EIEF WORKING PAPER SERIES

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 22/05

April 2022

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

By

Luca Coraggio
(University of Naples Federico II)

Marco Pagano
(University of Naples Federico II and EIEF)

Annalisa Scognamiglio
(University of Naples Federico II)

Joacim Tåg
(Research Institute of Industrial Economics (IFN))

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
IFN and Hanken School of Economics

11 December 2024

ABSTRACT

We develop a novel measure of job-worker allocation quality (*JAQ*) by exploiting employer-employee data with machine learning techniques. Based on our measure, the quality of job-worker matching correlates positively with individual labor earnings and firm productivity, as well as with market competition, non-family firm status, and employees' human capital. Management plays a key role in job-worker matching: when managerial hirings and firings persistently raise management quality, the matching of rank-and-file workers to their jobs improves. *JAQ* can be constructed from any employer-employee data set including workers' occupations, and used to explore research questions in corporate finance and organization economics.

Keywords: jobs, workers, matching, mismatch, machine learning, productivity, management.

JEL Codes: D22, J24, J31, L22, L23, M12, M54.

*We are grateful for insightful suggestions and comments from Ramin Baghai, Kirill Borusyak, Vicente Cuñat, Morten Grindaker, Daniel Halvarsson, Alex Xi He, Hans Hvide, Simon Ek, Isil Erel, Anastassia Fedyk, Jessica Jeffers, Tullio Jappelli, Camelia Kuhnen, Mikael Lindahl, Topi Miettinen, Ahmed Ameya Prapan, Raffaele Saggio, Elia Sartori, Tanja Saxell, Elena Simintzi, Rune Stenbacka, Marko Terviö and participants at the 2023 AFA Meetings, the 5th Bank of Italy-CEPR workshop on labour market policies and institutions, 2022 EFA Meetings, EEA-ESEM Congress and WFA 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, CBS, CSEF, EIEF, ETLA, Hanken, Helsinki GSE, IFN, Jönköping University, the NICE Conference, Swedish Conference in Economics, Ratio, UNC Kenan-Flagler Business School, and Örebro University. 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

Matching workers to their best possible job is of paramount importance for firms: hiring or firing the wrong people, or matching employees to the wrong task may sap a firm's productivity at least as much as it damages employees' careers, by slowing down their skill acquisition process, reducing their wage growth and possibly inducing them to switch to another employer. Indeed, the ability to match workers to the right jobs is a hallmark of a good manager, on a par with other management skills that have been shown to contribute to firm productivity (Bertrand and Schoar, 2003; Bennedsen et al., 2020; Bloom and Van Reenen, 2007, 2010; Bloom et al., 2013, 2019; Scur et al., 2021). However, managerial practices measured by existing research on human resource management focus on workers' incentives, and neglect choices regarding the allocation of workers to jobs. This shortcoming probably reflects that managerial practices have so far been measured by surveying managers regarding how they run their firms' operations, monitoring, incentives and targets, while it would be difficult to use surveys to inquire whether managers allocate workers to their best possible use within the firm.

This is where our paper comes in: we develop a novel measure of job allocation quality (*JAQ*, for short) by applying machine-learning (ML) techniques to administrative employer-employee data rather than building on responses to questionnaires, and validate our measure in a variety of ways, which also bear witness to its versatility. *JAQ* is built in four steps. First, we estimate via a ML algorithm how workers' characteristics map into jobs, using as a benchmark for match quality the allocation of workers to jobs in the most productive firms. This is similar in spirit to how Bloom and Van Reenen (2007) benchmark management practices against standards set by a leading management consulting firm and how Fredriksson et al. (2018) gauge mismatch of junior workers by the distance between their skills and those of senior workers holding the same job. Second, we predict worker suitability for each job based on the function estimated via the ML algorithm. Third, we build an indicator of whether an employee's actual job coincides with her most suitable one (*eJAQ*). Fourth, we average this indicator across all the employees of each firm to construct our firm-level measure of match quality (*JAQ*).

The ML algorithm on which *JAQ* is constructed maps employees' curricular characteristics such as education, age, gender and experience into a job assignment rule. Experience and education are the most important determinants of worker suitability to jobs, although their role appears to differ across occupations. For example, holding a relevant college degree is a key predictor of the suitability for legal professions, while for other jobs tenure and experience tend to have a much more prominent role. We decompose the variation in *eJAQ* between employees who stay in the firm and those who move to a new firm: most of the variation in *eJAQ* turns out to be accounted for by employees remaining with their current employer. This also applies at the firm level, where increases in *JAQ* are typically associated with employees who retain their initial job and improve their match quality as their experience accumulates.

We then validate the employee-level measure, testing whether workers' careers benefit from their job match quality. It is natural to expect employees to be assigned to more suitable jobs over their careers, as managers learn about their characteristics (Fredriksson et al., 2018), and employees themselves fine-tune their skill set via on-the-job training (Guvenen et al., 2020). Also, insofar as better job allocation enhances productivity, workers can be expected to appropriate at least part of the gain in the form of higher wages. Both of these predictions find support in our data. The goodness of worker-job matches (*eJAQ*) rises significantly over working lives, the largest gain occurring in the first few years. Moreover, workers allocated to their most suitable job earn significantly more than mismatched workers with the same characteristics or with the same job, and are less likely to switch to a new employer. Both of these findings dovetail with those reported by Fredriksson et al. (2018), despite differences in methodology and sample used.

Next, we validate the firm-level measure of job allocation quality (*JAQ*) by showing that it correlates positively with market competition, non-family firm status, workers' human capital and, most importantly, with productivity, just as good management practices do (Bloom and Van Reenen, 2007, 2010; Bloom et al., 2013, 2019). In particular, we show that *JAQ* has a significant, sizeable, and robust positive correlation with log value added and log sales per employee, even upon controlling for the firm-level variables generally associated with productivity (industry, capital and labor, ownership) and for the workers' characteristics used to predict *JAQ*. The correlation with productivity also persists using an adjusted *JAQ* measure that assumes the highest possible level of *JAQ* to be constrained by the employees and the positions available within the firm; hence, it does not consider the further potential improvement of match quality attainable via the hiring and firing process.

One may suspect the positive correlation between *JAQ* and productivity to be affected by circularity, as we first train the ML algorithm to assign workers to jobs based on data for the most productive firms, and then we investigate whether *JAQ* correlates with firm productivity. The first counter to this criticism is that the correlation between *JAQ* and productivity is estimated by 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 its estimation error may correlate by construction with firm productivity. We perform several exercises that dispel remaining circularity concerns, by training the ML algorithm on samples not based on firm productivity. First, we perform a placebo test replacing actual firm productivity with a noise variable distributed like the replaced variable and show that the relationship between *JAQ* and firm performance is not purely mechanical. Second, we retrain the algorithm on random subsamples of firms, to calibrate *JAQ* on the average firm in our sample rather than on top-productivity firms, and perform a Monte Carlo experiment to explore the robustness of the relationship between the resulting measures and firm productivity to sampling variability. Third, we construct an alternative data set to train the algorithm based on the residuals of an AKM wage regression model (Abowd et al., 1999), as these residuals capture the surplus from worker-job matches and therefore can be used to identify better-allocated workers. All of these robustness checks confirm our baseline results.

Finally, we investigate the role of managers in shaping a firm's job allocation quality by testing whether improvements in management quality bring about a better allocation of rank-and-file workers to tasks. Upon constructing two distinct *JAQ* measures for rank-and-file workers and for managers, the former turns out to be 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. In turn, the job allocation quality of the firm's managerial team has a sizable and robust correlation with firm productivity. Interestingly, the quality of rank-and-file workers' allocation improves significantly when the managerial allocation quality itself rises owing to new hirings and firings; this tends to occur in the wake of a deterioration in the allocation of rank-and-file workers. Conversely, when managerial hirings and firings result in a worse allocation of the firm's management, they are followed by a persistent disruption in the allocation of rank-and-file workers. These results persist even when the analysis is restricted to changes in management associated with the death of an incumbent manager, although in this subsample the estimates are imprecise due to the paucity of observations.

We additionally investigate the effects of the two types of managerial turnover events on the frequencies of employees' internal reallocation, firings and hirings, and on the allocation quality of rank-and-file workers. Better management improves the job allocation quality of incumbent rank-and-file workers ("stayers"), mostly by retaining them in their current positions and letting them accumulate job experience. In contrast, the arrival of worse managers lowers the job allocation quality of stayers via misguided job reassignments.

Thus, our contribution is twofold. First, we build a new firm-level measure that can be used to investigate the importance of worker-job matches for firm performance and can be constructed for any country and period where survey or administrative data on workers' characteristics and job titles exist, without requiring expensive targeted surveys (Bloom and Van Reenen, 2007; Bloom et al., 2019) or detailed expert evaluations of the skills required for each job (Lise and Postel-Vinay, 2020; Guvenen et al., 2020). Second, we contribute to the literature on the importance of managers in shaping firm performance, by highlighting the role that their quality and allocation play in matching rank-and-file workers to their jobs.

As already mentioned, the measure that we propose complements the research on the role of managerial practices and human capital for firm productivity (Scur et al., 2021). Most closely related are Ichniowski et al. (1997), Bloom and Van Reenen (2007), Bloom et al. (2013), Bloom et al. (2019) and Cornwell et al. (2021), who study how management practices relate to productivity, Bender et al. (2018), who investigate the relationship between productivity, management practices, and employee ability, the study by Fox and Smeets (2011) on the role of workers' quality in explaining the dispersion in productivity, and Minni (2023) on the role of managers for workers careers. Our distinctive contribution here is to focus on managerial policies governing the allocation of workers to jobs within the firm and relate job-worker mismatch to firm productivity. Our measure of mismatch is likely to be informative not only about the role of labor misallocation at the firm level, but also at higher levels of aggregation: for example, it can be used to investigate how technological innovation and regulatory changes influence economy-wide productivity and to assess how skill mismatch varies over the business cycle or during financial crises (Bowlus, 1995; Baley et al., 2022).

Furthermore, our paper relates to an emerging strand of research that exploits ML to address questions at the interface between labor and finance.[1] Examples include the appointment of board of directors (Erel et al., 2021), the screening of resumes in recruitment (Li et al., 2020), the measurement of corporate culture based on earnings call transcripts (Li et al., 2021), and the assessment of what managers do (Bandiera et al., 2020). Our results on the correlation between managerial quality and the match quality of rank-and-file workers directly relate to the research on managers' role in allocating human capital within the firm and in shaping workers' careers (Minni, 2023; Pastorino, 2024) and more generally in affecting firm performance (Bertrand and Schoar, 2003; Malmendier and Tate, 2005; Bennedsen et al., 2007; Malmendier and Tate, 2009; Kaplan et al., 2012; Lazear et al., 2015; Mullins and Schoar, 2016; Bandiera et al., 2018; Bennedsen et al., 2020). More broadly, this measure can be deployed to investigate several other questions in corporate finance, such as the role of human capital in private equity interventions and in mergers, where worker reallocation from combining the workforce of two firms may lead to higher productivity and lower costs. It can also shed light on the relationship between match quality and the financial returns on human capital investments, such as employee training, development programs, and

[1] 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).

recruitment strategies.

Finally, our work complements the vast research in labor economics on how workers match with firms (Jovanovic, 1979; Cahuc et al., 2006; Postel-Vinay and Robin, 2002; Chiappori and Salanié, 2016; Eeckhout and Kircher, 2018; Adenbaum, 2023; Pastorino, 2024) and with tasks (Perry et al., 2016; Lindenlaub, 2017; Deming and Kahn, 2018; Lise and Postel-Vinay, 2020; Guvenen et al., 2020; Ocampo, 2022), and on how managers match with firms (Terviö, 2008; Lippi and Schivardi, 2014; Benson et al., 2019; Bandiera et al., 2020). A particularly relevant 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. While their method only applies to the measurement of junior workers' match quality, our measure of job assignment quality applies to the allocation of all workers, which is key to evaluating how job assignment quality correlates with firm productivity.

The road map reads as follows. The next section describes our data, and Section 3 details how we construct our match quality measure and shows how it correlates with worker-level outcomes and firm characteristics. 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 over time the entire population of workers and firms in Sweden, including variables regarding workers' job histories, such as occupations and wages over their careers. 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 comes 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 to construct 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 of state ownership, listed status, and family ownership.[2] 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 data on board members and CEOs from the Swedish Companies Registrations office and the multi-generational register

[2]See Olsson and Tåg (2024) for details on the state-owned firm dummy.

on biological parent-child relationships at Statistics Sweden. Following Keloharju et al. (2023), a family firm is defined as one managed or owned by at least two members of the same family.

We identify jobs based on the international ISCO-88 (COM) 3-digit classification of occupations, based on data provided primarily by official wage statistics drawn from 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. In total, during our entire sample period, over 90% of workers are sampled at least once.[3]

In extracting our sample of firms from the LISA database, we apply two screens by firm size: we only retain firms whose median number of employees in the sample period is between 30 and 6,000. The lower bound is due to the sparsity of occupational information for firms with less than 30 employees: including these firms would introduce large noise in the estimation of the job-employee matching rule. These firms employ about 30% of the total reported workforce and on average report 2.9 employees per year. The upper bound of 6,000 employees excludes from the sample very large firms that may otherwise dominate the estimates of the job-employee matching rule, despite featuring a quite different structure from other firms, e.g., a more layered corporate hierarchy and a richer set of possible occupations. These firms account for 20% of the total reported workforce, but are few: out of a total of 945,385 firms in the database, there are only 80 such firms, which drop to 11 in the industries retained in our analysis.

Our sample includes firms in three industries: (i) manufacturing; (ii) real estate, renting and business activities; and (iii) wholesale and retail, which include 62% of the firms and 70% of the employees present in the LISA database over our sample period (after applying the screen based

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

on firm size), and thus employ far more workers than other industries.[4] Moreover, these industries feature the most complete and heterogeneous set of occupations in the Swedish economy: they include the greatest number of occupations, namely, 99%, 98% and 96% of the total 110 jobs, against a mean of 72% in other industries. The first two of these industries also feature a more diversified set of occupations than others, making the workers' assignment problem more relevant: the Herfindahl–Hirschman index measuring the concentration of occupations is 4% and 6% respectively, compared to a mean of 14% in other industries.[5] After applying these filters, our sample comprises 9,023 firms, employing a total of 1,541,343 employees over the 2001-10 period.

3 Measuring Job Assignment Quality

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. The allocation can vary depending on the firm's size and industry, and on workers' location and thus on features of the local labor market. However, firms may deviate from the most efficient assignment rule, incurring in errors that reduce their productivity, because of managerial shortcomings and/or information and learning frictions. As noted by Bloom and Van Reenen (2007), firms may also choose not to implement optimal management practices because doing so may be too costly for their managers. Furthermore, firms may adopt non-meritocratic personnel policies if these yield private benefits to their controlling shareholders in the form of power over the firm (Pagano and Picariello, 2024) or of utility from workplace discrimination (Becker, 1957). As external circumstances change, firms may also not immediately adjust to the new optimal assignment rule due to adjustment costs.

To assess a firm's job allocation quality (*JAQ*), it is necessary to estimate the rule that managers use to assign workers to jobs. In principle, this can be done using a random subsample of firms.

[4]The excluded industries are: agriculture, hunting, forestry and fishing; mining; utilities; construction; hotels and restaurants, transport, storage, and communications; financial intermediation; public administration and defense; education; health and social work; other service activities.

[5]This does not apply to the wholesale and retail industry, whose concentration index is 18%.

However, insofar as the rule maximizes productivity, the firms that apply it most rigorously should deviate less often from it and thus feature the highest productivity. Hence, the rule can be observed with the least noise for the most productive firms.

Thus, a key feature of our measure is that it is generated by benchmarking against the matches in the most productive firms. As such, it is consistent with any model of labor market search and worker assignment that predicts higher match quality in more productive firms, such as Moen (1997), Postel-Vinay and Robin (2002), and Cahuc et al. (2006). In these directed search models, more productive firms can afford to pay higher wages, thus attracting higher-quality workers and producing better match outcomes. Such productivity advantage may stem from managers being better at solving the multidimensional skill matching problem between the skills required by an occupation and workers' abilities to acquire those skills, as in Guvenen et al. (2020).

Accordingly, we use a machine learning (ML) algorithm to estimate the matching 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. Of course, the rule estimated by the ML algorithm is bound to be efficient only on average: firms are likely to condition their job-worker matches on more information than that available to us in estimating the algorithm. Hence, some of the observed firm-level deviations from the estimated rule may reflect firm-specific information not captured by the algorithm rather than firm-level errors in applying the optimal rule.

3.1 Mapping workers' characteristics to jobs via machine learning

In our framework, 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 firms' 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 with different characteristics rely on different rules.

For computational reasons, the sample is broken up in various subsamples to train the algorithm: this significantly reduces the estimation time compared to that required to estimate the algorithm on the full sample.[6] The sample split is based on firm characteristics: firms are sorted across the three industries described above and three size classes, resulting in 9 size-industry bins. The three size classes are based on firms' median number of employees, N, over the sample period: (i) small ($30 \leq N \leq 50$); (ii) medium ($51 < N \leq 250$); (iii) large firms ($N > 250$). The algorithm is estimated separately for the firms in each size-industry bin, taking into account that firms in different bins may use different rules to match their employees to occupations: for instance, larger firms typically have more layers in their hierarchy than smaller ones, and manufacturing firms have a greater variety of occupations than those in wholesale and retail trade. Hence, this approach amounts to estimating bin-specific conditional probabilities $P_z(J|X) \equiv P(J|X, Z = z)$, for $z \in \{1, \ldots, 9\}$, where Z is a variable identifying the size-industry bin a firm belongs to.

Within each size-industry bin, 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, to include in the learning sample only firms that are consistently more productive, for each size-industry bin 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 belonging to the top decile of this distribution. We then use 2010 data for 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 how firms allocate employees to jobs. Using data for these firms, we estimate bin-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 bin.

Table 1 compares the characteristics of the workers included in the main sample and in the learning sample: the workers included in the latter earn higher wages, are more educated, have longer tenure, and experience fewer days of unemployment than workers included in the former.

[6] Attempts to perform the estimation on the full sample exceeded the 1,000-hours limit to computation time set by the Statistics Sweden server.

These differences are consistent with the fact that the learning sample includes more productive firms, where workers can be expected to feature more productive matches and thus experience fewer separations.

Insert Table 1 here

Despite these differences, the two samples are sufficiently similar 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 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, even though the distribution of the learning sample places more weight on high predicted wages than that of the main sample. This evidence supports our assumption that the learning sample can be used to estimate an allocation rule that is relevant for workers in firms included in the main sample.

Insert Figure 1 here

Within the learning sample, we estimate the conditional probability of each occupation via the Random Forests algorithm (Breiman, 2001).[7] There are three advantages to using Random Forests (RF) in our setting: (i) they are among the best-performing algorithms for classification (Zhang et al., 2017);[8] (ii) they feature few-to-none tuning hyperparameters, dramatically reducing total estimation time;[9] (iii) they easily handle multi-class classification problems and mixed-type characteristics (continuous and categorical), which are relevant in our data.[10]

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

[8] 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).

[9] 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 node in random forest algorithms. Compared with other algorithms, such as neural networks, random forests require fewer parameters to be specified, making them 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.

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

As occupations are not all equally frequent in the sample, we adjust our estimation procedure by forming a balanced subsample via bootstrap, under-sampling more frequent occupations, and using 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 et al. (2008) and Balanced RF in Chen et al. (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.[11] The average F1 score is computed via a stratified 5-fold cross-validation: the learning sample is randomly partitioned into 5 subsamples, where each subsample has the same job frequencies as the initial sample, and the algorithm is trained using 4 subsamples and tested on the remaining one; the procedure is repeated until all of the 5 subsamples are used as a test set, so as to obtain a total of 5 pairs of weighted F1-scores (where each pair refers to a training set and to the corresponding test set); finally, these 5 weighted F1-scores are averaged. The average of the resulting F1 scores is 78% when computed for the training set and 68% when computed for the test set. This performance is reassuring, considering that a random allocation of workers 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 38[12], the maximal weighted F1 score resulting from a random allocation of workers to jobs in our sample would be at most 5.1%.[13]

To characterize our algorithm, we explore the role that worker characteristics play in the es-

[11] 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.

[12] This is because in the estimation we split the full sample in size-industry bins, and the minimal number of unique jobs in a bin is 38.

[13] 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 $s \neq k$.

timated job allocation rule. 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 a box plot of this measure for all the features used in the ML algorithm, which are listed on the horizontal axis.

Insert Figure 2 here

On the whole, Figure 2 highlights that both education and experience play a prominent role in job allocation. However, the type and level of education appear to matter more than tenure, experience, occupation-specific and industry-specific experience, suggesting that generic human capital is more important than firm-specific one in job-worker matching. Yet, the numerous outliers that can be observed for occupation-specific and industry-specific experience indicate that match quality in some industry-size bins is sensitive to experience in some jobs, such as computing professionals, legal professionals, writers, and creative performing artists, as well as metal and mineral products machine operators, building finishers, and office clerks. The same applies to the geographic location of employees: while typically a worker's municipality appears to play little role in determining match quality, it is quite important to match workers who live in a few specific areas, such as Stockholm, Gothenburg, and Malmö.

3.2 Job assignment quality at employee level

To predict the quality of worker-job matches in the main sample, we use the algorithm trained on the learning sample to construct an employee-level measure of job assignment quality (*eJAQ*, where "*e*" is a mnemonic for "employee"). 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 the firm level (*JAQ*), which is simply obtained by averaging *eJAQ* across the employees of the same firm in a given year.

While the *eJAQ* measure has the benefit of simplicity, it has two shortcomings: first, its changes capture changes in match quality only if workers switch to (or away from) their best possible match, thus neglecting any intermediate change in match quality; second, it does not take into account that workers may feature different suitability to the best possible match, depending on the specialization of their skill set: for instance, the probability that a position as computing professional is the top job for an electronic engineer may be, say, 90%, while a worker with a high-school degree may be well suited to several jobs as machine operator with equal probability of 30%. To overcome both limitations, we also construct a continuous measure of employee-level job match quality (*epJAQ*) by estimating the probability that the algorithm assigns to the actual job held by a worker ("*ep*" being a mnemonic for "employee probability"): formally, $epJAQ = P_z(J_i|X_i)$ for worker i. This alternative measure is a gauge of a worker's fit for her actual job compared to other jobs that she might perform, and as such it is able to capture the change in match quality associated with any job switch, as well as differences in the degree of specialization across workers. Indeed, the *epJAQ* measure ranges between zero and a worker-specific maximum, $\bar{p}_i$, defined as the highest predicted probability with which the algorithm assigns worker i to any job: formally, $\bar{p}_i = \max_{j \in \text{Jobs}} P_z(j|X_i)$. This upper bound measure is greater for workers with more specialized skill sets. In what follows, we shall denote the firm-level average of the *epJAQ* measure by *pJAQ*, which will thus be the continuous analog of the dichotomic *JAQ* measure.

3.2.1 Characterizing employee assignment

Table 2 provides evidence that illustrates the mapping from characteristics to jobs estimated by our algorithm. For a pool of selected jobs, the table shows the average characteristics of employees who, according to our ML algorithm, are very likely to be assigned to those jobs. The jobs included in the table are selected to provide information on the following broad occupation classes: 1) all managerial positions (ISCO codes: 121, 122, 123, 131), 2) professionals (242), 3) technicians and clerks (411), 4) skilled manual workers (723), 5) machine operators and assemblers (815), and

6) elementary occupations (932).[14] While the first occupation class includes all the managerial positions, for other occupation classes we focus on the jobs with the highest average predicted probability, $P(j|X)$.

Insert Table 2 here

The typical employee characteristics associated with each job are obtained by averaging over the subset of the 1,000 employees (990 for job 242 and 998 for job 723) with the highest predicted conditional probability for the corresponding job, $P(j|X)$, as estimated by the ML algorithm. The table shows that these are typically senior employees, as shown by the average age (around 50 for most jobs), as well as by their tenure and experience. Moreover, employees tend to have several years of experience in the same job, confirming not only the importance of general experience but also that of job-specific experience, probably due to on-the-job learning. This is consistent with the prominent role of experience and tenure in our ML algorithm according to Figure 2.

Legal professionals (242) are the only exception, as they are younger than other employees and feature lower experience and tenure but have greater and more specialized education: they all hold a college degree in the same subject, clearly a reflection of compulsory requirements to practice legal professions. These special characteristics enable the algorithm to identify these workers with almost perfect confidence (99.8%).

Female workers are a tiny minority in managerial positions (121, 122) and are also unlikely to work as other specialist managers (123), as well as in low-skill jobs (723, 815, 932), while they are almost equally represented as males in legal professions (242) and in small enterprises' managerial positions (131), and account for the totality of office secretaries and data entry operators (411).

Workers holding low-skill jobs (723, 815, and 932) feature a higher average number of unemployment days, likely associated with more frequent separations and/or longer unemployment spells. Moreover, these employees tend to work closer to their birthplace than employees in other jobs. Immigrant status does not seem to vary appreciably across jobs, consistently with its weak

[14] Skilled manual workers include service and shop sales workers, skilled agricultural and fishery workers, and craft and related trade workers. The exact job titles are indicated in the caption to Table 2.

role in the algorithm, already highlighted by Figure 2.

While Table 2 illustrates how individual characteristics of well-matched workers vary across occupations, it is also instructive to ask whether the likelihood of a good match varies across occupation groups. We do so by looking at the frequency of mismatches in the six job classes defined above. The upper panel of Figure 3 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 six job classes defined above. 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 (39%) and a considerably larger value for elementary occupations (65%): in the remaining classes, mismatches range from 56% for managers and 54% for technicians and clerks to 47% for skilled manual workers and 52% for machine operators and assemblers.

Insert Figure 3 here

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 job-worker mismatches may arise more easily than for other occupations. Second, fewer workers hold these jobs, so fewer observations inform their allocation rule. Indeed, elementary occupations account for a small fraction of jobs in the economy (6%), not dissimilar from that of managers (7%), while most workers hold jobs in intermediate classes, as shown in the lower panel of Figure 3. 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 smaller size.

3.2.2 Employee match quality, wages, and separations

We now provide evidence that validates *eJAQ* and *epJAQ* as measures of workers' job assignment quality. First, it is natural to expect that the likelihood of being assigned to a more suitable job increases along workers' careers, as managers learn about employees' characteristics (Fredriksson et al., 2018), and employees 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* and *epJAQ*. Third, separations should be less likely for workers matched to their most suitable job, as found by Fredriksson et al. (2018).

All these predictions find support in our data. Figure 4 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 rises significantly (from 35% to 57%) over a 50-year working life. The largest gain (about 12 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 4 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 columns 1-4 of 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. Panel B shows the estimates of the same specification, simply replacing $eJAQ_{it}$ with $epJAQ_{it}$.

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 0.024: a worker allocated to her most suitable job ($eJAQ_{it} = 1$) is estimated to earn 2.4% more than a mismatched worker with the same characteristics or with the same job ($eJAQ_{it} = 0$). The estimate of β increases slightly upon controlling for 2-digit industry dummies and firm characteristics (column 2), it decreases slightly upon considering only within-worker variation in $eJAQ_{it}$ (column 3), while it increases controlling for unobserved heterogeneity across firms (column 4): the estimated β in a specification that includes worker, jobs and year effects is 2.7% and highly statistically significant. These findings are in line with the -2% estimate of the coefficient of job mismatch in wage growth regressions reported in column 1 of Table 7 in Fredriksson et al. (2018), despite the differences in methodology and sample.

We also analyze the correlation of *epJAQ* with labor earnings in Panel B of Table 3 to provide a robustness check of the results obtained using the *eJAQ* indicator with a continuous measure of workers' suitability to jobs. The estimates shown in Panel B indicate that labor earnings are also positively and significantly correlated with this second measure of job match quality over workers' careers. The 0.05 coefficient estimate in column 1 indicates that a 10-percentage-point increase in a worker's *epJAQ* (amounting to half of its standard deviation) is associated with a 0.5% increase in labor earnings. This effect is qualitatively similar and precisely estimated in the specification with industry fixed effects and firm-level controls (column 2), and in those with worker fixed effects (column 3) and both worker and firm fixed effects (column 4).

Columns 5 through 8 show how the likelihood of a switch to a new employer is related to the two measures of job allocation quality, *eJAQ* and *epJAQ*, in Panel A and B, respectively. Specifically, we estimate a version of equation (1) where the outcome variable is a separation indicator, which equals 1 if worker i changes employer between year t and year $t+1$ and 0 otherwise.

Column 5 in Panel A documents that well-matched workers (i.e., those with $eJAQ_{it} = 1$) are 1.25 percentage points less likely to change employer than mismatched workers (i. e., those with

$eJAQ_{it} = 0$) with the same characteristics. The coefficient drops slightly upon adding firm controls and industry fixed effects (column 6). In the specification of column 7, we exploit only within-worker variation; namely, we ask how much less likely a given worker is to switch to a new employer when she goes from being mismatched to being well-matched: interestingly, in this case, the likelihood of separation drops by 2.7 percentage points. In column 8, we also control for unobserved heterogeneity in turnover rates across firms: in this specification, the likelihood of a separation reverts to being close to 1 percentage point lower for well-matched workers than for mismatched ones. Columns 5 through 8 in Panel B show that a 10-percentage-point increase in *epJAQ* is associated with a reduction in the likelihood of changing firm between 1.7 and 0.5 percentage points.

While Table 3 shows that labor earnings are positively correlated with worker match quality, even controlling for workers' unobserved heterogeneity, it is not informative about whether the relationship is symmetric, namely, whether increases in *eJAQ* are associated with wage rises and decreases in *eJAQ* are associated with wage penalties. We address this question by focusing on instances in which workers either increase or decrease their match quality relative to cases in which they do not: to the extent that *eJAQ* captures the match quality between a worker and her job, one would expect labor earnings to increase upon switching from a mismatched job ($eJAQ = 0$) to the most suitable one ($eJAQ = 1$), and decrease upon a switch in the opposite direction (from $eJAQ = 0$ to $eJAQ = 1$), relative to switches that do not entail changes in *eJAQ*. We perform this analysis for workers who switch across both occupations and employers, because this enables us to condition on homogeneous events.[15] We further require workers to be employed at least for two consecutive years in both the old and new job and estimate two distinct event-study regressions: the first compares the evolution of log labor earnings from two years before to one year after a switch from an occupation with $eJAQ = 0$ to one with $eJAQ = 1$, relative to switches across two occupations with $eJAQ = 0$; the second event-study regression compares switches from occupations with $eJAQ = 1$

[15] As the likelihood that occupation changes are associated with firm switches differs depending on the initial and final level of *eJAQ*, in a pooled sample containing all occupation switches the relationship between wage growth and changes in *eJAQ* would partly reflect the effect of separations, creating an omitted variable bias.

to occupations with $eJAQ = 0$, relative to switches across two occupations with $eJAQ = 1$. Figure 5 shows that in the first case, wages increase by about 3%, while in the second, they drop by about 10%, and in both cases, the change in relative log wages is statistically significant at the 1% level. Hence, when employees switch to a new job and a new employer, wages respond both to rises and drops in match quality, though more strongly to the latter. We find qualitatively similar but quantitatively smaller wage responses when considering only job changes that are not associated with separations.

Insert Figure 5 here

3.3 Job assignment quality at firm level

The next step in the analysis is to average *eJAQ* for all the employees of the same firm: we refer to the resulting firm-level measure as *JAQ*. By the same token, we average the *epJAQ* measure across the employees of each firm, to produce a firm-level continuous metric of job-workers match quality, *pJAQ*. As our approach builds on the assumption that firms differ in their ability to assign workers to jobs, we expect to observe heterogeneity in both of these variables across firms.

Insert Figure 6 here

The top panel of Figure 6 shows the kernel density estimate of firm-level *JAQ* for firms in the main sample and in the learning sample. The bottom panel shows the two corresponding densities for *pJAQ*. As expected, the density of both match quality measures 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 match quality measures 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 allocation rule estimated on the learning sample.

Instead, *pJAQ* is more concentrated around low values for the main sample than it is around high values for the learning sample. Again, this is for two reasons. First, firms in the main sample may assign workers to less suitable jobs because they deviate more often from the allocation rule estimated on the learning sample. Second, being based on the learning sample, the algorithm tends to predict jobs' conditional probabilities with lower confidence in the main sample, lowering the probability assigned to the most suitable job. Indeed, it turns out that on average, the algorithm places a 27% probability on the most suitable job for employees of main-sample firms, against 52% for employees of learning-sample firms.

To further corroborate the idea that firms differ systematically in their ability to assign workers to the correct job, we also investigate how changes in labor earnings and in *eJAQ* correlate with switches across firms with different *JAQ*. Indeed, one would expect that the individual worker-job match quality and (consequently) wages increase upon switching from a firm with low *JAQ* to one with high *JAQ* and vice-versa. Following Card et al. (2013), we restrict our sample to employer switchers with at least two consecutive years in both the origin and destination firm; next, we classify these firms by quartiles of the distribution of average *eJAQ* of current coworkers and partition the switches in cells based on the possible combinations of origin and destination quartiles. Finally, we compare changes of log labor earnings and *eJAQ* occurring around switches from firms in the bottom quartile of match quality to those in the top quartile, and viceversa: in the first case, both wages and *eJAQ* increase significantly relative to workers that remain in the same quartile, while in the second case wages do not change significantly and *eJAQ* drops (see Figure A1 in the Appendix). Hence, the ability of firms to allocate workers to jobs matters for both their wages and for their individual match quality.

3.4 What accounts for changes in job assignment quality?

In principle, the quality of the match between workers and firms can change for one of four reasons: employees (i) accumulate experience or are retrained to improve their fit with their existing jobs within the same firm; (ii) switch to more suitable positions within the same firm; (iii) switch to a new firm with which they have a better fit, but where they perform the same job; (iv) switch to a new task in a new firm, in which case the match quality improvement may stem from the change in the task and/or the employer. To account for the relative importance of these four sources of change in job allocation quality in our data, we focus on the instances in which a mismatched employee at year t ($eJAQ_{i,t} = 0$) becomes correctly allocated in the subsequent year $t+1$ ($eJAQ_{i,t+1} = 1$), or the opposite occurs. Figure 7 shows the frequency of these events in the main sample.

Insert Figure 7 here

Most changes in *eJAQ* (and thus in *JAQ*) appear to stem from employees improving their match from one year to the next, while retaining both their position and employer. Learning by doing is thus likely to be the prevalent reason behind improvements in match quality. Instead, employees changing jobs (either within the same firm or upon switching to a new one) account for the majority of cases in which we observe a deterioration in the match quality. However, these year-by-year changes may be transitory, so that even if changing jobs initially yields a poorer match, the match quality may subsequently recover: Section 5 shows that this is indeed the case, especially when the job change coincides with the replacement of a previous manager by a better one. Interestingly, the vast majority of the variation in *eJAQ* is not due to hiring (employees switching employers from one year to the next) but is accounted for by employees remaining with their current employer.

This is also confirmed by Figure 8, which breaks down the variance of yearly changes in the firm-level *JAQ* (denoted by ΔJAQ), into six components: the variance of changes due to employees retaining their initial job within the same firm; the variance of changes due to employees who switch to different jobs within the same firm; the variance of changes due to employee turnover; and the three corresponding covariances. The total variance of ΔJAQ is shown as the light gray background area in the figure, while the dark gray bars correspond to its components. The figure shows that most of the changes in *JAQ* over time stem from workers who retain their initial job and improve their match quality: these employees account for 48% of the total variance. However, sizable portions of the variance of changes in *JAQ* stem from workers' turnover (38%) and from job switches within the initial firm (19%).

Insert Figure 8 here

3.5 How does job assignment quality vary across firms?

The quality of management practices–defined as managers' ability to monitor performance, set targets, and incentivize employees–has been shown to be consistently higher in firms facing harsher product market competition, those run by non-family managers, and those with a better-educated workforce (Bloom and Van Reenen, 2007, 2010). These correlations have been respectively interpreted as reflecting the selection and incentive effects of competition, the inefficiencies stemming from dynastic succession in control, and the ability of better-managed firms to attract more skilled employees.

It is reasonable to expect similar correlations between these characteristics and the measures of job allocation quality presented in Section 3.3: product market competition should focus managers' attention on matching employees to the most suitable jobs; family management is more likely to promote family and friends irrespective of merit, compared to non-family management; finally, more educated workers may seek jobs in firms where they can expect to be correctly assigned, especially given the evidence in Section 3.2 that better matches are associated with higher earnings. The estimates in Table 4 are consistent with all three predictions.

Insert Table 4 here

The first four columns of the table present regressions of measures of job allocation quality on the Lerner index of market competition. The dependent variable is *JAQ* in columns 1 and 2 and *pJAQ* in columns 3 and 4. The Lerner index of competition is defined for each firm as

$1-$ profits/sales, lagged by two years to remove any potential contemporaneous feedback, and averaged across all firms in the same 2-digit industry excluding the firm itself. All specifications include year and industry dummies (where industries are manufacturing, real estate, renting and business activities, and wholesale and retail). The specifications of the even-numbered columns include the following additional controls: the share of employees with a college degree, log employment, log capital, log firm age, indicator for listed firms, and years of managerial experience averaged over employees in the firm. In all the specifications, firms operating in more competitive markets turn out to allocate their employees more closely in line with the estimated allocation rule, according to both our measures.

The last four columns of the table present regressions of measures of job allocation quality on a family firm dummy, which is based on family relations among major shareholders (called "owners" by tax authorities) and directors.[16] For each owner and director in a firm, we calculate the number of other family members who are directors or owners in the company. A company is a family firm if at least two family members are owners or board members, or at least one owner and one director come from the same family. The estimates show that family firms feature significantly lower job allocation quality in most specifications: based on the estimates in columns 5 and 6, in family firms, the probability that an employee is matched to his/her most suitable job is between 0.8 and 2.6 percentage points lower than in non-family ones. This finding complements existing evidence that family firms are more poorly managed than non-family firms (Bloom and Van Reenen, 2007, 2010; Bandiera et al., 2018; Lemos and Scur, 2019).

Finally, the coefficient of the share of employees with a college degree is positive and significantly different from zero in all the specifications where this variable is included among the explanatory variables: a 10-percent increase in this measure of employees' human capital is associated with about a 1 percentage point increase in the probability of a suitable job assignment. This correlation can either be seen as suggesting that better job-worker matching attracts more qualified employees or as indicating that managers pay more attention to the job assignment of employees

[16] An individual's family comprises his parents, grandparents, children, grandchildren, siblings, and partner(s). A partner is the person with whom the individual has a child.

with a college degree, or both.

4 Job Assignment Quality and Firm Performance

This section explores how the heterogeneity in *JAQ* and *pJAQ* 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 these two measures capture meaningful variation in the quality of workforce allocation, rather than just statistical noise or firm heterogeneity in 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.

4.1 Descriptive evidence

Figure 9 shows that firm-level productivity correlates positively with both *JAQ* and *pJAQ*. The figure shows partial regression plots of value added per employee against these two job-worker match quality variables, conditioning on year effects and 2-digit industry effects. The two top panels refer to main-sample firms, while the bottom two to learning-sample ones. The left-side panels of the figure show how value added per employee correlates with *JAQ*, and the right-side ones show how it correlates with *pJAQ*. 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.

Insert Figure 9 here

Specifically, the positive correlation with *JAQ* indicates that firms where workers are more often allocated to their most suitable job are more productive than others. Moreover, the positive correlation with *pJAQ* suggests that the specialization of a firm's workforce also plays a role. As explained in Section 3.2, a firm's employees may feature high *epJAQ* not only if they are well assigned within the set of jobs they can possibly hold (p_i close to $\bar{p}_i$), but also if they are highly specialized ($\bar{p}_i$ close to 1), in the sense that their characteristics make them highly suitable to a specific job profile.[17]

The two lower panels of Figure 9 instead show that no correlation between productivity and either *JAQ* or *pJAQ* is present for firms in the learning sample. This is to be expected: for these firms, variation in measures of match quality 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.

4.2 Regression analysis

Table 5 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: year dummies control for aggregate movements in productivity, while municipality ones control for productivity differentials across locations. The latter may arise not only from location-related technological advantage but also 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 5, column 1 reports the OLS estimates of a regression of log sales per employee on *JAQ* that only includes year dummies. We find a highly significant coefficient of 0.366, implying that a 10-percentage-point increase in *JAQ* is associated with a 3.66% increase in

[17] Indeed, we find that firms' productivity is positively and significantly correlated with firm-average $\bar{p}_i$, which can be regarded as a measure of specialization of its workforce, even after controlling for workers' characteristics.

sales per employee. Equivalently, a one-standard-deviation increase in *JAQ* (0.31) is associated with an 11.5% 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 stem from their different scope: *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. The difference may also be due to the most productive firms being excluded from our sample by construction.

Insert Table 5 here

In column 2 of Table 5, the dependent variable is the log of 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 1.95%, implying a 6% increase upon a one-standard-deviation increase in *JAQ*.

These results are not robust to the inclusion of firm fixed effects, possibly because of attenuation bias due to measurement error. However, we control for various possible sources of omitted variable bias, namely, firm characteristics, differences in firms' occupation structures, and differences in workers' quality across firms.

First, the correlation between productivity and our measures of job allocation quality is robust to the inclusion of 2-digit industry indicators, log number of employees, log capital, and the fraction of employees with at least a college degree among the regressors, as shown by the estimates in columns 4 and 5 of the table. The estimated coefficients of *JAQ* in columns 4 and 5 drop in magnitude, but remain positive and significantly different from zero.

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 differently: 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 address this issue, Panel B of Table 5 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 possible source of omitted variable bias is that firms with higher *JAQ* may feature higher-quality workers, irrespective of the job they are allocated to. Thus, in Panel C of Table 5, 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.[18]

In almost all of the specifications shown in Table 5, profitability, as measured by operating return on assets (OROA), is not significantly correlated with our measure of efficient job allocation, as shown in columns 3 and 6 of the table. A possible interpretation of this finding is that in Swedish

[18]The results reported in Table 5 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 9.

firms, the productivity gains afforded by better job-worker matches in 2001-10 translated mostly into higher wages rather than increases in firm profitability.

To check the robustness of these results, in Table 6, we repeat the estimation of the specifications shown in Table 5 upon measuring worker-job match quality by *pJAQ*. The estimated coefficient of this variable is positive and significantly different from zero in all the specifications of the productivity regressions. The baseline estimates shown in columns 1 and 2 of Panel A imply that a 10-percentage-point increase in firm-level suitability of workers to jobs is associated with a 10-percentage-point increase in log sales per employee and a 7-percentage-point increase in value added per employee. These results are qualitatively robust to the addition of other controls, even though they drop considerably in size. Differently from *JAQ*, columns 3 and 6 show that increases in *pJAQ* also weakly correlate with increases in profitability.

Insert Table 6 here

Finally, we note that *JAQ* is a combination of two components: allocation quality, i.e., whether firms allocate their current workers to the best possible jobs, and hiring quality, i.e., whether firms manage to hire the most suitable workers. The second of these two components may, to some extent, depend on firm productivity: for instance, as firms with more productive technologies may pay higher wages, they can attract higher-quality workers. Hence, it is worth investigating whether the first component alone correlates with productivity, by shutting down the hiring quality component. Panel C of Table 5 already addresses this point, as it controls for the average quality of the workers present in each firm and shows that firm productivity is still robustly correlated with *JAQ*. An additional exercise to address this question is to rely on an adjusted measure of job match quality, labeled *aJAQ*, where *a* is mnemonic for "adjusted": this measure is obtained by rescaling *JAQ* by the highest level that it can achieve in a given firm, based on its existing workers and occupations, i.e., $aJAQ \equiv JAQ/maxJAQ$. The scaling factor *maxJAQ* is obtained as follows: if firm f has K occupations, N_j positions in each occupation $j = \{1,\ldots,K\}$, and M_j employees for whom j is the most suitable job, then the firm can have at most $\min\{M_j,N_j\}$ employees with

$eJAQ = 1$. Hence the highest achievable JAQ for firm f is:

$$maxJAQ = \frac{\sum_{j=1}^{K} \min\{M_j, N_j\}}{\sum_{j=1}^{K} N_j}.$$

Hence, $aJAQ$ is meant to measure only job allocation quality within the firm, given its current workers and jobs. The resulting estimates, shown in Table A1, indicate that firm productivity is positively and significantly correlated with $aJAQ$ in almost all specifications, and also firm profitability is significantly correlated with this measure of job match quality. Hence, differences in firms' performance are associated with their ability to allocate workers to the most suitable positions available within the firm, not just with their ability to attract good hires.[19]

4.3 A circularity issue?

A possible concern with our methodology is the potential circularity in the construction and validation of JAQ and $pJAQ$ performed up to this point: 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 measures thus obtained correlate with firms' productivity. The obvious counter to this criticism is that the correlation between these measures 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.

In this section, we further dispel circularity concerns by constructing learning sets that do not rely on firm productivity and using them to retrain the ML algorithm. First, we perform a placebo test replacing actual firm productivity with a noise variable with the same distribution as firm productivity. Second, we retrain the algorithm on random subsamples of firms. Third, we construct the learning set based on the residuals of an AKM model (Abowd et al., 1999).

[19] We also perform an alternative exercise, redefining $eJAQ$ based on whether a worker is assigned to the most suitable job considering only those available within her/his firm, and then aggregating this measure at the firm level. This method yields results that are very close to those obtained using JAQ in Table 5, and are not reported for brevity.

Placebo test. The selection of firms into the learning sample based on productivity may create a mechanical correlation between our measures of job allocation quality and firm productivity: the out-of-sample prediction error may be systematically lower for firms that are more similar to those included in the learning set, driving a positive correlation between *JAQ* and productivity. To address this issue, we perform a placebo test: we replace firms' actual productivity (the log of 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, and is used to compute the *JAQ* and *pJAQ* measures again, as explained in Subsection 3.1. By construction, the new productivity variable used for the placebo test has the same distribution as the original one but is independent of the rest of the data.

Note that this placebo test leaves intact the relationship between employees' characteristics and task allocations and only alters the selection of firms in the learning sample. If the positive relationship between our measures of job assignment quality and firm productivity is indeed generated by this selection process, one should also expect to find a positive and significant correlation between them 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 be present. This is indeed what emerges from the top panels of Figure 10, which plot the results of the regression of the placebo productivity measure on *JAQ* (left-hand chart) and on *pJAQ* (right-hand chart). This lack of correlation contrasts with the positive and significant correlation obtained in the main estimation strategy and shown in the top panel of Figure 9.

Insert Figure 10 here

Random JAQ. In the regressions of Table 5 and Table 6, *JAQ* and *pJAQ* may spuriously correlate with productivity. Suppose, for instance, that the most productive firms are all located in large cities and that they only hire residents from those cities. Then, estimating the job allocation rule on the learn sample would generate a *JAQ* measure positively associated with hiring residents of large cities, creating a spurious correlation between *JAQ* and productivity in the main sample. To address this additional concern, we re-train our ML algorithm on a random subsample of firms to calibrate the reference rule based on the average firm in our sample rather than top-productivity firms and investigate whether the resulting measures of job assignment quality still correlate significantly with productivity across firms. Specifically, we redefine the learning sample used to train our ML algorithm as a 10% random draw of firms (in the same size-industry class) from our entire sample. We refer to the resulting measures of job allocation quality as JAQ^R and $pJAQ^R$, where the superscript R is a mnemonic for "random". Both JAQ^R and $pJAQ^R$ turn out to correlate positively and significantly with productivity, as shown by the two partial regression plots of the log of value added per employee (controlling for year and industry effects) in the bottom panel of Figure 10. The regressions shown in Table A2 in the Appendix show that these correlations are robust to the inclusion of firm and worker controls.

It is worth comparing the graphs shown in the two bottom panels of Figure 10 with the corresponding graphs in the top panel of Figure 9, which are based on the measures of job allocation quality calibrated on top-productivity firms. In the bottom panel of Figure 10 the positive correlation appears to be present especially for firms in the bottom and middle portion of the productivity distribution, rather than for the entire support of the distribution as in the top panel of Figure 9. As a result, the relationship between the two variables has an inverse-U shape. This is precisely as expected: since now the rule reflects the behavior of the average firm, the firms that adhere most closely to this rule (i.e., those with the highest value of JAQ^R and $pJAQ^R$) cannot hope to achieve more than an average productivity level. Still, conforming more closely to such an allocation rule is associated with productivity improvements for the typical firm in the sample, because it reduces firms' deviations from the estimated rule in allocating their employees. This applies particularly to firms at the bottom of the distribution, i.e., those that conform the least to the estimated rule: for such firms, an increase in JAQ^R and $pJAQ^R$ is associated with a steep productivity gain.

The results for JAQ^R and $pJAQ^R$ in the bottom panel of Figure 10 are not tied to a particular random draw of firms: they hold on average for any random draw of firms, as shown by the

Monte Carlo experiment illustrated in Figure 11. We sample 150 learn sets at random, calibrate the JAQ^R and $pJAQ^R$ measures for each of them, and estimate their coefficient point estimates in the productivity regressions, controlling for year and industry fixed effects.[20] Figure 11 shows the Monte Carlo distribution of these estimates, where the solid vertical lines indicate the analogous point estimate obtained using the 10% top-productivity firms as learning sample.

Insert Figure 11 here

All the densities place higher mass on positive values, indicating that *JAQ* measures positively correlate with productivity when the learn set is selected at random. However, the coefficient estimates obtained for the 10% top-productivity firm are significantly higher (solid vertical line) than the average estimate obtained using these random draws, as shown by the p-value reported in the top-right of each graph. This is perfectly in line with the aforementioned intuition that conforming to average firms typically leads to a productivity increase, but not as large as the increase obtainable by conforming to the rule used by top productivity firms.

AKM-based learning set. Yet another robustness check of our results is to produce an alternative measure of JAQ defining the learn set not based on firm productivity, but on the residuals of an AKM wage regression, which capture the surplus from worker-firm matches (Lachowska et al., 2018; Kugler et al., 2020). As long as this surplus at least partly arises from the fit between workers' skills and their occupations, observations with the highest AKM residuals should identify better-allocated workers and thus provide a valid learning sample. To implement this strategy, we (i) estimate an AKM model including worker fixed effects, firm fixed effects, year fixed effects, and workers' characteristics, (ii) define the learn set as the sample of workers featuring the top 10% residuals in 2010, (iii) train the ML algorithm on this sample, and (iv) finally correlate firm productivity with the resulting alternative measures of job allocation quality (JAQ^{AKM} and $pJAQ^{\text{AKM}}$), excluding all the firms that employ any of the workers belonging to the learn set. The

[20] Due to limitations in computing power, the procedure described above is carried out on a 15% random subsample of firms, drawn independently at each Monte Carlo iteration.

results, shown in Table 7, parallel very closely those reported in Panel A of Table 5: the coefficient estimate reported in column 2 of Panel A implies that a one-standard-deviation increase in JAQ^{AKM} (0.30) is associated with a 3.8% increase in value added per employee.

Insert Table 7 here

5 Impact of Manager Quality 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 the ML algorithm would actually assign its managers to managerial tasks in the most productive firms. Another, simpler measure of the quality of a firm's managers is their average work experience in managerial positions.

Hence, to investigate this issue, for each firm and each 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 jobs (*M-JAQ*). The first is the average *eJAQ* for all workers holding 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 8 presents the results of the corresponding regressions, which are based on data from 2003 to 2010: data for 2001 and 2002 are omitted 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 it only refers 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 relevant coefficient is positive and significantly different from zero in both regressions: a 10-percentage-point increase in the quality of managers' allocation is associated with an increase in the quality of rank-and-file workers' allocation ranging between 1 and 2 percentage points, depending on the specification. When the quality of managers' allocation only refers to the firm's top management, the coefficient approximately halves in size, indicating that middle management is also important for the correct allocation of workers to their jobs.

Insert Table 8 here

The specifications shown in columns 2 and 5 also include firm fixed effects and the average experience of the firm's 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, log number of employees and 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 large and precisely estimated. Importantly, these regressions are based only on within-firm variation in the relevant variables: they indicate that better matching of rank-and-file employees to jobs tends to occur when the firm improves its management's quality and experience.

These findings beg the question of whether managerial quality and expertise, by improving the matching of workers to jobs, contribute to the firm-level productivity differentials analyzed in Section 4. Table 9 shows that indeed this is the case: both the log of value added per employee and the log of sales per employee are positively and significantly correlated with the quality of managers' allocation, irrespective of whether the specification only contains year and municipality dummies (columns 1 and 2) or also includes industry dummies (columns 3 and 4). The specifications in columns 5 and 6 suggest that not only the allocation quality of managers but also their average experience contributes to firms' productivity. Hence, the evidence is consistent with the view that managerial quality and experience, via their effects on the matching of rank-and-file workers to jobs, account for the observed productivity differentials between firms. In terms of economic significance, higher managerial quality correlates with firm productivity almost as strongly as overall labor allocation quality: one-standard-deviation increases in *M-JAQ* and *JAQ* are respectively associated with a 4.8% and a 6% rise in the log of value added per employee, based on the coefficient estimates in column 2 of Table 9 and in the corresponding column of panel A of Table 5.[21]

Insert Table 9 here

Since it is natural to expect 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 when incumbent managers are replaced with more suitable ones, and worsens when they are 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.

The year-to-year change in managerial quality, denoted by $\Delta M\text{-}JAQ_\tau$, is the difference between the average quality of the new management team, $M\text{-}JAQ_\tau$, and that of the firm's previous management team $M\text{-}JAQ_{\tau-1}$, which in turn reflects both the average quality of the $N^s_{\tau-1}$ managers who "stay" in the firm and that of the $N^d_{\tau-1}$ managers who are "dismissed" in year τ:

$$\begin{aligned}\Delta M\text{-}JAQ_\tau \equiv M\text{-}JAQ_\tau - M\text{-}JAQ_{\tau-1} &= M\text{-}JAQ_\tau - \frac{\sum_{i=1}^{N^s_\tau} eJAQ^s_{i\tau-1} + \sum_{j=1}^{N^d_{\tau-1}} eJAQ^d_{j\tau-1}}{N_{\tau-1}}\\ &= \underbrace{M\text{-}JAQ_\tau - \frac{\sum_{i=1}^{N^s_\tau} eJAQ^s_{i\tau} + \sum_{j=1}^{N^d_{\tau-1}} eJAQ^d_{j\tau-1}}{N_{\tau-1}}}_{\Delta M\text{-}JAQ^{TO}} + \frac{\sum_{i=1}^{N^s_\tau}(eJAQ^s_{i\tau} - eJAQ^s_{i\tau-1})}{N^s_{\tau-1}}\frac{N^s_{\tau-1}}{N_{\tau-1}}. \quad (3)\end{aligned}$$

[21] The standard deviations of *M-JAQ* and *JAQ* are respectively 0.36 and 0.31. The rank-and-file allocation quality also features a similar correlation with firm productivity: replacing *M-JAQ* with *R&F-JAQ* in the specification of column 2 of Table 9, we find that its estimated coefficient is 0.174 (significant at 1% level), implying that a one-standard-deviation increase (0.32) in *R&F-JAQ* is associated with a 5.5% increase in the log of value added per employee.

To identify in this expression the change in managerial quality related to turnover, i.e., to external hiring or firing, one must shut down the part of the variation in $\Delta M\text{-}JAQ_\tau$ induced by internal promotions or demotions, and by changes in job allocation quality of incumbent managers. This is done in the last step of equation (3), by adding and subtracting the term $\sum_{i=1}^{N_\tau^s} eJAQ_{i\tau}^s/N_{\tau-1}$, i.e., the time-$\tau$ average job assignment quality of employees who stay in the firm with a managerial position, including employees who are internally promoted to such positions. Then in the first term, $\Delta M\text{-}JAQ^{TO}$, the change in the quality of the managerial team is purged of that contributed by the "stayers", and thus measures the change in managerial quality only due to turnover. Instead, the last term measures the change in the average quality of retained managers between years τ and $\tau-1$, including that stemming from any demotions of former managers: this is the portion of the change in the firm's managerial quality that does not arise from turnover. Indeed, $\Delta M\text{-}JAQ^{TO}$ is zero by construction if no managers are dismissed ($N_{\tau-1}^d = 0$) and no managers are hired from outside the firm.

We then define a "positive turnover event" to occur for a given firm in year τ if in that year $\Delta M\text{-}JAQ^{TO}$ in (3) turns positive for the first time for that firm, and this rise in managerial quality persists over time, i.e., is not subsequently reversed or more than reversed. Symmetrically, a "negative turnover event" occurs in year τ if in that year $\Delta M\text{-}JAQ^{TO}$ 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, 1,451 firms (19.9% of the total) experienced positive turnover events, 1,668 (22.9%) experienced negative ones, and the remaining 4,173 (57.2%) experienced 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 et al. (2021), Sun and Abraham (2021), Athey and Imbens (2022), and Baker et al. (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 12 shows the estimated dynamic treatment effects on rank-and-file workers around managerial turnover events, respectively associated with an increase (left chart) or a decrease (right chart) in the *JAQ* of the relevant firm's management.

Insert Figure 12 here

The chart on the left shows that the 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 is followed by a significant improvement over the subsequent five years, starting at about 3 percentage points at the time of the event, and eventually vanishing. Conversely, the chart on the right indicates that the replacement of incumbent managers with worse ones tends to occur in firms starting from a normal level of rank-and-file workers allocation quality, but are associated with a strong and persistent deterioration in the allocation of rank-and-file workers, especially in the first three years.[22] Overall, this evidence suggests that persistent changes in managerial quality are an important driver of changes in workers' allocation and, therefore, of organizational change within firms.[23] To address causality, we repeat the estimation using only the 350 managerial turnover events associated with the death of the incumbent management, and find qualitatively similar results, although the effects are barely statistically significant due to the paucity of observations (see Figure A2 in the Appendix). On the whole, this evidence rhymes well with the finding in Bender et al. (2018) that firms with better management have workers with higher human capital.

[22] We also repeat the estimation restricting managerial turnover events to top executives only, and investigate their impact separately on the match quality of middle management and on that of rank-and-file workers. We find that only negative changes in the top managers' quality are significantly associated with subsequent decreases in the average match quality of middle managers and rank-and-file workers. Hence, the data suggest that decreases in the match quality of top managers due to external turnover events "trickle down" to both middle management and rank-and-file workers (see Figure A3).

[23] We also find limited evidence that increases in managerial quality lead to higher productivity if one focuses on positive turnover events triggering a large impact effect on the job allocation quality of rank-and-file workers (i.e., an increase in *R&F-JAQ* above the 75^{th} percentile). The corresponding results are shown in Figure A4 of the Appendix.

In principle, the organizational changes brought about by new management may involve the reallocation of existing employees to different tasks, changes to the composition of the firm's workforce via new hires and/or dismissals, or both of these. Moreover, the extent to which the new managerial team relies on each of these two strategies may differ depending on whether it is better or worse than the preexisting one, according to our metric. To investigate these points, in Table 10 we present the Callaway-Sant'Anna estimates of the impact effect (i.e., the time-0 parameter) of the positive and negative managerial turnover events on several outcome variables.

Insert Table 10 here

The first three rows of the table illustrate the effects that managerial turnover events have on the frequency of employees' internal reallocation, firing and hiring. Upon the arrival of better management, firms feature a significant drop in the internal reshuffling of employees across jobs, a marked decrease in the frequency of hirings, and an increase in the frequency of firings. In contrast, a worse incoming management team engages in a significantly larger reallocation of existing workers and into substantial firings, while its hiring activity is negligible. On balance, however, the more moderate activism displayed by the new managerial team in the first case appears to be much more beneficial to the job allocation quality of rank-and-file workers than its greater activism in the second case: on impact, $R\&F\text{-}JAQ$ increases by 3% upon a positive managerial turnover event, while it drops by 9% upon a negative one. Most of the decrease in job allocation quality triggered by negative managerial turnover events is accounted for by a drop in the job allocation quality of workers who "stay" in the same firm, i.e., $R\&F\text{-}JAQ^s$.

Next, we consider the impact of managerial hirings and firings on the change of job allocation quality of rank-and-file employees relative to the previous year, i.e., $\Delta R\&F\text{-}JAQ$, rather than on its level. It turns out that positive managerial turnover events trigger a statistically significant 8.4 percentage points increase in the change of match quality of rank-and-file workers relative to the previous year, while the corresponding outcome for negative managerial events is close to zero and insignificant.

Finally, we partition this overall change into a component associated with turnover (hires and fires) and one reflecting the change in allocation quality of retained employees, analogously to equation (3) for managers' allocation quality.[24] Relying on notation similar to that used in equation (3), we denote by $\Delta R\&F\text{-}JAQ^{TO}$ the change in the allocation quality of rank-and-file workers stemming from employee turnover, and by $\Delta R\&F\text{-}JAQ^{s}$ the change in the job allocation quality of "stayers". The figures in the last two rows of the table show that positive managerial turnover events raise the change in match quality of the relevant firms' workforce both via the turnover of employees and via the change in match quality of existing employees. The result that a better managerial team raises the change in the match quality of existing employees is not in contrast with the above-noted fact that such a team reduces internal reallocation, as the match quality of "stayers" may simply improve as a result of the increased job experience of employees that do not switch jobs. In contrast, negative managerial turnover events do not affect the change in match quality of rank-and-file workers via the turnover channel, but trigger a strong and significant decrease both in the level and in the change of match quality of "stayers": hence, they appear to engage in misguided reshuffling of employees. One possible explanation for the asymmetric impact of positive and negative events is that achieving a good job allocation requires time and effort, whereas it can be easily disrupted. This intuition is in line with our finding in Section 3.4 that most year-to-year improvements in workers' allocation quality accrue from stayers retaining their jobs, while most of their drops stem from misguided reshuffling or hiring and firing.

6 Conclusions

This paper proposes a novel measure of job-worker allocation quality (*JAQ*) by combining employer-employee data with machine learning techniques and validates it 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

[24]However, this decomposition is not as precise as that shown in equation (3), because the allocation quality of employees who "stay" in the firm is not re-weighted by their fraction over the total workforce.

over their careers, and that workers better matched to their jobs are less likely to switch to a new employer. Job allocation quality is found to vary systematically across firms: companies that operate in more competitive markets, those that are not family-managed and those with a more educated workforce do a better job at matching their employees to jobs. Most importantly, firm productivity correlates robustly with our measure of job-worker allocation quality.

Hence, our measure correlates with key firm characteristics in the same way as management practices do, uncovering a hitherto unmeasured dimension of management's ability. Indeed, we find that the quality of management plays a key role in the efficient assignment of workers to jobs: rank-and-file workers' allocation improves significantly when managerial hirings and firings lead to a better-assigned and more experienced management team, and deteriorates significantly in the opposite case.

The measure proposed in this paper can be constructed for any 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 a vast range of research questions in organizational economics and corporate finance, as witnessed by the evidence provided by this paper.

References

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

Abowd, J. M., Kramarz, F., Margolis, D. N., 1999. High Wage Workers and High Wage Firms. Econometrica 67(2), 251–333.

Adenbaum, J., 2023. Endogenous Firm Structure and Worker Specialization. Technical report, University of Edinburgh.

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

Athey, S., Imbens, G. W., 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., Larcker, D. F., Wang, C. C., 2022. How Much Should We Trust Staggered Difference-in-Differences Estimates? Journal of Financial Economics 144(2), 370–395.

Baley, I., Figueiredo, A., Ulbricht, R., 2022. Mismatch Cycles. Journal of Political Economy 130(11), 2943–2984.

Bandiera, O., Lemos, R., Prat, A., Sadun, R., 2018. Managing the Family Firm: Evidence from CEOs at Work. The Review of Financial Studies 31(5), 1605–1653.

Bandiera, O., Prat, A., Hansen, S., Sadun, R., 2020. CEO Behavior and Firm Performance. Journal of Political Economy 128(4), 1325–1369.

Becker, G. S., 1957. The economics of discrimination. Chicago: University of Chicago Press.

Bender, S., Bloom, N., Card, D., Van Reenen, J., Wolter, S., 2018. Management Practices, Workforce Selection, and Productivity. Journal of Labor Economics 36(S1), S371–S409.

Bennedsen, M., Nielsen, K. M., Lez, F. P. R.-G., Wolfenzon, D., 2007. Inside the Family Firm: The Role of Families in Succession Decisions and Performance. Quarterly Journal of Economics 122(2), 647–691.

Bennedsen, M., Pérez-González, F., Wolfenzon, D., 2020. Do CEOs Matter? Evidence from Hospitalization Events. The Journal of Finance 75(4), 1877–1911.

Benson, A., Li, D., Shue, K., 2019. Promotions and the Peter Principle. The Quarterly Journal of Economics 134(4), 2085–2134.

Bertrand, M., Schoar, A., 2003. Managing with Style: The Effect of Managers on Firm Policies. The Quarterly Journal of Economics 118(4), 1169–1208.

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

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., Roberts, J., 2013. Does Management Matter? Evidence from India. The Quarterly Journal of Economics 128(1), 1–51.

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

Bloom, N., Van Reenen, J., 2010. Why do management practices differ across firms and countries? Journal of Economic Perspectives 24(1), 203–24.

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

Bowlus, A. J., 1995. Matching Workers and Jobs: Cyclical Fluctuations in Match Quality. Journal of Labor Economics 13(2), 335–350.

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

Cahuc, P., Postel-Vinay, F., Robin, J.-M., 2006. Wage Bargaining with On-the-Job Search: Theory and Evidence. Econometrica 74(2), 323–364.

Callaway, B., Sant'Anna, P. H., 2021. Difference-in-differences with multiple time periods. Journal of Econometrics 225(2), 200–230.

Card, D., Heining, J., Kline, P., 2013. Workplace Heterogeneity and the Rise of West German Wage Inequality*. The Quarterly Journal of Economics 128(3), 967–1015.

Chen, C., Liaw, A., Breiman, L., 2004. Using random forest to learn imbalanced data. University of California at Berkeley Working Paper No. 666.

Chiappori, P.-A., Salanié, B., 2016. The Econometrics of Matching Models. Journal of Economic Literature 54(3), 832–861.

Cornwell, C., Schmutte, I. M., Scur, D., 2021. Building a Productive Workforce: The Role of Structured Management Practices. Management Science 67(12), 7308–7321.

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

Deming, D., Kahn, L. B., 2018. Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals. Journal of Labor Economics 36(S1), S337–S369.

Eeckhout, J., Kircher, P., 2018. Assortative Matching With Large Firms. Econometrica 86(1), 85–132.

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

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

Fox, J. T., Smeets, V., 2011. Does Input Quality Drive Measured Differences in Firm Productivity? International Economic Review 52(4), 961–989.

Fredriksson, P., Hensvik, L., Skans, O. N., 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.

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

Ichniowski, C., Shaw, K., Prennushi, G., 1997. The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. The American Economic Review 87(3), 291–313.

Jovanovic, B., 1979. Job Matching and the Theory of Turnover. Journal of Political Economy 87(5), 972–990.

Kaplan, S. N., Klebanov, M. M., Sorensen, M., 2012. Which CEO Characteristics and Abilities Matter? The Journal of Finance 67(3), 973–1007.

Keloharju, M., Knüpfer, S., Tåg, J., 2023. CEO health. The Leadership Quarterly 34(3), 101672.

Kugler, A. D., Muratori, U., Farooq, A., 2020. Do unemployment insurance benefits improve match and employer quality? evidence from recent u.s. recessions. NBER Working Peper No. 27574.

Lachowska, M., Mas, A., Woodbury, S. A., 2018. Sources of Displaced Workers' Long-Term Earnings Losses. American Economic Review 110(10), 3231—-66.

Lazear, E. P., Shaw, K. L., 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.

Lazear, E. P., Shaw, K. L., Stanton, C. T., 2015. The Value of Bosses. Journal of Labor Economics 33(4), 823–861.

Lemos, R., Scur, D., 2019. The ties that bind: implicit contracts and management practices in family-run firms. CEPR Discussion Paper No. DP13794.

Li, D., Raymond, L., Bergman, P., 2020. Hiring as exploration. NBER Working Paper 27736.

Li, K., Mai, F., Shen, R., Yan, X., 2021. Measuring Corporate Culture Using Machine Learning. The Review of Financial Studies 34(7), 3265–3315.

Lindenlaub, I., 2017. Sorting Multidimensional Types: Theory and Application. The Review of Economic Studies .

Lippi, F., Schivardi, F., 2014. Corporate control and executive selection. Quantitative Economics 5(2), 417–456.

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

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

Malmendier, U., Tate, G., 2005. CEO Overconfidence and Corporate Investment. The Journal of Finance 60(6), 2661–2700.

Malmendier, U., Tate, G., 2009. Superstar CEOs. Quarterly Journal of Economics 124(4), 1593–1638.

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.

Minni, V., 2023. Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs .

Moen, E. R., 1997. Competitive Search Equilibrium. Journal of Political Economy 105(2), 385–411.

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

Mullins, W., Schoar, A., 2016. How do CEOs see their roles? Management philosophies and styles in family and non-family firms. Journal of Financial Economics 119(1), 24–43.

Ocampo, S., 2022. A Task Based Theory of Occupations. Technical report, University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers.

Olsson, M., Tåg, J., 2024. What is the Cost of Privatization for Workers? The Journal of Finance (forthcoming).

Pagano, M., Picariello, L., 2024. Corporate governance, meritocracy and careers. Review of Finance (forthcoming).

Pastorino, E., 2024. Careers in Firms: The Role of Learning and Human Capital. Journal of Political Economy 132(6).

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

Postel-Vinay, F., Robin, J.-M., 2002. Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. Econometrica 70(6), 2295–2350.

Robnik-Šikonja, M., 2004. Improving Random Forests. In: Boulicaut, J.-F., et al. (Eds.), ECML 2004, *LNAI*, volume 3201, p. 359–370, Springer.

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

Scur, D., Sadun, R., Van Reenen, J., Lemos, R., Bloom, N., 2021. The world management survey at 18: Lessons and the way forward. Oxford Review of Economic Policy 37(2), 231–258.

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

Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics 225(2), 175–199.

Terviö, M., 2008. The Difference That CEOs Make: An Assignment Model Approach. American Economic Review 98(3), 642–668.

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

Tåg, J., Åstebro, T., Thompson, P., 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., Dahabreh, I. J., 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., Liu, C., Zhang, X., Almpanidis, G., 2017. An up-to-date comparison of state-of-the-art classification algorithms. Expert Systems with Applications 82, 128–150.

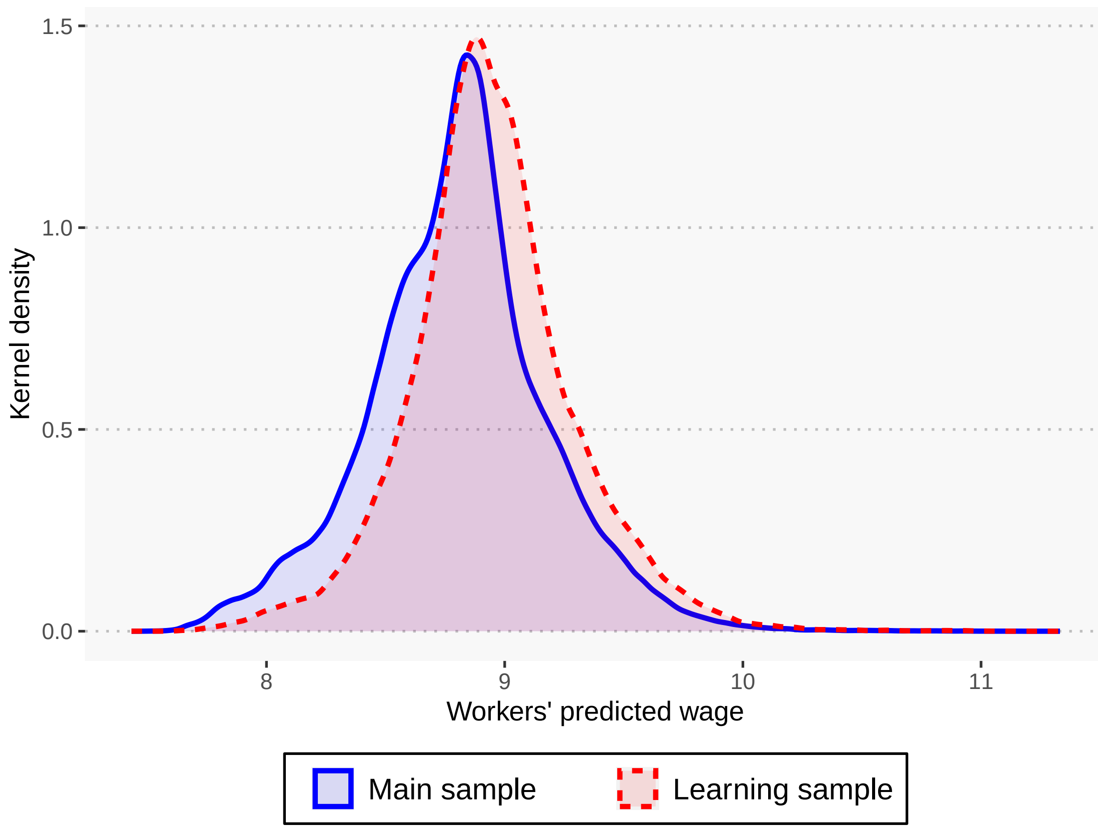


Figure 1: Common support of worker characteristics in the main and the learning samples
This figure shows the distributions of the predicted wages for workers in the learning sample (red line) and the main sample (blue line). For both samples, the predictions are obtained from wage regressions estimated 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 almost perfectly overlaps.

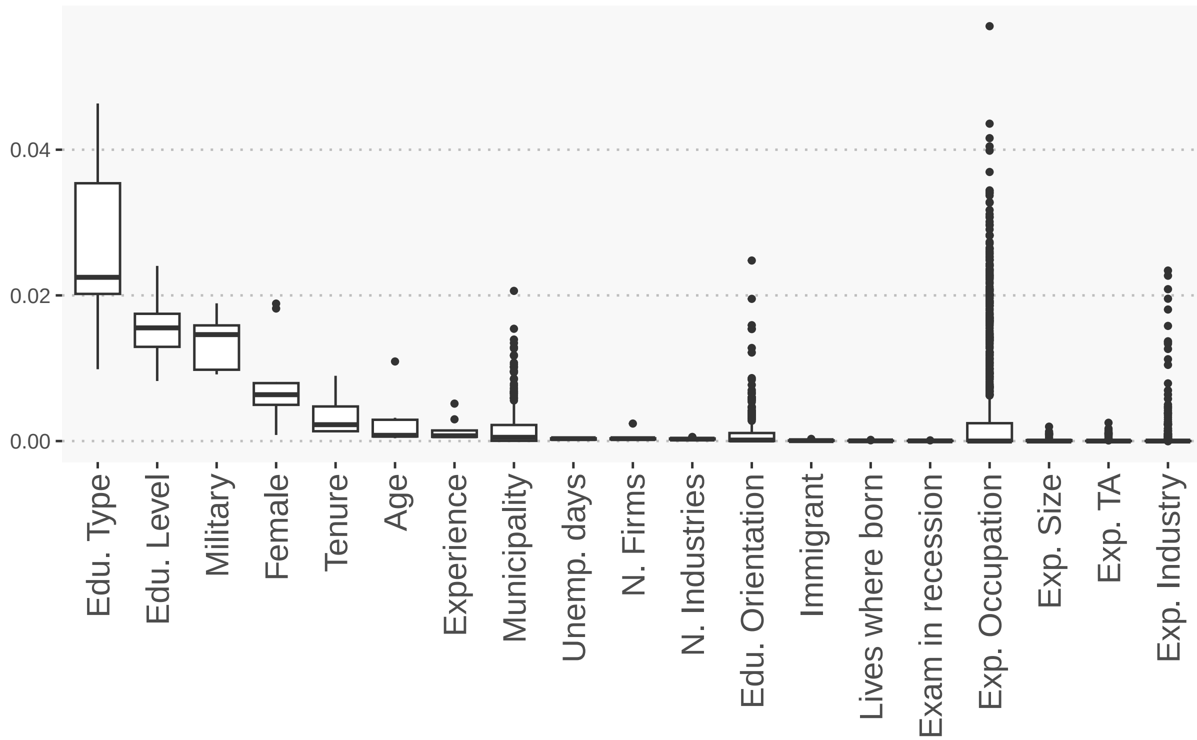


Figure 2: Importance of workers' features in the random forest algorithm
The graph plots the maximum explanatory power of all the workers' features used in the random forest algorithm. Features are listed on the horizontal axis, and the importance of each feature—defined as in Robnik-Šikonja (2004) and 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. Type" 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, "Unemp. 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.

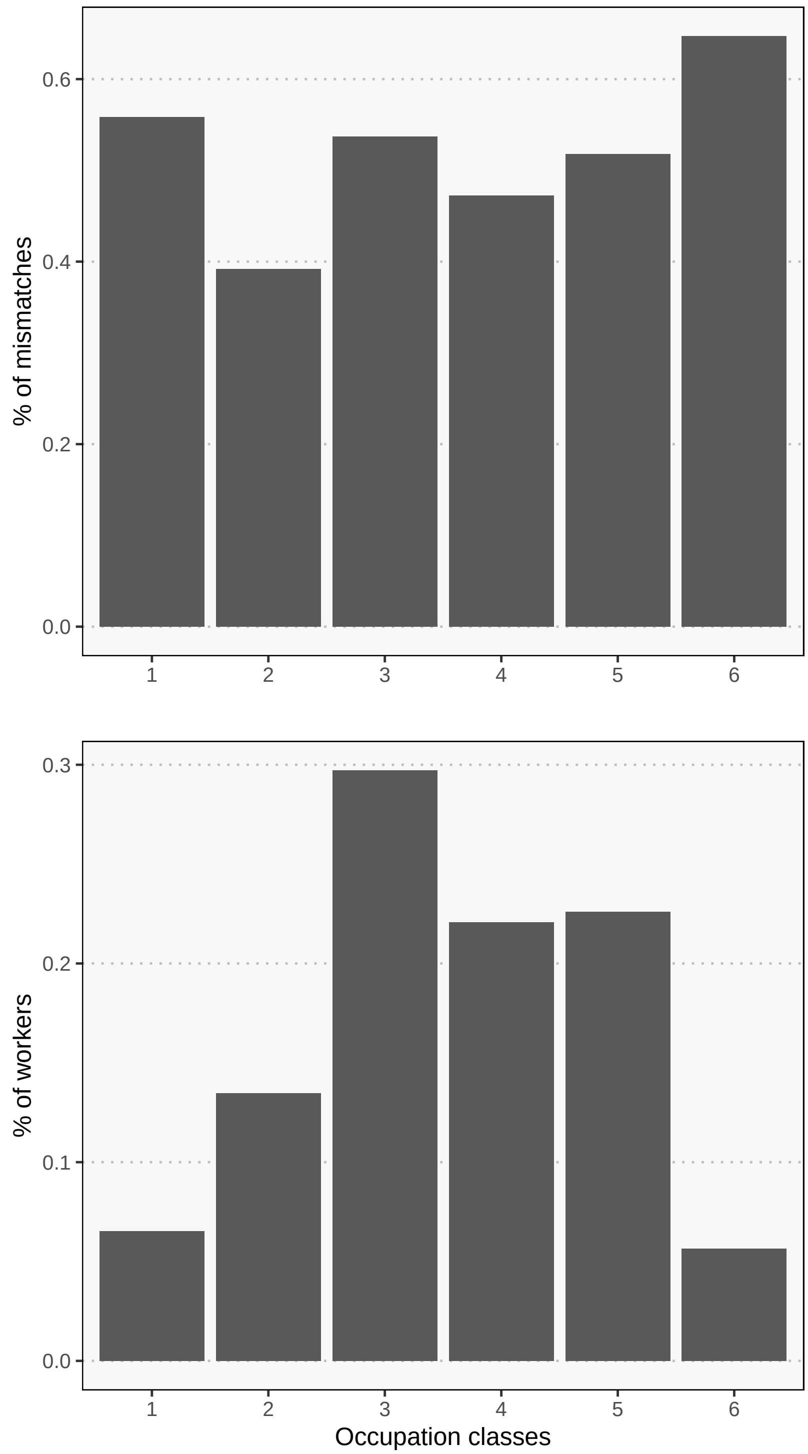


Figure 3: Distribution of mismatches and workers by occupation classes
The top graph shows the percentage of mismatches in each occupation class in the main sample. A mismatch occurs when an employee's observed job differs from the job predicted by the estimated allocation rule. The bottom graph shows 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.

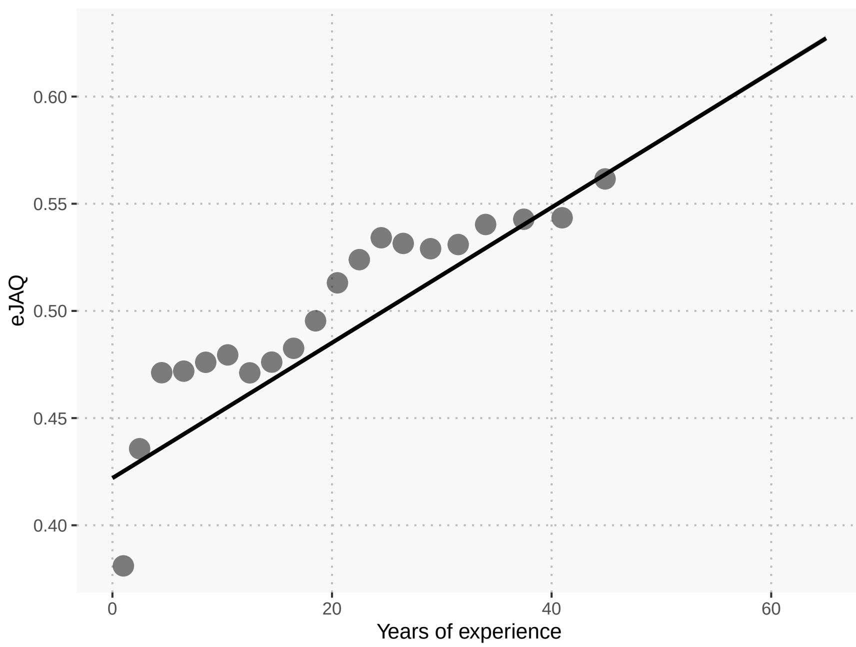


Figure 4: Worker-level job allocation quality (*eJAQ*) by labor market experience
This figure shows the binned scatter plot of an indicator for being assigned to the job predicted by the ML algorithm (*eJAQ*) against years of labor market experience.

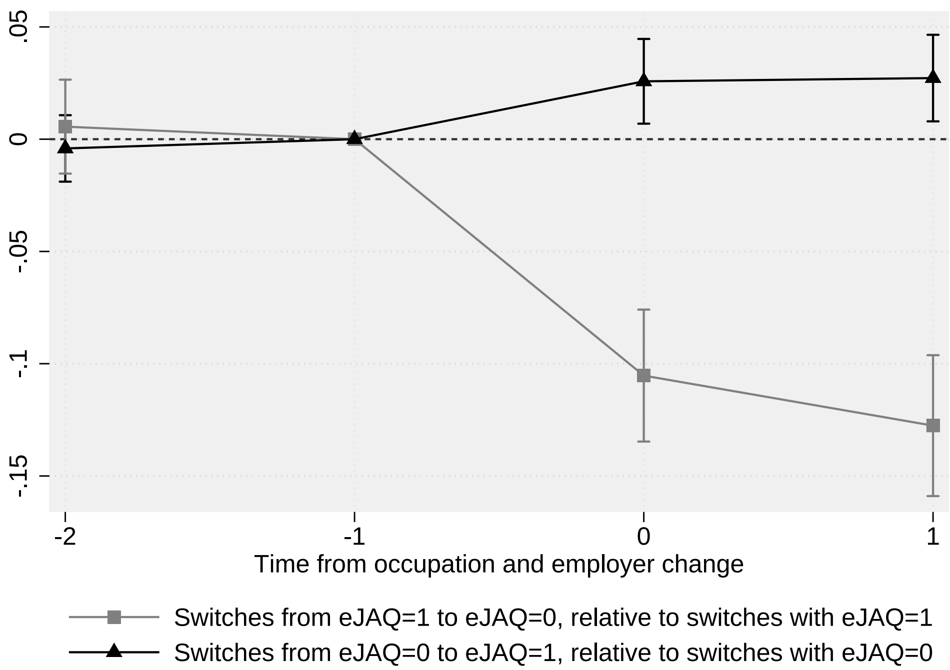


Figure 5: Response of earnings to employer and occupation switches
The figure shows the coefficient estimates and the 95% confidence intervals of two distinct event studies: the first compares changes of log labor earnings from 2 years before to 1 year after a switch from an occupation (and an employer) with $eJAQ = 0$ to one with $eJAQ = 1$, relative to the changes occurring around switches across two occupations with $eJAQ = 0$; the second compares changes of log labor earnings from 2 years before to 1 year after a switch from an occupation with $eJAQ = 1$ to occupations with $eJAQ = 0$, relative to switches across two occupations with $eJAQ = 1$.

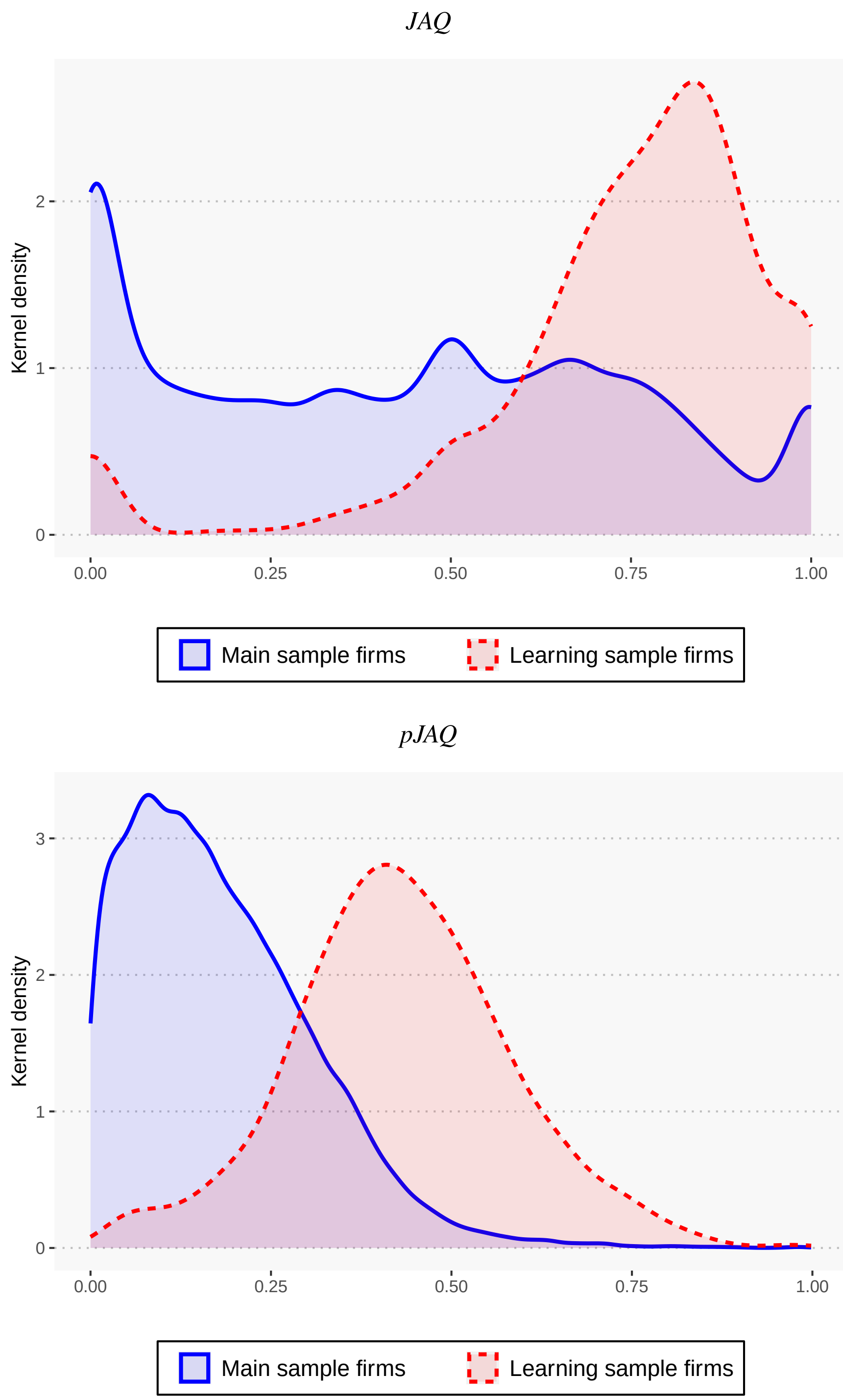


Figure 6: Distribution of match quality

The top graph shows the kernel density estimate of *JAQ* for firms in the main sample (blue line) and in the learning sample (red line). The bottom graph presents the corresponding kernel density estimates for *pJAQ*.

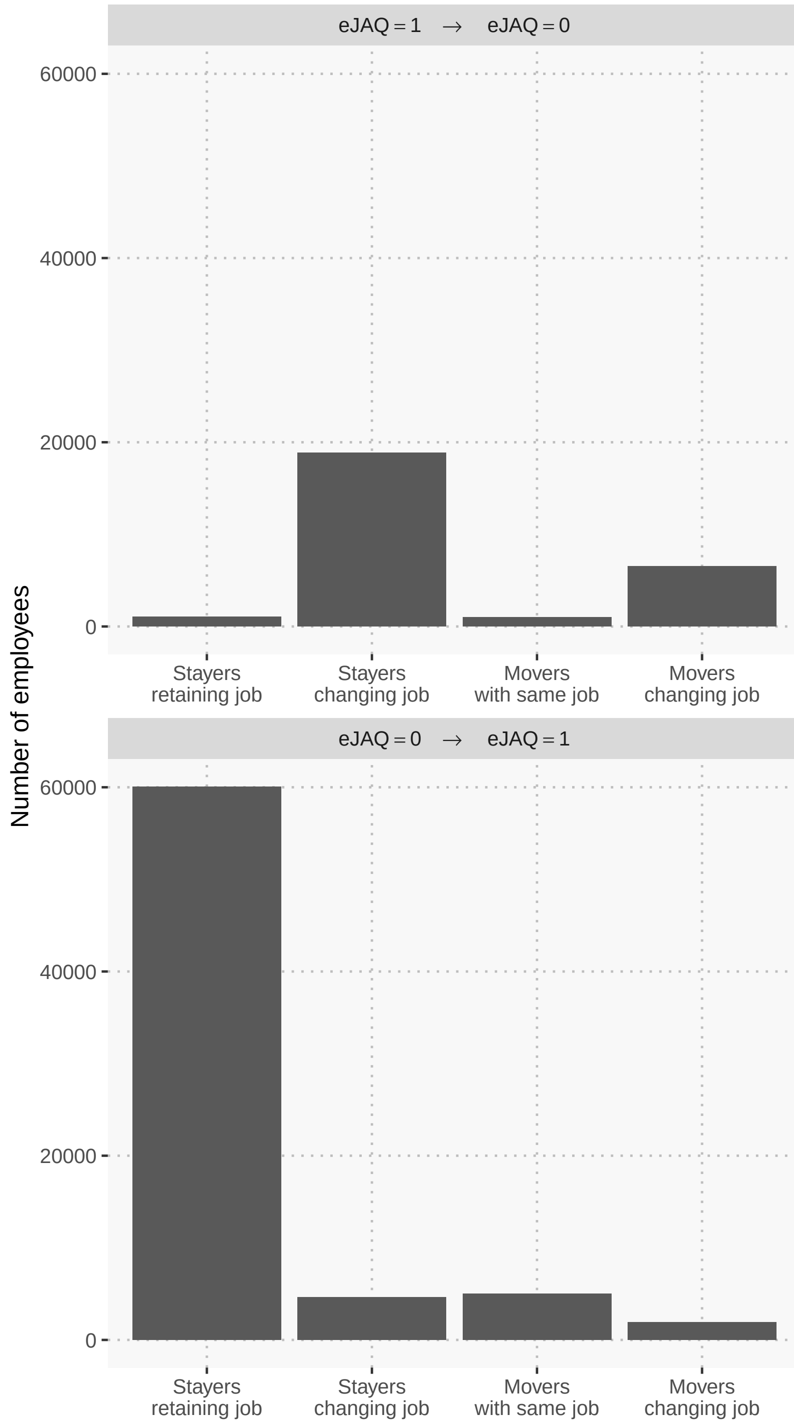


Figure 7: Accounting for changes in employee-level match quality (*eJAQ*)
The top graph shows the number of well matched employees ($eJAQ_{i,t} = 1$) who become mismatched in the subsequent year ($eJAQ_{i,t+1} = 0$), while the bottom graph shows the number of mismatched employees who become well matched in the subsequent year. In each of the two graphs, the four bars respectively refer to employees who (i) keep their jobs in the same firm; (ii) switch to a new job within the same firm; (iii) switch to a new firm while retaining the same job; (iv) switch to a new job and a new firm.

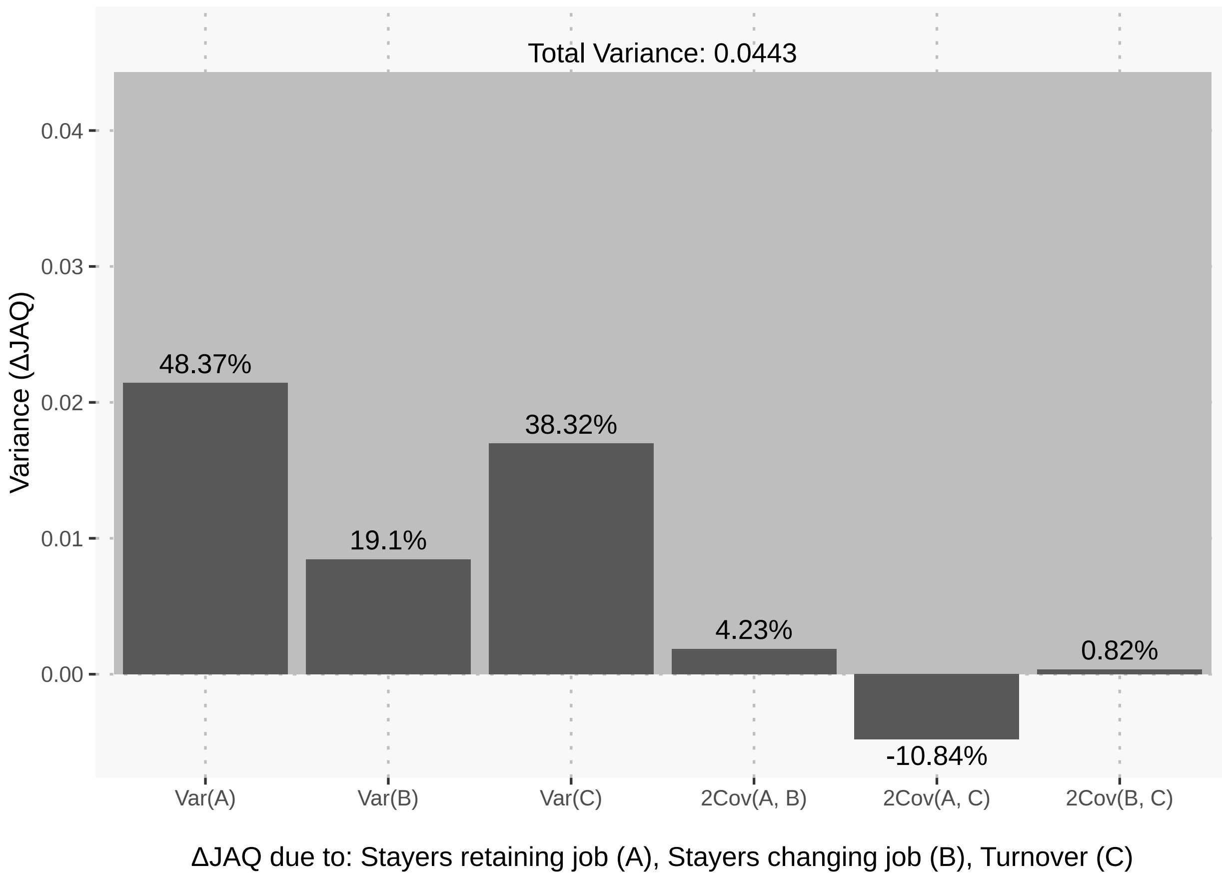


Figure 8: Accounting for changes in firm-level match quality
The figure shows a variance decomposition of year-to-year changes in *JAQ*, ΔJAQ, decomposed into six components: the variance of changes due to employees remaining with the same firm who retain their job (A); the variance of changes due to employees remaining with the same firm who switch to different jobs (B); changes in *JAQ* due to employee turnover (C); and the three corresponding covariances. The light gray area indicates the total variance of ΔJAQ.

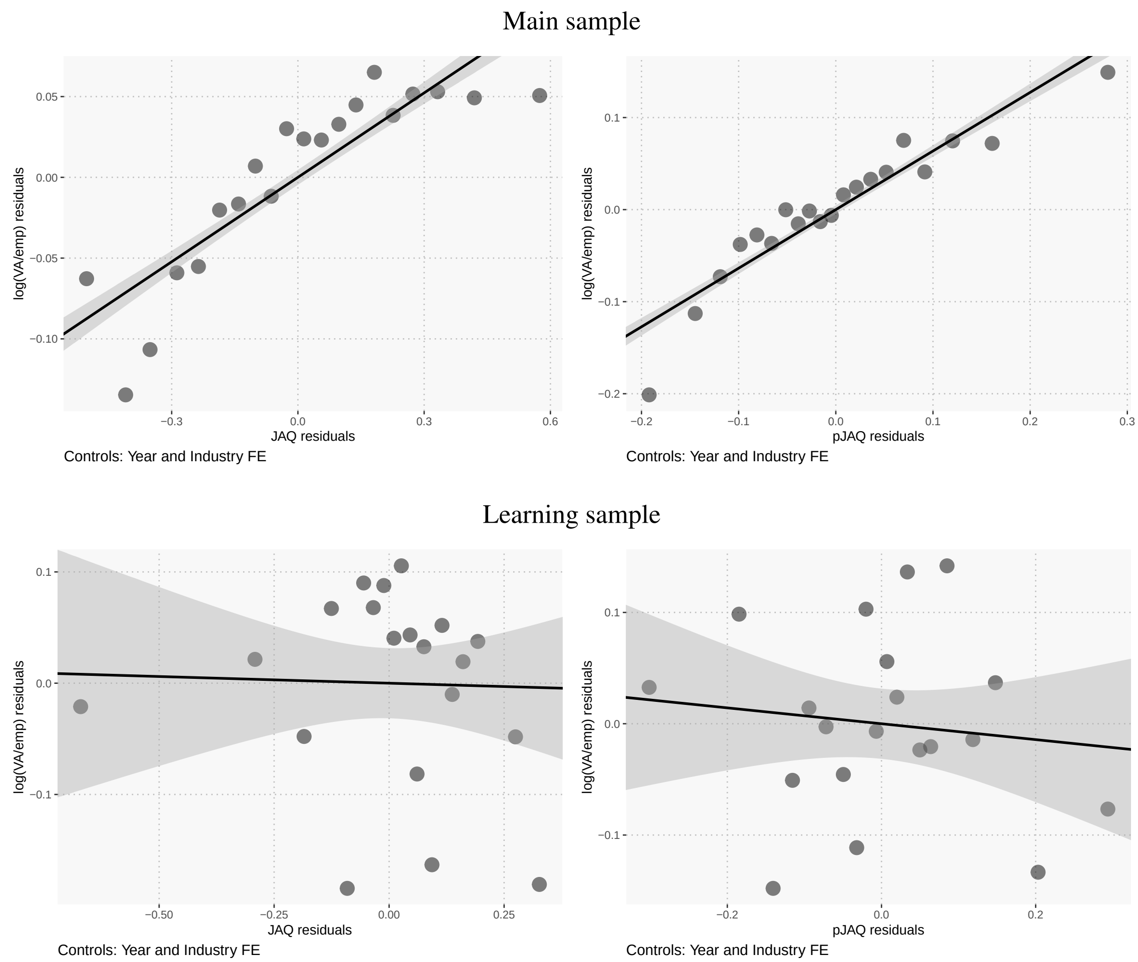


Figure 9: Correlation between productivity and job allocation quality

The figure shows binned scatter plots of productivity, as measured by log value added per employee, against job allocation quality, as measured by *JAQ* in the left charts and by *pJAQ* in the right charts, in each case controlling for year and industry fixed effects. The two top charts refer to the main sample and the two bottom ones to the learning sample. The regression lines are shown together with the respective 95% confidence intervals (shaded area). The points shown in the graphs represent the residuals of the partial regression plots and are computed as follows: residuals are first split into 20 equal-sized bins on the horizontal axis; points in a bin are represented with a unique point, with coordinates given by the average of the coordinates of the points in that bin. The regression line shown fits the residuals and not the binned points.

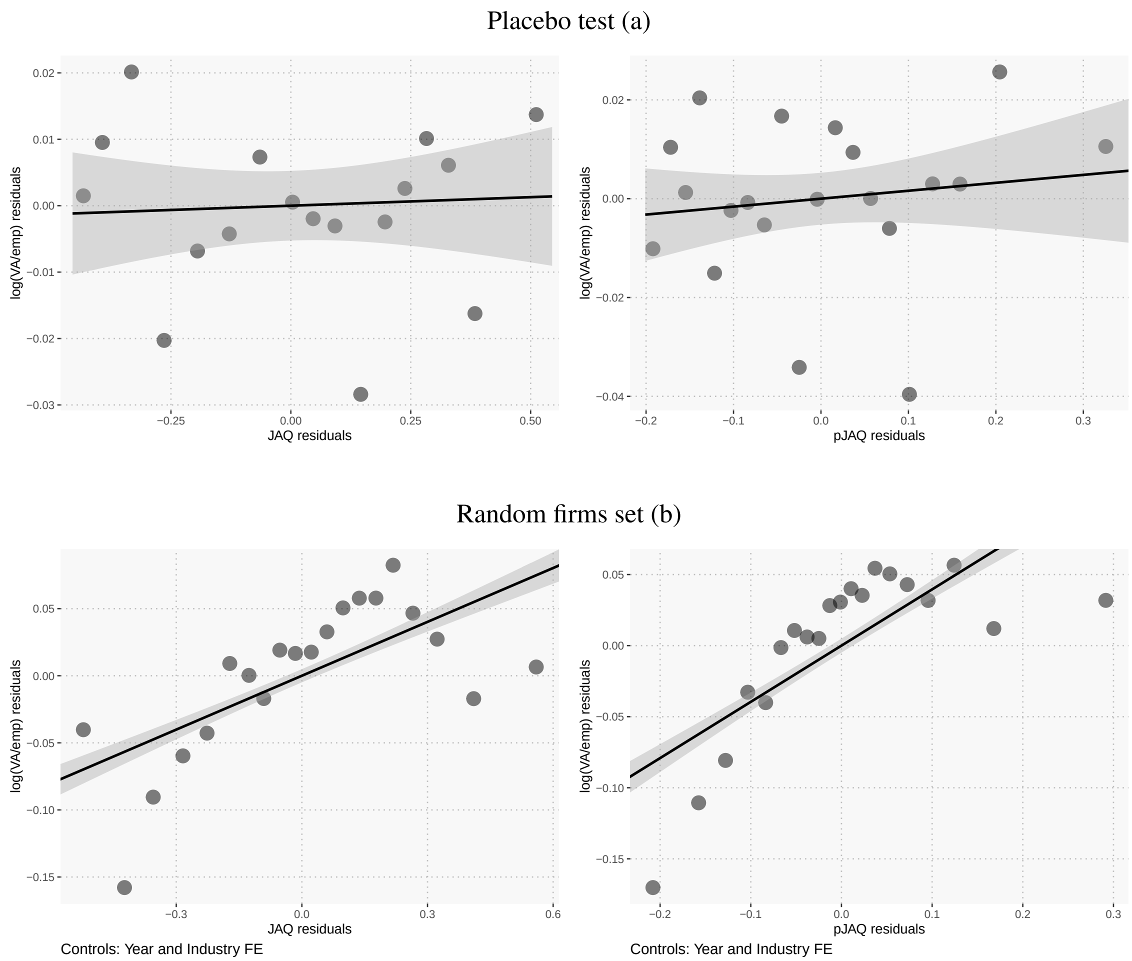


Figure 10: Addressing the circularity issue
Panel (a) shows partial regression plots of placebo productivity on *JAQ* in the left-hand chart and *pJAQ* on the right-hand chart, controlling for a constant. Placebo productivity is a noise variable obtained by randomly reshuffling the original log value added per employee across firms. Panel (b) shows partial regression plots of productivity, as measured by log value added per employee, against job allocation quality, as measured by JAQ^R in the left-hand chart and by $pJAQ^R$ in the right-hand chart. JAQ^R and $pJAQ^R$ are measures of job allocation quality and workers' suitability to their actual job respectively obtained by estimating a ML algorithm on a 10% randomly chosen sample of firms, controlling for year and industry fixed effects. All the regression lines are shown together with the respective 95% confidence intervals (shaded area). The points shown in the two graphs represent the residuals of the partial regression plots and are computed as follows: residuals are first split into 20 equal-sized bins on the horizontal axis; each point in the figure represents the points in a bin, and its coordinates are determined by the average coordinates of the points in that bin. The regression line shown is fitted on the residuals and not on binned points.

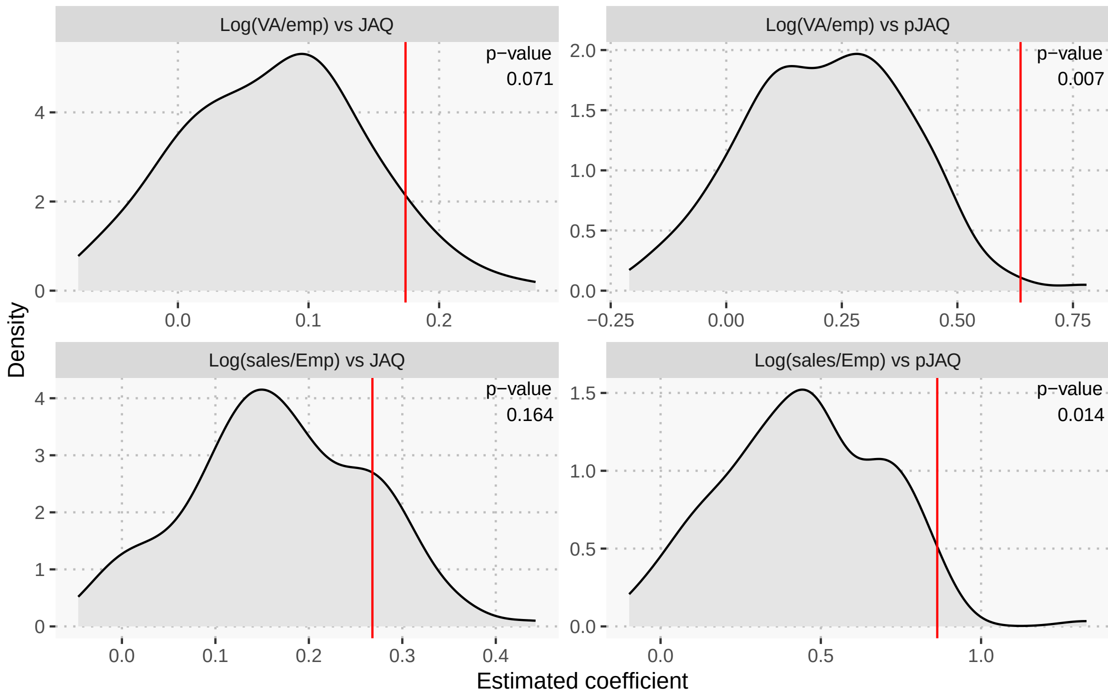


Figure 11: Monte Carlo distribution of JAQ^R coefficient in productivity regressions
Each panel in the figure plots the Monte Carlo distribution of the coefficients estimated regressing log value added per employee (top panel) and log sales per employee (bottom panel) on JAQ^R (left column) or $pJAQ^R$ (right column), controlling for 2-digit industry and year fixed effects. The Monte Carlo simulation is performed by training our ML algorithm on 150 randomly drawn learning samples, estimating JAQ^R and $pJAQ^R$ from each of them, and finally estimating the corresponding 150 above-described regressions. The vertical line in each graph indicates the *JAQ* and *pJAQ* coefficient estimated in the corresponding productivity regression, where *JAQ* and *pJAQ* are obtained using the top decile of the firms' productivity distribution as learning sample, while the p-value in each graph is the probability of observing a higher value under the relevant Monte Carlo distribution.

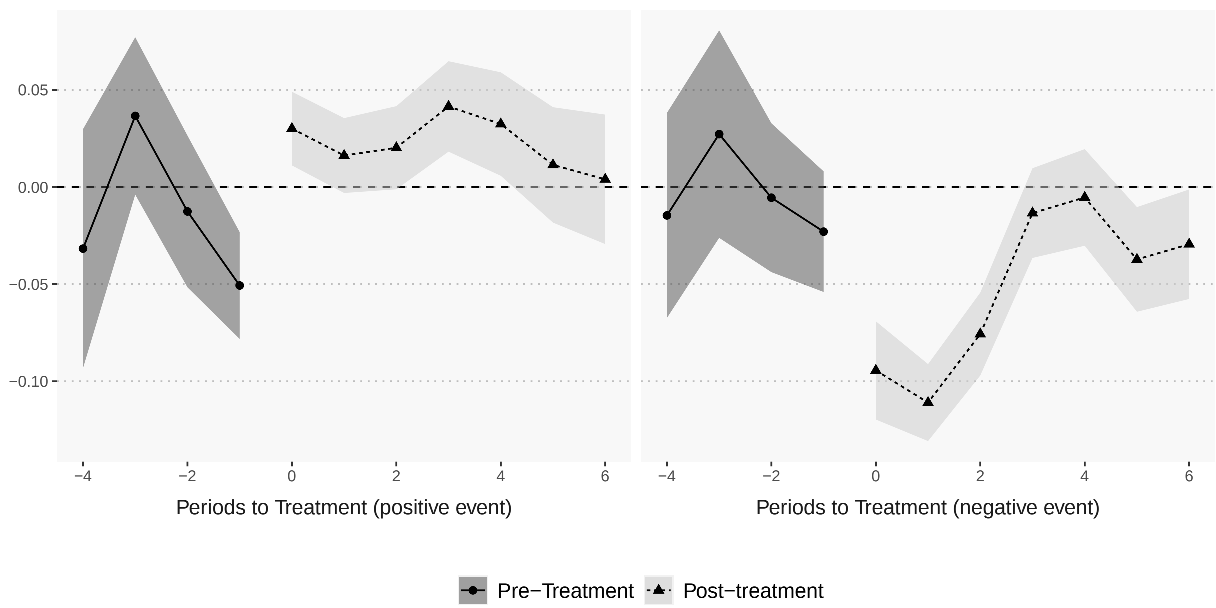


Figure 12: Response of rank-and-file workers' *JAQ* to positive (left) or negative (right) managerial turnover events

The figure shows the behavior of the *JAQ* of rank-and-file workers 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). The event study coefficients are estimated using the method proposed by Callaway and Sant'Anna (2021).

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 active at some point between 2001 and 2010, reporting a yearly median number of employees between 30 and 6000, and 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 5,901,551 observations at the individual level and the learning sample 66,684 observations.

	Mean	P50	P10	P25	P75	P90	SD
Panel A: Main sample							
Labor income (TSEK 2019)	351.991	324.377	168.806	258.118	409.305	541.715	197.704
University degree	0.13	0.00	0.00	0.00	0.00	1.00	0.34
Age	40.64	40.00	25.00	31.00	50.00	58.00	11.96
Female	0.34	0.00	0.00	0.00	1.00	1.00	0.47
Immigrant	0.13	0.00	0.00	0.00	0.00	1.00	0.34
Mobility (lives where born)	0.65	1.00	0.00	0.00	1.00	1.00	0.48
Labor market experience	19.66	19.00	3.00	9.00	29.00	39.00	12.84
Tenure	5.29	4.00	0.00	1.00	8.00	13.00	5.14
# industries worked in	2.28	2.00	1.00	1.00	3.00	4.00	1.26
# jobs held	2.29	3.00	1.00	2.00	4.00	5.00	1.69
# unemployment days since '92	181.90	0.00	0.00	0.00	224.00	599.00	338.90
Panel B: Learning sample							
Labor income (TSEK 2019)	474.627	415.484	270.307	333.533	535.107	724.454	301.928
University degree	0.22	0.00	0.00	0.00	0.00	1.00	0.41
Age	43.24	43.00	29.00	35.00	52.00	58.00	10.88
Female	0.31	0.00	0.00	0.00	1.00	1.00	0.46
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	21.36	21.00	5.00	11.00	31.00	39.00	12.28
Tenure	7.81	6.00	0.00	2.00	12.00	20.00	6.54
# industries worked in	2.71	3.00	1.00	1.00	4.00	5.00	1.49
# jobs held	3.29	3.00	1.00	2.00	4.00	6.00	1.92
# unemployment days since '92	164.37	0.00	0.00	0.00	192.00	530.00	321.09

Table 2: Average employees' characteristics for selected occupations

The table shows the average employees' characteristics (rows) for a pool of selected jobs (columns). Jobs are labeled by their ISCO code, as follows: directors and chief executives (121); production and operations managers (122); other specialist managers (123); managers of small enterprises (131); legal professionals (242); office secretaries and data entry operators (411); machinery mechanics and fitters (723): chemical-processing-plant operators (815); manufacturing labourers (932). The jobs included in the table are selected to provide information on the following broad occupation classes: 1) managers (121, 122, 123, 131), 2) professionals (242), 3) technicians and clerks (411), 4) skilled manual workers (723), 5) machine operators and assemblers (815), and 6) elementary occupations (932). For each job, the average characteristics are obtained averaging over the subset of the 1000 employees (990 and 998, for jobs 242 and 723, respectively) with the highest predicted conditional probability for the corresponding job, $P(j|X)$, as estimated by the ML algorithm. While occupation class 1 includes all managerial positions, for other occupation classes we select jobs with the highest average predicted probability, $P(j|X)$. For numerical variables, standard deviations are reported in parentheses; for categorical ones (Education level, Education type, Education orientation) the table shows the most frequent code with relative share in square brackets.

	Jobs								
	121	122	123	131	242	411	723	815	932
Characteristics									
$P(j\|X)$	0.63	0.683	0.656	0.683	0.998	0.848	0.931	0.896	0.837
	(0.03)	(0.038)	(0.032)	(0.073)	(0.001)	(0.052)	(0.018)	(0.032)	(0.062)
Age	50.22	49.54	46.61	44.11	32.85	51.62	40.13	45.10	44.24
	(7.26)	(6.88)	(6.72)	(9.76)	(5.16)	(8.93)	(10.50)	(9.47)	(10.90)
Female	0.01	0.07	0.28	0.55	0.48	1.00	0.00	0.04	0.07
	(0.11)	(0.26)	(0.45)	(0.50)	(0.50)	(0.00)	(0.00)	(0.19)	(0.26)
Lives where born	0.46	0.48	0.49	0.68	0.42	0.67	0.85	0.88	0.76
	(0.50)	(0.50)	(0.50)	(0.47)	(0.49)	(0.47)	(0.36)	(0.33)	(0.43)
Immigrant	0.21	0.09	0.14	0.07	0.08	0.10	0.02	0.07	0.14
	(0.41)	(0.29)	(0.35)	(0.25)	(0.28)	(0.29)	(0.13)	(0.26)	(0.35)
Education level[a]	4	4	4	2	4	2	2	2	2
[share]	[0.86]	[0.52]	[0.82]	[0.71]	[1.00]	[0.80]	[0.99]	[0.77]	[0.54]
Education type[b]	2	6	2	0	1	2	6	6	0
[share]	[0.53]	[0.96]	[0.94]	[0.40]	[1.00]	[0.85]	[0.99]	[0.69]	[0.55]
Education orient.	340a	521b	340a	010a	380a	346x	525c	010a	010a
[share]	[0.49]	[0.22]	[0.88]	[0.21]	[1.00]	[0.32]	[0.77]	[0.23]	[0.46]
Unemployment days	22.17	23.68	45.92	153.74	51.60	64.26	92.87	33.95	125.26
	(84.20)	(84.56)	(111.13)	(261.94)	(114.13)	(199.56)	(258.25)	(135.74)	(304.29)
Tenure	7.09	11.46	7.03	5.72	3.23	10.50	8.04	14.50	12.37
	(5.14)	(5.80)	(5.13)	(4.74)	(2.16)	(5.61)	(5.48)	(4.26)	(5.87)
Experience	24.20	25.99	20.98	25.03	6.90	30.86	21.33	26.67	26.29
	(7.85)	(8.01)	(7.83)	(10.37)	(4.99)	(9.80)	(10.68)	(9.90)	(11.58)
N. industries	2.80	2.24	2.74	2.20	2.59	1.81	1.76	1.30	1.54
of employment	(1.53)	(1.19)	(1.34)	(1.17)	(1.04)	(0.96)	(0.88)	(0.69)	(0.85)
Military service	0.23	0.07	0.17	0.26	0.20	0.04	0.01	0.03	0.05
	(0.21)	(0.11)	(0.17)	(0.16)	(0.22)	(0.09)	(0.05)	(0.07)	(0.10)
Experience in job	3.12	5.05	4.74	2.38	3.52	5.88	6.65	6.55	5.47
	(1.79)	(1.45)	(1.89)	(1.44)	(1.63)	(1.53)	(1.62)	(1.13)	(1.59)

[a] Education level: Basic (1); High school (2); Vocational (3); University (4).
[b] Education type: No specialization (0); Law (1); Business and economics (2); Health and medicine (3); Natural science (4); Teaching (5); Engineering (6); Social sciences (7); Services (8); Other specialization (9).

Table 3: Labor earnings, separations and match quality

This table displays the relationship between the log of labor earnings and match quality in columns (1) to (4), and between a separation indicator (which equals 1 if a worker changes employer between time $t-1$ and t and 0 otherwise) and match quality in columns (5) to (8). Match quality is measured by *eJAQ* in Panel A and by *epJAQ* in Panel B 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.

	Log(labor earnings)				Separation indicator			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
eJAQ	0.0240***	0.0266***	0.0174***	0.0272***	-0.0125***	-0.0119***	-0.0270***	-0.0103***
	(0.0004)	(0.0004)	(0.0004)	(0.0013)	(0.0003)	(0.0004)	(0.0005)	(0.0024)
Panel B								
epJAQ	0.0497***	0.0574***	0.0539***	0.0649***	-0.0697***	-0.0582***	-0.1723***	-0.0878***
	(0.0011)	(0.0012)	(0.0014)	(0.0074)	(0.0011)	(0.0012)	(0.0017)	(0.0133)
Year and job FE	✓	✓	✓	✓	✓	✓	✓	✓
Worker controls	✓	✓			✓	✓		
Industry FE		✓				✓		
Firm controls		✓				✓		
Worker FE			✓	✓			✓	✓
Firm FE				✓				✓
Observations	5,901,551	3,909,445	5,901,551	5,526,718	4,484,975	3,415,510	4,484,975	4,262,039

Table 4: Job allocation quality, market competition and firm ownership

This table shows regressions of measures of job allocation quality (*JAQ*) and of workers' suitability to their actual job (*pJAQ*) on the Lerner index of market competition in columns 1 to 4, and on a family firm status dummy in columns 5 to 8. The Lerner index for each firm is defined as 1 − profits/sales lagged by 2 years and averaged across all firms in the same 2-digit industry, excluding the firm itself. All specifications include year and industry dummies (where industries are manufacturing, real estate, renting and business activities, and wholesale and retail). The specifications of the even-numbered columns control for the share of employees with a college degree, and include the following additional controls (whose coefficients are not shown for brevity): log employment, log capital, log firm age, indicator for listed firms, years of managerial experience averaged over employees in the firm.

	JAQ		*pJAQ*		*JAQ*		*pJAQ*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lerner index	0.003***	0.003***	0.003***	0.002***				
(2-year lagged)	(0.000)	(0.000)	(0.000)	(0.000)				
Family firm					-0.026***	-0.008	-0.024***	-0.008***
					(0.007)	(0.007)	(0.003)	(0.003)
Share emp. w/ college		0.070***		0.110***		0.112***		0.125***
		(0.019)		(0.011)		(0.020)		(0.011)
Year dummies	✓	✓	✓	✓	✓	✓	✓	✓
Industry dummies		✓		✓		✓		✓
Firm controls		✓		✓		✓		✓
Observations	33,254	33,254	33,254	33,254	29,947	29,541	29,947	29,541
No. Firms	6,269	6,269	6,269	6,269	6,372	6,294	6,372	6,294
y Mean	0.496	0.496	0.211	0.211	0.483	0.484	0.210	0.210
y St. Dev.	0.296	0.296	0.128	0.128	0.300	0.299	0.131	0.130

Table 5: Firm performance and *JAQ*

This table displays results from regressions of firm productivity and profitability on *JAQ* and control variables. 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.

	Log(sales/emp) (1)	Log(VA/emp) (2)	OROA (3)	Log(sales/emp) (4)	Log(VA/emp) (5)	OROA (6)
Panel A						
JAQ	0.366***	0.195***	-0.005	0.102***	0.080***	0.005
	(0.023)	(0.014)	(0.005)	(0.014)	(0.010)	(0.005)
log(cap/emp)				0.413***	0.237***	-0.020***
				(0.012)	(0.009)	(0.002)
log(emp)				0.003	-0.005	-0.003*
				(0.007)	(0.005)	(0.002)
Share emp w/ college				0.108***	0.337***	0.013
				(0.031)	(0.022)	(0.010)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Observations	48,116	47,743	48,116	48,116	47,743	48,116
No. Firms	7,875	7,827	7,875	7,875	7,827	7,875
LHS mean	7.306	6.140	0.079	7.306	6.140	0.079
LHS SD	0.793	0.534	0.179	0.793	0.534	0.179
Panel B						
JAQ	0.169***	0.116***	0.005	0.080***	0.067***	0.012*
	(0.017)	(0.012)	(0.005)	(0.017)	(0.013)	(0.007)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Firm controls				✓	✓	✓
Observations	48,116	47,743	48,116	29,947	29,742	29,947
No. firms	7,875	7,827	7,875	6,372	6,339	6,372
LHS mean	7.306	6.140	0.079	7.339	6.173	0.089
LHS SD	0.793	0.534	0.179	0.790	0.527	0.183
Panel C						
JAQ	0.091***	0.062***	0.004	0.044***	0.045***	0.010
	(0.017)	(0.012)	(0.005)	(0.018)	(0.014)	(0.007)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Observations	48,116	47,743	48,116	29,947	29,742	29,947
No. firms	7,875	7,827	7,875	6,372	6,339	6,372
LHS mean	7.306	6.140	0.079	7.339	6.173	0.089
LHS SD	0.793	0.534	0.179	0.790	0.527	0.183

Table 6: Firm performance and worker suitability

This table displays results from regressions of firm productivity and profitability on firm-level *pJAQ* (defined as the average of worker-level *pJAQ* for a given firm) and control variables. 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.

	Log(sales/emp) (1)	Log(VA/emp) (2)	OROA (3)	Log(sales/emp) (4)	Log(VA/emp) (5)	OROA (6)
Panel A						
pJAQ	0.995***	0.724***	0.000	0.289***	0.266***	0.025*
	(0.064)	(0.036)	(0.014)	(0.040)	(0.028)	(0.014)
log(cap/emp)				0.412***	0.235***	-0.020***
				(0.012)	(0.009)	(0.002)
log(emp)				0.005	-0.003	-0.003*
				(0.007)	(0.005)	(0.002)
Share emp w/ college				0.088***	0.317***	0.011
				(0.031)	(0.022)	(0.010)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Observations	48,116	47,743	48,116	48,116	47,743	48,116
No. firms	7,875	7,827	7,875	7,875	7,827	7,875
LHS mean	7.306	6.140	0.079	7.306	6.140	0.079
LHS SD	0.793	0.534	0.179	0.793	0.534	0.179
Panel B						
pJAQ	0.545***	0.385***	0.021	0.242***	0.217***	0.039**
	(0.050)	(0.033)	(0.014)	(0.048)	(0.034)	(0.019)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Size-industry FE	✓	✓	✓	✓	✓	✓
Firm controls				✓	✓	✓
Observations	48,116	47,743	48,116	29,947	29,742	29,947
No. firms	7,875	7,827	7,875	6,372	6,339	6,372
LHS mean	7.306	6.140	0.079	7.339	6.173	0.089
LHS SD	0.793	0.534	0.179	0.790	0.527	0.183
Panel C						
pJAQ	0.261***	0.190***	0.029*	0.150***	0.156***	0.042*
	(0.056)	(0.037)	(0.017)	(0.055)	(0.038)	(0.022)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Size-industry FE	✓	✓	✓	✓	✓	✓
Observations	48,116	47,743	48,116	29,947	29,742	29,947
No. firms	7,875	7,827	7,875	6,372	6,339	6,372
LHS mean	7.306	6.140	0.079	7.339	6.173	0.089
LHS SD	0.793	0.534	0.179	0.790	0.527	0.183

Table 7: ***JAQ*** **and productivity with AKM-based selection of the learn set**

The table reports the estimated relationship between productivity and the three firm-level measures of job allocation quality, when the ML algorithm is estimated on the subsample of firms within the top-decile of the residuals' distribution of an AKM model. JAQ^{AKM} and $pJAQ^{\mathrm{AKM}}$ are the two resulting job-allocation quality measures. Panels A and B respectively replicate panels A in Table 5 and Table 6. Standard errors clustered at firm level are shown in parentheses.

	Log(sales/emp) (1)	Log(VA/emp) (2)	OROA (3)	Log(sales/emp) (4)	Log(VA/emp) (5)	OROA (6)
Panel A						
JAQ^{AKM}	0.238***	0.127***	-0.007	0.120***	0.049***	-0.001
	(0.037)	(0.024)	(0.008)	(0.021)	(0.018)	(0.009)
log(cap/emp)				0.423***	0.265***	-0.015***
				(0.015)	(0.012)	(0.004)
log(emp)				0.013	-0.012	-0.013***
				(0.010)	(0.008)	(0.004)
Share emp w/ college				0.075*	0.344***	0.014
				(0.041)	(0.031)	(0.014)
Panel B						
$pJAQ^{\mathrm{AKM}}$	0.744***	0.581***	0.016	0.393***	0.222***	0.038
	(0.117)	(0.072)	(0.031)	(0.068)	(0.060)	(0.034)
log(cap/emp)				0.422***	0.264***	-0.015***
				(0.015)	(0.012)	(0.004)
log(emp)				0.016	-0.012	-0.014***
				(0.010)	(0.008)	(0.004)
Share emp w/ college				0.047	0.326***	0.010
				(0.042)	(0.031)	(0.014)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Observations	16,592	16,403	16,592	8,405	8,322	8,405
No. Firms	3,696	3,656	3,696	2,474	2,450	2,474
LHS Mean	7.264	6.147	0.069	7.282	6.185	0.083
LHS St. Dev.	0.828	0.603	0.208	0.797	0.594	0.224

Table 8: 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, *M-JAQ* and Manager exp refer 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 are clustered at firm level and are shown in parentheses.

	R&F-JAQ					
	(1)	(2)	(3)	(4)	(5)	(6)
M-JAQ	0.191***	0.115***	0.062***	0.115***	0.062***	0.019**
	(0.007)	(0.006)	(0.008)	(0.007)	(0.006)	(0.008)
Manager exp		0.018***	0.019***		0.009***	0.008***
		(0.002)	(0.002)		(0.002)	(0.002)
Industry FEs			✓			✓
Municipality FEs			✓			✓
Year FEs	✓	✓	✓	✓	✓	✓
Firm FEs		✓	✓		✓	✓
Firm controls			✓			✓
Observations	36,230	36,230	22,830	22,821	22,821	14,149
No. Firms	7,680	7,680	6,084	6,454	6,454	4,712

Table 9: Role of management in firm productivity

This table displays the estimated relationship between productivity and the quality of managers' allocation. Productivity is measured either as log sales per employee or log value added per employee. The regressions are based on data from 2003 to 2010. All specifications include municipality and year fixed effects; those in columns 3 to 6 include industry fixed effects. The specifications in columns 5 and 6 also control for experience in managerial jobs. Standard errors are clustered at firm level and are shown in parentheses.

	Log(Sales/emp) (1)	Log(VA/emp) (2)	Log(Sales/emp) (3)	Log(VA/emp) (4)	Log(Sales/emp) (5)	Log(VA/emp) (6)
M-JAQ	0.213***	0.132***	0.158***	0.081***	0.109***	0.060***
	(0.019)	(0.015)	(0.015)	(0.013)	(0.017)	(0.013)
Managers exp					0.030***	0.013***
					(0.004)	(0.004)
Industry FEs			✓	✓	✓	✓
Year FEs	✓	✓	✓	✓	✓	✓
Municipality FEs	✓	✓	✓	✓	✓	✓
Observations	35,971	35,823	35,971	35,823	35,971	35,823
No. Firms	7,592	7,559	7,592	7,559	7,592	7,559
LHS Mean	7.408	6.163	7.408	6.163	7.408	6.163
LHS St. Dev.	0.779	0.577	0.779	0.577	0.779	0.577

Table 10: Decomposing the impact of managerial turnover

The table shows the impact of positive and negative managerial turnover events on the allocation of employees staying in the firm. The coefficients in columns 1 and 2 measure the estimated average treatment on the treated (ATT) for the outcome variable shown at the beginning of the respective row of the table at the event time (standard errors in parentheses) using the Callaway-Sant'Anna method, for positive and negative managerial turnover events, respectively. Internal reallocation is the number of workers that switch jobs within the same firm as a fraction of the total number of employees remaining in the firm in the same year. Firings is the fraction of workers fired at the time of managerial turnover out of the total number of previous year's employees. Hirings is the fraction of workers hired at the time of managerial turnover out of the total number of employees in the same year. $R\&F\text{-}JAQ$ is the average allocation quality of rank-and-file employees. $\Delta R\&F\text{-}JAQ^{TO}$ is the change in the allocation quality of rank-and-files employees net that of rank-and-file workers who stay in the firm (defined by applying equation (3) to rank-and-file workers rather than managers). $R\&F\text{-}JAQ^{s}$ is the fraction of correctly allocated rank-and-files employees among those retained by the firm when the event occurs. $\Delta R\&F\text{-}JAQ^{s}$ is the change in the allocation quality of rank-and-file employees who stay in the firm.

	Positive event (1)	Negative event (2)
Internal reallocation	-0.049***	0.146***
	(0.015)	(0.017)
Firings	0.044***	0.056***
	(0.016)	(0.019)
Hirings	-0.154***	0.007
	(0.021)	(0.019)
$R\&F\text{-}JAQ$	0.03***	-0.094***
	(0.01)	(0.013)
$R\&F\text{-}JAQ^{s}$	-0.017	-0.081***
	(0.012)	(0.014)
$\Delta R\&F\text{-}JAQ$	0.084***	0.005
	(0.02)	(0.026)
$\Delta R\&F\text{-}JAQ^{TO}$	0.067***	0.002
	(0.015)	(0.013)
$\Delta R\&F\text{-}JAQ^{s}$	0.039**	-0.104***
	(0.015)	(0.018)

Appendix

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

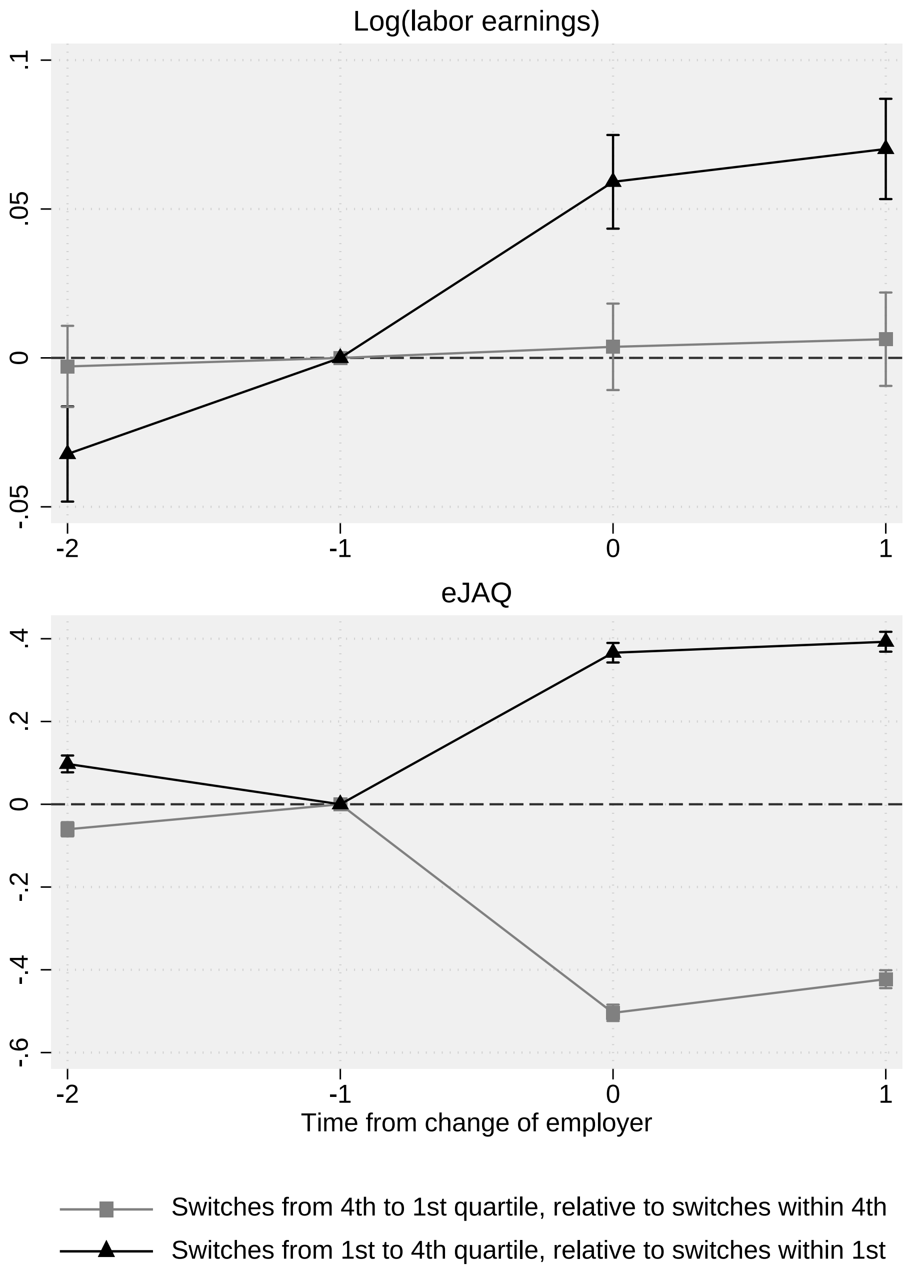


Figure A1: Response of earnings and *eJAQ* to employer switches
The graphs show the coefficient estimates and the 95% confidence intervals of two distinct event studies: the first plots changes of log labor earnings (top graph) or *eJAQ* (bottom graph) from 2 years before to 1 year after a switch from an employer belonging to the first quartile of the co-worker *eJAQ* distribution to one in the fourth quartile, relative to the changes occurring around switches across two employers in the first quartile of the co-worker *eJAQ* distribution; the second plots changes of log labor earnings (top panel) or *eJAQ* (bottom panel) from 2 years before to 1 year after a switch from a firm in the fourth quartile of the co-worker *eJAQ* distribution to one in the first quartile, relative to switches across two employers belonging to the fourth quartile.

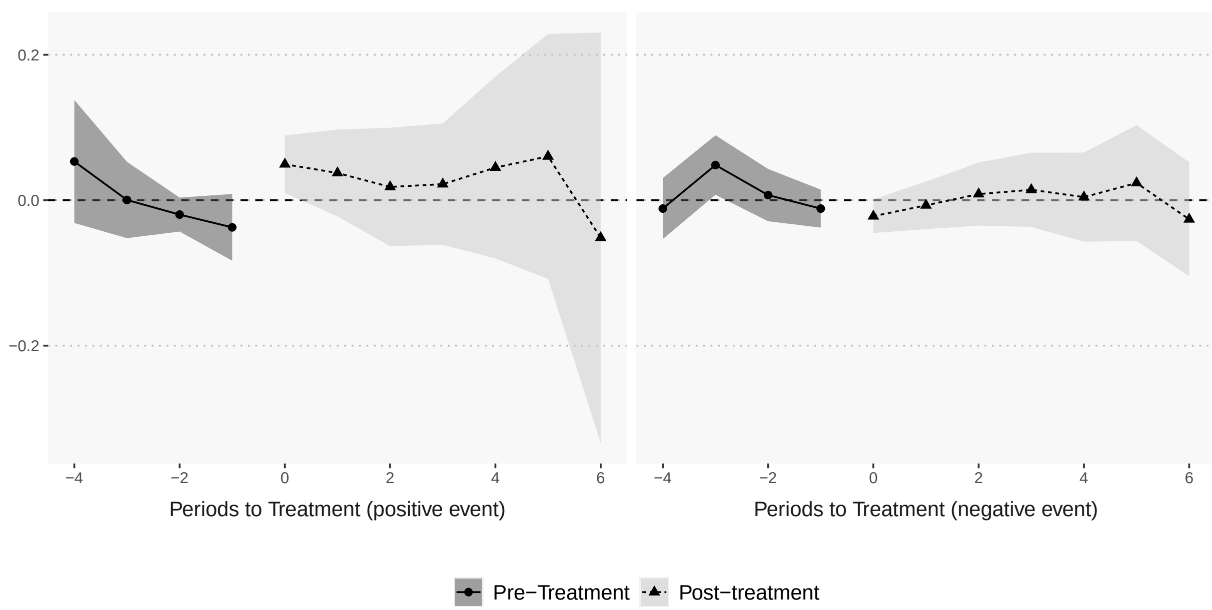
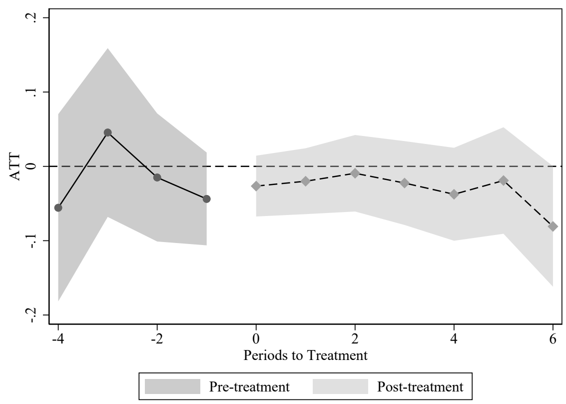
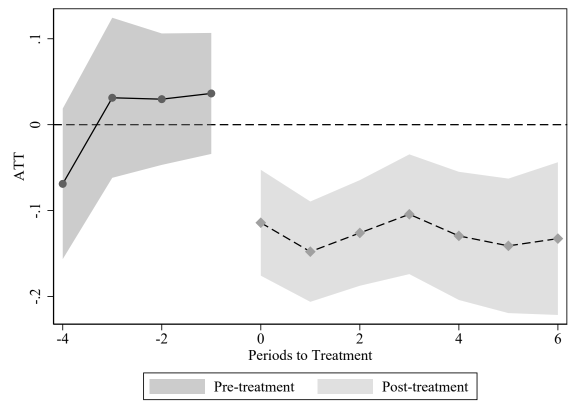


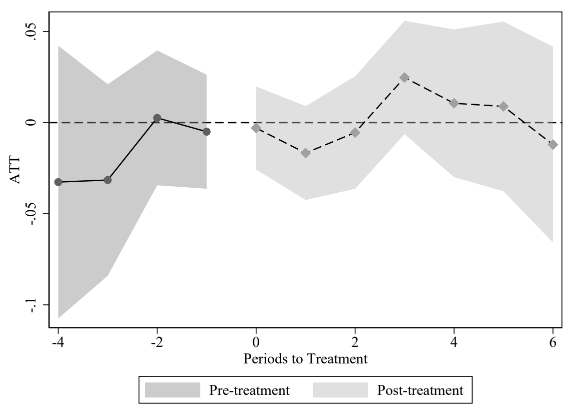
Figure A2: Response of rank-and-file workers' *JAQ* to positive (left) or negative (right) managers' death events The figure shows the event study estimated using the method by Callaway and Sant'Anna (2021) relating the JAQ of rank-and-file workers with managerial turnover events associated with the death of (at least) one member of the incumbent management team. The left graph refers to positive events, i.e., events associated with an increase in managerial *JAQ*, whereas the right graph refers to negative events, i.e., those associated with a decrease in managerial JAQ.



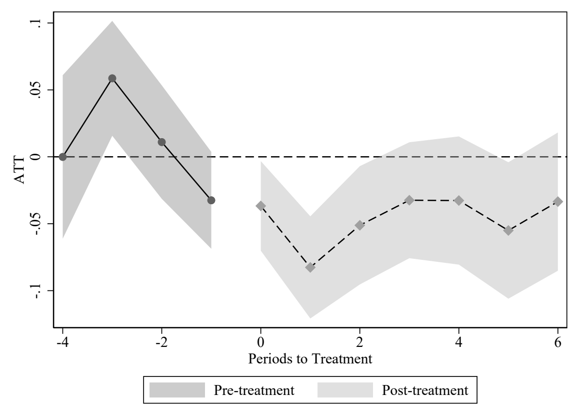
Impact of top-manager positive turnover event on middle management's *JAQ*



Impact of negative top-manager turnover event on middle management's *JAQ*



Impact of top-manager positive turnover event on *R&F-JAQ*



Impact of top-manager negative turnover event on *R&F-JAQ*

Figure A3: Impact of top-manager turnover events on middle management's *JAQ* and on *R&F-JAQ*
The figure shows the impact of top-manager turnover events on the match quality of middle management and on that of rank-and-file workers. The event study coefficients are estimated using Callaway and Sant'Anna's method (2021). The top-manager turnover events are defined as in Section 5, but are restricted to consider turnover of top executives only. Panels on the left and right columns report estimates for positive and negative turnover events, respectively. The top panels show the impact on middle managers' allocation quality, and the bottom ones show the impact on that of rank-and-file employees.

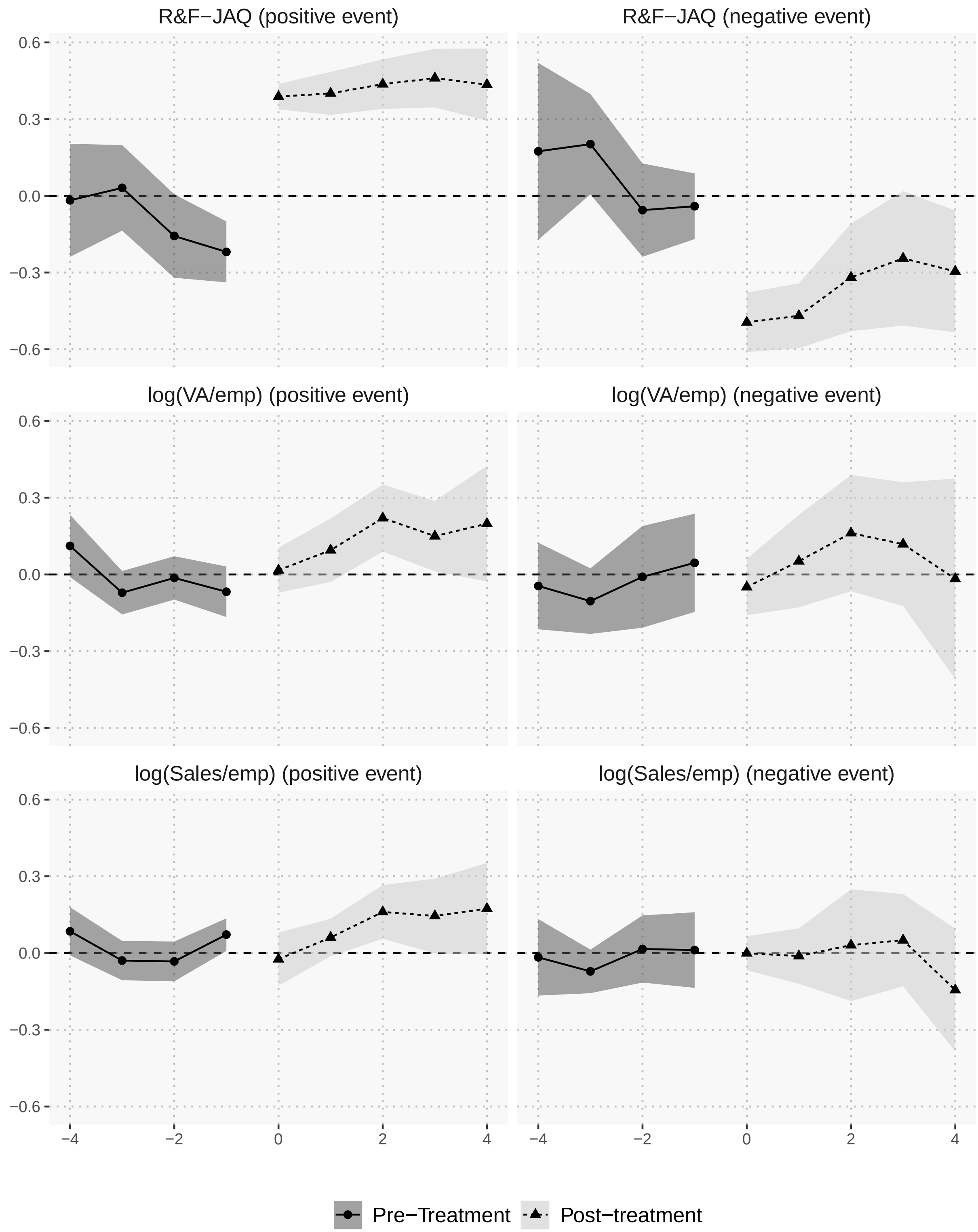


Figure A4: ***JAQ*** **and productivity response to events leading to extreme** ***R&F-JAQ*** **changes**
The left and right panels of this figure respectively show the response of rank-and-file workers' *JAQ*, log(VA/emp) and log(Sales/emp) to positive and negative managerial turnover events. The event study coefficients are estimated using Callaway and Sant'Anna's method (2021). The managerial turnover events are defined as in Section 5, but are restricted to events associated with large impact effects on *R&F-JAQ*, defined as those in the top quartile of positive events (left panel) and in the bottom quartile of negative events (right panel).

Table A1: Firm performance and *aJAQ*

The table reports the estimated relationship between productivity and *aJAQ*. 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 are clustered at firm level and are shown in parentheses.

	Log(sales/emp)	Log(VA/emp)	OROA	Log(sales/emp)	Log(VA/emp)	OROA
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
aJAQ	0.096***	0.052***	0.023***	0.059***	0.077***	0.022***
	(0.027)	(0.016)	(0.006)	(0.015)	(0.012)	(0.006)
log(cap/emp)				0.399***	0.228***	-0.020***
				(0.013)	(0.009)	(0.002)
log(emp)				0.002	-0.002	-0.002
				(0.007)	(0.005)	(0.002)
Share emp w/ college			0.136***	0.394***	0.011	
				(0.035)	(0.025)	(0.013)
Industry dummies				✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Observations	42,545	42,232	42,545	42,545	42,232	42,545
No. Firms	7,531	7,485	7,531	7,531	7,485	7,531
LHS Mean	7.351	6.157	0.078	7.351	6.157	0.078
LHS SD	0.773	0.526	0.179	0.773	0.526	0.179
Panel B						
aJAQ	0.094***	0.091***	0.023***	0.065***	0.087***	0.026***
	(0.018)	(0.013)	(0.006)	(0.022)	(0.018)	(0.010)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Firm controls				✓	✓	✓
Observations	42,545	42,232	42,545	27,209	27,030	27,209
No. Firms	7,531	7,485	7,531	6,154	6,121	6,154
LHS Mean	7.351	6.157	0.078	7.374	6.186	0.089
LHS SD	0.773	0.526	0.179	0.773	0.521	0.185
Panel C						
aJAQ	-0.010	0.025*	0.021***	0.016	0.055***	0.020*
	(0.019)	(0.014)	(0.007)	(0.024)	(0.020)	(0.011)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Observations	42,545	42,232	42,545	27,209	27,030	27,209
No. Firms	7,531	7,485	7,531	6,154	6,121	6,154
LHS Mean	7.351	6.157	0.078	7.374	6.186	0.089
LHS SD	0.773	0.526	0.179	0.773	0.521	0.185

Table A2: *JAQ^R* and productivity

The table reports the estimated relationship between productivity and the three firm-level measures of job allocation quality, when the ML algorithm is estimated on a random subsample of firms: JAQ^R (panel A) and $pJAQ^R$ (panel B). The panels replicate panel C from Table 5 and Table 6, respectively. The regressions include controls for firms' occupation structure (the fraction of workers in firm f assigned to job j in year t), 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), and worker characteristics (listed in the notes to Table 3). Standard errors are clustered at firm level and are shown in parentheses.

	Log(sales/emp)	Log(VA/emp)	OROA	Log(sales/emp)	Log(VA/emp)	OROA
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
JAQ^R	0.086***	0.045***	0.006	0.065***	0.042***	0.016**
	(0.018)	(0.013)	(0.006)	(0.018)	(0.015)	(0.007)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Observations	48,519	48,172	48,519	29,988	29,804	29,988
No. Firms	7,865	7,820	7,865	6,338	6,308	6,338
LHS mean	7.374	6.211	0.085	7.410	6.246	0.095
LHS SD	0.822	0.590	0.188	0.820	0.585	0.194
Panel B						
$pJAQ^R$	0.161***	0.071*	0.032*	0.171***	0.087**	0.041*
	(0.051)	(0.037)	(0.019)	(0.045)	(0.036)	(0.021)
Industry dummies	✓	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓	✓
Municipality dummies	✓	✓	✓	✓	✓	✓
Occupations	✓	✓	✓	✓	✓	✓
Workers X	✓	✓	✓	✓	✓	✓
Firm Z				✓	✓	✓
Size-industry bin dummies	✓	✓	✓	✓	✓	✓
Observations	48,519	48,172	48,519	29,988	29,804	29,988
No. Firms	7865.000	7820.000	7865.000	6,338	6,308	6,338
LHS mean	7.374	6.211	0.085	7.410	6.246	0.095
LHS SD	0.822	0.590	0.188	0.820	0.585	0.194