JAQ of All Trades: Job Mismatch, Firm Productivity and Managerial Quality*

Luca Coraggio
University of Naples Federico II

Marco Pagano
University of Naples Federico II

Annalisa Scognamiglio
University of Naples Federico II

Joacim Tåg
Hanken School of Economics and IFN

October 2022

ABSTRACT

Does the matching between workers and jobs help explain productivity differentials across firms? To address this question we develop a job-worker allocation quality measure (*JAQ*) by combining employer-employee administrative data with machine learning techniques. The proposed measure is positively and significantly associated with labor earnings over workers' careers. At firm level, it features a robust positive correlation with firm productivity, and with managerial turnover leading to an improvement in the quality and experience of management. *JAQ* can be constructed for any employer-employee data including workers' occupations, and used to explore the effect of corporate restructuring on workers' allocation and careers.

Keywords: jobs, workers, matching, mismatch, machine learning, productivity, management.
JEL Codes: D22, D23, D24, G34, J24, J31, J62, L22, L23, M12, M54.

*We are grateful for insightful suggestions and comments from Ramin Baghai, Kirill Borusyak, Vicente Cuñat, Daniel Halvarsson, Alex Xi He, Hans Hvide, Simon Ek, Anastassia Fedyk, Jessica Jeffers, Tullio Jappelli, Camelia Kuhnen, Mikael Lindahl, Ahmed Ameya Prapan, Raffaele Saggio, Elia Sartori, Elena Simintzi and participants at the 5th Bank of Italy-CEPR workshop on labour market policies and institutions, 2022 EFA Meetings, EEA-ESEM Congress and WF Conference, 2021 Brucchi Luchino Workshop and ES Winter Meetings, the Labor and Finance Online Seminar, the Bristol-Exeter-Lancaster seminar, and at Ca' Foscari University of Venice, CSEF, EIEF, IFN, Ratio, and UNC Kenan-Flagler Business School. Marco Pagano acknowledges funding from the Einaudi Institute for Economics and Finance (EIEF) and the Italian Ministry for University and Research (MUR). Work on this paper by Annalisa Scognamiglio has been supported by the Modigliani Research Grant awarded by the UniCredit Foundation. Joacim Tåg thanks the Marianne and Marcus Wallenberg Foundation (2015.0048, 2020.0049), Torsten Söderbergs Stiftelse (E31/18, ET2/20), and Jan Wallanders och Tom Hedelius stiftelse samt Tore Browaldhs stiftelse (P22-0094) for financial support. E-mails: lucacoraggio@hotmail.com, pagano56@gmail.com, annalisa.sco@gmail.com, and joacim.tag@ifn.se.

1 Introduction

Why are some firms more productive than others, and why do such productivity differentials persist over time? Research has investigated several avenues to answer this question, exploring the potential role of the quality of capital and labor, innovation, competition, and lately managerial practices as determinants of productivity (Syverson, 2011). In particular, the literature on managerial practices (Bloom and Van Reenen, 2007; Bloom, Brynjolfsson, Foster, Jarmin, Patnaik, Saporta-Eksten, and Van Reenen, 2019) has highlighted the contribution of managerial decisions concerning human resources to productivity. However, the managerial practices related to human resources analyzed in this literature focus on workers' incentives, via pay for performance, promotions and monitoring, and neglect managerial policies governing the allocation of workers to jobs within the firm, which in principle can be a very important determinant of productivity. This shortcoming probably reflects the fact that the measurement of managerial practices in these studies is based on replies to surveys regarding the way managers run their firms' operations, monitoring, incentives and targets; however, it would be very difficult to use self-reported information to evaluate whether managers allocate workers to their best possible use within the firm.

In this paper, we shed light into this black box, by combining administrative employer-employee matched data with machine-learning (ML) techniques, and bringing them jointly to bear on the measurement of the quality of job-worker matches within firms. We then validate the resulting measure of job assignment quality (*JAQ*) by exploring its correlation with firm performance in terms of productivity and profitability, as well as with workers' wages over their careers.

Specifically, we proceed in four steps. First, we estimate via a ML algorithm how workers' characteristics map into jobs in the most productive firms (which should be better at allocating workers, other things equal). Second, we predict worker suitability for each job based on the function estimated via the ML algorithm. Third, we measure whether a worker's actual job coincides with the job that she is predicted to be most suited for within the firm (*eJAQ*) and show that *eJAQ* is positively and significantly associated with workers wages and with labor market experience. Fourth, we build *JAQ* by aggregating worker level indicators of match quality at firm level.

Using this firm-level measure of worker-job match quality, we obtain two main findings. First, we show that *JAQ* has a significant, sizeable and robust positive correlation with measures of firm performance, even when controlling for the firm-level variables generally associated with productivity (industry, capital and labor, ownership) and the workers' characteristics used to predict *JAQ*. A possible concern about this finding is that it may be vitiated by circularity, as we first train the ML algorithm to assign workers to jobs based on data for the most productive firms, and then investigate whether *JAQ* correlates with firm productivity. The first counter to this criticism is that the correlation between *JAQ* and productivity is estimated dropping the observations used to train the ML algorithm. However, the correlation between *JAQ* and productivity may still arise because the assignment rule is estimated on the most productive firms, so that its estimation error may correlate by construction with firm productivity. To face this concern, we perform a placebo test where firms' actual productivity is replaced by a noise variable, and use it in place of the original productivity measure to re-estimate the algorithm. If the positive relationship between *JAQ* and firm productivity were indeed driven by spurious correlation, one should also expect to find a positive and significant correlation between the firms' *JAQ* and the noise productivity measure in the placebo test. Instead, this correlation turns out to be not significantly different from zero.

Our second main finding is that managerial quality plays a key role in the extent to which firms are able to achieve good worker-job matches in the allocation of their human resources. Upon constructing two distinct *JAQ* measures for rank-and-file workers and for managers, we show that the former is positively and significantly correlated with the latter, as well as with the average experience of the firm's management team, even when only within-firm variation is exploited. Interestingly, the quality of rank-and-file workers' allocation rises significantly when turnover leads to an improvement of the allocation of management, and this tends to occur in the wake of a deterioration in the allocation of rank-and-file workers. Conversely, when managerial turnover results in a worse assignment of managers, it is associated with a persistent disruption in the allocation of rank-and-file workers.

The measure proposed in this paper not only uncovers a hitherto neglected dimension of managerial practices and shows that it correlates with firm productivity, but can be applied more widely in testing predictions in corporate finance, organization theory, labor economics, and corporate governance. Moreover, it can be constructed for any country and time period where employer-employee administrative data exist and include workers' job assignments, without requiring either expensive surveys (Bloom and Van Reenen, 2007; Bloom et al., 2019) or detailed expert evaluations of the skills required for each job, such as those contained in the O*Net data (Lise and Postel-Vinay, 2020; Guvenen, Kuruscu, Tanaka, and Wiczer, 2020). Our proposed measure may also provide a benchmark for human resource practitioners wishing to assess the quality of firms' job assignment in a given industry and firm size class. [1]

Our paper is related to an emerging strand of literature that exploits ML to address research questions in labor and finance,[2] such as the appointment of board of directors (Erel, Stern, Tan, and Weisbach, 2021), the screening of resumes in recruitment (Li, Raymond, and Bergman, 2020), the measurement of corporate culture based on earnings call transcripts (Li, Mai, Shen, and Yan, 2020), and internal labor markets and hierarchies in matched employer-employee data (Huitfeldt, Kostøl, Nimczik, and Weber, 2021). In contrast, we use ML to generate a measure of the quality of job-worker matches within firms to test predictions on the correlation between corporate events and the internal organization of firms.

This paper also relates to the literature on the mismatch between workers and jobs and its consequences for the evolution of workers' careers (Perry, Wiederhold, and Ackermann-Piek, 2016; Fredriksson, Hensvik, and Skans, 2018; Lise and Postel-Vinay, 2020; Guvenen et al., 2020). Perhaps most closely related to our study is Fredriksson et al. (2018), who investigate the impact of job mismatch on starting wages and subsequent labor market outcomes, measuring mismatch as the absolute distance between senior workers' and new hires' talent. Our measure of job assignment quality differs from that used in this study in two main respects: (i) we rely on a ML-estimated function rather than on average characteristics of senior workers to determine the efficient alloca-

[1] We are planning on providing an easy to use statistical package to construct our measures.

[2] For surveys on how ML can be applied to economics research in general, see for instance Varian (2014), Mullainathan and Spiess (2017), or Abadie and Kasy (2019), Athey (2019).

tion of workers across jobs; (ii) by the same token, our method applies just as well to the allocation of senior workers as to that of junior ones, while the other method only applies to junior ones. The latter point is key to our objective of evaluating the correlation between job assignment quality and firm productivity, as this exercise requires measuring job assignment quality for *all* the employees of each firm.

The road map reads as follows. The next section describes our data and Section 3 details how we construct *JAQ*. Section 4 relates *JAQ* to firm performance, and Section 5 explores the relationship between the quality of rank-and-file worker-job matches and the quality of management, especially in the wake of managerial turnover. The last section concludes.

2 Data

To develop and estimate the *JAQ* measure proposed in this paper we use Swedish registry data. This data set is ideal for our purposes for at least two reasons. First, it allows us to observe for a relatively long period the entire population of workers and firms in Sweden, including a number of variables regarding workers' job histories, such as occupations and wages over their career. Second, despite their institutional differences, labor markets are surprisingly similar in their functioning in Scandinavian countries, Belgium, France, Germany, Italy, the Netherlands and the United States (Lazear and Shaw, 2009), which bodes well for the external validity of our results.

The bulk of our data come from the Statistics Sweden LISA database that covers the whole Swedish population of individuals who are at least 16 years old and reside in Sweden at the end of each year. This longitudinal matched employer-employee database integrates information from registers held by various government authorities. We have data for the 1990–2010 interval but our analysis focuses on the 2001–10 interval since occupation information is not available prior to 2001. However, we draw on 1990–2000 data in constructing worker job histories.

The estimation of a worker's suitability for a given job is based on the same type of information that would typically be included in individual resumes available to managers assigning workers to jobs, namely, background information, education, and past work experience. Background information, drawn from LISA, includes age, gender, an indicator for immigrant status, residence municipality and a mobility indicator equal to one for workers employed in a county different from the county of birth. As for education, we observe both the education level (basic, high school, vocational, or university) and the education subject (no specialization, law, business and economics, health and medicine, natural sciences, teaching, engineering, social sciences, services, or other specializations). Finally, past work experience is captured by labor market experience (measured as years since graduation), tenure at the current firm, number of firms and number of two-digit industries where an individual previously worked, total number of unemployment days since 1992 (when the unemployment data starts in LISA), years of experience in each occupation, years of experience in each 2-digit industry, and years of experience in each decile of the distribution of firms' number of employees or total assets.

The firm-level variables drawn from LISA are firm age, 2-digit industry, size (measured by the number of employees), sales, and total assets, as well as ownership categories measured by indicators for the firm being a state-owned firm, a listed firm, or a family firm. Information on listed status is drawn from the Statistics Sweden's FRIDA database, and the indicator of family firm status is obtained by combining information on firm ownership from FRIDA with info on board members and CEOs from the Swedish Companies Registrations office and the multi-generational register on biological parent-child relationships. Following Keloharju, Knüpfer, and Tåg (2020), a family firm is defined as one managed or owned by at least two members of the same family.

We identify jobs based on workers international ISCO-88 (COM) classification of occupations, based on data provided primarily by official wage statistics yearly surveys of around 11,000 companies. Companies with more than 500 workers are surveyed every year and the remainder is a random sample of firms. Occupation data is gathered for around a million workers each year. The second source is a yearly survey sent out by mail to around 30,000–47,000 companies that are not selected for inclusion in the official wage statistics survey (a total of around 150,000 private sector companies per year). The surveys are sent out on a rolling basis: all 150,000 companies are surveyed at least once in five years time. In total, over our entire sample period over 90% of workers are sampled at least once.[3]

Our main sample includes firms with at least 30 employees, that are active at some point between 2001 and 2010 and report positive total assets and sales. Since information about a worker's specific occupation is not always available, we restrict the sample to firms with at least 10 workers for whom we do observe the current occupation.

3 Measuring *JAQ*

Suppose that managers strive to allocate workers to jobs so as to maximize productivity, by picking a job assignment function that maps observable worker and firm characteristics to jobs within the firm. However, firms may deviate from the most efficient assignment rule, incurring in errors that reduce their productivity, because of managerial shortcomings and/or internal information frictions. For brevity, hereafter we refer to this assignment function as the "rule".

In order to judge the job allocation quality (*JAQ*) of a firm it is necessary to estimate this rule. In principle, this can be done using a random subsample of firms. However, insofar as the rule maximizes productivity, the firms that apply it most rigorously should incur in fewer errors and feature the highest productivity. Hence, the rule can be observed with the least noise for the most productive firms.

Accordingly, we use a machine learning (ML) algorithm to estimate the rule using only observations that refer to firms in the top decile of the productivity distribution. The benchmark provided by this ML prediction enables us to measure how close the job allocation adopted by any given firm is to that predicted by the estimated rule.

[3]See Tåg (2013) and Tåg, Åstebro, and Thompson (2016) for additional details and descriptive statistics on occupations and hierarchical structures within firms.

3.1 Mapping workers' characteristics to jobs via machine learning

In our framework, talented managers use the job assignment rule $J = g(X,Z)$ to identify the job J to which each worker is best assigned, based on workers' observable characteristics, X, and on firm's characteristics Z. We do not observe g, but we can recover it by estimating the conditional probabilities $P(J|X,Z)$ for firms that are likely to adhere most closely to the rule, i.e., the most productive firms.

We do not impose any particular restriction or parametric form on g, and allow for the possibility that firms in different classes rely on different rules. Hence, the conditional probabilities to be estimated are denoted by $P(J|X,Z) = P_Z(J|X)$. Firms are sorted in 9 classes, according to their size and industry. There are three size classes, based on firms' median number of employees N over the sample period: (i) small ($N \leq 50$); (ii) medium ($51 < N \leq 250$); (iii) large firms ($N > 250$); and three industries: (i) manufacturing; (ii) wholesale and retail; (iii) real estate, renting and business activities. Hence, the firm's characteristics Z reduce to a variable identifying to which of the 9 size-industry classes a firm belongs.

Within each size-industry class, we define the "learning sample" used to estimate the conditional probabilities $P_Z(J|X)$ as the subsample of firms in the top decile of the productivity distribution. More precisely, in order to include in the learning sample only firms that are consistently more productive, for each size-industry class we (i) estimate a model of value added per employee with firm fixed effects and calendar year effects, (ii) consider the distribution of fixed effects for firms present in the 2010 subsample and (iii) select firms that belong to the top decile of this distribution. We then use 2010 data referring to these firms to train our algorithm: being the last year available in our sample, it contains the longest job histories that can be exploited to learn the way firms allocate employees to jobs. Using data for these firms, we estimate class-specific conditional probabilities $\hat{P}_Z(J|X)$ to predict workers' allocation to jobs in remaining firms – referred to as the "main sample" – within the corresponding class.

Table 1 compares the characteristics of the workers included in the learning sample and in the main sample: the workers included in the former sample earn higher wages, are more educated and have fewer days of unemployment and longer tenure than workers included in the latter sample. These differences are consistent with the fact that the learning sample includes more productive firms, where workers can be expected to be of better quality and feature more productive matches, hence fewer separations.

Insert Table 1 here

Despite these differences, the two samples are sufficiently similar as to have common support: this is shown by Figure 1, which displays the distributions of the predicted wages for workers in the two samples for each size-industry class. For both samples, the predictions are obtained from wage regressions estimated on the main sample, whose explanatory variables are the worker characteristics included in the ML algorithm. The figure shows that the support of the two distributions overlaps considerably for all size-industry classes, even though the distribution of the learning sample places more weight on high predicted wages than that of the main sample. The evidence just described supports our assumption that the learning sample can be used to estimate an allocation rule that is relevant also for workers in firms included in the main sample.

Insert Figure 1 here

Within each of the nine learning samples, one for each size-industry class, we estimate the conditional probabilities via the Random Forests algorithm (Breiman, 2001).[4] There are three advantages to using Random Forests (RF) in our setting: (i) they are among the best performing algorithms for classification (Zhang, Liu, Zhang, and Almpanidis, 2017);[5] (ii) they feature few-to-none tuning hyperparameters, dramatically reducing total estimation time;[6] (iii) they easily handle

[4] As implemented by Robnik-Šikonja and Savicky (2020) with the R language.

[5] Although some of the measures we build rely on the full set of estimated conditional probabilities $P_Z(J|X)$, our main measure of job assignment quality relies solely on workers' classification into their most suitable jobs. Moreover, we use a bagging procedure for estimation, which significantly mitigates possible calibration issues related to the estimation of conditional probabilities (Wallace and Dahabreh, 2012).

[6] Hyperparameters are parameters set by the researcher to control the learning process, such as the number of trees and the number of features selected at each a node in random forest algorithms. Compared with other algorithms, such as neural networks, random forests require fewer parameters to be specified, making them relatively easier to tune. This reduces the overall estimation time due to the limited need to estimate multiple models in order to choose the best performing one.

multi-class classification problems and mixed-type characteristics (continuous and categorical), which are relevant in our data.[7]

Occupations are not all equally frequent in the sample, i.e., they suffer from unbalancedness. Hence we adjust our estimation procedure by forming a balanced subsample via bootstrap, under-sampling more frequent occupations, and use this subsample to train a random forest with 50 trees. This is repeated 100 times and the results from the 100 random forests are averaged together—a strategy that combines ideas from EasyEnsemble proposed by Liu, Wu, and Zhou (2008) and Balanced RF in Chen, Liaw, and Breiman (2004).

We evaluate the performance of our algorithm via an average of the F1 scores, computed across jobs (labeled as the macro F score in Sokolova and Lapalme (2009)), with weights equal to job frequencies to address the unbalancedness of the sample.[8] Table 2 shows the average F1 scores computed via a stratified 10-fold cross-validation: the learning sample is randomly partitioned in 10 subsamples, where each subsample has the same job frequencies as the initial sample, and the algorithm is trained using 9 subsamples and tested on the remaining one; the procedure is repeated until all of the 10 subsamples are used as a test set, so as to obtain a total of 10 pairs of weighted F1-scores (where each pair refers to a training set and to the corresponding test set). These 10 weighted F1-scores are used to compute the mean and standard errors in Table 2.

Insert Table 2 here

The upper panel of the table shows that the weighted F1 score of the algorithm on the training sets is between 71% and 81%. The bottom panel refers to the test sets and shows an average F1 score ranging between 61% and 77%. In both cases, the algorithm performs better for employees of large firms. This performance is reassuring, considering that a random allocation of workers

[7]To deal with categorical variables with a high number of levels, we use the coding proposed by Micci-Barreca (2001).

[8]The F1 score for a given class is computed as the harmonic mean of the estimator's precision and recall scores for such class. The precision score is defined as the ratio between the number of instances correctly identified as belonging to the class and the total number of instances that the estimator attributes to the class: it indicates the ability to estimate the class "precisely". The recall score is defined as the ratio between the number of instances correctly identified as belonging to the class and the total number of instances belonging to the class: it indicates the ability of the estimator to retrieve instances of that class.

to jobs would at most achieve an average weighted F1-score of $2/(K+1)$, where K is the total number of jobs. Since the minimal number of jobs in our training and test sets is 46, the maximal weighted F1 score resulting from a random allocation of workers to jobs in our sample would be 4.3% at most.[9]

To characterize our algorithm, we explore the role that each worker characteristic plays in identifying the allocation of jobs across workers. To this purpose, we compute the explanatory power of each of the workers' features used in the random forest algorithm, i.e., its discriminatory power in the correct classification of the instances, as described in Robnik-Šikonja (2004) and Robnik-Šikonja and Savicky (2020). Figure 2 displays this measure for the features listed on the horizontal axis, in manufacturing (panel a), wholesale and retail services (panel b), and real estate, renting and business activities (panel c), distinguishing between large, medium and small firms in each industry.

Insert Figure 2 here

On the whole, Figure 2 highlights that experience by occupation and industry plays an important role in job allocation in all sectors, and education and tenure are more relevant in large firms than in medium and small ones, suggesting that both generic and firm-specific human capital are more important in the former than in the latter. Other features, instead, differ considerably in their importance across industries. In manufacturing, the worker's municipality is the single most important variable in large and medium firms, and the second most important one in small firms, while it is much less important in other sectors, possibly because skilled manufacturing workers tend to be more concentrated in specific locations than workers in other sectors. In contrast, gender plays little role in manufacturing, while it is quite prominent in the service sector and in the real estate and business sector, especially in small firms.

[9] This can be seen as follows. Denote job frequencies by π_k, $k = 1 \dots K$. If the algorithm were to assign workers to jobs at random with equal probability, the probability of assigning a worker to a given job is $1/K$. Hence, the precision and recall for class k, in large samples, are roughly π_k and $1/K$, respectively, so that the F1 score for class k is $2\pi_k/(K\pi_k+1)$, and the weighted F1 score is $2\sum_k \pi_k^2/(K\pi_k+1)$. The maximal value of this expression is $2/(K+1)$, which is achieved when $\pi_s = 1$ for some s and $\pi_k = 0$ for $s \neq k$.

3.2 Job assignment quality at employee level

To predict each worker's suitability to each job in the main sample, we re-estimate the algorithm on the whole learning sample, and use the resulting mapping of employee characteristics to jobs to construct an employee-level measure of job assignment quality (*eJAQ*). This measure equals 1 if the employee's job coincides with the most suitable one, i.e., the job to which the algorithm assigns the highest conditional probability for that worker, and 0 otherwise: formally, if $\hat{J}_i$ is the job predicted for worker i and J_i is the actual job held by that worker, then $eJAQ_i = \mathbf{1}_{\{J_i=\hat{J}_i\}}$. This indicator is the key building block of our measure of job assignment quality at firm level (*JAQ*), which is simply obtained by averaging *eJAQ* across the employees of the same firm in a given year.

In this section we corroborate the validity of *eJAQ* as a measure of workers' job assignment quality. First, it is natural to expect that the likelihood of being assigned to a suitable job increases over time, as managers learn about workers' characteristics (Fredriksson et al., 2018), and workers themselves adapt their skills via on-the-job training (Guvenen et al., 2020). Second, insofar as an improvement in job allocation generates productivity gains, these are likely to be partly appropriated by workers in the form of higher wages. Hence, one can expect wages to be positively related to *eJAQ*.

Both predictions find support in our data. Figure 3 shows the binned scatter plot of *eJAQ* against labor market experience: the likelihood of being assigned to the job predicted by the ML algorithm increases with experience, as the goodness of worker-job matches almost doubles (from 35% to 63%) over the span of a 50-year working life. The largest gain (about 17 percentage points) occurs in the first 5 years of a worker's career: this accords with the intuition that learning is faster for junior workers, and that their reallocation to more suitable jobs is easier than for senior employees (Farber and Gibbons, 1996)

Insert Figure 3 here

Moreover, better matches between workers and jobs are systematically associated with higher compensation, suggesting that assigning workers to the right jobs brings about efficiency gains.

This is shown in Table 3, where Panel A reports the estimates of the following earnings regression:

$$w_{it} = \alpha_j + \beta eJAQ_{it} + \gamma X_{it} + \delta Z_{f(i,t)} + \lambda_t + u_{it} \tag{1}$$

where w_{it} is the logarithm of annual earnings of worker i in year t; α_j are job indicators; $eJAQ_{it}$ is a dummy variable that equals 1 if worker i is allocated to her most suitable job in year t, and 0 otherwise; X_{it} are all the workers' characteristics included in the ML algorithm; $Z_{f(i,t)}$ are the characteristics of the firm f that employs worker i in year t (e.g., 2-digit industry dummies, firm age, indicators for family firm, listed company, presence of a human resources manager), and λ_t are year dummies.

Insert Table 3 here

Column 1 of Panel A reports the estimate of β in a version of equation (1) that includes only job and year effects and the machine learning variables. The resulting estimate is 0.014: a worker allocated to her most suitable job ($JAQ_{it} = 1$) is estimated to earn 1.4% more than a mismatched worker with the same characteristics or with the same job ($eJAQ_{it} = 0$). The estimate of β drops to 0.7% in a specification that also controls for 2-digit industry dummies and firm characteristics (column 2). A similar result is obtained considering only within-worker variation in $eJAQ_{it}$ (column 3): the estimated β in a specification that includes worker, jobs and year effects is 0.8% and highly statistically significant. These findings are broadly in line with the -1% estimate of the coefficient of job mismatch in earnings regressions reported in Table 7 in Fredriksson et al. (2018), despite the differences in the methodology and the sample used.

We also construct a continuous measure of employee-level job match quality (*Suitability*) by estimating the probability that the algorithm assigns to the actual job held by worker i, $\text{prob}(J_i)$, and explore its correlation with labor earnings in Panel B of Table 3 to provide a robustness check of the results obtained using the *eJAQ* indicator. This alternative measure is a gauge of a worker's fit for her actual job compared to other jobs that she might perform, while *eJAQ* is a binary measure of whether the employee is assigned to her best possible job or not. To better understand the difference between the two measures, suppose that a worker's suitability to a position as "travel attendant" and as "cashier" are 0.5 and 0.48, respectively: for such a worker *eJAQ* equals 1 if assigned to the first position, and 0 if assigned to the second, although the worker is almost equally suited to both positions; in contrast, the *Suitability* measure captures how fit the worker is to perform a job that is not her ideal one.

The estimates shown in Panel B of Table 3 indicate that labor earnings are also positively and significantly correlated with this second measure of job match quality over workers' careers. The 0.18 coefficient estimate in column 1 indicates that a 10 percentage points increase in a worker's *Suitability* is associated with a 1.8% increase in labor earnings. This effect drops to 0.3% in the specification with industry fixed effects and firm-level controls shown in column 2, and 0.05% in that with worker fixed effects, but in both cases is precisely estimated.

As the quality of worker-job matches is positively and significantly associated with labor earnings, it is worth asking which types of jobs are more often assigned to the wrong workers according to our algorithm, thus forgoing attainable increases in labor earnings. The upper panel of Figure 4 shows the percentages of instances in which workers fail to be allocated to their most suitable job in the main sample, averaging such percentages within each of the following six job classes: 1) managers, 2) professionals, 3) technicians and clerks, 4) skilled manual workers, 5) machine operators and assemblers, and 6) elementary occupations.[10] Thus, for each job class, the corresponding bar in the figure indicates the frequency of cases in which a worker holding a job in that class should have been allocated to a different job according to our algorithm. The graph shows that the frequency of mismatches is quite uniform across job classes, except for a slightly lower value for professionals (35%) and a considerably larger value for elementary occupations (60%): in the remaining classes, mismatches range from 49% for managers and 46% for technicians and clerks to 48% for skilled manual workers and 40% for machine operators and assemblers.

Insert Figure 4 here

[10]Skilled manual workers comprise service and shop sales workers, skilled agricultural and fishery workers, and craft and related trade workers.

The greater frequency of mismatches for elementary occupations may be due to two concomitant reasons: first, these are low-skill jobs and as such they do not require very specific worker profiles, so that job-worker mismatches may arise easily than for other occupations; second, fewer workers hold these jobs, so that there are fewer observations to inform their allocation rule. Indeed, elementary occupations account for a relatively small fraction of jobs in the economy (6%), not dissimilar from that of managers (7%), while the bulk of workers hold jobs in intermediate classes, as shown in the lower panel of Figure 4. Hence, the absolute frequencies of mismatches in the extreme job classes is much lower than in the intermediate ones: the inefficiency arising from the misallocation in the two extreme classes is mitigated by their relatively lower size.

3.3 Job assignment quality at firm level: *JAQ*

The next step in the analysis is to average $eJAQ_{it}$ for all the employees i of any firm f in a given year t: we refer to the resulting firm-level measure of job allocation quality as *JAQ*. As our approach builds on the assumption that firms differ in their ability to assign workers to jobs, we expect to observe heterogeneity in *JAQ* across firms.

Insert Figure 5 here

Figure 5 shows the kernel density estimate of firm-level *JAQ* for firms in the main sample and in the learning sample. As expected, the density of *JAQ* in the main sample assigns greater probability mass to lower values than the corresponding density for the learning sample. Moreover, the dispersion in *JAQ* across firms in the main sample exceeds that in the learning sample. This is as expected, for two reasons. First, the learning sample is used to train the ML algorithm at the core of our *JAQ* measure, so that by construction this sample features a better fit between firms' observed choices and the estimated allocation rule. Second, our learning sample is formed by firms in the top productivity decile: insofar as their higher productivity results from fewer mistakes in applying the most efficient allocation rule, they should feature more concentrated *JAQ* than firms in the main sample. In the limit, if there were no noise in the estimation procedure, the learning sample should feature no dispersion in *JAQ* (i.e., we should observe $JAQ = 1$ for all firms), while there should be dispersion in *JAQ* in the main sample, reflecting deviations from the efficient rule.

4 *JAQ* and Firm Performance

This section explores how the heterogeneity in *JAQ* correlates with firm performance, as measured by sales per employee, value added per employee, and operating return on assets (OROA): we wish to determine whether *JAQ* captures meaningful variation in the quality of workforce allocation, rather than just statistical noise or heterogeneity along observable determinants of firm productivity. Our exercise parallels the approach used by Bloom et al. (2019) to validate their measure of structured management practices, by investigating their correlation with various indicators of firm performance.

Insert Figure 6 here

Figure 6 shows that firm-level productivity correlates positively with *JAQ*. The estimates in the two top panels refer to main-sample firms, and those in the bottom two to learning-sample ones. The left-side panels of the figure show the binned scatter plot of value added per employee against *JAQ*, and the right-side ones show log sales per employee against *JAQ*, in each case controlling for year and 2-digit industry dummies. A positive relationship is evident in the two top graphs, providing preliminary evidence that main-sample firms tend to feature higher productivity insofar as they allocate employees more closely along the rule estimated from the learning sample. The two lower panels instead show that no correlation between productivity and *JAQ* emerges for firms in the learning sample. This is to be expected, as for these firms variation in *JAQ* should only reflect sampling variability stemming from random deviations from the estimated allocation rule. This can be easily illustrated by considering an extreme example: if firms in the learning set were to adhere perfectly to a common deterministic allocation rule, then *JAQ* would equal 1 for all of them, and would feature no relationship with productivity. To the extent that the variation in *JAQ* detected in the learning sample reflects firms' random deviations from the same allocation rule, one would not expect it to feature a systematic relationship with firm productivity.

Table 4 explores further the firm-level correlation between productivity (as well as profitability) and *JAQ*, controlling for other determinants of productivity. All the specifications presented in the table include year dummies and municipality dummies: the first control for aggregate movements in productivity, the second for productivity differentials across locations. The latter may arise not only from location-related technological advantages, but also from those stemming from access to deeper and more diversified local labor markets. Hence, the relationship between productivity and *JAQ* captured by our estimates is not driven by differences in the availability of workers or labor market conditions across firms' locations.

In Panel A of Table 4, column 1 reports the OLS estimates of a regression of log sales per employee on *JAQ* controlling only for year dummies. We find a highly significant coefficient of 0.540, implying that a 10-percentage-point increase in *JAQ* is associated with a 5.5% increase in sales per employee. Equivalently, a one-standard-deviation increase in *JAQ* (0.27) is associated with a 15.7% increase in sales per employee. To put this estimate in perspective, Bloom et al. (2019) find that a one-standard-deviation increase in their management score is associated with a 26.2% increase in sales per employee. The difference in magnitude between the two estimates may reflect the fact that *JAQ* focuses on the gains stemming from the efficient allocation of employees, while the score constructed by Bloom et al. (2019) is a broader synthetic indicator of management practices.

Insert Table 4 here

In column 2 of Table 4, the dependent variable is value added per employee, and the coefficient of *JAQ* is again positive and highly significant: a 10-percentage-point increase in *JAQ* is associated with an average increase in value added per employee of 1.85%, the sample mean of the dependent variable being 0.64. Instead, profitability, as measured by operating return on assets, is not significantly correlated with our measure of efficient job allocation, as shown in column 3 of the table.

One may suspect this positive firm-level correlation between productivity and *JAQ* to be spurious for a variety of reasons, namely omitted variables such as firm characteristics, differences in firms' occupation structures and in workers' quality across firms. In what follows, we test the robustness of our results to each of these possible concerns.

First, the correlation may reflect other firm characteristics such as their size, sector or input mix. However, this is not the case, as shown by the estimates in columns 4, 5 and 6 of the table, which refer to specifications that control for 2-digit industry indicators, log number of employees, log capital, and the fraction of employees with at least a college degree. The estimated coefficients of *JAQ* in columns 5 and 6 drop in magnitude, but remain positive and significantly different from zero.

To check whether the estimates reported in Table 4 are driven by unobserved heterogeneity across firms, we estimate the models in columns 1-3 including firm fixed effects, in addition to the year effects: the relationship of log sales per employee with *JAQ* is robust to the inclusion of firm fixed effects. Instead this does not apply to its relationship with value added per employee, possibly because adding firm fixed effects greatly reduces the variation in *JAQ* (the standard deviation of the residual variation in *JAQ* is 0.08, against a total standard deviation of 0.27), and exacerbates the attenuation bias from measurement error. For brevity, we do not report these additional estimates.

A second possible concern in the previous specifications is that the firms being compared may have different occupation structures. Two otherwise comparable firms may structure their internal hierarchy in a different fashion: if for instance a firm has an inefficiently large number of managerial positions relative to technical ones compared to other firms in its industry, and those managerial positions are harder to fill with suitable employees, it is likely to end up both with lower productivity and lower *JAQ*, creating a spurious correlation between the two variables. To take this concern into account, Panel B of Table 4 reports the estimates of the following specification:

$$y_{ft} = \theta_0 + \theta_1 JAQ_{ft} + \theta_2 F_{jft} + \theta_3 Z_{ft} + \lambda_t + \gamma_h + u_{ft} \tag{2}$$

where y_{ft} is log(sales/employees), value added per employee or operating return on assets, F_{jft} is the fraction of workers in firm f assigned to job j in year t; Z_{ft} are the characteristics of firm f in year t, namely their age, family firm status, state-owned status, listed firm status, an indicator for the presence of a human resources (HR) manager, its log number of employees and its log of total assets; λ_t are year dummies, and γ_h are 2-digit-industry dummies. In columns 1, 2 and 3 of Panel B this specification is estimated omitting the firm characteristics Z_{ft}, while in columns 4, 5 and 6 these are also included. The results are qualitatively similar to those in Panel A: the estimated coefficients of *JAQ* drop in magnitude, but remain positive and statistically significant in columns 1, 2, 4 and 5.

A third source of spurious correlation is that firms with higher *JAQ* may feature higher-quality workers, irrespective of the job they are allocated to, thus creating a spurious correlation between *JAQ* and productivity. To address this concern, in Panel C of Table 4 we augment specification (2) with the workers' characteristics included in the machine learning algorithm, averaged across all workers employed in firm f in year t. In columns 1, 2 and 3 we control for year effects, occupation structure and workers' characteristics. Columns 3, 4, and 5 also add industry dummies and firm characteristics. The coefficient of *JAQ* remains positive and statistically significant also in these very conservative specifications, even though in some of them it drops further in magnitude.[11]

To check the robustness of these results, in Table 5 we repeat the estimation of the specifications shown in Table 4 upon measuring worker-job match quality by the firm-level average of the *Suitability* variable described in Section 3.2. The estimated coefficient of this variable is positive and significantly different from zero in all the specifications of the productivity regressions, but not of the profitability regressions, in line with the results of the previous table. The baseline estimates shown in columns 1 and 2 of Panel A imply that a 10 percentage points increase in firm-level suitability of workers to jobs is associated with a 13 percentage points increase in log sales per employee and a 3.2 percentage points increase in value added per employee. These results are

[11]The results reported in Table 4 are obtained using the main sample. Upon estimating the same specifications with the learning sample, no robust relationship between *JAQ* and productivity emerges, consistently with what is shown in Figure 6.

qualitatively robust to the addition of other controls, even though they drop considerably in size.

Insert Table 5 here

One last concern is that the construction and validation of *JAQ* performed up to this point may be vitiated by circularity: as explained in Subsection 3.1, we train the ML algorithm to assign workers to jobs in firms from the top decile of the productivity distribution, and then check whether the *JAQ* measure thus obtained correlates with firms' productivity. The obvious counter to this criticism is that the correlation between *JAQ* and productivity is tested on the main sample, and not on the learning sample used to train the algorithm, and indeed the correlation is present only for the main and not for the learning sample, as shown in Section 4. However, one may still fear that *JAQ* correlates positively with productivity for spurious reasons: as the assignment rule is estimated on the decile of the most productive firms (by value added per employee), its estimation error may correlate by construction with firm productivity, thus contaminating the regressions in Table 5.

To address this concern, we perform the following placebo test: we replace firms' actual productivity measure (value added per employee) with a noise variable, obtained by randomly reshuffling the original variable across firms, and use it in place of the original productivity measure to re-estimate the algorithm. That is, the learn set is now built using the top 10% firms in terms of the noise variable, as explained in Subsection 3.1, and is used to compute again the *JAQ* measure. By construction, the new productivity variable constructed for the placebo test has the same distribution as the original one, but is independent from the rest of the data.

Note that this placebo test leaves intact the relationship between employees' CVs and task allocations, and only alters the selection of firms into the learn set. If the positive relationship between *JAQ* and firm productivity is indeed driven by spurious correlation, one should also expect to find a positive and significant correlation between the firms' *JAQ* and the noise productivity measure in the placebo test, for the main sample. If instead there is no mechanical relationship induced by the learn set selection process, no significant relationship should emerge. This is indeed what emerges from Figure 7, which plots the results of the regression of the placebo productivity measure on *JAQ* (panel b). The lack of correlation between these two variables contrasts with the positive and significant correlation obtained in the main estimation strategy (panel a). These results are robust to the inclusion of year and industry effects (Figure 7) and also to the inclusion of additional controls like firms' size and municipality dummies (unreported for brevity).

Insert Figure 7 here

5 Impact of management quality and turnover on job-worker matches

The results presented so far are consistent with our ML algorithm capturing a best-practice rule to allocate workers to jobs, whose adoption is correlated with higher firm-level productivity. Why don't all firms in our sample follow such a best-practice rule? As workers' hiring, assignment to jobs and promotions are typically management decisions, it is natural to inquire whether workers' assignment to jobs is systematically related to managerial quality in our data. This immediately begs another question, namely, how to measure managerial quality based on the observed characteristics of managers. In line with the approach of this paper, a synthetic measure of a firm's managerial quality should be the frequency with which they are assigned to their managerial task in the most productive firms. Another, simpler measure of the quality of firm's managers is their average work experience in managerial positions.

Hence, to investigate this issue, for each firm and date we split *JAQ* into its two components, one measuring the quality of rank-and-file employees' assignment to jobs (*R&F-JAQ*) and the other measuring the quality of managers' allocation to their respective jobs (*M-JAQ*). The first is the average *eJAQ* for all workers that hold non-managerial positions in a given firm at a given date, while the latter is the average *eJAQ* for the corresponding firm's managers. Next, we investigate the firm-level relationship between these two variables, as well as *R&F-JAQ* and managerial experience.

Table 6 presents the results of the corresponding regressions, which are based on data from 2003 to 2010: data for 2001 and 2002 are omitted in order to enable the *JAQ* measure to condition on at least two years of experience for all workers. In columns 1 to 3 *M-JAQ* refers both to top managers (CEOs and firm directors) and to middle managers, whereas in columns 4 to 6 they only refer to top managers. Columns 1 and 4 display results from baseline regressions whose dependent variable is the job allocation quality of rank-and-file employees (*R&F-JAQ*), and whose explanatory variable is the allocation quality of managers (*M-JAQ*), including only year effects. The correlation is positive and significantly different from zero in both regressions: 10 percentage points increase in the quality of managers' allocation is associated with a 1.5 percentage points increase in the quality of rank-and-file workers' allocation; when the quality of managers' allocation refers only to the firm's top management, the coefficient halves in size, indicating that also middle management is quite important for the correct allocation of workers to their jobs.

Insert Table 6 here

The specifications shown in columns 2 and 5 also include firm fixed effects and the average experience of the firms' managers (Manager exp), and those shown in columns 3 and 6 additionally include industry fixed effects, municipality fixed effects and firm controls (age, family firm status, state-owned status, listed firm status, an indicator for the presence of a human resources manager, its log number of employees and its log of total assets). In both of them, managerial experience appears to contribute positively and significantly to *R&F-JAQ*, but the coefficient of *M-JAQ* remains as large as in the baseline regressions. Importantly, these regressions are based only on within-firm variation in the relevant variables: they indicate that a firm's improvements in allocating of rank-and-file employees to jobs tend to occur when the firm improves its management's quality and experience.

Since it is natural to expect marked improvements in managerial quality and experience to result from the hiring of better managers and/or the dismissal of incompetent ones, our next step is to test whether the allocation of rank-and-file workers improves upon incumbent managers being replaced with more suitable ones, and worsens upon them being replaced with less suitable ones.

To perform this test, the first step is to measure the change in managers' quality associated with their turnover, relative to the counterfactual level of managerial quality associated with no turnover.

Letting τ denote a year in which managerial turnover occurs in a given firm (meaning that at least one of its managers changes), we measure the concomitant change in managerial quality, denoted by $\Delta M\text{-}JAQ_{\tau}$, as the difference between the average quality of the firm's new management team, $M\text{-}JAQ_{\tau}$, and the weighted average of the mean quality of retained managers in year τ and the mean quality of dismissed managers in year $\tau-1$. Formally, the change in managerial quality associated with turnover is

$$\Delta M\text{-}JAQ_{\tau} = M\text{-}JAQ_{\tau} - \frac{\sum_{i=1}^{N_{\tau}^{r}} eJAQ_{i\tau}^{r} + \sum_{j=1}^{N_{\tau-1}^{d}} eJAQ_{j\tau-1}^{d}}{N_{\tau}^{r} + N_{\tau-1}^{d}}, \tag{3}$$

where $eJAQ_{i\tau}^{r}$ denotes the quality of retained manager i in year τ, $eJAQ_{j\tau-1}^{d}$ denotes the quality of dismissed manager j in year $\tau-1$, N_{τ}^{r} the number of managers retained in year τ and $N_{\tau-1}^{d}$ the number of managers dismissed in year $\tau-1$. The fractional term in expression (3) measures the counterfactual level of managerial quality in the absence of managerial turnover, which is based on the assumption that the average quality of dismissed managers would have remained the same if they had not been dismissed. Importantly, this measure is designed so as to only track changes in the quality of the managerial team associated with changes in its composition: it disregards the change in the average quality of retained managers between years τ and $\tau-1$, as this change would occur irrespective of managerial turnover. Indeed, $\Delta M\text{-}JAQ_{\tau}$ is zero by construction if no managers are dismissed ($N_{\tau-1}^{d} = 0$) and no managers are hired.

We then define a "positive turnover event" to occur for a given firm in year τ if in that year expression (3) turns positive for the first time for that firm, and this rise in managerial quality is persistent over time, i.e., is never subsequently reversed, or more than reversed. Symmetrically, a "negative turnover event" occurs in year τ if in that year expression (3) turns negative, and this drop in managerial quality is persistent over time. This is done to purge the event of interest from the confounding effects of sequences of transitory changes in managerial quality associated with turnover. In our data, 1,360 firms (20.8% of the total) experience positive turnover events, 3,124 (47.7%) experience negative ones, and the remaining 2,065 (31.5%) experience none.

Our final step is to investigate whether such positive and negative managerial turnover events are associated with significant changes in the allocation quality of rank-and-file workers. To this purpose, we estimate the parameters of the treatment effects of these managerial turnover events on *R&F-JAQ*, exploiting variation in treatment timing. To estimate the dynamic treatment effects of interest, we employ the estimator proposed in Callaway and Sant'Anna (2021). This estimator bypasses the pitfalls related to the interpretation of the TWFE estimators – see for instance de Chaisemartin and D'Haultfœuille (2020), Goodman-Bacon (2021), Borusyak, Jaravel, and Spiess (2021), Athey and Imbens (2022), Sun and Abraham (2021), and Baker, Larcker, and Wang (2022). It is particularly well-suited to our setting because it focuses on recovering treatment effect dynamics with variation in the timing of the treatment. Figure 8 show the estimated dynamic treatment effects on rank-and-file workers around managerial turnover events, respectively associated with an increase (chart a) or a decrease (chart b) in the *JAQ* of the relevant firm's management.

Insert Figure 8 here

The chart on the left shows that replacement of incumbent managers with better ones tends to occur in the wake of sharp and statistically significant deterioration in the allocation of rank-and-file workers to jobs (by about 5 percentage points on average), and are followed by a significant and persistent improvement over the subsequent five years, starting at 10 percentage points at the time of the event, and increasing further subsequently. Conversely, the chart on the right indicates that replacement of incumbent managers with worse ones tend to occur in firms starting from a normal level of rank-and-file workers allocation quality, but are associated with a strong, persistent and statistically significant deterioration in the allocation of rank-and-file workers–by about 15 percentage points on impact and 10 percentage points after three years. Overall, this evidence suggests that persistent changes in managerial quality are an important driver of changes in workers' allocation, and therefore – for better or for worse – of organizational change within firms.

In principle, the organizational changes brought about by new management may consist mainly of reallocating of the existing employees to different tasks or rather of changes in the composition of the firm's workforce via new hires and/or dismissals. Moreover, the reliance on one or the other of these strategies may differ depending on whether the new managerial team is better or worse than the preexisting one according to our metric. To investigate this point, we define three comparable outcome variables: (i) the change in the average allocation quality of rank-and-file employees ($\Delta R\&F\text{-}JAQ$); (ii) the change in the fraction of correctly allocated employees within the set of retained employees ($\Delta R\&F\text{-}JAQ^r$), and (iii) the change in the fraction of correctly allocated employees among retained employees that change jobs in year t ($\Delta R\&F\text{-}JAQ^c$). While the first of these three variables is simply the change in the variable $R\&F\text{-}JAQ$ used as dependent variable in the estimates of Figure 8, the other two are computed by averaging the *eJAQ* of the relevant set of workers around positive and negative turnover event dates. For each of the three variables, Table 7 shows the estimates obtained with the Callaway-Sant'Anna method at event-time (i.e. the time-0 parameter) of the treatment effects of the positive and negative managerial turnover events, respectively.

Insert Table 7 here

While the estimates reported in the first column of Table 7 show that positive and negative managerial turnover events respectively trigger a 14-percentage-points rise and a 13-percentage-points drop in the overall allocation quality of rank-and-file workers, the second column shows that about two thirds of these changes in allocation quality stem from the change in the allocation quality of retained workers, the remaining third then being attributable to new hires and/or dismissals. The estimates shown in the third column indicate that these changes in the quality of the allocation of retained workers are the result of workers being reassigned to new jobs, rather than from the change in the match quality spontaneously arising from workers' learning by doing. Indeed, in the wake of positive turnover events all of the improvement in retained workers' match quality stems from management reassigning these workers to different jobs, and in the wake of negative events the deterioration in match quality due to mistaken reallocation of retained workers is twice as large as the total deterioration of match quality, i.e. it also eats into the spontaneous improvement in match quality due to workers' learning by doing. In conclusion, most of the changes in the ability of allocating firm's employees is due to the improved or worsened ability in reallocating pre-existing workforce across occupations, both when the new management is better than the previous one and when it is worse.

6 Conclusions

This paper contributes new results to the issue why some firms are persistently more productive than others. We focus on managerial policies governing the allocation of workers to jobs within the firm and develop a job-worker allocation quality measure by combining administrative employer-employee matched data with ML techniques.

We validate this measure by exploring its correlation with workers' wages over their careers, firm performance, and with managerial turnover. Our evidence shows that workers earn significantly more as they are better allocated to jobs over their careers, and that firms are more productive if they achieve better worker-job matches. The quality of management appears to play a key role in the efficient assignment of workers to jobs: rank-and-file workers' allocation improves significantly when managerial turnover leads to better assigned and more experienced managers, while the opposite occurs when turnover leads to lower-quality management. In both cases, most of the improved ability in workforce allocation is due to better or worse ability of the management in re-allocating retained employees.

The measure proposed in this paper can be constructed for any linked employer-employee data that include workers' occupations, without requiring either expensive surveys or detailed expert evaluations of the skills required for each job, and can be applied to explore the effect of corporate restructurings (such as those resulting from private equity interventions, mergers and acquisitions, changes in corporate control or initial public offerings) on workers' allocation and on their careers.

References

Abadie, A. and M. Kasy (2019). Choosing Among Regularized Estimators in Empirical Economics: The Risk of Machine Learning. *The Review of Economics and Statistics 101*(5), 743–762.

Athey, S. (2019). The impact of machine learning on economics. In A. Agrawal, J. Gans, and A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda*, pp. 507–547. University of Chicago Press.

Athey, S. and G. W. Imbens (2022). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics 226*(1), 62–79. Annals Issue in Honor of Gary Chamberlain.

Baker, A. C., D. F. Larcker, and C. C. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2019). What Drives Differences in Management Practices? *American Economic Review 109*(5), 1648–1683.

Bloom, N. and J. Van Reenen (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Borusyak, K., X. Jaravel, and J. Spiess (2021). Revisiting event study designs: Robust and efficient estimation. Technical report, arXiv:2108.12419.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230. Themed Issue: Treatment Effect 1.

Chen, C., A. Liaw, and L. Breiman (2004). Using random forest to learn imbalanced data. Technical Report No. 666, University of California, Berkeley.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–96.

Erel, I., L. H. Stern, C. Tan, and M. S. Weisbach (2021). Selecting Directors Using Machine Learning. *The Review of Financial Studies 34*(7), 3226–3264.

Farber, H. S. and R. Gibbons (1996). Learning and Wage Dynamics. *The Quarterly Journal of Economics 111*(4), 1007–1047.

Fredriksson, P., L. Hensvik, and O. N. Skans (2018). Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility. *American Economic Review 108*(11), 3303–3338.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277. Themed Issue: Treatment Effect 1.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer (2020). Multidimensional Skill Mismatch. *American Economic Journal: Macroeconomics 12*(1), 210–244.

Huitfeldt, I., A. R. Kostøl, J. Nimczik, and A. Weber (2021). Internal Labor Markets: A Worker Flow Approach. *IZA DP No. 14637*.

Keloharju, M., S. Knüpfer, and J. Tåg (2020). CEO Health. *SSRN Electronic Journal*.

Lazear, E. P. and K. L. Shaw (2009). Wage Structure, Raises and Mobility: An Introduction to International Comparisons of the Structure of Wages Within and Across Firms. In *The Structure of Wages: An International Comparison*, pp. 1–57. University of Chicago Press.

Li, D., L. Raymond, and P. Bergman (2020). Hiring as Exploration. *NBER Working Paper 27736*.

Li, K., F. Mai, R. Shen, and X. Yan (2020). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies*, 3265–3315.

Lise, J. and F. Postel-Vinay (2020). Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review 110*(8), 2328–2376.

Liu, X.-Y., J. Wu, and Z.-H. Zhou (2008). Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 39*(2), 539–550.

Micci-Barreca, D. (2001). A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems. *ACM SIGKDD Explorations Newsletter 3*(1), 27–32.

Mullainathan, S. and J. Spiess (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives 31*(2), 87–106.

Perry, A., S. Wiederhold, and D. Ackermann-Piek (2016). How Can Skill Mismatch be Measured? New Approaches with PIAAC. *Methods, Data, Analyses 8*(2), 137–74.

Robnik-Šikonja, M. (2004). Improving random forests. In J.-F. Boulicaut et al. (Eds.), *ECML 2004*, Volume 3201 of *LNAI*, pp. 359–370. Springer.

Robnik-Šikonja, M. and P. Savicky (2020). *CORElearn: Classification, Regression and Feature Evaluation*. R package version 1.54.2.

Sokolova, M. and G. Lapalme (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management 45*(4), 427–437.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199. Themed Issue: Treatment Effect 1.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature 49*(2), 326–365.

Tåg, J. (2013). Production Hierarchies in Sweden. *Economics Letters 121*(2), 210–213.

Tåg, J., T. Åstebro, and P. Thompson (2016). Hierarchies and Entrepreneurship. *European Economic Review 89*, 129–147.

Varian, H. R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives 28*(2), 3–28.

Wallace, B. C. and I. J. Dahabreh (2012). Class probability estimates are unreliable for imbalanced data (and how to fix them). In *2012 IEEE 12th international conference on data mining*, pp. 695–704. IEEE.

Zhang, C., C. Liu, X. Zhang, and G. Almpanidis (2017). An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications 82*, 128–150.

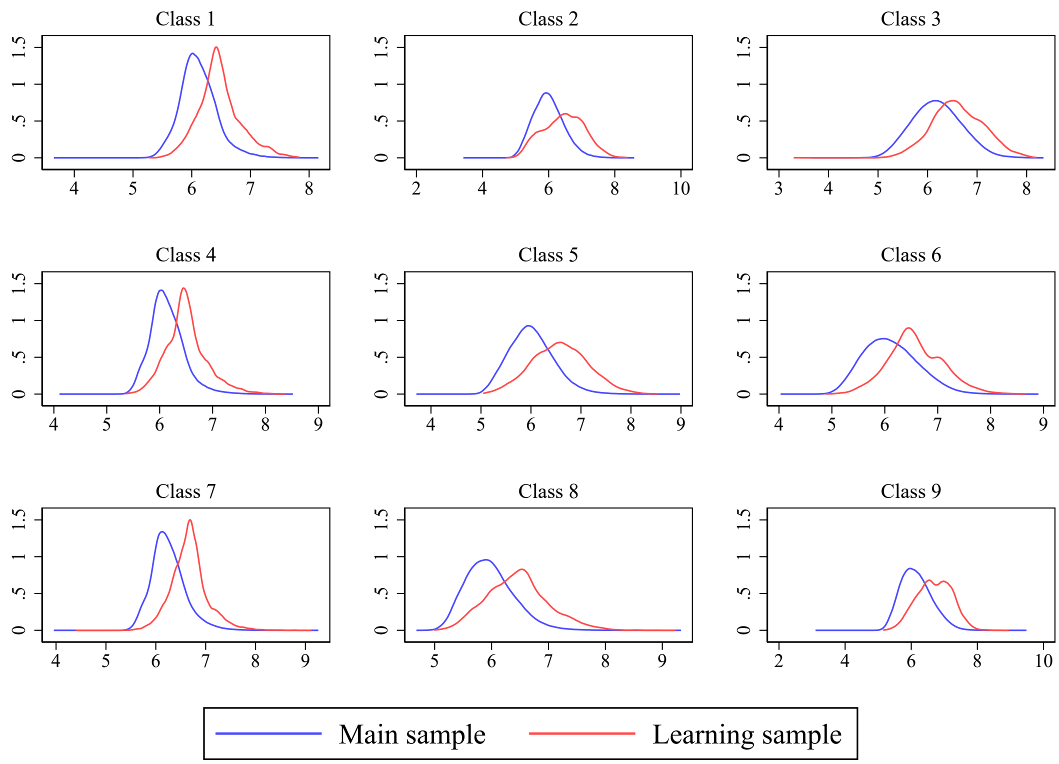


Figure 1: Common support of worker characteristics for each size-industry class
These figures show the distributions of the predicted wages for workers in the learning sample (blue lines) and the main sample (red lines) for each size-industry class. Classes 1, 2 and 3 refer to small firms respectively in Manufacturing, Retail and Real Estate; classes 4, 5 and 6 correspond to medium firms and classes 7, 8 and 9 to large firms in each of the three industries. For both samples, the predictions are obtained from wage regressions estimated on the main sample using as explanatory variables the worker characteristics included in the ML algorithm. These are age, gender, an indicator for immigrant status, residence municipality, a mobility indicator equal to one for workers employed in a county different from the county of birth, education level (basic, high school, vocational, or university), education subject (no specialization, law, business and economics, health and medicine, natural sciences, teaching, engineering, social sciences, services, or other specializations), labor market experience (measured as years since graduation), tenure at the current firm, number of firms and number of two-digit industries where an individual previously worked, total number of unemployment days since 1992 (when the unemployment data starts in LISA), years of experience in each occupation, years of experience in each 2-digit industry, and years of experience in each decile of the distribution of firms' number of employees or total assets. The figure shows that the support of the two distributions overlaps considerably for all size-industry classes.

(a) Manufacturing

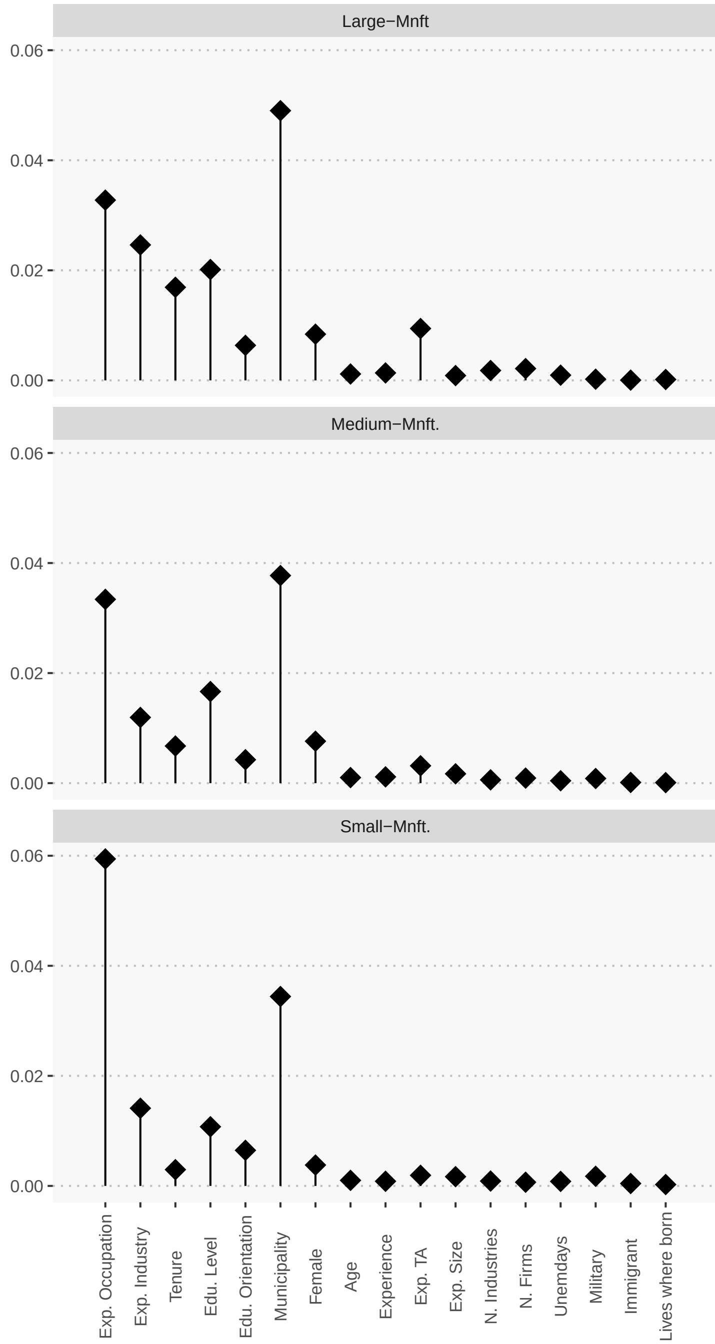


Figure 2: Importance of workers' features in the random forest algorithm, by size-industry bins
The graphs plots the maximum explanatory power of all the workers' features used in the random forest algorithm, in manufacturing (panel a), wholesale and retail services (panel b) and real estate, renting and business activities (panel c), separately for large, medium and small firms in each industry. Features are listed on the horizontal axis, and the importance of each feature—defined as in Robnik-Šikonja (2004) and Robnik-Šikonja and Savicky (2020)—measures its discriminatory power in the correct classification of the instances. Some features are aggregated under a single label: "Exp. Occupation" aggregates the years of experience in each occupation, "Exp. Industry" those in each industry, "Exp. TA" those in firms with given total assets, and "Exp. Size" those in firms with given number of employees. "Edu. Orientation" aggregates features related to education specialization. "Tenure" is the years of employment in the current firm, "Municipality" codes the worker's residence, "Female" and "Experience" are the worker's gender and years of experience, "N. Industries" and "N. Firms" the number of industries and firms where a worker was employed, "Unem. days" the number of unemployment days. "Military", "Immigrant" and "Lives where born" are dummy variables indicating whether the worker performed military service, is an immigrant and lives in his/her birthplace, respectively.

(b) Wholesale and Retail Services

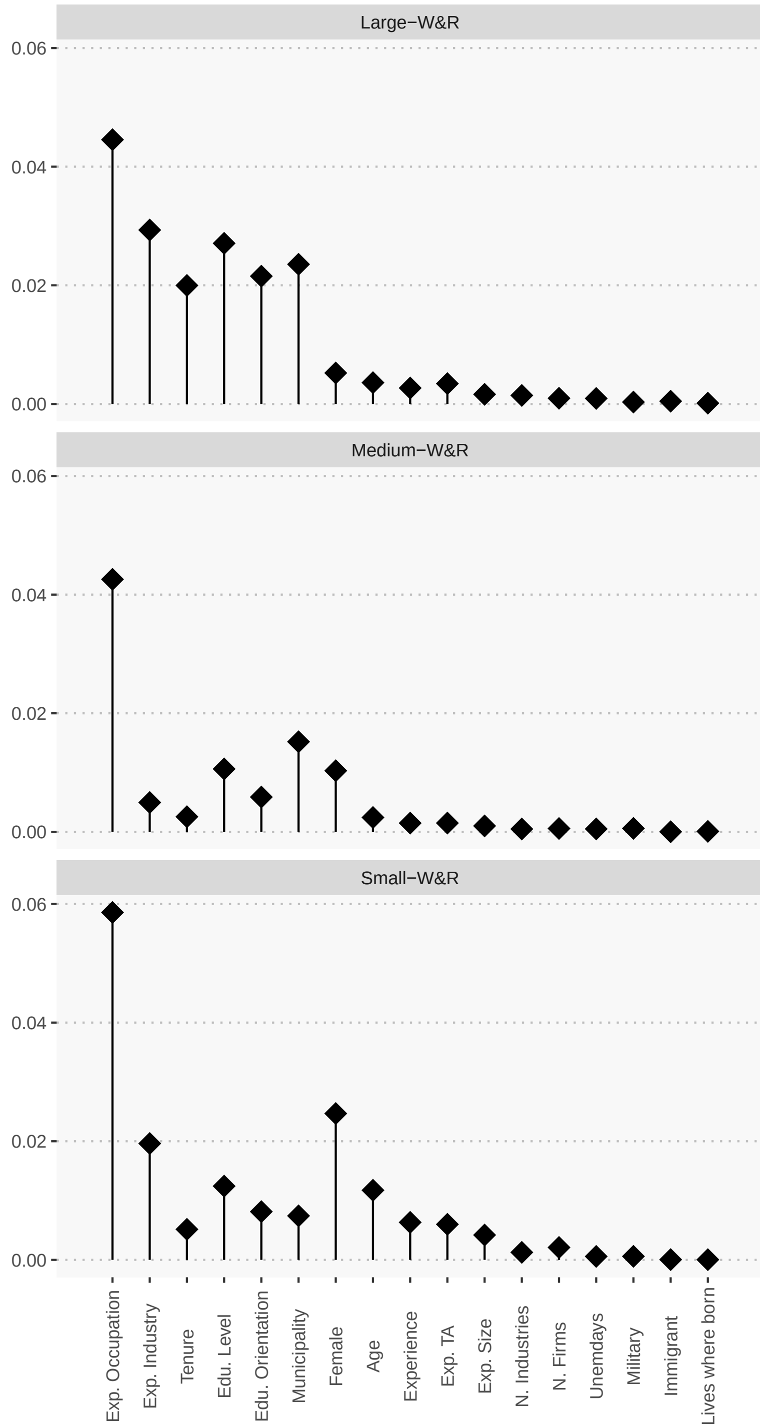


Figure 2: continued

(c) Real Estate, Renting and Business Activities

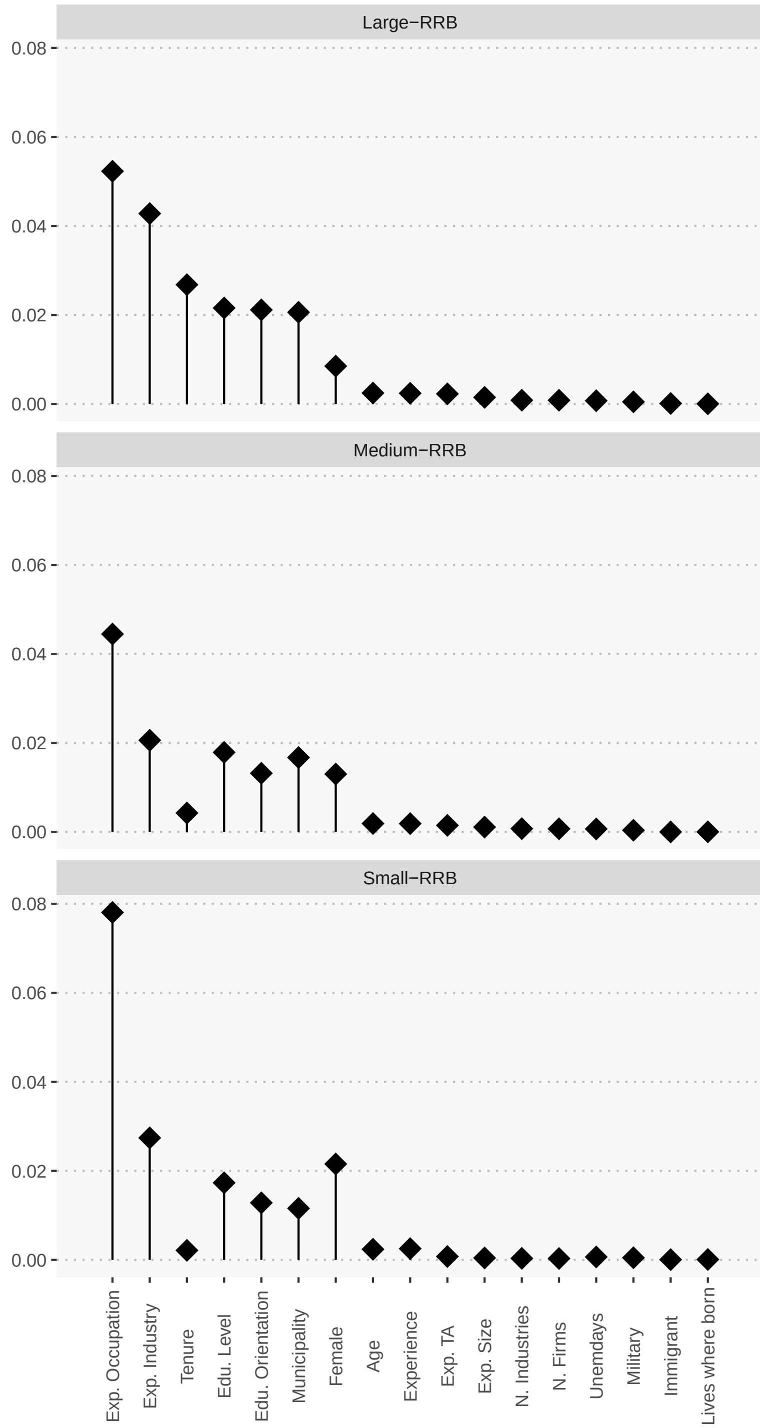


Figure 2: continued

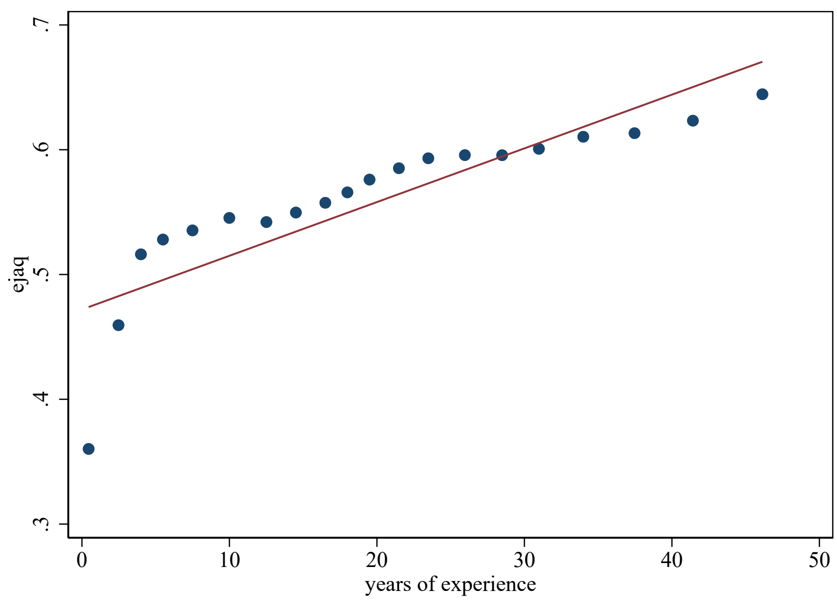


Figure 3: Worker-level job allocation quality (*eJAQ*) by labor market experience
This figure shows the binned scatter plot of an indicator for being assigned to the job predicted by the ML algorithm (*eJAQ*) against years of labor market experience.

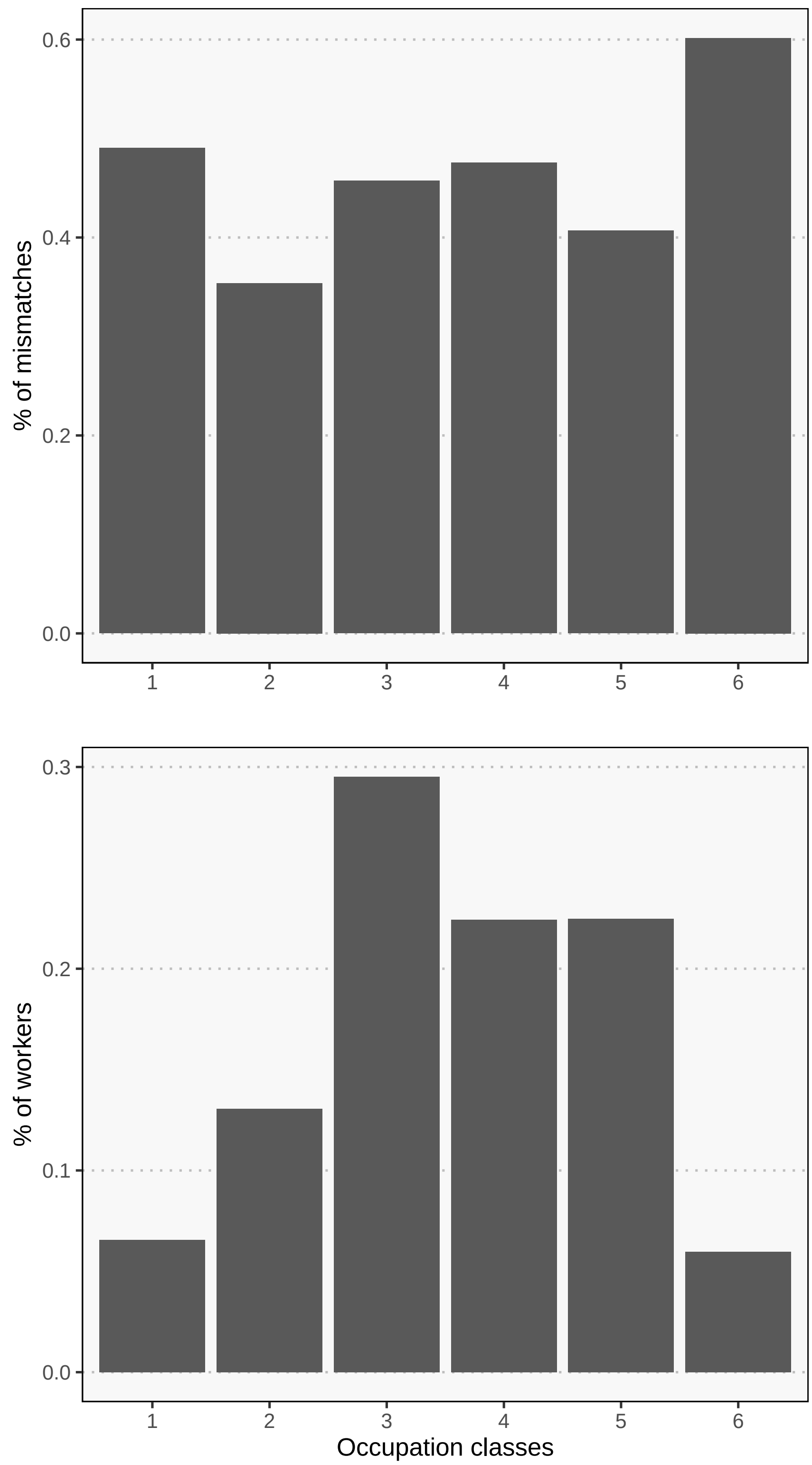


Figure 4: Distribution of mismatches and workers by occupation classes
The top graph shows the percentage of mismatches in each occupation class in the main sample. A mismatch occurs when an employee's observed job differs from the job predicted by the estimated allocation rule. The bottom graph shows the percentage of workers by occupation classes in the main sample. Occupation classes are defined as follows: 1) managers, 2) professionals, 3) technicians and clerks, 4) skilled manual workers , 5) machine operators and assemblers, and 6) elementary occupations.

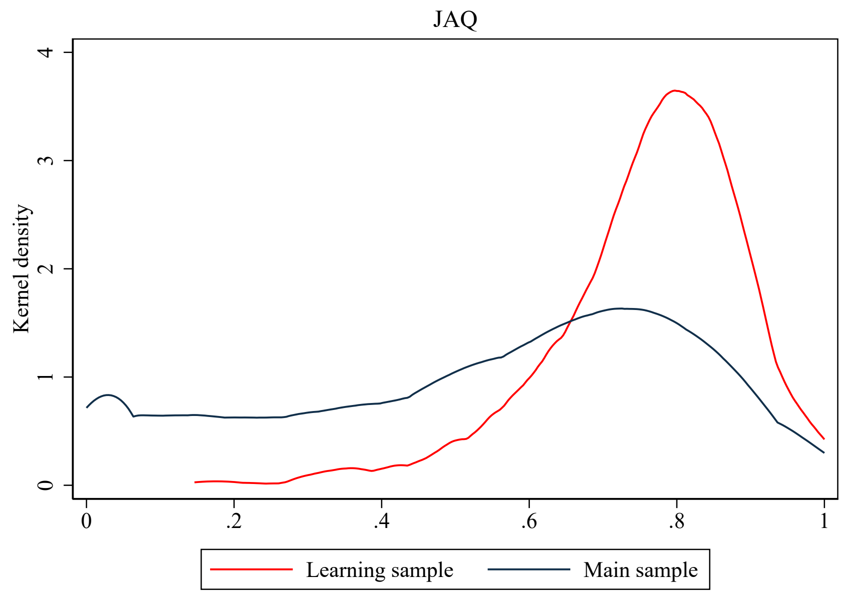
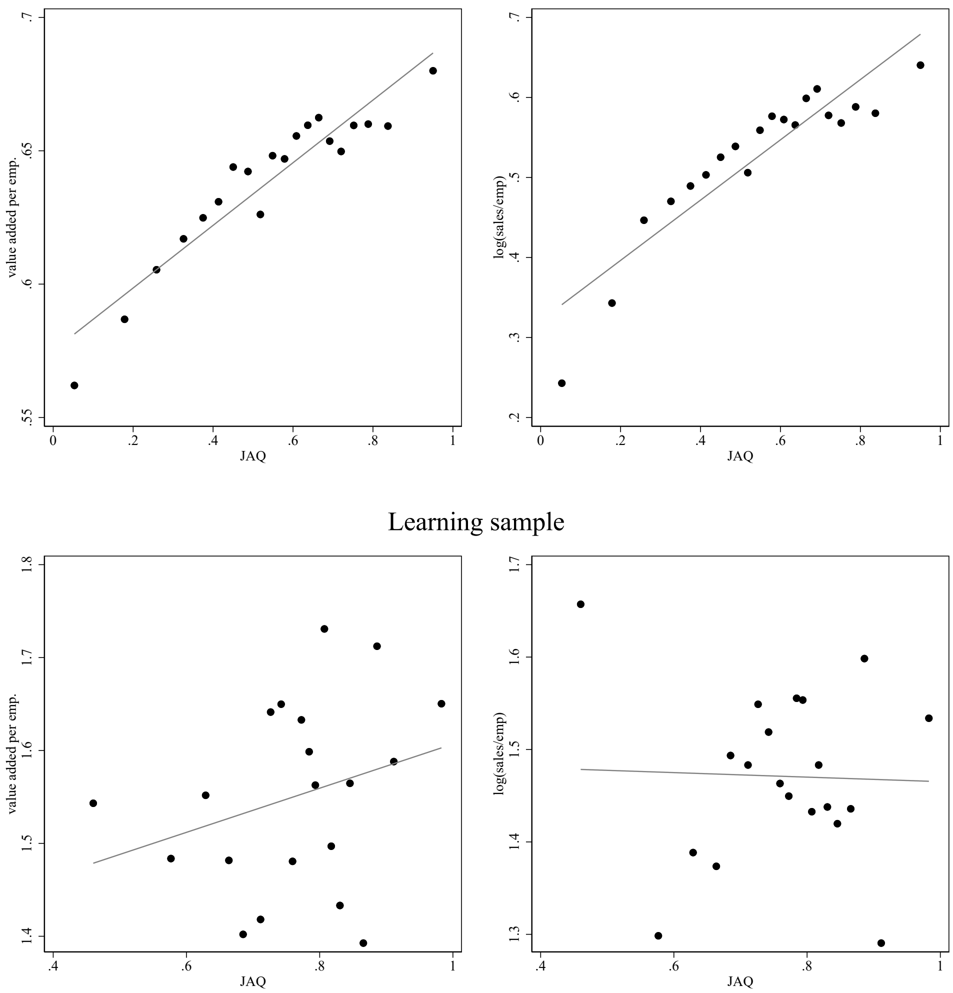


Figure 5: The distribution of *JAQ*
This figure shows the kernel density estimate of *JAQ* for firms in the main sample (blue line) and in the learning sample (red line).

Main sample

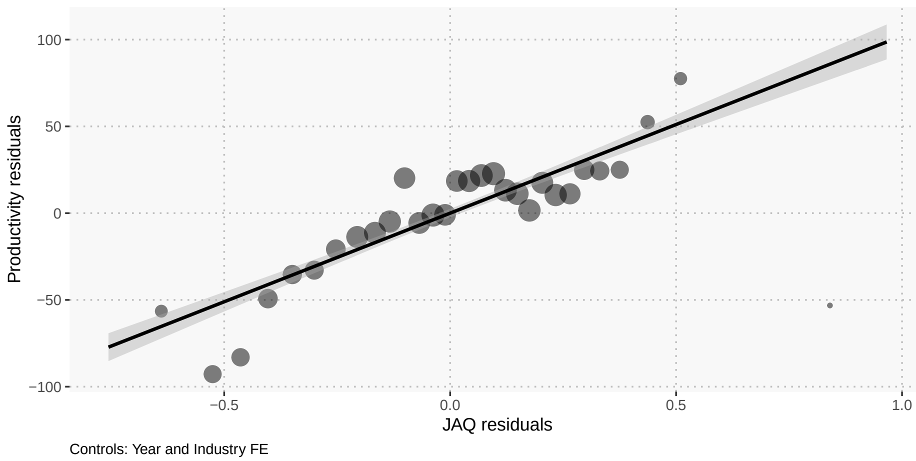
Learning sample



Controls: year and 2digits-industry dummies

Figure 6: The correlation between productivity and ***JAQ***
The figure shows binned scatter plots of productivity (y-axis) against *JAQ* (x-axis) controlling for year and industry fixed effects. The two top figures refer to the main sample and the two bottom ones to the learning sample. The left-hand graphs measure productivity using value added per employee and the right-hand side figures using the log of sales per employee. Both measures of productivity show a strong correlation with *JAQ* in the main sample.

(a) Top firms set



(b) Placebo test

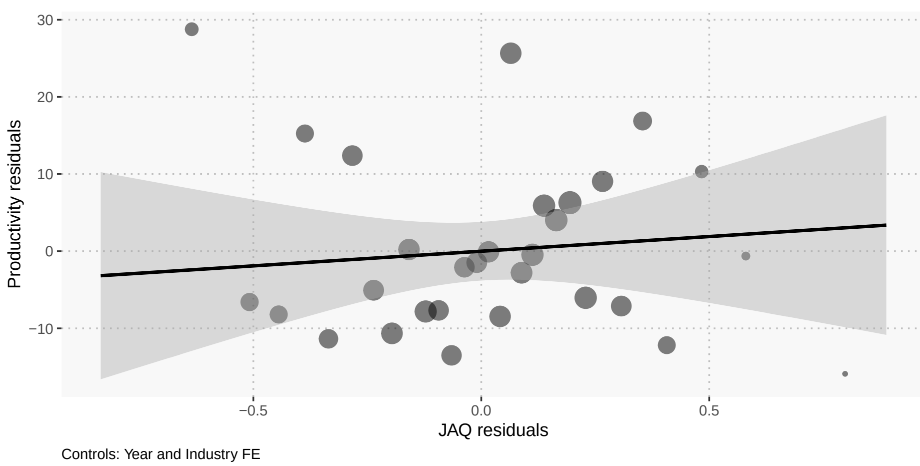


Figure 7: Placebo test

The figure shows the partial regression plots between (placebo) productivity and *JAQ* for the allocation rule estimated using firms in the top decile by actual productivity (panel a) and by placebo productivity (panel b), respectively. The regression lines are shown together with the respective 95% confidence intervals (shaded area); the controls used in both regressions are years and industry effects. The points shown in the two graphs represents the residuals of the partial regression plots and are computed as follows: residuals are first clustered via a k-means algorithm on the horizontal axis; centroids determine the abscissa of the points; the ordinate is the average productivity values of the points in a cluster, and their size is proportional to the clusters' size. The regression line shown is fitted on true residuals and not on clustered points.

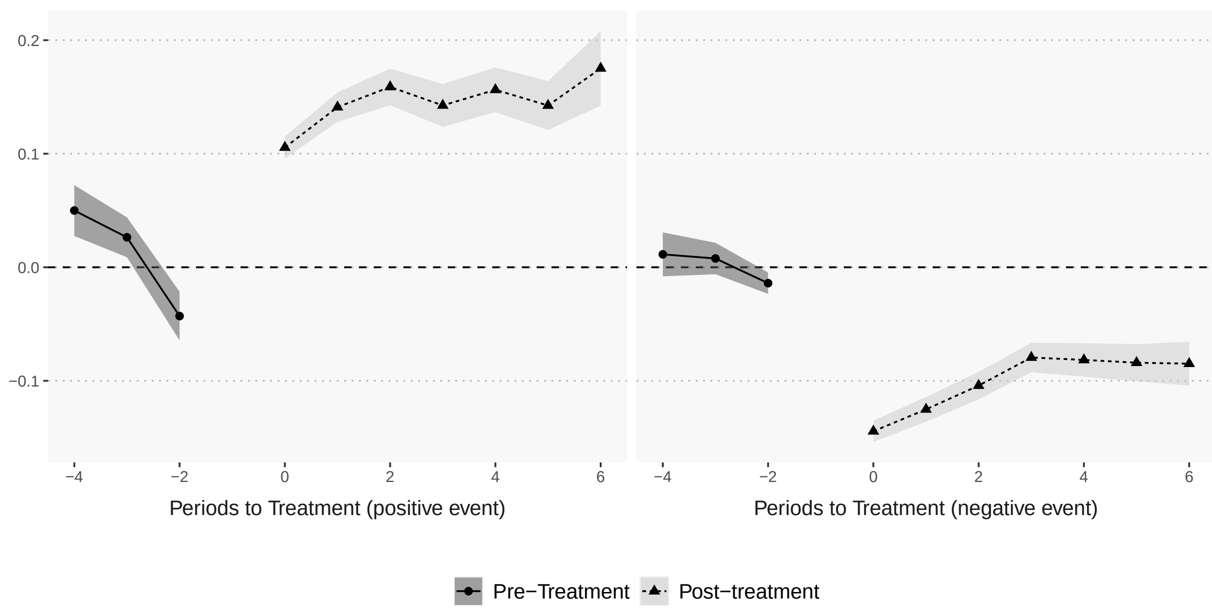


Figure 8: Response of rank-and-file workers' *JAQ* to positive (left) or negative (right) managerial turnover events

The figure shows shows the coefficients on event time dummies estimated with the method by Callaway and Sant'Anna (2021) explaining the *JAQ* of rank-and-file workers with event time dummies around managerial turnover events respectively associated with a persistent increase in the *JAQ* of the relevant firm's management (left panel) and with a persistent decrease in the *JAQ* of the relevant firm's management (right panel), controlling for year and firm fixed effects.

Table 1: Descriptive statistics

This table reports the summary statistics of the individuals included in the main sample and in the learning sample. Our total sample includes firms with at least 30 employees, that are active at some point between 2001 and 2010 and report positive total assets and sales. Since information about a worker's specific occupation is not always available, we restrict the sample to firms with at least 10 workers for whom we do observe the current occupation. The main sample contains 7,180,365 observations at the individual level and the learning sample 81,390 observations.

	Mean	P50	P10	P25	P75	P90	SD
Panel A: Main sample							
Labor income (TSEK 2017)	330.05	305.59	132.78	234.08	390.54	521.77	206.03
University degree	0.13	0.00	0.00	0.00	0.00	1.00	0.34
Age	40.49	40.00	24.00	31.00	50.00	58.00	12.19
Female	0.35	0.00	0.00	0.00	1.00	1.00	0.48
Immigrant	0.14	0.00	0.00	0.00	0.00	1.00	0.34
Mobility (lives where born)	0.64	1.00	0.00	0.00	1.00	1.00	0.48
Labor market experience	19.46	18.00	3.00	8.00	29.00	39.00	13.05
Tenure	4.99	3.00	0.00	1.00	8.00	13.00	5.08
# industries worked in	2.30	2.00	1.00	1.00	3.00	4.00	1.27
# jobs held	2.99	3.00	1.00	2.00	4.00	5.00	1.72
# unemployment days since '92	192.46	0.00	0.00	0.00	243.00	626.00	349.44
Panel B: Learning sample							
Labor income (TSEK 2017)	444.78	389.02	221.95	309.10	505.68	691.86	315.90
University degree	0.21	0.00	0.00	0.00	0.00	1.00	0.41
Age	42.45	42.00	27.00	34.00	51.00	58.00	11.36
Female	0.32	0.00	0.00	0.00	1.00	1.00	0.47
Immigrant	0.14	0.00	0.00	0.00	0.00	1.00	0.35
Mobility (lives where born)	0.62	1.00	0.00	0.00	1.00	1.00	0.49
Labor market experience	20.55	20.00	4.00	10.00	30.00	39.00	12.62
Tenure	6.95	5.00	0.00	2.00	11.00	19.00	6.41
# industries worked in	2.64	2.00	1.00	1.00	4.00	5.00	1.44
# jobs held	3.42	3.00	1.00	2.00	5.00	6.00	1.96
# unemployment days since '92	182.10	0.00	0.00	0.00	220.00	576.00	344.07

Table 2: Performance of the algorithm

Weighted average of job-specific F1 scores (standard errors in parentheses), weighted by job frequencies, for each size-industry class. Training and test set results are obtained via a stratified 10-fold cross validation; mean and standard errors are computed across the 10 cross validated pairs of scores obtained for the training and the test sets, respectively. The weighted F1 score of the algorithm on the training sets is between 71% and 81% and on the test sets between 61% and 77%. This performance can be benchmarked against that resulting from a a random allocation of workers to jobs in our sample, which would result in a 4.3% maximal weighted F1 score.

	Size	Manufacturing	Wholesale	Real Estate
Training	Large	0.8014 (0.0014)	0.7702 (0.0039)	0.8133 (0.0039)
	Medium	0.7244 (0.0015)	0.7042 (0.0034)	0.7795 (0.0031)
	Small	0.7195 (0.0035)	0.7532 (0.0018)	0.7896 (0.0034)
Test	Large	0.7678 (0.0076)	0.7253 (0.0089)	0.7678 (0.012)
	Medium	0.6676 (0.0072)	0.6264 (0.0074)	0.7156 (0.0039)
	Small	0.6084 (0.0092)	0.6579 (0.0108)	0.6774 (0.0099)

Table 3: Individual labor earnings and *eJAQ*

This table displays the relationship between *eJAQ* and labor earnings for workers in the main sample. The worker controls (used in the ML algorithm) are age, gender, an indicator for immigrant status, residence municipality, a mobility indicator equal to one for workers employed in a county different from the county of birth, education level (basic, high school, vocational, or university), education subject (no specialization, law, business and economics, health and medicine, natural sciences, teaching, engineering, social sciences, services, or other specializations), labor market experience (measured as years since graduation), tenure at the current firm, number of firms and number of two-digit industries where an individual previously worked, total number of unemployment days since 1992 (when the unemployment data starts in LISA), years of experience in each occupation, years of experience in each 2-digit industry, and years of experience in each decile of the distribution of firms' number of employees or total assets. The firm controls are firm age, size (measured by the number of employees), sales, and total assets, as well as ownership categories measured by indicators for the firm being a state-owned firm, a listed firm, or a family firm. Standard errors clustered at worker level are shown in parentheses and three stars denote statistical significance at the one percent level.

	(1) Log(earnings)	(2) Log(earnings)	(3) Log(earnings)
Panel A			
eJAQ	0.014***	0.007***	0.008***
	(0.000)	(0.000)	(0.000)
Panel B			
Suitability	0.180***	0.032***	0.005***
	(0.001)	(0.001)	(0.001)
Year and job FE	✓	✓	✓
Worker controls (used in ML)	✓	✓	
Industry FE		✓	
Firm controls		✓	
Worker FE			✓
Observations	7,180,365	7,180,365	7,180,365
LHS mean	6.326	Mean earnings	330.05
LHS SD	0.658	SD earnings	206.03

Table 4: Firm performance and *JAQ*

This table displays results from regressions on the association between productivity and *JAQ*. Panel A refers to our baseline specification. The results in Panel B control for firms' occupation structure (the fraction of workers in firm *f* assigned to job *j* in year *t*) and firm controls (age, family firm status, state-owned status, listed firm status, an indicator for the presence of a human resources (HR) manager, its log number of employees and its log of total assets). Panel C adds controls for worker characteristics (listed in the notes to Table 3). Standard errors clustered at firm level are shown in parentheses and three stars denote statistical significance at the one percent level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(sales/emp)	VA/emp	OROA	Log(sales/emp)	VA/emp	OROA
Panel A						
JAQ	0.521***	0.125***	-0.003	0.159***	0.039***	0.007
	(0.030)	(0.009)	(0.006)	(0.016)	(0.007)	(0.005)
log(cap/emp)				0.400***	0.107***	-0.009***
				(0.011)	(0.005)	(0.002)
log(emp)				0.017***	0.018***	0.005***
				(0.006)	(0.003)	(0.002)
Share emp w/ college				0.194***	0.301***	-0.006
				(0.040)	(0.021)	(0.013)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Panel B						
JAQ	0.209***	0.061***	0.006	0.109***	0.028***	0.003
	(0.020)	(0.008)	(0.005)	(0.016)	(0.008)	(0.006)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Class FE	✓	✓	✓	✓	✓	✓
Firm controls				✓	✓	✓
Panel C						
JAQ	0.093***	0.038***	0.004	0.055***	0.025***	0.004
	(0.019)	(0.008)	(0.005)	(0.016)	(0.008)	(0.006)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Class dummies	✓	✓	✓	✓	✓	✓
Observations	50,107	50,107	50,107	50,107	50,107	50,107
No. firms	7,232	7,232	7,232	7,232	7,232	7,232
LHS mean	0.525	0.639	0.067	0.525	0.639	0.067
LHS SD	0.788	0.304	0.178	0.788	0.304	0.178

Table 5: Firm performance and worker suitability

This table displays results from regressions on the relationship between productivity and firm-level suitability, defined as the average of worker-level suitability for a given firm. Panel A refers to our baseline specification. The results in Panel B control for firms' occupation structure (the fraction of workers in firm f assigned to job j in year t) and firm controls (age, family firm status, state-owned status, listed firm status, an indicator for the presence of a human resources (HR) manager, its log number of employees and its log of total assets). Panel C adds controls for worker characteristics (listed in the notes to Table 3). Standard errors clustered at firm level are shown in parentheses and three stars denote statistical significance at the one percent level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(sales/emp)	VA/emp	OROA	Log(sales/emp)	VA/emp	OROA
Panel A						
Suitability	1.219***	0.316***	-0.011	0.436***	0.097***	0.021*
	(0.060)	(0.019)	(0.012)	(0.035)	(0.016)	(0.012)
log(cap/emp)				0.397***	0.107***	-0.009***
				(0.011)	(0.005)	(0.002)
log(emp)				0.022***	0.019***	0.006***
				(0.006)	(0.003)	(0.002)
Share emp w/ college				0.152***	0.292***	-0.008
				(0.040)	(0.021)	(0.013)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Panel B						
Suitability	0.591***	0.138***	0.011	0.339***	0.058***	0.007
	(0.042)	(0.019)	(0.012)	(0.034)	(0.017)	(0.012)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Class FE	✓	✓	✓	✓	✓	✓
Firm controls				✓	✓	✓
Panel C						
Suitability	0.294***	0.078***	0.010	0.231***	0.054***	0.011
	(0.043)	(0.020)	(0.013)	(0.036)	(0.019)	(0.013)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Class dummies	✓	✓	✓	✓	✓	✓
Observations	50,107	50,107	50,107	50,107	50,107	50,107
No. firms	7,232	7,232	7,232	7,232	7,232	7,232
LHS mean	0.525	0.639	0.067	0.525	0.639	0.067
LHS SD	0.788	0.304	0.178	0.788	0.304	0.178

Table 6: Role of management in the allocation quality of rank-and-file employees

This table displays results from regressions whose dependent variable is the job allocation quality of rank-and-file employees (*R&F-JAQ*) and whose explanatory variables are the allocation quality of managers (*M-JAQ*) and their experience in managerial jobs (Manager exp). In columns 1 to 3 Manager-*JAQ* refers both to top managers (CEOs and firm directors) and to middle managers, whereas in columns 4 to 6 they only refer to top managers. The regressions are based on data from 2003 to 2010. All specifications include year fixed effects; those in columns 2, 3, 5 and 6 include firm fixed effects, and those in columns 3 and 6 include industry fixed effects, municipality fixed effects and firm controls (age, family firm status, state-owned status, listed firm status, an indicator for the presence of a human resources (HR) manager, its log number of employees and its log of total assets). Standard errors clustered at firm level are shown in parentheses and three stars denote statistical significance at the one percent level.

	(1) *R&F-JAQ*	(2) *R&F-JAQ*	(3) *R&F-JAQ*	(4) *R&F-JAQ*	(5) *R&F-JAQ*	(6) *R&F-JAQ*
M-JAQ	0.149***	0.144***	0.142***	0.077***	0.065***	0.063***
	(0.004)	(0.007)	(0.007)	(0.003)	(0.006)	(0.006)
Manager exp		0.034***	0.033***		0.019***	0.019***
		(0.002)	(0.002)		(0.002)	(0.002)
Industry FEs			✓			✓
Municipality FEs			✓			✓
Year FEs	✓	✓	✓	✓	✓	✓
Firm FEs		✓	✓		✓	✓
Firm controls			✓			✓
Observations	31,488	31,488	31,475	22,391	22,391	22,386
No. Firms		5,677	5,677		4,822	4,822

Table 7: Decomposition of managerial turnover impact on stayers

The table shows the impact of managerial turnover events on allocations of employees staying in the firm. The coefficients measure the estimated average treatment effect on treated at event time (standard errors in parentheses) obtained using the Callaway-Sant'Anna method, for positive and negative managerial turnover events, respectively. The first column shows the estimates obtained when the dependent variable is $\Delta R\&F\text{-}JAQ$, i.e., the change in the average allocation quality of rank-and-file employees. The second column shows the estimates obtained when the dependent variable is $\Delta R\&F\text{-}JAQ^r$, i.e., the change in the fraction of correctly allocated employees among those retained by the firm when the event occurs. The third column shows the the estimates obtained when the dependent variable is $\Delta R\&F\text{-}JAQ^c$, i.e., the change in the fraction of correctly allocated employees for those retained by the firm when the event occurs and reassigned to different jobs.

	$\Delta R\&F\text{-}JAQ$	$\Delta R\&F\text{-}JAQ^r$	$\Delta R\&F\text{-}JAQ^c$
Positive event	0.1443 (0.023)	0.0935 (0.019)	0.0974 (0.0765)
Negative event	−0.1317 (0.0171)	−0.0953 (0.0128)	−0.1897 (0.0519)