ESRI Discussion Paper Series No.352

Management Practices Meet Labor Market Outcomes

Ryo Kambayashi
Taisuke Kameda
Takuma Kawamoto
Shigeru Sugihara
Mari Tanaka

November 2019

Economic and Social Research Institute
Cabinet Office
Tokyo, Japan

Management Practices Meet Labor Market Outcomes*

Ryo Kambayashi† Taisuke Kameda‡ Takuma Kawamoto§
Shigeru Sugihara¶ Mari Tanaka‖

November, 2019

Abstract

There has been a longstanding interest in explaining labor market outcomes by firm management practices. However, direct empirical evidence of the link between them remains limited. We combine a large scale management survey of Japanese establishments covering two periods with the corresponding employee data on wages and working hours. We obtain two main findings. First, workers become more likely to work more than moderate hours of overtime in establishments that introduce more structured bonuses and promotion practices. However, workers become less likely to work excessively long overtime hours in establishments that adopt structured monitoring and targeting practices. Second, there is a reduction in the hourly wage gap by tenure, especially among long-tenured male and short-tenured female workers, in those establishments more inclined to quickly dismiss or reassign underperforming workers. Overall, we find that the adoption of structured management practices is associated with the narrowing of disparities in working hours and wages within establishments.

Keywords: Human Resource Management, Management Practices, Hours of work, Wage

*Acknowledgement: The authors would like to thank Nick Bloom, Catherine Buffington, Renata Lemos, and Atushi Ohyama for valuable input in designing the Management and Organizational Practice Survey in Japan (JP-MOPS), and conference and seminar participants at Kochi University of Technology, the Kansai Labor Seminar, the 2018 Transpacific Labor Seminars, the 2018 Empirical Management Conference, Tokyo Labor Workshop, Contract Theory Workshop, and Western Economic Association International 2019 International Conference for their useful comments and suggestions.

†Hitotsubashi University, Institute of Economic Research.
‡Cabinet Office of Japan.
§Cabinet Office of Japan.
¶Nihon University, College of Economics.
‖Hitotsubashi University, Graduate School of Economics.

1 Introduction

Labor economists have long interpreted stable relationships between labor market outcomes as evidence for key economic mechanisms. For instance, the positive tenure-wage profile has been considered as evidence supporting human capital accumulation, agency, and job-matching models (Becker, 1962; Lazear, 1981; Jovanovic, 1979; Baker, Gibbs and Holmstrom, 1994a,b). In reality, these statistical relationships are driven by company management practices, such as training, payment, recruitment and retention policies. However, evidence linking labor market outcomes and direct measures of management practices remains very limited, and this lack of evidence prevents managers from learning how to improve labor outcomes by changing their practices (Lazear, 1999; Ichniowski and Shaw, 2003). Recently, there has been methodological developments to obtain direct measures of management practices using large scale surveys of firm managers (Bloom and Reenen, 2007; Bloom et al., 2019). Nevertheless, these are rarely linked to worker behavior.

This paper connects the two strands of this literature by matching a large panel data of establishment management practices with the hours worked and wages earned by their individual employees. The data on establishment-level management practices is from the manufacturers' sample of the Management and Organizational Practice Survey in Japan (JP-MOPS), conducted by the Japanese Cabinet Office in 2017. This survey is a part of the international MOPS project and its questionnaire is carefully harmonized with that of the 2015 US-MOPS (Bloom et al., 2019). As in the 2015 US-MOPS, the JP-MOPS contains 16 management questions on monitoring and targets (component of non-human resource management, non-HRM) and incentives (component of human resource management, HRM) and asks establishments about their practices in 2010 and 2015. This management practice data is then matched with the information on the wages and hours of individual workers working in these same establishments around these years from the Basic Survey on Wage Structure (BSWS). The matched sample includes observations of multiple workers within establishments due to the two-stage sampling frame of the BSWS.

Exploiting the establishment-level panel structure in the employee-year-level data, we obtain two main results. First, the adoption of more structured management practices is associated with the concentration of the distribution of overtime hours around the average level for the establishment. This apparently arises through the interplay between HRM (in particular, bonus and promotion) and non-HRM (monitoring and targeting) practices. The adoption of more structured bonuses and promotion practices that evaluate workers based on individual performance predicts the higher

probability of working above short-to-medium overtime hours. This relationship is stronger among short-tenured workers with potentially larger promotion incentives. Conversely, the introduction of more structured monitoring and targeting practices predicts the lower likelihood of working excessively long hours, presumably because these practices identify and reduce those problems triggering long overtime (e.g. production bottlenecks).[1] Accordingly, this empirical relationship is stronger among workers whose characteristics predict a greater propensity to work excessively long hours (e.g., male short-tenured workers). Second, establishments using more structured worker displacement practices tend to reduce the wage gaps across their workers. In particular, we reveal a flattening tenure-wage profile in those establishments more inclined to quickly displace under-performing workers, as predicted by standard models of the tenure-wage profile (e.g., Lazear, 1981). In addition, in these establishments, the wages of long-tenured male workers decrease, while the wages of short-tenured female workers tend to weakly increase, resulting in a narrowing gender-wage gap.

Overall, the results show that better management practices are accompanied with less disparity among workers within establishments in terms of both hours and wages. This is consistent with the hypothesis in laboratory experiments that better management technologies lead to leveling workloads and earnings across workers (Bewley, 2000; Bartling and von Siemens, 2011). This leveling of workload across workers may improve firm productivity, assuming that the marginal product of labor of an individual diminishes as working hours increase, as empirically supported by Brachet, David and Drechsler (2012); Pencavel (2015, 2016), and Collewet and Sauermann (2012). This could then at least partly explain why structured management practices improve firm performance as shown in previous studies (e.g., Bloom et al., 2013). Moreover, we show that not only HRM (structured incentives) but also non-HRM (structured monitoring and targets) help explain individual working hours. This calls for greater attention to the non-HRM practices of firms as a means of achieving some labor market policy goals (e.g., reducing excessive overtime work).

The remainder of the paper proceeds as follows. Section 2 reviews the literature as a means to clarify the contribution of this analysis and Section 3 presents the data. Sections 4 describes our analysis of working hours and management practices, and Section 5 details the results for wages and management practices. Section 6 provides additional evidence on within-establishment

[1] Work design theories of management literature describe such a link between management practices and working hours. For example, see the literature survey by Grant and Parker (2009).

distributional outcomes, and section 7 concludes.

2 Related literature

This paper broadly relates to three strands of the literature. The first is the literature on labor demand-side factors determining working hours. Most previous studies in this area have focused on employers' responses to tax and regulations on working hours and compensation schedules (e.g., Hamermesh, 1993; Trejo, 1991; Crépon and Kramarz, 2002).[2][3] However, empirical studies relating working hours to firm HRM are limited. Among these, Gicheva (2013) and Frederiksen, Kato and Smith (2018) show that longer working hours predict the increased probability of future promotion and wage growth, which they interpret as evidence supporting firm usage of certain promotion incentive mechanisms. Our contributions to this literature are threefold. First, we link direct measures of HRM practices (bonus and promotion policies) obtained from an establishment survey to working hours. Second, we examine the roles of non-HRM practices (monitoring and targets) in determining working hours. Lastly, we analyze the changes in the distribution of working hours *within* establishments, a matter rarely examined in extant empirical studies.

Second, we build on an existing rich empirical literature exploring the labor demand-side determinants of wages. Within this, a group of studies closely related to ours has documented stable positive relationships between wages and tenure and interpreted it as evidence supporting models of human capital accumulation, agency problems, and job matching (Becker, 1962; Lazear, 1981; Jovanovic, 1979; Baker, Gibbs and Holmstrom, 1994a,b). However, the empirical evidence linking wages and direct measures of management practice remains limited to those focusing on either a small number of firms or firms within a specific industry (e.g., Lazear, 2000). We contribute to this literature by linking direct measures of managerial practices to the within-establishment wage gap, especially tenure-wage curves and gender-wage gaps, for a large sample of establishments. Our results are consistent with some theoretical predictions in the literature (e.g., Lazear, 1981). Another group of studies decomposes the variation in wages into firm- and worker-specific components using large employer-employee matched data (Abowd, Kramarz and Margolis, 1999; Song et al., 2019). For the most part, these studies demonstrate the importance of both firm- and worker-

[2] In Japan, the standard workweek is 40 hours per week. An overtime premium of at least 25% must then be paid for overtime of less than 60 hours per month, with an overtime premium of at least 50% for overtime of more than 60 hours.

[3] There are also studies concerning the business cycle as a determinant of hours of work (Prescott, 2004; Rogerson, 2006).

specific components as determinants of wages, thereby explaining trends in wage inequality across and within firms. Our results complement this evidence by showing that management practice, as a presumably important establishment-specific component, also explains the variation in wages across and within establishments.[4]

Finally, our paper builds on empirical studies using direct measures of both HR and non-HR management practices using large-scale surveys of firm managers (Bloom and Reenen, 2007; Bloom et al., 2019). These studies show that management practices explain a range of firm performance measures, including productivity, profitability, and growth. There is also a substantial body of evidence concerning the causal effect of this relationship using field experiments of certain sets of firms (e.g., Bloom et al., 2013; Bruhn, Karlan and Schoar, 2018). Nevertheless, works that relate management practices to worker-level outcomes are limited. An exception is Bender et al. (2018) using cross-sectional data on management practices matched with employee earning records, which shows a positive association between management scores and average worker skills mostly explained by managerial human capital. We add to this literature some evidence newly using a panel of data on management practices matched with employee working hours and wages.

3 Data

The Management and Organizational Practices Survey (MOPS) is an internationally coordinated governmental survey that collects information on the management practices of business. Since the first survey was conducted in 2011 by the US Census Bureau for the manufacturing sector (Buffington et al., 2017), several other countries have joined and conducted near identical surveys, including Pakistan in 2015 (Lemos et al., 2016), the second wave in the US in 2016 (Buffington et al., 2017; Bloom et al., 2019), and in the UK in 2017 (Office for National Statistics, 2018). In Japan, the Economic and Social Research Institute (ESRI) of the Japanese Cabinet Office funded and directed the Japanese version of the MOPS (JP-MOPS) administered in January 2017.

The original US-MOPS contained 16 management questions concerning (non-HRM) monitoring and targets, and (HRM) incentives as in the World Management Survey (WMS) initiated by Bloom and Reenen (2007). Each question in the JP-MOPS is the precise Japanese language translation of those in the 2015 US-MOPS. The questionnaire asks about practices in the base year (2015 in the case of JP-MOPS) and retrospectively about practices for the five years before (2010 in the case of

[4] As discussed later, one limitation of our study is that the structure of the Japanese data does not permit us to decompose wages into separate components determined by firms and workers.

the JP-MOPS). Responses are obtained from persons responsible for overall management practices in the establishment, typically the president, executive officers, plant managers, and general or department managers (see Table A.1 in the appendix). Less than 0.4% of respondents are from human resource-related departments, and around 90% of respondents have tenures longer than 5 years with a median tenure of about 21 years.

One difference in survey design to the original US-MOPS is that the JP-MOPS only includes establishments employing 30 or more workers. This is mainly because the MOPS was expected to be matched with other governmental surveys to minimize respondent burden; detailed information about establishments in the Japanese Census of Manufacturers is also only collected for establishments with 30 or more workers. The sample is based on the 2015 Development of Establishment Frame Database, which is a complete census of Japanese establishments in 2015, where some 56,237 establishments met the criterion for private manufacturing with 30 or more employees. We conducted stratified sampling by two-digit industry and size, with the questionnaire being sent to 36,052 establishments in the second week of January 2017. After two follow-up calls, we finally received 11,405 responses, representing 31.6 percent. In general, the establishments in the JP-MOPS are relatively larger than the establishments in the 2014 Japanese Census of Manufacturers (see Table A.2 in the appendix).

The (non-HRM) monitoring and targets section in the survey asks establishments about their collection and utilization of information about the production process.[5] A question about monitoring asks "What best describes what happened at this establishment when a problem in the production process arose?", and the response options include "We fixed it but did not take further action". The other questions on monitoring are about the variety of key performance indicators (KPIs) collected, the frequency with which these were reviewed, and the existence of display boards indicating the KPIs. Questions about targets ask "How easy or difficult was it for this establishment to achieve its production targets", for which the available responses range from "Possible to achieve without much effort" to "Only possible to achieve with extraordinary effort". The remaining questions about targets ask firms about the time frame of the production targets and who is aware of the targets.

For the HRM questions, the survey asks whether there are any performance bonuses, and if so, the determinants of the bonuses (either individual, team, establishment, and/or firm performance)

[5]For the full MOPS questionnaire, see `https://www.census.gov/programs-surveys/mops/technical-ocumentation/questionnaires.html` for the US-MOPS and `http://www.esri.go.jp/jp/prj/current_research/service/manage/chosa_seizo.pdf` for the JP-MOPS.

and what proportion of workers are covered by the bonuses. The questions on promotion concern the primary manner in which non-managers and managers are promoted, either solely on performance or ability and/or other factors, such as tenure or family connections. Finally, the questions about displacement ask "When was an underperforming non-manager reassigned or dismissed at this establishment?", to which each respondent selects a response from "Within 6 months of identifying non-manager under performance", "After 6 months of identifying non-manager under performance", and "Rarely or never". The same question was also asked about managers.

Following the literature, we first score the responses to each question (for details, see Table A.4 in the appendix). For the monitoring sections, higher score is assigned when the management uses more KPIs, more frequently reviews them, and continuously improves production process. The scores for production targets are higher if they are a composite of long- and short-term perspectives with a moderate level of difficulty to be achieved. The scores for the bonus sections are higher when bonuses cover a wider range of workers and depend on individual rather than team, establishment, or company performance. For the promotion section, scores are higher if promotion is based solely on performance and ability rather than other factors such as tenure and family connections. The scores for the questions about displacement are higher if the establishment is more inclined to reassign or dismiss underperforming workers quickly.

Elsewhere, Bloom et al. (2019) employ a single score referred to as structural management, which is the unweighted average of the scores of the 16 questions. However, while aggregation may be appropriate for the analysis of firm performance, it may not be ideal when analyzing worker outcomes. Given HRM practices have been considered as the main channel through which management affects worker outcomes in the literature of personnel economics, we separate the responses into HRM and non-HRM practices. We then further disaggregate the HRM practices into those related to wages and employment. More concretely, we use three scores by taking the unweighted average of the scores across three groups of questions: monitoring and targets (MOPS Questions 1-8), bonus and promotion (MOPS Questions 9-14), and displacement (MOPS Questions 15-16).[6]

We match the JP-MOPS data to the Basic Survey on Wage Structure (BSWS) using the establishment identifiers provided in the government census. The BSWS is an employer-employee matched survey conducted annually by the Japanese Ministry of Labor, Health, and Welfare

[6]We show that further disaggregation provides qualitatively similar results, although the statistical power is reduced.

(MLHW) on June 30. The main information we use from the BSWS is the employee-level wage and hours worked. The sample size is large, covering around one million workers across about fifty thousand establishments each year. The sampling is done in two stages. First, establishments are selected by a uniform sampling method from private establishments with 5 employees or more. Second, employees are selected by a uniform sampling method within each establishment selected in the first stage. The survey is a repeated cross-section, covering around 5% of establishments across Japan each year. The survey asks employees about scheduled/overtime salary for the month of June and the total amount of bonuses for the previous year.[7] It also collects information about the scheduled hours worked as well as the overtime hours worked for the month of June. Thus, we can compute the hourly wage precisely by dividing the total payment per month by the actual hours worked in the month. In addition, the survey collects worker attributes such as education, gender, tenure, and age. Throughout our analysis, we limit our attention to full-time workers strictly less than 60 years of age. We also exclude employees in managerial positions, anticipating differing effects on management to other employees.[8] Furthermore, we match the data to the Census of Manufacturer, which is a survey of all manufacturing establishments with four or more employees by the Japanese Ministry of Economy, Trade, and Industry. This link to the manufacturing census enables us to control for establishment-level demand shocks using production data such as material, fuel, and electricity cost.

We construct panel data for the management-employee matched data as follows.[9] The JP-MOPS includes retrospective questions about management practices in 2010 in addition to those in 2015, allowing us to construct establishment-level panel data of management practices. We then match the JP-MOPS data on management practices at the end of 2010 and 2015 to the two years of employee outcomes (in the BSWS) during June 2010-2011 and 2015-2016, respectively.[10] Among the original 225,984 establishment-year observations in the BSWS in 2010, 2011, 2015, and 2016, there are 40,789 manufacturing and private establishments. Among these, we can match 5,771 establishment-year observations to the JP-MOPS, yielding the matching ratio of 14% from the BSWS.[11] From the JP-MOPS side, as the original sample size is 11,427 establishments, the

[7] BSWS is essentially a copy of firm payroll records. In Japan, there is a set of necessary information each labor contract must include and employers are legally obliged to retain this information as the payroll record.

[8] These are employees whose title is recorded as either *Kacho* (boss or supervisor) or *Bucho* (department, section, or division head) in the BSBW. They account for around 3% of employees with strictly less than 60 years of age.

[9] The data has a panel structure only at the establishment and not the employee level.

[10] We assume that management practices tend to be stable over half a year.

[11] This corresponds to 202,755 employee-year observations in the BSWS successfully matched to the JP-MOPS data, yielding a matching ratio of about 19%. Note that the BSWS includes establishments with less than 30 employees.

matching ratio is about 51%. In the end, we have a sample of 78,994 employee-year observations in the BSWS for the 940 establishments whose management practices (JP-MOPS) are observed during both first (2010-2011) and second (2015-2016) periods. We examine the differences in labor outcomes between the matched and unmatched samples and find no systematic differences unexplained by worker characteristics (see Table A.3 in the appendix). We then further match the Census of Manufactures from 2010-2011 and 2015-2016 to use the information on the total cost of material, fuel, and electricity for controlling establishment-level demand shocks.

Table A.5 in the appendix provides summary statistics of the establishment-level variables for the balanced sample of establishments. The mean of the overall management score in 2015 is 0.503 with a standard deviation of 0.171. Between 2010 and 2015, this improved by about 0.045 on average, although some 33% of establishments reported no change in their average management score. It is also notable that a not negligible proportion (11%) of establishments reported a deterioration in their management scores (for the distribution, see Figure A.1 in the appendix). Furthermore, there is variation in the directions and magnitudes of the changes in the monitoring and targets, bonus and promotion, and displacement scores within establishments, while they are positively correlated overall (see Figure A.5 in the appendix).[12]

What factors do drive the changes in management practices? Although this question is not our main interest in this paper, it is a background useful for understanding our main results. Table A.6 in the appendix details the results of a decomposition analysis of the changes in overall management practices by industry and regions. Some 17% of the variation is explained by industry fixed effects, while prefecture fixed effects explain just 2% of these changes, suggesting that a possibility of knowledge spillovers within industries. In the sample of establishments matched with intermediate inputs in the Census of Manufacturers, the changes in the intermediate inputs explain little of the changes in management practices, after controlling for the industry and prefecture fixed effects. A similar result is obtained when we regress the changes in management practices on production shocks in the baseline years (2009–2011) measured by the changes in the log of shipment values in these years (see Table A.7 in the appendix for the results).

Summary statistics for the employee-level variables are in Table A.8 in the appendix.

[12]The displacement score is substantially lower in Japan than in the US: only 0.24 in the JP-MOPS compared to 0.51 in the US-MOPS (Buffington et al., 2017). Strikingly, only about 66.3% (73.0%) of the JP-MOPS sample responded with a "rarely or never" when asked whether an underperforming non-manager (manager) was reassigned or dismissed, compared to 33.2% (42.8%) in the US-MOPS. Given such a stark difference, together with the widespread long-term employment in the Japanese employment system (Hashimoto and Raisian, 1985), it is possible that establishment characteristics relating to worker displacement play a distinct role in determining wages.

4 Overtime work and management practices

We first examine the association between changes in management practice scores and changes in the distribution of working hours in the establishment. From 2010/2011 to 2015/2016, owing to a period of economic expansion in Japan, average overtime hours increased by about 1.41 hours. However, there is significant variation in the length of overtime work across workers (see Table A.9 in the appendix). While many workers work less than 10 hours (45% in 2015), there is also a non-negligible number of workers working more than 50 hours (8% in 2015). The latter has been lately considered a major policy concern, partly as a possible cause of health problems and work-related stress, which in the worst case, may result in suicides. In general, junior workers tend to work longer overtime hours than senior workers: the median overtime of workers with less than 10 years of tenures is 13 hours, whereas that for workers with longer tenures is 10 hours. Male workers work longer overtime hours than female workers.

To begin, we simply regressing overtime hours on management practice scores using the employee-establishment matched data. Standard errors are clustered at the establishment level. Table 1 shows the results. Column (1) regress overtime hours on the overall management score, i.e., unweighted average of the 16 scores, controlling only for year fixed effects and basic worker characteristics that are measured by age, age squared, tenure, tenure squared, a female dummy, and four education dummies. The standard errors are clustered at the level of establishments. The results show that the usage of more structured management practices positively and significantly predicts longer overtime hours. However, as shown in column (2), additionally controlling for establishment fixed effects drastically reduces the coefficient and increases the standard error, making the estimated coefficient negative and statistically insignificant. Noting that management practices could be correlated with establishment-level demand shocks, in column (3) we additionally control for the amount of flexible inputs used in production, as measured by the log of the total cost of intermediate inputs (material, fuel, and electricity). This further reduces the size of the coefficient. These results highlight the importance of controlling for establishment fixed effects for understanding within-establishments variation in management practices and hours.

The insignificant results may be potentially explained by the aggregation of the management scores as each type of practices may influence working hours differently. In column (4), we disaggregate the overall management score into its three components, namely, the monitoring and targets, bonus and promotion, and displacement scores. As expected, the monitoring and target score coef-

ficient is negative, while the bonus and promotion score coefficient is positive, and both coefficients are relatively large. Nonetheless, the standard errors also remain large. This is potentially because of heterogeneous effects of these management practices across workers, particularly by the level of overtime hours worked. In particular, even if management practices do not alter the *average* overtime hours in the establishment, they may affect the *distribution* of overtime hours within the establishment. For example, if structural management practices identify and reduce production bottlenecks, they may affect only the upper tail of the overtime hour distribution. Therefore, we now examine the changes in the distribution of overtime hours, not just the changes in the mean overtime hours.

We start by graphically illustrating how the changes in the distribution of overtime hours relate to the changes in management practices of establishments. To do this, we divide the sample of establishments into two groups according to whether the change in overall management score from 2010-2015 is below or above the median of the change in the sample. Figure 1 depicts the distribution of overtime hours among male workers[13] according to the two groups of establishments. For both groups of establishments, overtime hours increased on average during this period, possibly reflects overall economic expansion.[14] However, there are differences in the way hours increased across the two groups. In establishments with relatively more improved management scores, the increase in positive overtime hours mostly corresponds to an increase in the share of workers working between 10 and 50 hours overtime. In contrast, in establishments with less improved management scores, the increase in positive overtime hours largely corresponds to an increase in individuals working overtime for more than 50 hours, combined with a slight increase in workers working overtime for 20 and 50 hours.

We next explore this relationship in a regression framework controlling for observed worker and establishment characteristics. As the graphical illustration suggests, note that the marginal effect of the management practice score on overtime hours can be heterogeneous depending on the level of overtime hours. Therefore, we use an indicator of working overtime for more than k hours, where $k = 5, 10, \cdots, 50$, for individual i in establishment j in year t, denoted by $\mathbb{1}(OH_{ijt} > k)$, as the dependent variable. We estimate the following equation:

[13] In this graphical illustration, we focus on the sample of male workers because the incidence of long working hours seems concentrated among male workers (as shown in Table A.9 in the appendix).

[14] The cases of zero overtime hours are included in Figure 1 as 0-10 overtime hours. The share of workers doing any positive hours of overtime increased by 0.042 in establishments with less-improved management scores and by 0.034 in establishments with more-improved management scores.

$$\mathbb{1}(OH_{ijt} > k) = Management'_{jt}\beta^{hk} + X'_{ijt}\gamma^k + \delta^k Demand_{jt} + \phi^k_j + \phi^k_t + \mu^k_{ijt} \quad (1)$$

for each $k = 5, 10, \cdots, 50$. $Management_{jt}$ is a vector of establishment j's management scores in year t.[15] X_{ijt} is a vector of worker i's attributes including age, age squared, tenure, tenure squared, dummy variables for educational attainment, and gender. As in Table 1, we control for establishment demand shocks ($Demand_{jt}$) using the log of the cost for intermediate inputs, measured by the total cost for material, fuel and electricity, as a proxy of flexible inputs. We also control for year and establishment fixed effects to control for unobserved economy wide shocks and time-invariant establishment-level factors affecting overtime hours. The standard errors are clustered at the level of establishments.

Table 2 provides the results. The estimated coefficients of the bonus and promotion score are positive and statistically significant for overtime of more than 5–35 hours. These magnitudes are large, implying that an increase in the bonus and promotion score by 0.1 (roughly equivalent to a one-standard deviation of the changes in the score[16]) is associated with a 7% increase in the share of workers working more than 20 hours. The magnitudes of the coefficients peak around working overtime above 10–20 hours, become smaller for longer hours, and finally become insignificant for overtime of more than 40 hours. Conversely, the coefficients of monitoring and targets score are negative for all ranges. The coefficients are close to zero for overtime above 5-10 hours, and become larger for longer hours. In particular, for overtime of above 40–50 hours, the coefficients are statistically significant at a 10% level. The coefficients' sizes are economically large: an improvement in the monitoring and targets score by 0.1 (roughly equivalent to a one-standard deviation of the changes in the score[17]) is associated with a 8% reduction in each of the share of workers working more than 50 hours.

One interpretation of the results for the bonus and promotion score is that the introduction of individual performance-based bonus and promotion schemes induces some workers to exert greater effort in the form of increasing working hours. This is consistent with the literature showing

[15] The subscript t indicates one of the four years, 2010–2011 and 2015–2016. The management practice scores in 2011 (2016) are the same as those in 2010 (2015).

[16] For example, a change in the bonus and promotion score by 0.1 can be roughly achieved if an establishment changed the non-managers' bonus system to be based on team performance rather than establishment performance and modified the non-managers' promotion system to be based solely on individual performance rather than both individual performance and tenure.

[17] For example, a change in the monitoring and target score by 0.1 can be achieved if an establishment increased the number of KPIs collected from 1–2 to 3–9 and started to show the KPIs not only in one display board but in multiple display boards.

that performance-based pay improves productivity, presumably by inducing greater worker effort (e.g. see the literature survey by Bloom and Reenen (2011)). An interesting point to note in our results is that the bonus and promotion score coefficients are smaller for working longer hours and become insignificant for excessively long hours. This could be because the marginal product of labor decreases in hours (as per evidence by Brachet, David and Drechsler, 2012; Pencavel, 2015, 2016; and Collewet and Sauermann, 2012), which predicts that payment schemes based on individual output would not encourage workers to work beyond that level where the marginal effort cost exceeds the marginal increase in payment. This implies that workers who have already worked long hours may not respond to changes in the incentive scheme.

For structured monitoring and targets, two key features of the results facilitate our interpretation. First, note that all of the coefficients are negative. One explanation is that these practices eliminate waste in production. For instances, collecting and reviewing detailed data on production progress may alert managers of problems like bottlenecks that could trigger overtime. Continuously improving a production system could also reduce the frequency of problems causing overtime. Second, the coefficients for shorter hours are small and become larger for longer hours. An interpretation of this result is that these practices may also allow leveling of workload across workers and time. In a simple world where all workers have the same concave individual-level production function of working hours, the most efficient workload allocation is for every worker to work the same hours everyday. However, actual workload can be heterogeneous across workers and days for various reasons such as miscommunication. In such a case, leveling of workload across workers and time improves allocative efficiency. In practice, setting and sharing targets at a moderate level of difficulty would help adjust workloads, thereby reducing excessively long overtime.

Alternative explanation is that the higher monitoring score is associated with better monitoring of working hours.[18] If some workers overreport hours, better monitoring of hours may reduce the reported hours to the level of actual hours. Recall that our data on working hours is in principal a payroll record of the establishment about the sampled workers, hence, it indicates the hours based on which the establishment paid to the workers. Therefore, under this alternative explanation, the results implies that an improvement of monitoring score eliminates a part of establishment's wage bill that used to be paid for producing nothing. However, we argue that this alternative

[18]Higher monitoring score should not be directly interpreted as better monitoring of hours. Note that questions on monitoring are asked about production, but not about workers. In addition, the respondents are mostly top-level managers of the establishments, and there were only few respondents from human resource management departments. Even so, there is a possibility that monitoring practices about production and hours are correlated.

story does not explain our whole results. The reason is that, under this story, the coefficients of monitoring and targets should be universal across all levels of working hours (because workers should have incentive to upwardly misreport at any range of working hours), while we observe large and significant coefficients only for long overtime hours.

The results are robust to alternative specifications. First, we find similar results when we change the variable controlling for demand shock from intermediate input cost to the log of establishment's sales (see Table A.10 in the appendix). Second, we test if the results are driven by recall errors in the retrospective questions on practices in 2010. According to Bloom et al. (2019), a comparison of management practice scores in 2010 using 2010 and 2015 US-MOPS, which asked the same retrospective questions as ours, shows that the tenure at the establishment of the manager responding to the survey is an important determinant of recall errors. In particular, they show that response quality is high if the respondent had been in the establishment since at least one year before the period of the recall. We find that our results are qualitatively unchanged when we restrict the sample to the JP-MOPS respondents whose tenure is longer than 7 years (see Table A.11 in the appendix). Lastly, we estimate an alternative specification by quantile regressions. Note that, while our baseline specification allows us to examine the changes in the likelihood of working above certain levels of hours, it does not fully speak to the changes in the shapes and locations of within-establishment distributions of overtime hours. Quantile regressions shed light on the latter aspects. Following the quantile regression method with panel fixed effects proposed by Canay (2011), we use the residuals after mean regression of individual overtime hours on establishment fixed effects and the log of the total cost for intermediate inputs. We then regress the residual on the change in overall management scores from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term by a quantile regression. The estimated coefficients on the interaction term, summarized in Figure A.2 in the appendix, indicate that an improvement in overall management scores is associated with small increases in overtime hours at around 45 and 50 percentiles of the distributions and large decreases at around 80–95 percentiles.[19] When we repeat a similar process for estimating quantile regression for each of the three scores (Figure A.4 in the appendix for the results), we find that the former is driven by the changes in bonus and promotion scores, and the latter is mostly driven by monitoring and targets scores, consistent with our baseline results.

Next, we explore the heterogeneity in the results across workers. First, we examine the dif-

[19]Since the distributions of overtime hours are truncated at 0, it is noteworthy that non-zero changes at lower percentiles of the distributions are observed only for a limited fraction of establishments. Therefore, the results for lower quantiles may not be informative for selection issues (see Figure A.3 in the appendix for more details).

ferences by tenure. In general, the incentive for promotion is likely to be highest in earlier career stages and, therefore, should affect shorter tenured workers most. To examine this possibility, in Panel A of Table 3, we estimate the equation for overtime above 10 ("short"), 30 ("medium"), and 50 ("long") hours using the sample divided by tenure group, with tenure being less than the median (10 years) ("junior") or not ("senior") . As expected, the bonus and promotion score coefficients are larger for shorter tenured group for all levels of overtime. For the short-tenured group, the bonus and promotion coefficient is positive and significant, even for overtime exceeding 50 hours and of sizable magnitude, indicating that an increase in score by 0.1 is associated with a 10% increase in the share of workers working these long hours. As a robustness check, we run the same specification by disaggregating the bonus (MOPS Questions 9–12) and promotion scores (Questions 13–14), and find that the coefficients of promotion are larger for the short-tenured group, and sizable for the overtime exceeding 30 hours, although insignificant (see Table A.12 in the appendix).[20]

Interestingly, the negative coefficient of monitoring and targets score is large and significant for overtime of above 50 hours for junior workers but smaller and insignificant for senior workers. The coefficient for junior workers implies that an increase in the score by 0.1 is associated with a 12% reduction in the share of workers working these hours. One interpretation is that firm managers can control the working hours of workers who tend to overreact to incentives by using monitoring and targets to restrain them from working excessively long hours. When we alternatively divide the sample by ages, we also find that younger workers are more likely to work long hours and the negative coefficient of monitoring and targets for long hours is larger for them (see Table A.13 in the appendix).

Second, we divide the sample by gender. Panel B in Table 3 provides the results. Overall, female workers tend to work less overtime across the full range of overtime hours. The bonus and promotion coefficients are similar for male and female workers for short overtime, while they are larger for female than male workers for medium and long overtime hours. The coefficients for female workers are sizable, for example, indicating that an increase in the bonus and promotion score by 0.1 is associated with a 23% (37%) increase in the share of workers working more than medium (long) hours, respectively. As a result, the change is associated with narrower gender gap in the share of workers working medium and long overtime. One explanation is that performance-based

[20]The insignificant results for the promotion score are possibly because of the high correlation between the changes in bonus and promotion scores. For example, regressing the change in promotion scores on the change in the bonus score in the MOPS establishment-level sample, the coefficient is 0.23 with a standard error of 0.06. However, regressing the change in the displacement score on the change in the bonus score gives a much smaller coefficient of 0.14 with only a slightly higher standard error of 0.07.

bonuses and promotion policies help break the gender glass ceiling which may prevail otherwise, for example under tenure-based policies. Another interpretation is that working hours signal the workers' commitment to the firm, and such signal is stronger for female workers, consistent with the model and evidence provided by (Kato, Ogawa and Owan, 2017). As for the monitoring and targets score, both male and female coefficients are negative and large for long overtime, although the coefficient of female workers is erroneous and statistically insignificant possibly due to a small share of female workers working long overtime (3%).

Overall, the heterogeneous results thus far suggest that results for monitoring and targets appear to be stronger among workers tending to work more overtime hours (i.e., short-tenured workers). We explore this possibility further by dividing the sample of workers by their propensities for working overtime over 10, 30, and 50 hours. By estimating the probability of working overtime above 10, 30, and 50 hours conditional on gender, age, tenure, and education level, we exploit the full variation in workers' observed characteristics. For this exercise, we use the 2009 BSWS data, being the year prior to our baseline year, and estimate probit models. The results yield some basic characteristics of workers working overtime hours (see Table A.14 in the appendix). Consistent with previous results, junior and male workers tend to work longer overtime hours. In addition, the probability of working overtime is increasing but strongly concave in age, and college graduates tend to work less overtime than high school graduates. We use these estimated probit models to predict the propensity of working overtime above these hours for all workers in our analysis sample of 2010–2011 and 2015–2016. We then divide the sample by the median of the estimated propensity scores. Columns (1)–(2) and (3)–(4) in Panel C of Table 3 describe the results for working more than short and medium overtime, respectively. The coefficients are similar for the two groups. The results for long overtime hours are in columns (5)–(6). The monitoring and targets score coefficients are negative and significant for workers with a greater propensity to work long hours and insignificant for the other workers. The coefficient is sizable for workers with a greater propensity to work long hours, suggesting that a 0.1 increase in the monitoring and targets score is associated with a 9.7% reduction in the share. This result supports our conjecture that structured monitoring and targeting practices control for excessive amounts of overtime by workers who tend to do more of it.

5 Wages and Management Practices

We next examine the association between management practices and wages using panel regression analysis. We first estimate the following wage regression with management scores:

$$w_{ijt} = Management_{jt}^{'}\beta^{w} + X_{ijt}^{'}\gamma^{w} + \delta^{w}Demand_{jt} + \xi_{j} + \xi_{t} + \epsilon_{ijt} \tag{2}$$

where w_{ijt} is the log of the hourly wage of employee i in establishment j in year t. We calculate the hourly wage by dividing the annual monthly salary including monthly bonus payments by the total hours of work. $Management_{jt}$ is a vector of management scores of establishment j in year t. $X_{ijt}$are human capital variables that include the same worker attributes as the analysis of working hours. We also control for establishment-specific demand shocks by including the amount of flexible inputs used, measured by the log of the total cost for material, electricity, and fuel; δ_j are establishment-fixed effects and δ_t are year fixed effects. The standard errors are clustered at the establishment level.

To start, column (1) in Table 4 shows the results of estimating equation (2) without controlling for establishment fixed effects. The estimated coefficient is positive (0.213) and statistically significant (se=0.032), implying that hourly wages are positively correlated with the overall management scores of the establishments. This result may reflect certain patterns of cross-sectional matching and sorting of workers across establishments: for example, establishments using more structured management practices may attract workers with higher unobserved skills. This result is consistent with Bender et al. (2018), who find similar evidence in Germany.

From column (2) onwards, we focus on the within-establishment variations in management scores and wages over five years by controlling for establishment fixed effects. We find that the estimated coefficient of the overall management score is small (-0.048) and insignificant (se=0.049), respectively. As in our analysis of hours, in column (3), we decompose the overall score into three components: monitoring and target, bonus and promotion, and displacement scores. The result in column (3) shows that the coefficient on displacement score is negative and significant, and the rest of the coefficients are small and insignificant. This result gives us a conjecture that separation policy may matter for wage structure, possibly through deferred compensation schemes.

Figure 2 illustrates the change in tenure-wage profiles by the change in the displacement scores. We divide the sample of establishments into two groups according to the degree in which the displacement score changed from 2010 to 2015. The graph on the right plots the establishments in

which changes in the displacement scores are higher than the sample median, and that on the left uses the remaining sample. We plot the residuals of the log of hourly wages of male workers after controlling for age, age squared, and education dummies across ten equal-sized bins of tenure and for two periods (2005–2006 and 2010–2011). In establishments with less-improved displacement score (the plot on the left), the tenure-wage profile changes only a little over time. However, in establishments with more-improved displacement score (the plot on the right), the tenure-wage profile becomes flatter.

This result is consistent with the HRM literature. For example, both Lazear (1979) and Frederiksen and Takats (2011) predict that deferred compensation requires employment security. Intuitively, once an employer changes its displacement policy to fire workers more frequently, junior workers may become dissatisfied with a steep tenure-wage profile. We examine this relationship for workers in different tenure groups and confirm this using regression analysis. We divide employees into four groups by tenure so that the groups are roughly equal sized. Columns (4)-(7) in Table 4 estimate the same equation as in column (3) but by dividing the sample into the four tenure groups based on quantiles. The results indicate that the negative association of the displacement score and wages is statistically significant among the group of long tenured workers (columns (6)–(7)). The coefficient for longest tenured group implies that an increase in displacement score by 0.2 (roughly equivalent to a one-standard deviation of the changes in the score[21]) is associated with a reduction in wage by 1.2% (=[Exp(-0.06×0.2)-1]×100). In contrast, among shortest-tenured workers, the estimated correlation is positive and small (column (4)), although insignificant. The results remain qualitatively the same when we alternatively control for the log of establishment sales instead of the cost of intermediate goods (see Table A.15 in the appendix). Overall, the results imply that the wage-tenure profile becomes flatter in establishments that are inclined to displace workers more quickly.

Lastly, we divide the sample further by gender, and Table 5 shows the results. The negative coefficient for displacement score is strongest among male workers, particularly among those workers with long tenures (column (3) and (4)). In contrast, the coefficients of displacement score are overall smaller among female workers, and positive for short tenured female workers, although insignificant (column (5)). Note that on average the longest-tenured male workers receive the highest wages and shortest-tenured female workers receive the lowest wages among the tenure-

[21]For example, an increase in displacement score by 0.2 is roughly equivalent to an increase in the score for an establishment that changed displacement policies of under-performing non-managers from "rarely reassigning or dismissing" to "reassigning or dismissing 6 months of identifying non-manager underperformance".

gender groups. Therefore, this suggests that an improvement in the displacement score is associated with a reduction in the within-establishment wage gaps across both gender and tenure. The gender difference in the results can be explained by the steeper tenure-wage profile of male workers than female workers, as indicated by the average log of wages at the bottom of Table 5. One explanation for the steeper tenure-wage profile of male workers is their longer expected length of tenure. As the establishment utilizes outside-labor market more, the equilibrium steepness of tenure-wage profile may become flatter, especially for male workers whose expected length of tenure decline more.

An alternate possibility is that these results are affected by the changes in the composition of workers through hiring and separation, which may be partly driven by the changes in displacement policy. Using the employee data from the BSWS, we compute the share of workers hired in the past five years using the information on tenure. We then decompose these into new graduate or mid-career hires by age. The remainder, those whose tenure is more than five years, are those who have stayed in the establishment for the preceding five years. One hypothesis is that the faster displacement occurs, we see fewer workers staying in the establishments, in which case there may be negative or positive selection of those remaining in terms of wages (Jovanovic, 1979). However, the results of regressing the share of "stayers" (those whose tenure is longer than five years) on the management practices scores does not support this hypothesis, with the estimated coefficients for the displacement and other scores being insignificant (see Table A.16 in the appendix). Furthermore, the results do not suggest that establishments' changing management practices are more or less likely to hire young or mid-career workers.[22] As an additional exercise, we use the BSWS establishment level survey asking about the hiring of new graduates. Using the number of new graduates hired for each establishment and those by gender, we cannot identify a significant association with the management practice scores (see Table A.17 in the appendix). Overall, our results are best interpreted as short-run relationships between management practices and labor outcomes before substantial compositional adjustments take place.

6 Inequality of hours and wages within establishments

Previous results hint that improving management scores are associated with declining inequality of hours and wages within establishments. We directly test this hypothesis in two alternative ways.

[22]The coefficients are all small and insignificant, but the coefficients on monitoring and targets are negative for the fraction of hired workers, while it is positive for the fraction of stayers. This result may reflect a potential weak relationship between an introduction of these practices and a reduction in turnover rates.

The first is to use quantile regressions. For overtime hours, we have describes the results in section 4 that improvements in overall management scores are associated with a large decline of overtime hours at 90 and 95 percentiles of within-establishment distributions, accompanied by moderate increases in overtime hours between 65 and 80 percentiles. For hourly wages, the results using the same procedure of quantile regressions indicate that improvements in overall management scores are associated with declines of wages between 70 and 95 percentiles of within-establishment distributions and an increase at 5 percentile (for the results, see Figure A.6 in the appendix).

The second way to test implications on within-establishment inequality is to regress establishment-year level inequality indicators on the overall management score, controlling for year, establishment fixed effects, and demand shocks. In Table 6, we first regress the average overtime hours worked in the establishment on the overall management score in column (1). The coefficient is positive and insignificant consistent with the results in Table 1. In column (2), the dependent variable is the difference in overtime hours between the 95 and 5 percentile in the establishment.[23] The coefficient is negative (-6.47) and insignificant with a large standard error (9.25). The large standard error is possibly because of heterogeneity in the level of overtime hours across establishments. To address this issue, in column (3) we use an alternate variable obtained by dividing the gap by the average overtime hours in the establishment. The coefficient for the overall management score is again negative (-2.23) but statistically significant with a lower standard error (1.01) than before. We conduct a similar exercise for wages. Column (4) shows the results of regressing the average of $log(wage)$ in the establishment on the overall management score. The coefficient is small and insignificant. In column (5), the coefficient for the 5-95 percentile gap of $log(wage)$ is negative and significant. Dividing this gap by the average of log wage in the establishment, the results are qualitatively identical. Overall, the results suggest that establishments that improved management scores narrowed the gaps between high and low tails of distributions of both overtime hours and wages.

7 Concluding remarks

There has long been an interest in explaining observed labor market outcomes using firm management practices. However, to date, empirical evidence linking labor market outcomes and direct measures of management practices has been limited to studies focusing on a small number of firms

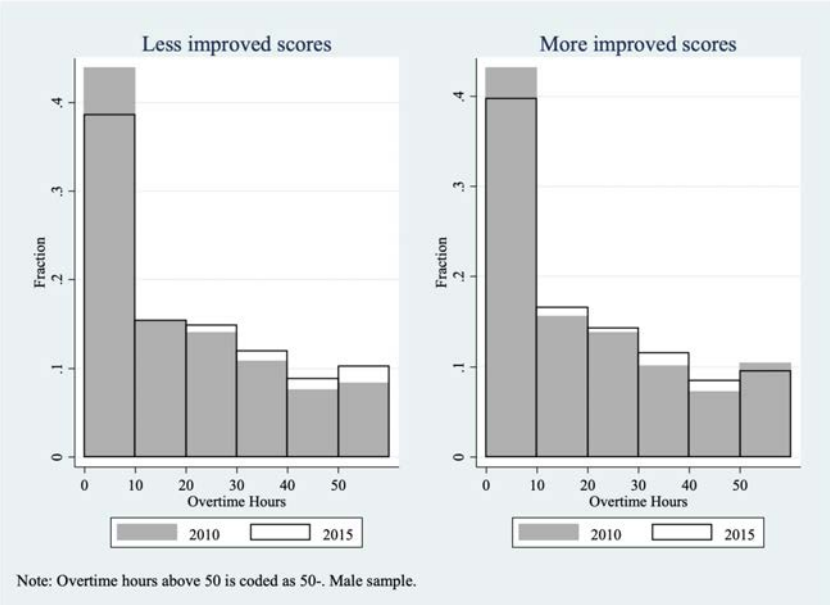
[23] In columns (2) and (3), we restrict the sample to establishments where there is at least one worker working overtime in the sample.

or a specific industry. This is despite recent empirical studies using large scale surveys of firm managers showing that these practices strongly explain firm performance. In this paper, we employed a comparable large panel data on the management practices of establishments in the Japanese manufacturing sector and matched it with the data of their employees on working hours and wages.

We obtain two main results. First, the adoption of more structured bonuses and promotion practices (HRM) is associated with increasing a moderate length of overtime. This is consistent with standard theoretical predictions that compensation and promotion schemes tied to individual performance induce worker effort. However, we also find that the introduction of more structured monitoring and target practices (non-HRM) is associated with reductions in excessively long overtime. One explanation is that collecting and reviewing detailed data and continuous improvements in the production process allow leveling of workload across workers and time and reducing those problems triggering long overtime. In addition, setting production targets at a moderate level of difficulty and sharing them across workers should help adjust workloads across workers and time, also reducing excessively long overtime. Overall, a combination of HRM and non-HRM practices would seem to play a key role in narrowing the disparity in hours across workers within establishments.

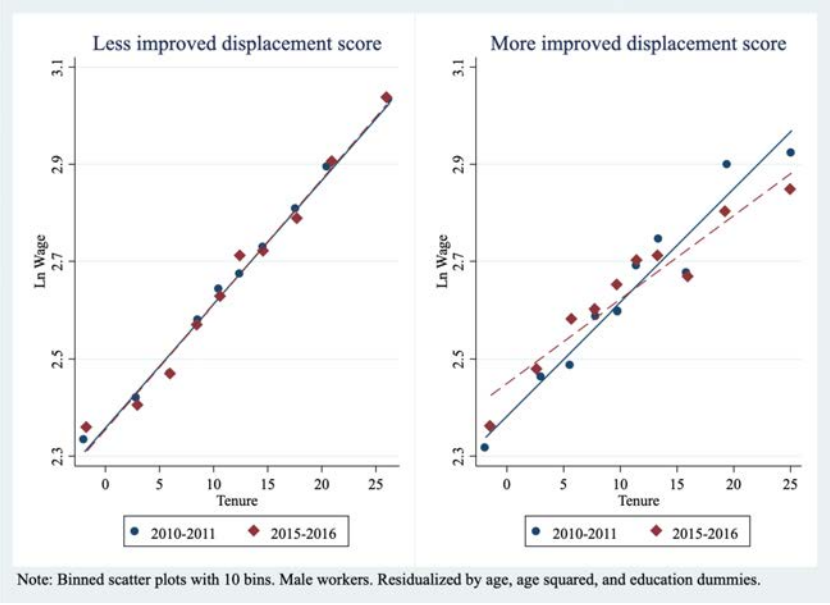
Second, firms using more structured displacement practices tend to experience declining within-establishment wage gaps, especially by tenure and gender. In particular, we find that the tenure-wage profile becomes flatter in establishments more inclined to quickly displace underperforming workers, as predicted by standard models of tenure-wage profiles (e.g., Lazear, 1979). In addition, in these establishments, we identify the strongest wage decline among long-tenured male workers, accompanied with a weak wage increase among short-tenured female workers. Combined together, we find that the adoption of more structured management practices is associated with reduced wage inequality within establishments.

Figure 1: Distribution of overtime hours by changes in management scores



Notes: We divide the sample of establishments into two groups according to whether the change of overall management scores from 2010 to 2015 is bellow (left figure) or above (right figure) the median of the changes in the sample. The figures shows the histograms of overtime hours among male workers by the two groups of establishments and by periods, 2010–2011 (light gray) and 2015–2016 (transparent). All workers working more than 50 hours are top-coded as above 50 hours.

Figure 2: Tenure-wage profile and displacement score



Notes: We divide the sample of establishments into two groups according to whether the change of displacement scores (based on question 15 and 16 in JP–MOPS) from 2010 to 2015 is bellow (left figure) or above (right figure) the median of the changes in the sample. We plot the mean of residuals of male workers' wages after regressed on age, age squared, education dummies, and year dummies by equally-sized 10 bins of tenure, by the two groups of establishments, and by periods, 2010–2011 (solid line with round points) and 2015–2016 (dashed line with diamond points).

Table 1. Average overtime hours and management practices

	(1)	(2)	(3)	(4)
Variable	Overtime hours			
Overall Management	7.822	-0.675	-1.936	
	(1.412)	(4.497)	(4.757)	
Monitoring–target				-5.603
				(3.655)
Bonus–promotion				9.162
				(4.096)
Displacement				-1.264
				(2.327)
Year FE	Yes	Yes	Yes	Yes
Worker attributes controls	Yes	Yes	Yes	Yes
Establishment FE	No	Yes	Yes	Yes
Demand shock control	No	No	Yes	Yes
Observations	159,927	159,923	117,207	116,374
Mean dep. var.	17.79	17.79	18.40	18.31

Notes: Standard errors are clustered at the level of establishments. Column (1) controls only for year fixed effects and worker attributes (age, age squared, tenure, tenure squared, female dummy, and four education dummies). Columns (2)–(4) additionally control for establishment fixed effects. Columns (3)–(4) additionally control for the log of intermediate input cost.

Table 2. Distribution of overtime hours and management practices

Variables	(1) 𝟙(OH≥5)	(2) 𝟙(OH≥10)	(3) 𝟙(OH≥15)	(4) 𝟙(OH≥20)	(5) 𝟙(OH≥25)	(6) 𝟙(OH≥30)	(7) 𝟙(OH≥35)	(8) 𝟙(OH≥40)	(9) 𝟙(OH≥45)	(10) 𝟙(OH≥50)
Monitoring–target	-0.006	-0.027	-0.071	-0.114	-0.134	-0.132	-0.117	-0.096	-0.080	-0.068
	(0.070)	(0.077)	(0.088)	(0.094)	(0.093)	(0.083)	(0.067)	(0.058)	(0.048)	(0.041)
Bonus–promotion	0.218	0.258	0.257	0.261	0.235	0.179	0.140	0.097	0.057	0.044
	(0.085)	(0.084)	(0.093)	(0.100)	(0.102)	(0.092)	(0.073)	(0.062)	(0.050)	(0.044)
Displacement	-0.072	-0.042	-0.004	-0.014	-0.013	-0.016	-0.034	-0.028	-0.035	-0.021
	(0.057)	(0.055)	(0.055)	(0.051)	(0.050)	(0.046)	(0.040)	(0.036)	(0.032)	(0.030)
Observations	116,374	116,374	116,374	116,374	116,374	116,374	116,374	116,374	116,374	116,374
Mean dep. var.	0.638	0.544	0.464	0.388	0.316	0.256	0.200	0.154	0.109	0.081

Notes: Standard errors are clustered at the level of establishments. All regressions control for year fixed effects, establishment fixed effects, the log of intermediate input cost, and worker attributes (age, age squared, tenure, tenure squared, female dummy, and four education dummies).

Table 3. Distribution of overtime hours and management practices by workers' characteristics

Panel A: By Tenure

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥50)	$\mathbb{1}$(OH≥50)
Sample: tenure	Junior	Senior	Junior	Senior	Junior	Senior
Monitoring–target	-0.052	-0.026	-0.194	-0.078	-0.101	-0.046
	(0.097)	(0.081)	(0.100)	(0.086)	(0.049)	(0.045)
Bonus–promotion	0.332	0.184	0.246	0.132	0.091	0.010
	(0.100)	(0.092)	(0.110)	(0.100)	(0.052)	(0.060)
Displacement	-0.067	-0.022	-0.018	-0.026	-0.047	-0.003
	(0.068)	(0.060)	(0.062)	(0.046)	(0.043)	(0.028)
Observations	54,859	61,486	54,859	61,486	54,859	61,486
Mean dep. var.	0.574	0.517	0.277	0.237	0.091	0.072

Panel B: By Gender

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥50)	$\mathbb{1}$(OH≥50)
Sample: gender	Male	Female	Male	Female	Male	Female
Monitoring–target	-0.029	-0.044	-0.137	-0.107	-0.072	-0.050
	(0.079)	(0.128)	(0.086)	(0.117)	(0.049)	(0.037)
Bonus–promotion	0.238	0.345	0.136	0.288	0.015	0.111
	(0.088)	(0.118)	(0.098)	(0.103)	(0.052)	(0.056)
Displacement	-0.049	-0.008	-0.015	-0.008	-0.019	-0.023
	(0.056)	(0.081)	(0.055)	(0.054)	(0.039)	(0.026)
Observations	89,989	26,263	89,989	26,263	89,989	26,263
Mean dep. var.	0.596	0.367	0.294	0.124	0.096	0.030

Panel C: By Propensity to Work Overtime above the Threshold

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥50)	$\mathbb{1}$(OH≥50)
Sample: propensity	High $\hat{P}$(OH≥10|X)	Low $\hat{P}$(OH≥10|X)	High $\hat{P}$(OH≥30|X)	Low $\hat{P}$(OH≥30|X)	High $\hat{P}$(OH≥50|X)	Low $\hat{P}$(OH≥50|X)
Monitoring–target	-0.005	-0.071	-0.155	-0.107	-0.108	-0.027
	(0.085)	(0.091)	(0.104)	(0.080)	(0.055)	(0.034)
Bonus–promotion	0.281	0.256	0.159	0.216	0.038	0.052
	(0.093)	(0.097)	(0.110)	(0.086)	(0.054)	(0.046)
Displacement	-0.075	-0.007	-0.004	-0.024	-0.037	-0.010
	(0.064)	(0.063)	(0.061)	(0.044)	(0.043)	(0.024)
Observations	61,548	54,800	61,390	54,958	60,568	55,779
Mean dep. var.	0.645	0.430	0.327	0.177	0.111	0.049

Notes: Standard errors are clustered at the level of establishments. In Panel A, columns indicated "Junior" use the sample of workers whose tenure is less than 10 years (=median of tenure in the sample), and columns indicated "Senior" use the rest. In Panel B, columns indicated "Male" use the sample of male workers, and columns indicated "Female" use the female sample. The propensity function $\hat{P}$(OH≥ $h|X$) for each h (=10, 30 and 50) is estimated by probit model using the BSBW data in 2009 (see Appendix Table A.3 for the estimation result), where X is the set of workers' attributes. All regressions control for year fixed effects, establishment fixed effects, the log of intermediate input cost, and female dummy (except for Panel B), tenure, tenure squared, age, age squared, and four education dummies.

Table 4. Wages and management practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable				Ln Wage			
Sample: by tenure	All	All	All	≤3	4–9	10–19	≥20
Overall Management	0.213	-0.058					
	(0.032)	(0.058)					
Monitoring–target			0.012	0.065	-0.013	0.022	0.024
			(0.044)	(0.064)	(0.053)	(0.054)	(0.065)
Bonus–promotion			-0.016	0.012	-0.051	0.008	-0.078
			(0.042)	(0.049)	(0.051)	(0.056)	(0.081)
Displacement			-0.063	0.007	-0.055	-0.079	-0.060
			(0.025)	(0.034)	(0.037)	(0.032)	(0.032)
Observations	117,209	117,207	116,374	25,206	29,487	28,462	32,833
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker attributes controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demand shock control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment FE	No	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are clustered at level of establishments. The demand shock control is the log of intermediate input cost. Worker attributes controls are age, age squared, tenure and tenure squared (only in (1)–(3)), gender, and four education dummies.

Table 5. Wages and management practices by gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable					Ln Wage			
Sample: Gender	Male	Male	Male	Male	Female	Female	Female	Female
Sample: Tenure	≤3	4–9	10–19	≥20	≤3	4–9	10–19	≥20
Monitoring–target	0.064	-0.018	0.018	0.027	-0.006	0.035	0.008	0.009
	(0.080)	(0.060)	(0.057)	(0.067)	(0.072)	(0.074)	(0.086)	(0.094)
Bonus–promotion	0.019	-0.054	-0.001	-0.070	-0.078	-0.099	-0.002	-0.101
	(0.057)	(0.057)	(0.062)	(0.083)	(0.068)	(0.065)	(0.071)	(0.124)
Displacement	-0.014	-0.041	-0.096	-0.069	0.051	-0.034	-0.042	-0.067
	(0.043)	(0.041)	(0.030)	(0.033)	(0.063)	(0.056)	(0.056)	(0.057)
Observations	18,325	22,320	21,713	27,008	6,317	6,610	6,187	5,299

Notes: Standard errors are clustered at the level of establishments. All regressions control for year fixed effects, establishment fixed effects, the log of intermediate input cost, and worker attributes that are age, age squared, and four education dummies.

Table 6. Distributions of Hours and Wages

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Hours	Hours	Hours	Wage	Wage	Wage
Statistics	Mean	5p-95p	(5p-95p)/mean	Mean	5p-95p	(5p-95p)/mean
Overall Management Score	1.964	-6.466	-2.231	0.031	-0.194	-0.072
	(4.599)	(9.249)	(1.012)	(0.056)	(0.096)	(0.034)
Observations	2,757	2,627	2,627	2,757	2,757	2,757
Mean dep. var.	16.91	42.34	3.262	2.896	0.913	0.316

Notes: This table uses the sample aggregated at the level of establishment-year. Standard errors are clustered at the level of establishments. All regressions control for year fixed effects, establishment fixed effects, and the log of intermediate input cost. "5p-95p" denotes the difference between 95 percentile and 5 percentile. "(5p-95p)/mean" denotes the difference between 95 percentile and 5 percentile divided by the average of the establishment.

References

Abowd, John M., Francis Kramarz, and David Margolis. 1999. "High Wage Workers and High Wage Firms." Econometrica, 67(2): 251–333.

Baker, G., M. Gibbs, and B. Holmstrom. 1994a. "The Internal Economics of the Firm: Evidence from Personnel Data." The Quarterly Journal of Economics, 109(4): 881–919.

Baker, G., M. Gibbs, and B. Holmstrom. 1994b. "The Wage Policy of a Firm." The Quarterly Journal of Economics, 109(4): 921–955.

Bartling, Björn, and Ferdinand A. von Siemens. 2011. "Wage inequality and team production: An experimental analysis." Journal of Economic Psychology, 32(1): 1–16.

Becker, Gary S. 1962. "Investment in Human Capital : A Theoretical Analysis." Journal of Political Economy, 70(5): 9–49.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter. 2018. "Management Practices, Workforce Selection, and Productivity." Journal of Labor Economics, 36(1): 371–409.

Bewley, Trumen F. 2000. Why Wages Don't Fall During a Recession. Cambridge, MA and London:Harvard University Press.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." Quarterly Journal of Economics, 122(4): 1351–1408.

Bloom, Nicholas, and John Van Reenen. 2011. "Chapter 19 - Human Resource Management and Productivity." In . Vol. 4 of Handbook of Labor Economics, , ed. David Card and Orley Ashenfelter, 1697 – 1767. Elsevier.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." Quarterly Journal of Economics, 128(1): 1–51.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron S. Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What Drives Differences in Management Practices?" American Economic Review, 109(5): 1648–1683.

Brachet, Tanguy, Guy David, and Andrea M. Drechsler. 2012. "The Effect of Shift Structure on Performance." American Economic Journal: Applied Economics, 4(2): 219– 246.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." Journal of Political Economy, 126(2): 635–687.

Buffington, Catherine, Lucia Foster, Ron Jarmin, and Scott Ohlmacher. 2017. "The Management and Organizational Practices Survey (MOPS): An Overview." Center for Economic Studies, U.S. Census Bureau.

Canay, Ivan A. 2011. "A Simple Approach to Quantile Regression for Panel Data." Econometrics Journal, 14: 368–386.

Collewet, Marion, and Jan Sauermann. 2012. "Working hours and productivity." Labour Economics, 47: 96– 106.

Crépon, Bruno, and Francis Kramarz. 2002. "Employed 40 Hours or Not Employed 39: Lessons from the 1982 Mandatory Reduction of the Workweek." Journal of Political Economy, 110(6): 1355–1389.

Frederiksen, Anders, and Elod Takats. 2011. "Promotions, Dismissals, and Employee Selection: Theory and Evidence." The Journal of Law, Economics, and Organization, 27(1): 159.

Frederiksen, Anders, Takao Kato, and Nina Smith. 2018. "Working Hours and Top Management Appointments: Evidence from Linked Employer-Employee Data." IZA DP No. 11675.

Gicheva, Dora. 2013. "Working Long Hours and Early Career Outcomes in the High-End Labor Market." Journal of Labor Economics, 31(4): 785–824.

Grant, Adam M., and Sharon K. Parker. 2009. "Redesigning Work Design Theories: The Rise of Relational and Proactive Perspectives." The Academy of Management Annals, 3(1): 317–375.

Hamermesh, Daniel S. 1993. Labor Demand. Princeton University Press, New Jersey.

Hashimoto, Masanori, and John Raisian. 1985. "Employment Tenure and Earnings Profiles in Japan and the United States." American Economic Review, 75(4): 721–735.

Ichniowski, Casey, and Kathryn Shaw. 2003. "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." Journal of Economic Perspectives, 17(1): 155–180.

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

Kato, Takao, Hiromasa Ogawa, and Hideo Owan. 2017. "Working Hours, Promotion, and the Gender Gap in the Workplace." mimeo.

Lazear, Edward P. 1979. "Why Is There Mandatory Retirement?" Journal of Political Economy, 87(6): 1261–1284.

Lazear, Edward P. 1981. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." American Economic Review, 71(4): 606–620.

Lazear, Edward P. 1999. "Personnel Economics: Past Lessons and Future Directions Presidential Address to the Society of Labor Economists, San Francisco, May 1, 1998." Journal of Labor Economics, 17(2): 199–236.

Lazear, Edward P. 2000. "Performance Pay and Productivity." American Economic Review, 90(5): 1346–1361.

Lemos, Renata, Ali Choudhary, John Van Reenen, and Nicholas Bloom. 2016. "Management in Pakistan: First Evidence from Punjab." International Growth Center.

Office for National Statistics. 2018. "Management Practices and Productivity in British Production and Services Industries - Initial Results from the Management and Expectations Survey: 2016." https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/labourproductivity/articles, Office for National Statistics.

Pencavel, John. 2015. "The Productivity of Working Hours." Economic Journal, 125: 2052–2076.

Pencavel, John. 2016. "Recovery from Work and the Productivity of Working Hours." Economica, 83: 545–563.

Prescott, Edward C. 2004. "Why Do Americans Work So Much More Than Europeans?" Federal Reserve Bank of Minneapolis Quarterly Review, 28: 2–12.

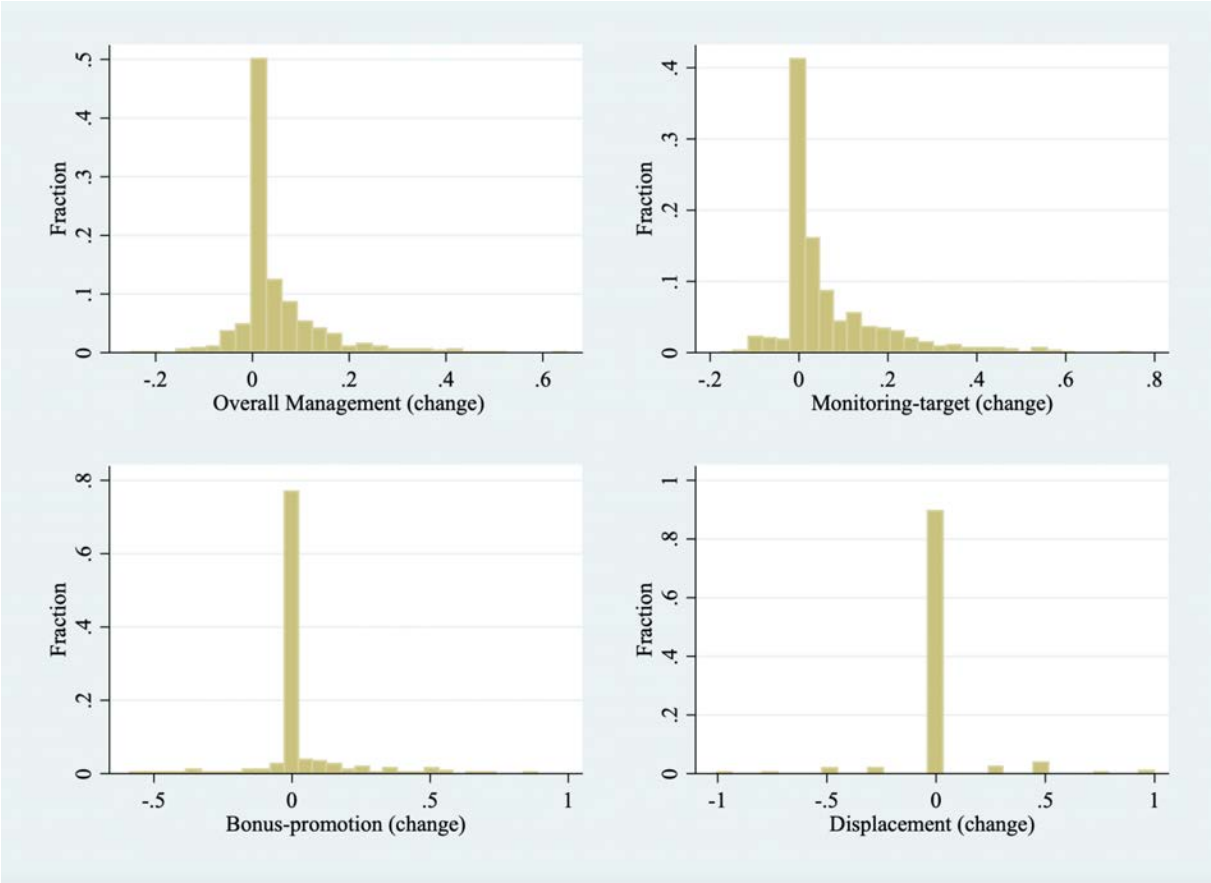
Rogerson, Richard. 2006. "Understanding Differences in Hours Worked." Review of Economic Dynamics, 9: 365–409.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2019. "Firming Up Inequality." Quarterly Journal of Economics, 134(1): 1–50.

Trejo, Stephen J. 1991. "The Effects of Overtime Pay Regulation on Worker Compensation." American Economic Review, 81(4): 719–740.

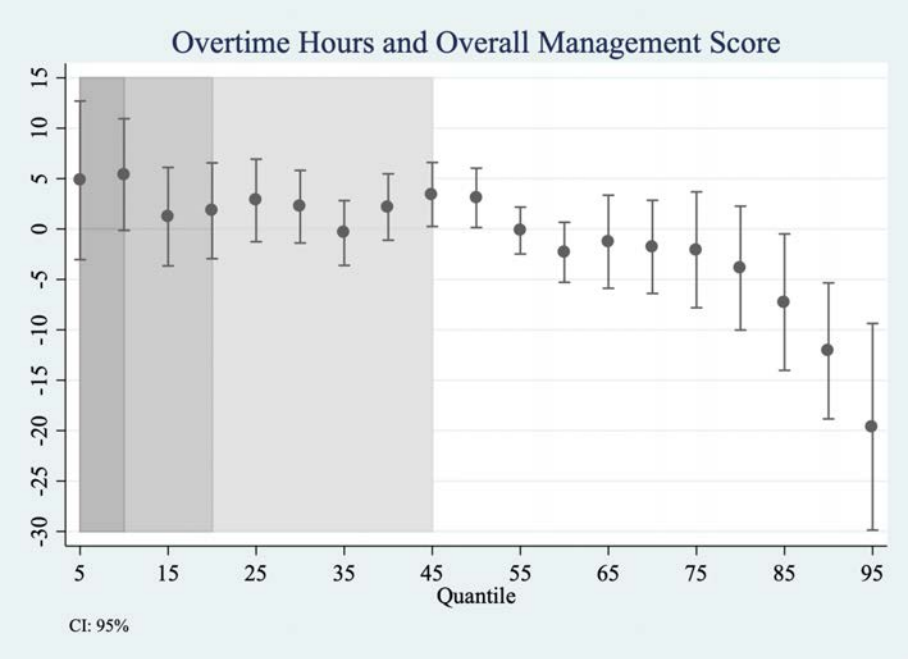
Appendix

Figure A.1. Distribution of changes in scores



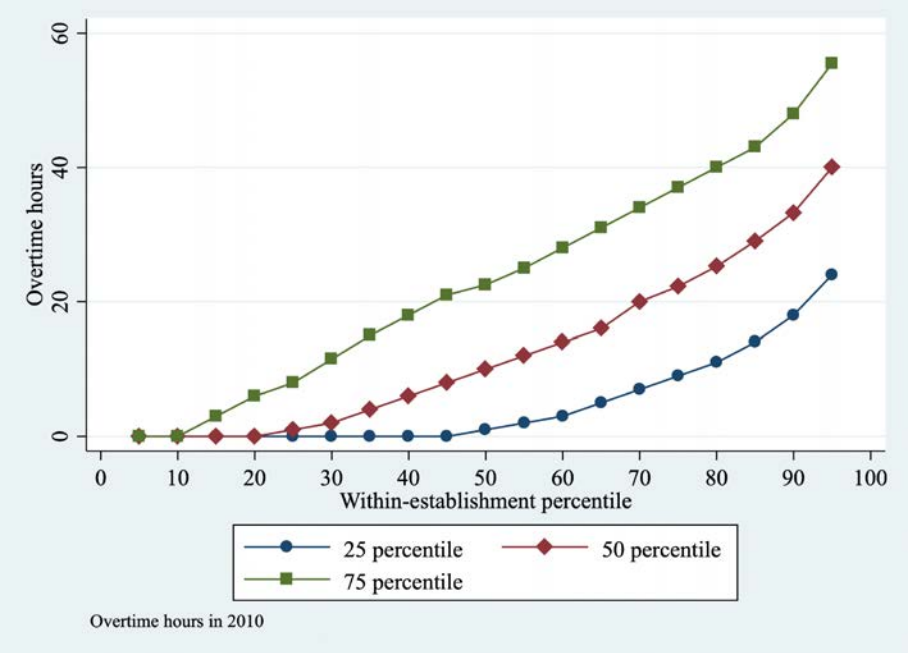
Notes: The figures distribution of establishment-level changes in the management scores using the sample of establishments in 2015 JP-MOPS matched with BSBW data for at least one year during each period of 2010–2011 and 2015–2016.

Figure A.2. Quantile regressions for overtime hours



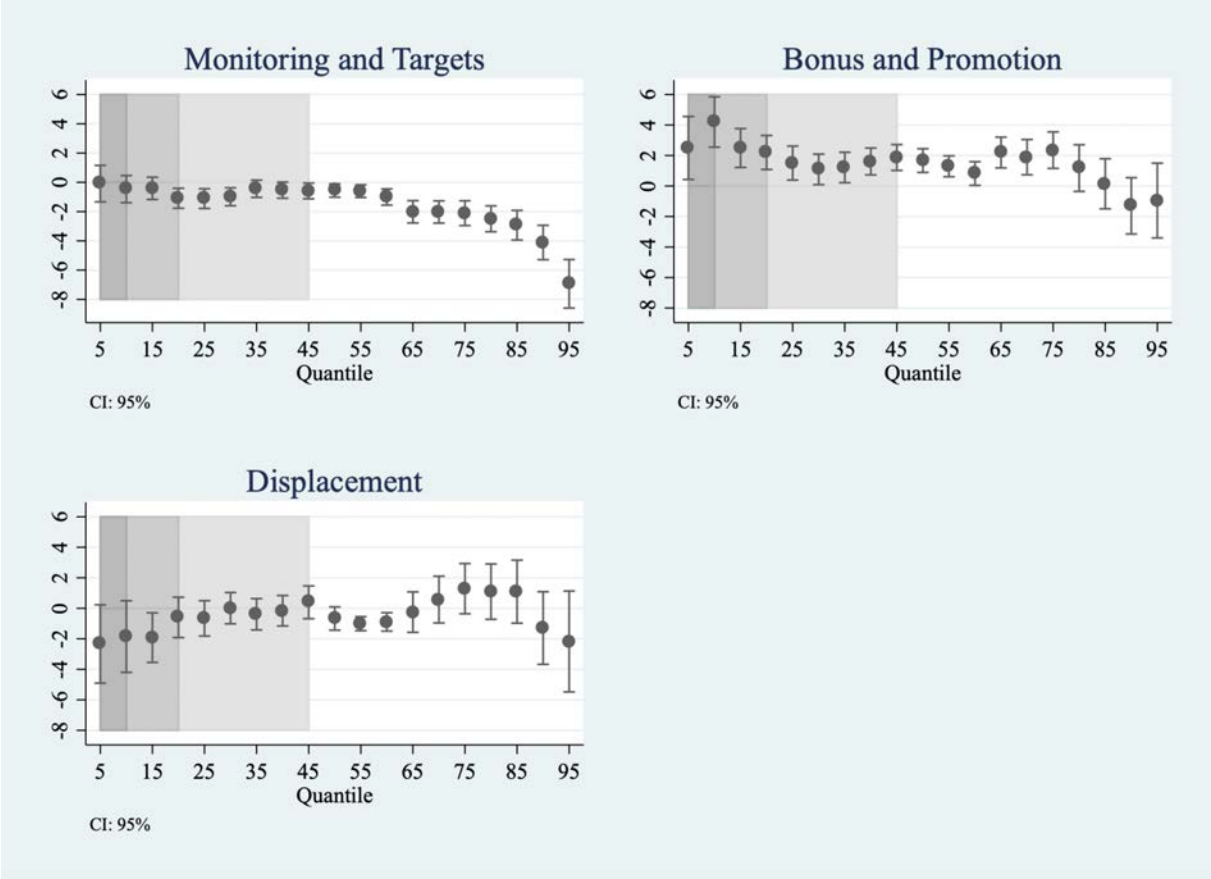
Notes: The figure summarizes the results of quantile regressions for overtime hours. We first obtain the residuals after mean regression of individual overtime hours on establishment fixed effects and the log of the total cost for intermediate goods. We then regress the residuals on the change in the overall management scores from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term by quantile regressions. The estimated coefficients on the interaction term are shown in the figure. 95 percent confidence intervals are calculated based on standard errors estimated by bootstrap with 500 replications. Since the distributions of overtime hours are truncated at 0, non-zero changes at lower percentiles of the distributions are observed only for a limited fraction of establishments. The gray areas indicate the magnitude of this issue by colors. As shown in the figure A.3, 1) about three quarters of establishments have 0 as a 10 percentile of the establishment's distribution of overtime hours, 2) about a half of establishments have 0 as a 20 percentile of the establishment's distribution of overtime hours, and 3) about a quarter of establishments have 0 as a 45 percentile of the establishment's distribution of overtime hours.

FIGURE A.3. WITHIN- AND ACROSS-ESTABLISHMENTS DISTRIBUTION OF OVERTIME HOURS



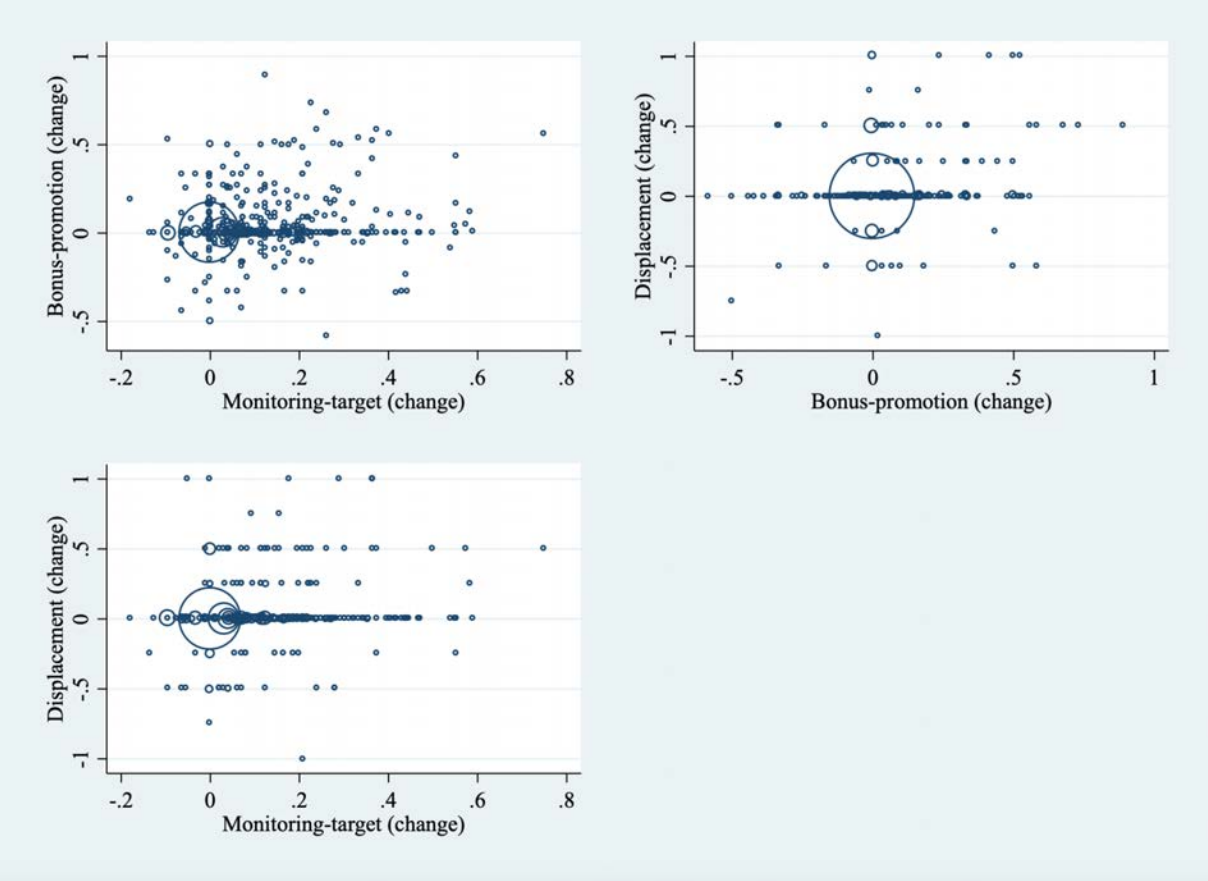
Notes: To make this figure, we first identify overtime hours at 5, 10, ..., and 95 percentiles in each establishment, and then for each percentile shown in x-axis, we plot 25, 50, and 75 percentile of the across-establishment distribution. The figure indicates that 1) about three quarters of establishments have 0 as a 10 percentile of the establishment's distribution of overtime hours, 2) about a half of establishments have 0 as a 20 percentile of the establishment's distribution of overtime hours, and 3) about a quarter of establishments have 0 as a 45 percentile of the establishment's distribution of overtime hours. The figure use the BSBW data of overtime hours in 2010.

Figure A.4. Quantile regressions for overtime hours



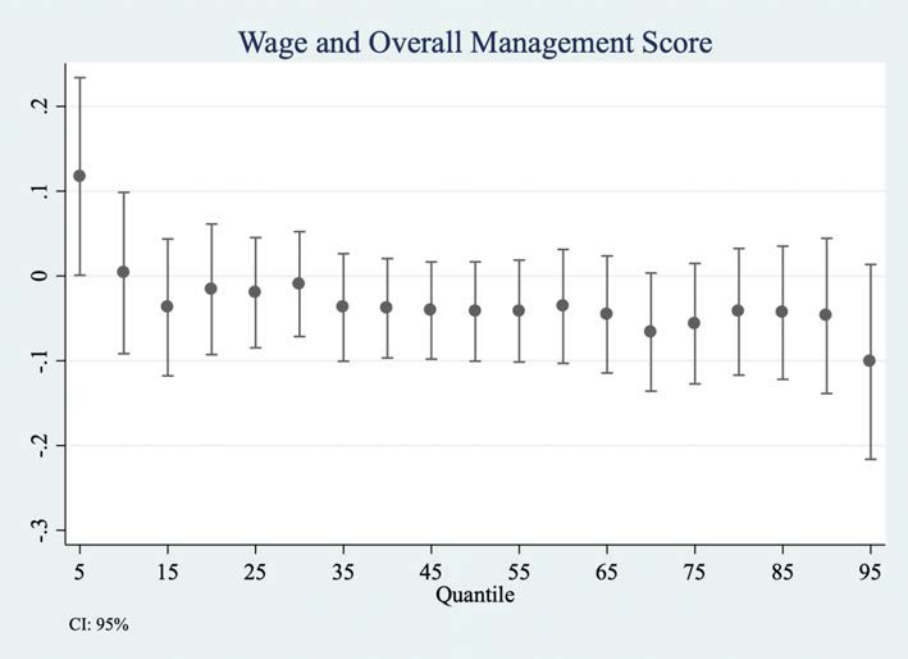
Notes: The figure summarizes the results of quantile regressions for overtime hours for each score. We first obtain the residuals after mean regression of individual overtime hours on establishment fixed effects and the log of intermediate input cost. We then regress the residuals on a dummy variable that takes 1 if the score improved from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term by quantile regressions. The estimated coefficients on the interaction term are shown in the figure. 95 percentile confidence intervals are calculated based on standard errors estimated by bootstrap with 500 replications. Since the distributions of overtime hours are truncated at 0, non-zero changes at lower percentiles of the distributions are observed only for a limited fraction of establishments. The gray areas indicate the magnitude of this issue by colors. As shown in the figure A.3, 1) about three quarters of establishments have 0 as a 10 percentile of the establishment's distribution of overtime hours, 2) about a half of establishments have 0 as a 20 percentile of the establishment's distribution of overtime hours, and 3) about a quarter of establishments have 0 as a 45 percentile of the establishment's distribution of overtime hours.

FIGURE A.5. SCATTERED PLOTS OF CHANGES IN SCORES



Notes: The figures distribution of establishment-level changes in the management scores using the sample of establishments in 2015 JP-MOPS matched with BSBW data for at least one year during each period of 2010–2011 and 2015–2016. The size of the circle reflects the number of observations in the point.

FIGURE A.6. QUANTILE REGRESSIONS FOR WAGES



Notes: The figure summarizes the results of quantile regressions for the log of hourly wage. We first obtain the residuals after mean regression of individual overtime hours on establishment fixed effects and the log of intermediate input cost. We then regress the residuals on the change in the overall management scores from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term by quantile regressions. The estimated coefficients on the interaction term are shown in the figure. 95 percent confidence intervals are calculated based on standard errors estimated by bootstrap with 500 replications.

Table A.1. Titles of JP-MOPS contacts by HR and non-HR related departments

Title	Non-HR	HR	Share of the title
President	566	0	5%
Director/Chairman/Executive Officer	2,601	9	24%
Plant Manager	1,332	4	12%
Manager (e.g. General Manager, Department Manager, excluding Plant Manager)	2,386	70	23%
Section Chief/Head	2,116	185	21%
Others	1,470	108	15 %
Total	10,471	376	100 %

Notes: The table uses the data from JP-MOPS meta data on titles of contacts, asked in a free-answer format. We classified the original titles into the six mutually exclusive groups as indicated in the table. For each title, the column "HR" shows the number of responses from departments/sections related to Human Resource, and "Non-HR" shows those from any of the other categories.

Table A.2. Characteristics of MOPS sample in comparison with the Census of Manufacturers

	MOPS sample (N=9,983)			Non MOPS sample (N=34,390)		
	Mean	S.D.	Median	Mean	S.D.	Median
Number of employees	140	349	70	122	286	60
Paid-up capital	490,030	2,716,451	5,000	303,562	2,264,533	4,212
Tangible assets	178,296	821,277	42,357	134,907	638,852	34,705
Value of production	741,724	3,738,793	144,261	523,660	3,123,739	115,031
Per-employee wage	413	158	396	391	194	374
Fraction of exports	2.98	11.18	0.00	2.55	10.82	0.00

Notes: This table uses the data of the 2014 Japanese Census of Manufacturers that asks about production to establishments with above 30 employees in 2015 Economic Census for Business Frame. The columns under "MOPS sample" indicate the characteristics of the establishments matched with JP-MOPS. Employees include part-time workers and contractual workers. Tangible assets consist of lands, buildings, machines, and equipment. Wage includes basic salary and bonuses.

Table A.3. Differences in wage and hours regressions by matched and unmatched samples

VARIABLES	Ln(Wage)	Ln(Wage)	Ln(Wage)	Hour	Hour	Hour
Sample	All	Male	Female	All	Male	Female
Matched with JP-MOPS	-0.002	-0.001	-0.003	0.105	0.196	-0.180
	(0.004)	(0.004)	(0.004)	(0.199)	(0.224)	(0.214)
Female	-0.289			-7.249		
	(0.002)			(0.097)		
Tenure	0.032	0.033	0.028	-0.113	-0.219	-0.099
	(0.000)	(0.000)	(0.000)	(0.014)	(0.017)	(0.017)
Tenure^2	-0.037	-0.045	-0.022	-0.149	0.135	0.017
	(0.001)	(0.001)	(0.001)	(0.031)	(0.037)	(0.045)
Education (middle school)	-0.107	-0.112	-0.081	2.380	2.359	1.089
	(0.003)	(0.003)	(0.004)	(0.164)	(0.194)	(0.274)
Education (2-year college/technical)	0.081	0.066	0.100	-2.039	-2.160	-1.201
	(0.002)	(0.002)	(0.003)	(0.117)	(0.148)	(0.139)
Education (4-year college)	0.235	0.216	0.267	-5.067	-5.480	-0.892
	(0.002)	(0.002)	(0.004)	(0.116)	(0.124)	(0.180)
Age	0.041	0.047	0.024	0.590	0.910	-0.084
	(0.000)	(0.000)	(0.001)	(0.023)	(0.027)	(0.030)
Age^2	-0.049	-0.052	-0.036	-0.970	-1.416	0.033
	(0.000)	(0.001)	(0.001)	(0.026)	(0.031)	(0.035)
Ln(Employment)	0.088	0.090	0.086	1.057	0.972	1.164
	(0.001)	(0.001)	(0.002)	(0.081)	(0.090)	(0.085)
Observations	746,245	577,121	169,124	746,245	577,121	169,124
R-squared	0.618	0.573	0.526	0.113	0.108	0.061

Notes: Sample of employees in 2010-2011 and 2015-2016 BSWS are used, restricting to those in establishments with more than 30 employees. The table shows the results of regressing the log of hourly wage and overtime hours on workers' characteristics and a dummy variable indicating that the observation is matched with JP-MOPS ("Matched with JP-MOPS"). All regressions control for year fixed effects, industry fixed effects (3-digit codes in BSWS), and prefecture fixed effects. "Employment" is the total number of regular and temporary employment in the establishment.

Table A.4. Scoring MOPS Survey Questions

Question 1: What best describes what happens at your firm when a problem in the production process arises?	
Response	**Score**
We fixed it but did not take further action	1/3
We fixed it and took action to make sure that it did not happen again	2/3
We fixed it and took action to make sure that it did not happen again, and had a continuous improvement process to anticipate problems like these in advance	1
No action was taken	0
Question 2: How many key performance indicators are monitored in your firm?	
Response	**Score**
1-2 key performance indicators	1/3
3-9 key performance indicators	2/3
10 or more key performance indicators	1
No key performance indicators	0
Question 3: How frequently are key performance indicators typically reviewed by managers at your firm?	
Response	**Score**
Yearly	1/6
Quarterly	1/3
Monthly	1/2
Weekly	2/3
Daily	5/6
Hourly or more frequently	1
Never	0
Question 4: How frequently are key performance indicators typically reviewed by non-managers at your firm?	
Response	**Score**
See question 3	See question 3
Question 5: Where are display boards showing service quality, output and other key performance indicators located in your firm?	
Response	**Score**
All display boards were located in one place (e.g. in the store back office or at the end of the production line)	1/2
Display boards were located in multiple places (e.g. at multiple places in the store or establishment)	1
We did not have any display boards	0
Question 6: What best describes the time frame of operational targets at your firm?	
Response	**Score**
Main focus was on short-term (less than one year) targets	1/3
Main focus was on long-term (more than one year) targets	2/3
Combination of short-term and long-term targets	1
No targets	0
Question 7: How easy or difficult is it in your firm for people to typically achieve their operational targets?	
Response	**Score**
Possible to achieve without much effort	0
Possible to achieve with some effort	1/2
Possible to achieve with normal amount of effort	3/4
Possible to achieve with more than normal effort	1
Only possible to achieve with extraordinary effort	1/4
Question 8: Who was aware of the operational targets at your firm?	
Response	**Score**
Only senior managers	0
Most managers and some workers	1/3
Most managers and most workers	2/3
All managers and most workers	1

Table A.4. Scoring MOPS Survey Questions (continued)

Question 9: What are non-managers' performance bonuses usually based on in your firm?	
Response	**Score**
Their own performance	1
Their team or shift performance	3/4
Their local establishment or branch's performance	1/2
Their entire company's performance	1/4
No performance bonuses	0
Question 10: When targets are met, what percent of non-managers received performance bonuses?	
Response	**Score**
0%	1/5
1-33%	2/5
34-66%	3/5
67-99%	4/5
100%	1
Targets not met	0
Question 11: What were managers' performance bonuses usually based on in your firm?	
Response	**Score**
See question 9	See question 9
Question 12: When production targets are met, what percent of managers at your firm received performance bonuses?	
Response	**Score**
See question 10	See question 10
Question 13: What is the primary way non-managers are promoted in your firm?	
Response	**Score**
Promotions are based solely on performance and ability	1
Promotions are based partly on performance and ability, and partly on other factors (for example, tenure or family connections)	2/3
Promotions are based mainly on factors other than performance and ability (for example, tenure or family connections)	1/3
Non-managers are normally not promoted	0
Question 14: What is the primary way managers are promoted in your firm?	
Response	**Score**
See question 13 (Replace "non-manager" with "manager")	See question 13
Question 15: When is an under-performing non-manager usually reassigned or dismissed?	
Response	**Score**
Within 6 months of identifying non-manager underperformance	1
After 6 months of identifying non-manager underperformance	1/2
Rarely or never	0
Question 16: When an under-performing manager is usually reassigned or dismissed?	
Response	**Score**
See question 15 (Replace "non-manager" with "manager")	See question 15

Notes: Management practices are scored by 0-1 scale following Bloom et al. (2019). Questions 3, 4 and 5 are scored at 0 if missing, which typically arises from firms reporting "no performance indicators" to question 2 and skipping to question 6. The rationale for this is that firms with no performance indicators have no managerial or non- managerial review of performance indicators, and have no performance display boards. For questions with multiple possible responses (those with "mark all that apply") the average value was used. Only establishments with at least 10 scored responses were included.

Table A.5. Basic statistics (establishment' level variables)

Variable	Mean	Std. Dev.	Min.	Max.	N
2010 Management practice scores					
Overall Management	0.503	0.171	0	0.937	940
Monitoring–target	0.5	0.223	0	0.957	940
Bonus–promotion	0.617	0.221	0	1	939
Displacement	0.225	0.341	0	1	934
Employment	311.57	543.41	11	6440.362	940
Cost of intermediate inputs	11218.609	48977.826	0	831907.813	869
2015 Management practice scores					
Overall Management	0.548	0.166	0.033	0.952	940
Monitoring–target	0.567	0.22	0	0.957	940
Bonus–promotion	0.64	0.211	0	1	939
Displacement	0.243	0.349	0	1	934
Employment	300.741	526.259	12	6527.674	940
Cost of intermediate inputs	10586.419	43984.231	0	765063.438	898
Changes in Management practice scores 2010-2015					
Overall Management (change)	0.045	0.087	-0.25	0.653	940
Monitoring-target (change)	0.066	0.115	-0.178	0.75	940
Bonus-promotion (change)	0.023	0.122	-0.583	0.89	939
Displacement (change)	0.018	0.163	-1	1	934
Employment (change)	-10.829	138.86	-1826	756	940
Cost of intermediate inputs (change)	-453.522	13442.682	-277463.062	124737.281	867

Notes: This table shows the statistics of establishment-level variables. The sample are the balanced panel of establishments in JP-MOPS (2015) that are observed in BSBW data for at least one year during each period of 2010–2011 and 2015–2016. "Cost of intermediate inputs" is the total cost for material, fuel and electricity in 1 million JPY.

Table A.6. Decomposing changes in management practices

	(1)	(2)	(3)	(4)	(5)	(6)
	Changes in Management practice scores 2010-2015					
Industry FE	Yes	No	No	Yes	Yes	Yes
Size bin FE	No	Yes	No	Yes	Yes	Yes
Prefecture FE	No	No	Yes	Yes	Yes	Yes
Change in flexible input use	No	No	No	No	No	Yes
Sample	All	All	All	All	Input observed	Input observed
Observations	940	940	940	940	834	834
R-squared	0.169	0.023	0.066	0.252	0.294	0.294

Notes: This table examines how much of the changes in management practice scores are explained by industry, firm size, prefecture, and firm-specific demand. Size bin takes one of seven categories of firm size in BSBW. The change in flexible input from 2010 to 2015, measured by the change in the log of the total cost for material, fuel and electricity, is included only in column (6) as a proxy for firm-specific demand. Column (5) restricts the sample to those establishments with non-missing values for the change in flexible input. Industry fixed effects are included at the level of 3-digit industry code in BSBW.

Table A.7. Production shocks and changes in Management Practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Changes in scores from 2010-2015							
	Overall management		Monitoring and targets		Bonus and promotion		Displacement	
D Log(Shipment)	0.0117		0.0201		0.0057		-0.0001	
	(0.0108)		(0.0171)		(0.0142)		(0.0141)	
D Log(Cost of intermediate inputs)		0.0059		0.0091		0.0024		-0.0000
		(0.0075)		(0.0111)		(0.0097)		(0.0100)
Observations	860	827	860	827	859	826	854	821
R-squared	0.0018	0.0009	0.0032	0.0012	0.0002	0.0001	0.0000	0.0000

Notes: This table regress the changes in management practice scores on the production shocks around the baseline years and a change of managers. The production shocks in the baseline years are measured by the average of yearly establishment shipment growth and intermediate input cost (both measured in log differences) from 2009 to 2011. This measure is rescaled to mean 0 and standard deviation 1 to ease the interpretation of the coefficients.

Table A.8. Basic statistics (employees' level variables)

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Wage (hourly, 100 JPY)	21.451	10.263	3.913	468.222	78998
Ln(Wage)	2.966	0.445	1.364	6.149	78998
Overtime hours	18.018	20.275	0	195	78998
Female	0.223	0.416	0	1	78998
Age	38.784	10.893	16	59	78998
Tenure	13.456	10.814	0	44	78998
Middle school	0.032	0.177	0	1	78998
High school	0.675	0.468	0	1	78998
Junior/Technical college	0.096	0.295	0	1	78998
4-years college	0.197	0.398	0	1	78998

Notes: This table shows the statistics of employee-level variables. The sample are the employees in BSBW in the balanced panel of establishments that are matched to the JP-MOPS at least one year in each period of 2010–2011 and 2015–2016.

Table A.9. Overtime hours basic statistics

	Mean	SD	Median	P(OH≥10)	P(OH≥50)	N
Year=2010, 2011	17.32	20.2	10	0.51	0.078	39991
Year=2015, 2016	18.73	20.32	13	0.55	0.083	39007
Tenure≤ 9	19.23	20.79	13	0.56	0.091	36441
Year=2010, 2011	18.77	21.06	13	0.54	0.093	18601
Year=2010, 2011	18.77	21.06	12	0.58	0.088	18601
Tenure≥ 10	16.98	19.77	10	0.51	0.072	42557
Year=2010, 2011	16.06	19.34	9	0.49	0.065	21390
Year=2010, 2011	16.06	19.34	9	0.53	0.079	21390
Male	20.2	21.09	15	0.58	0.096	61369
Year=2010, 2011	19.38	21.06	14	0.56	0.092	31006
Year=2015, 2016	21.04	21.08	16	0.61	0.099	30363
Female	10.42	14.82	4	0.35	0.028	17629
Year=2010, 2011	10.25	14.89	3	0.34	0.029	8985
Year=2015, 2016	10.6	14.74	4	0.36	0.026	8644

Notes: This table shows the statistics of overtime hours. The sample are the employees in BSBW in the balanced panel of establishments that are matched to the JP-MOPS at least one year in each period of 2010–2011 and 2015–2016.

Table A.10. Controlling for sales as an alternative demand shock

	(1)	(2)	(3)
Variables	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥50)
Control	Ln(Sales)	Ln(Sales)	Ln(Sales)
Monitoring–target	-0.031	-0.129	-0.067
	(0.079)	(0.085)	(0.040)
Bonus–promotion	0.225	0.170	0.048
	(0.085)	(0.088)	(0.041))
Displacement	-0.026	-0.015	-0.022
	(0.057)	(0.045)	(0.029)
Observations	118,497	118,497	118,497
Mean dep var	0.542	0.254	0.0801

Notes: Standard errors are clustered at the level of establishments. The regressions control for the log of sales of the establishment in addition to establishment fixed effects and worker attributes that are age, age squared, tenure, tenure squared, female dummy, and four education dummies.

Table A.11. Overtime hours and management practices restricting to respondents who were in the firm in 2010

	(1)	(2)	(3)
Variables	$\mathbb{1}$(OH≥10)	$\mathbb{1}$(OH≥30)	$\mathbb{1}$(OH≥50)
Monitoring–target	-0.023	-0.131	-0.081
	(0.091)	(0.097)	(0.045)
Bonus–promotion	0.232	0.141	0.044
	(0.096)	(0.104)	(0.048)
Displacement	-0.050	-0.052	-0.017
	(0.061)	(0.049)	(0.034)
Observations	93,459	93,459	93,459
Mean dep var	0.539	0.251	0.078

Notes: Standard errors are clustered at the level of establishments. The sample is restricted to the JP-MOPS respondents whose tenure is longer than 7 years. All regressions control for establishment fixed effects, the log of intermediate input cost, and worker attributes that are age, age squared, tenure, tenure squared, female dummy, and four education dummies.

Table A.12. Overtime hours and management practices by tenure, disaggregating bonus and promotion scores

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	𝟙(OH≥10)	𝟙(OH≥10)	𝟙(OH≥30)	𝟙(OH≥30)	𝟙(OH≥50)	𝟙(OH≥50)
Sample: by tenure	Junior	Senior	Junior	Senior	Junior	Senior
Monitoring–target	0.010	-0.001	-0.149	-0.006	-0.113	-0.017
	(0.108)	(0.083)	(0.096)	(0.081)	(0.057)	(0.045)
Bonus	0.245	0.202	0.185	0.117	0.103	-0.013
	(0.088)	(0.075)	(0.099)	(0.076)	(0.069)	(0.055)
Promotion	0.062	-0.057	0.086	-0.042	0.015	0.011
	(0.095)	(0.074)	(0.084)	(0.094)	(0.048)	(0.062)
Displacement	-0.049	-0.007	-0.023	-0.009	-0.048	0.001
	(0.068)	(0.058)	(0.063)	(0.046)	(0.045)	(0.029)
Observations	51,400	58,171	51,400	58,171	51,400	58,171
Mean dep var	0.577	0.523	0.279	0.238	0.0904	0.0713

Notes: Standard errors are clustered at the level of establishments. Columns indicated "Junior" use the sample of workers whose tenure is less than 10 years, and columns indicated "Senior" use the rest. All regressions control for year fixed effects, establishment fixed effects, the log of intermediate input cost, and worker attributes that are age, age squared, female dummy, and four education dummies.

Table A.13. Overtime hours and management practices by age

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	𝟙(OH≥10)	𝟙(OH≥10)	𝟙(OH≥30)	𝟙(OH≥30)	𝟙(OH≥50)	𝟙(OH≥50)
Sample: by age	Young	Mid-career	Young	Mid-career	Young	Mid-career
Monitoring–target	-0.016	-0.053	-0.148	-0.112	-0.083	-0.048
	(0.091)	(0.079)	(0.106)	(0.077)	(0.049)	(0.039)
Bonus–promotion	0.283	0.255	0.181	0.200	0.073	0.025
	(0.094)	(0.092)	(0.110)	(0.088)	(0.050)	(0.049)
Displacement	-0.048	-0.038	-0.009	-0.028	-0.026	-0.016
	(0.060)	(0.059)	(0.052)	(0.048)	(0.033)	(0.030)
Observations	56,813	59,538	56,813	59,538	56,813	59,538
Mean dep var	0.601	0.490	0.294	0.219	0.0958	0.0666

Notes: Standard errors are clustered at the level of establishments. Columns indicated as "Young" use the sample of workers whose age is equal or less than 38 years old (sample median), and columns indicated as "Mid-career" use the rest. All regressions control for year fixed effects, establishment fixed effects, the log of intermediate input cost, and worker attributes that are age, age squared, tenure, tenure squared, female dummy, and four education dummies.

Table A.14. Characteristics of workers working overtime (estimates of propensity score)

	(1)	(2)	(3)
Variables	1(OH≥10)	1(OH≥30)	1(OH≥50)
Age	0.027	0.030	0.038
	(0.002)	(0.003)	(0.004)
Age^2	-0.051	-0.052	-0.057
	(0.003)	(0.003)	(0.005)
Female	-0.380	-0.503	-0.541
	(0.007)	(0.010)	(0.015)
Tenure	-0.006	-0.007	-0.010
	(0.001)	(0.001)	(0.002)
$Tenure^2$	0.007	0.004	0.002
	(0.003)	(0.004)	(0.005)
Middle school	0.035	0.049	0.051
	(0.013)	(0.016)	(0.023)
Junior/Technical college	-0.075	-0.077	-0.048
	(0.010)	(0.013)	(0.018)
College	-0.207	-0.136	-0.143
	(0.007)	(0.009)	(0.013)
Constant	-0.499	-1.261	-2.060
	(0.038)	(0.048)	(0.068)
Observations	222,293	222,293	222,293

Notes: Probit models are estimated for the probability of working overtime above x hours, (OH$\geq x$) for $x = 10, 30, 50$. Full-time workers at non-managerial positions in BSBW 2009 data are used.

Table A.15. Wages and management practices controlling for sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable				Ln Wage			
Sample: by tenure	All	All	All	≤3	4–10	11–21	≥22
Overall Management	0.124	-0.065					
	(0.031)	(0.052)					
Monitoring–target			0.003	0.058	-0.019	0.015	0.018
			(0.042)	(0.063)	(0.052)	(0.050)	(0.064)
Bonus–promotion			-0.018	0.015	-0.037	-0.015	-0.066
			(0.037)	(0.046)	(0.047)	(0.048)	(0.078)
Displacement			-0.058	0.011	-0.043	-0.066	-0.061
			(0.023)	(0.032)	(0.035)	(0.029)	(0.031)
Log(Sales)	0.087	0.037	0.037	0.017	0.051	0.048	0.033
	(0.003)	(0.008)	(0.008)	(0.011)	(0.010)	(0.010)	(0.011)
Observations	117,341	117,339	115,151	24,961	29,083	28,208	32,511
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker attributes controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demand shock control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Establishment FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep var	2.967	2.967	2.963	2.617	2.857	3.013	3.284

Notes: This table reestimate Table 4 by alternatively controlling for the log of establishment sales as a demand shock control variable. Standard errors are clustered at level of establishments. Worker attributes controlled are age, age squared, tenure and tenure squared (only in (1)–(3)), gender, and four education dummies.

Table A.16. Compositional changes of workers

	(1)	(2)	(3)	(4)
	Share of employees who are...			
Tenure	Hired	Hired	Hired	Stayer
Age		Young	Mid-career	
Monitoring–target	-0.052	-0.006	-0.046	0.052
	(0.048)	(0.040)	(0.034)	(0.048)
Bonus–promotion	0.052	0.032	0.020	-0.052
	(0.054)	(0.035)	(0.040)	(0.054)
Displacement	-0.019	-0.023	0.004	0.019
	(0.029)	(0.023)	(0.018)	(0.029)
Observations	3,940	3,940	3,940	3,940
Number of establishments	2,239	2,239	2,239	2,239
Mean dep var	0.271	0.156	0.115	0.729

Notes: Standard errors are clustered at the level of establishments. This table uses the employees sample of BSBW aggregated at the level of establishment-year observations. All regressions control for year fixed effects, establishment fixed effects, and the log of intermediate input cost. The dependent variables are within-establishment shares of employees having the specified characteristics among the employees observed in the BSBW. "Hired" indicates that the employee's tenure is shorter than 5 years. "Stayer" indicates that the employee's tenure is equal or longer than 5 years. "Young" indicates that the worker's age is equal or less than 30 years old. "Mid-career" indicates that the worker's age is above 30 years old.

Table A.17. Hiring of new graduates

	(1)	(2)	(3)
	Number of hiring		
	Total	Male	Female
Monitoring–target	-1.008	-0.564	-0.444
	(2.240)	(1.950)	(0.598)
Bonus–promotion	-0.478	-0.191	-0.287
	(1.563)	(1.155)	(0.756)
Displacement	-1.339	-0.359	-0.980
	(1.195)	(0.834)	(0.700)
Observations	4,817	4,817	4,817
Number of establishments	3,037	3,037	3,037
Mean dep var	5.131	4.100	1.031

Notes: Standard errors are clustered at the level of establishments. This table uses the establishment-year level sample of BSWS in 2010-2011 and 2015–2016. All regressions control for year fixed effects, establishment fixed effects, and the log of intermediate input cost. The dependent variables are the total number of the establishment's hiring of new graduates under regular contracts, and those by gender.