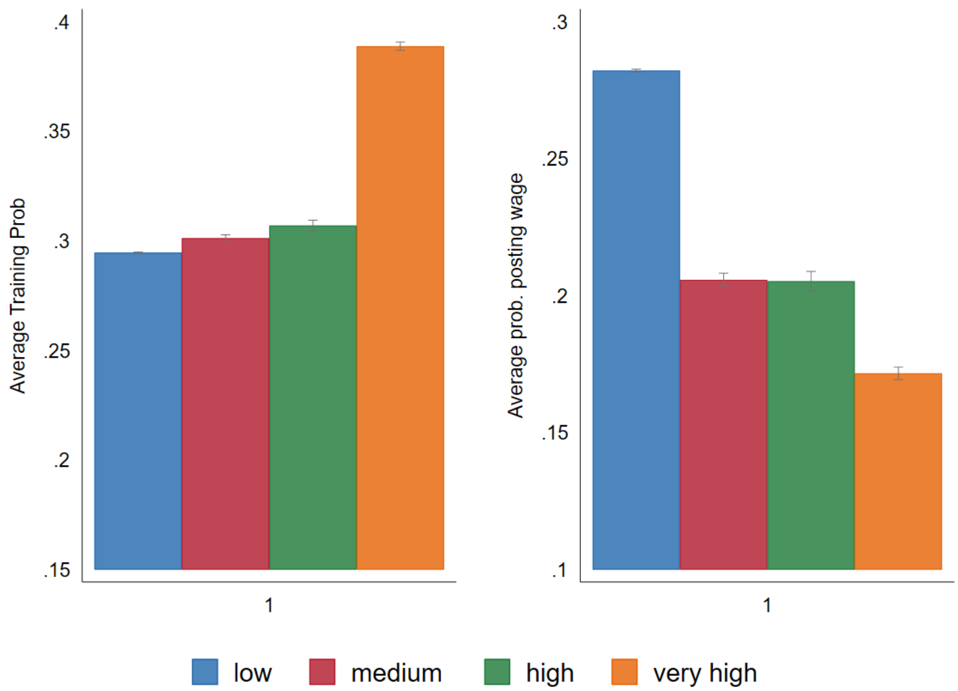
Note: This Figure compares various on-the-job training measures. Our new measure obtained from BGT job ads is on the X axis, while Panels have the SIPP measure on the y-axis. We look at the correlation across states. Weighted by the number of job ads posted in 2019.

Figure F6: AUC-ROC graph



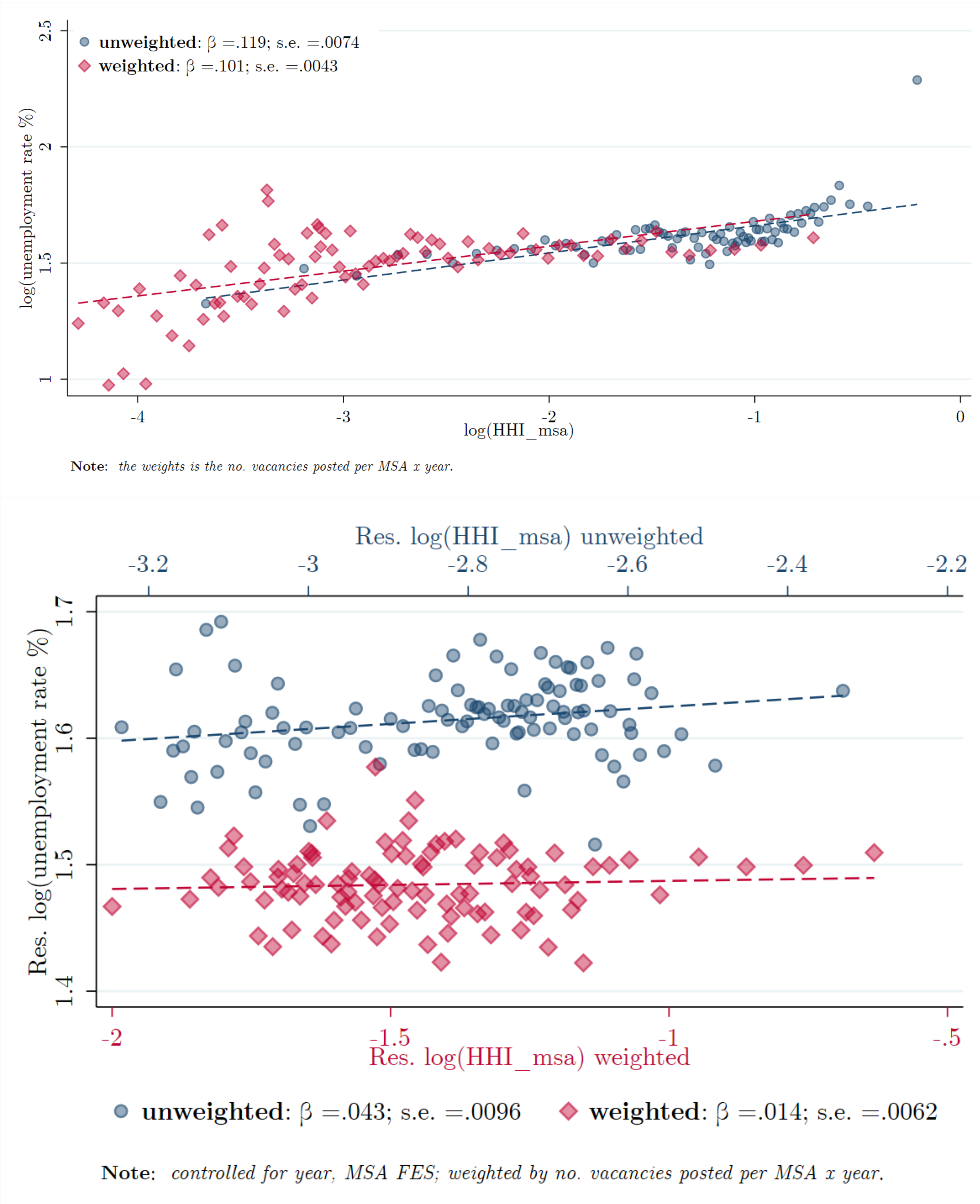
Note: The figure plots ROC curve and the AUC measure of the Logistic regression classifier for predicting training offers. The ROC (Receiver Operating Characteristic) Curve displays the percentage of true positives predicted by the model as the prediction probability cutoff is lowered from 1 to 0. The higher the AUC (area under the curve), the more accurately our model is able to predict outcomes. The True positive rate is defined as TP/(TP+FN), while the False Positive rate as FP/(TN+FP). A AUC of 0.86 means that that given a randomly-seleted training-offering vacancy and a non-training-offering vacancy, there is a 86% change that our model ranks correctly the training vacancy with more predicted probability than the non-training one.

Figure F7: Average training offer and wage posted across HHI classes



Note: Using BGT 2019 data on our 10% random sample of employers, the Figure (left) shows the average training probability across different HHI groups following DOJ/FTC (2010) categories, whereas the (right) Figure displays the share of vacancies posting the wage offered across the different HHI groups.

Figure F8: Binned scatter plots: Log. unemployment rate and HHI_MSA



Note: Binned scatter plots between the LAUS unemployemnt rate and log HHI_MSA, for the years 2013-2019. An observation is a combination between a year, and MSA. Weighted by the number of vacancies posted in a MSA × year.

D Extra Tables

Table T1: First Stage

	log(HHI)	log(HHI)	log(HHI)
instrument log (HHI)	0.1282***	0.1244***	0.1185***
	(0.0035)	(0.0035)	(0.0032)
Controls-IV	✓	✓	✓
Year FE	✓	✓	✓
MSA FE	✓	✓	✓
SOC_6d FE	✓	✓	✓
NAICS_2d FE		✓	✓
Employer_FE			✓
R^2	0.836	0.828	0.843
N	7,444,415	6,428,522	6,401,726

Table T2: Estimates of labor market concentration on predicted training probability, excluding the years 2018 and 2019

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
log(HHI)	0.0076***	0.0078***	0.0048***	0.0088***	0.0099***	0.0071***
	(0.0009)	(0.0009)	(0.0006)	(0.0033)	(0.0035)	(0.0027)
Year FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
SOC_6d FE	✓	✓	✓	✓	✓	✓
NAICS_2d FE		✓	✓		✓	✓
Employer_FE			✓			✓
Controls-IV				✓	✓	✓
MDV	0.291	0.300	0.300	0.282	0.291	0.291
mean(HHI)	0.077	0.080	0.080	0.072	0.074	0.074
std(log(HHI))	1.240	1.231	1.231	1.252	1.244	1.244
R2	.229	.242	.472	.	.	.
F	.	.	.	927	976	1,067
no employers	33,689	28,030	20,216	29,287	24,150	16,879
N	7,081,951	6,514,483	6,506,669	4,499,991	4,115,728	4,108,457

Sample: all vacancies posted between 2013-2017, by a random sample of 10% of all the employers posting vacancy 2019.

Notes: Each observation consists in a vacancy. This table reports the TSLS and OLS regression outputs using as dependent variable *Training*, which defines the estimated probability that that vacancy is offering on-the-job training. The independent variable is the log of the employment HHI, measured at the occupation (6-digit SOC code), Metropolitan Statistical Area (MSA), and year level. The ”Control-IV” identifies the exposure, vacancy growth, and predicted vacancy growth controls as described in Section 5. MDV reports the Mean of the Dependent Variable. F shows the Kleibergen-Paap F Statistics from the regression. Standard errors are clustered at the market level, which consists of the combination between MSA, occupation, and year. ***, **, and * indicate significance level at the 1%, 5%, and 10% level, respectively.