

Managing Long Working Hours: Evidence from a Management Practice Survey

Mari Tanaka* Taisuke Kameda† Takuma Kawamoto‡

Shigeru Sugihara§ Ryo Kambayashi**

Abstract: We investigate the relationship between management practices and long working hours by combining large-scale establishment panel data on management practices with the corresponding employee data on overtime hours in the manufacturing sector. We find that the adoption of more structured bonus and promotion practices is correlated with an increasing probability of workers working more than short-to-medium overtime hours. In addition, the adoption of more structured production monitoring and targeting practices is associated with a lower probability of workers working long overtime hours, resulting in narrowing disparities in overtime hours across workers within establishments.

JEL: J22, J23, M10, M50

Keywords: Working hour, Management practice

* Mari Tanaka (corresponding author) is an associate professor of economics at Hitotsubashi University, Institute of Economic Research. Email: mari.tanaka@r.hit-u.ac.jp.

Acknowledgments: The authors thank Nick Bloom, Catherine Buffington, Renata Lemos, and Atsushi Ohyama for their valuable inputs in designing the Management and Organizational Practice Survey in Japan (JP-MOPS). We are also grateful for the helpful comments from two anonymous referees and the participants at the 2018 Empirical Management Conference, the 2018 Transpacific Labor Seminars, and the 2019 EALE Conference. This work was supported by JSPS KAKENHI Grant Numbers JP18H03633 and JP19H00592. The ESRI had the right to review the paper prior to its circulation. The authors have no relevant or material financial interests that relate to the research described in this paper. This paper was previously titled “Management Practices Meet Labor Market Outcomes.” The data used in the analysis are the microdata of government statistics, which we are not allowed to make public. Researchers can gain access to the data by submitting a written application to the Japanese government ministries: Ministry of Health, Labor and Welfare for the Basic Survey of Wage Structure, Social Research Institute (ESRI) of the Cabinet Office for the Management and Organizational Practice Survey in Japan, and Ministry of Economy, Trade and Industry for the Census of Manufacturers. Researchers can gain access to the data by submitting a written application to the above ministries.

† Taisuke Kameda is Deputy Director for International Economic Affairs, Cabinet Office, Government of Japan

‡ Takuma Kawamoto is a researcher at Economic and Social Research Institute (ESRI), Cabinet Office, Government of Japan.

§ Shigeru Sugihara is a professor at Nihon University, College of Economics.

** Ryo Kambayashi is a professor of economics at Hitotsubashi University, Institute of Economic Research.

doi:10.3368/jhr.0421-11605R2

This open access article is distributed under the terms of the CC-BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0>) and is freely available online at: <http://jhr.uwpress.org>

1. Introduction

Overwork can have negative consequences on workers' welfare and labor productivity. From the workers' perspective, it may worsen work-life balance, health (Pega et al. 2021; Cygan-Rehm and Wunder 2018; Berniell and Bietenbeck 2020), and overall life satisfaction (Hamermesh, Kawaguchi, and Lee 2017). Work-life balance is a serious issue in countries such as Japan, where the percentage of employees working more than 50 hours per week exceeds the OECD average of 11% (OECD 2021). Overwork may impair workers' health, and in the worst-case scenario, it may result in death (*karoshi*), which can also result in the employer being sued. The WHO and the ILO estimate that 745,000 deaths from ischemic heart disease or stroke globally in 2016 were attributed to working very long hours, defined as more than 55 hours per week (Pega et al. 2021). Overwork can also be harmful on the production side. Several empirical studies show that the marginal output of an extra hour of work substantially declines after the worker works for a number of hours (e.g., Pencavel 2015, 2016). To avoid these issues, firm managers are generally considered responsible for preventing employees from overworking while inducing a moderate amount of effort.

However, how firms' management practices are related to working hours remains largely unknown. In the economics literature, the determinants of working hours have been discussed in a simple demand-and-supply framework featuring concave utility and production functions to derive short-run equilibrium hours of work (e.g., Hamermesh 1993). In this simple model, actual working hours can be long, depending on the functional forms of utility and production. More recently, bonus and promotion schemes, that is, human resource management (HRM) practices, for coping with moral hazard and hidden ability, have been shown to be important determinants of long working hours (e.g., Gicheva 2013; Sousa-Poza and Ziegler 2003; Landers, Rebitzer, and Taylo 1996). However, not only human resource management (HRM) practices, but also non-

human resource management (non-HRM) practices such as production monitoring and targeting may affect working hours. For example, inappropriate monitoring and targeting practices can cause inefficient allocation of tasks across workers and time, which results in working long hours for some workers. In such a case, improving monitoring and targeting practices could reduce the number of working hours, especially for workers working long hours. The relationship between management practices and working hours has been difficult to empirically investigate because of limited availability of large-scale data on management practices that can be linked to working hours.

To fill this gap, this study investigates the empirical relationship between management practices and working hours by matching a large panel of Japanese firm management practice data with the overtime hours of employees. The firm management practice data were obtained from the Management and Organizational Practice Survey in Japan (JP-MOPS), conducted by the Japanese Cabinet Office in 2017. The Management and Organizational Practice Survey (MOPS) is an internationally coordinated project that collects information on establishments' management practices in 2010 and 2015 regarding non-HRM practices, such as production monitoring and targeting practices, and HRM practices, such as bonuses, promotions, and displacement. The JP-MOPS 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). We matched the JP-MOP data in the manufacturing sector with information on the overtime hours of multiple workers in the same establishments around the same years, from another government survey. The establishment-level panel structure of our matched data enables us to focus on changes in management practices by controlling for unobserved establishment fixed characteristics affecting both management practices and overtime hours.

We find that the HRM and non-HRM practices play different roles in the management of overtime hours. On one hand, the adoption of more structured bonus and promotion (HRM) practices that evaluate workers based on individual performance predicts a higher probability of working above short-to-medium overtime hours. A plausible explanation is that such high-powered incentive schemes prevent moral hazard and induce some workers to exert greater effort, as widely discussed in the literature on pay for performance and career concerns (Bloom and Van Reenen 2011; Holmström 1999). On the other hand, the introduction of more structured production monitoring and targeting (non-HRM) practices predicts a lower probability of workers working for long hours, in particular, those with more than 45 hours of overtime (equivalent to around 50 hours of work per week). This can be attributed to the introduction of these practices, which enables identification and reduction of problems that trigger long overtime hours.¹ Hence, our results suggest that the adoption of more structured monitoring and targeting practices is associated with a concentration of the distribution of overtime hours to zero. This evidence calls for greater attention to firms' non-HRM practices as a means of achieving labor market policy goals, such as reducing excessive overtime work and improving labor productivity. In addition to these results, we find only a weak and insignificant association between changes in management practices and labor turnover. Hence, we interpret our results as short-run outcomes before significant compositional adjustments occur.

This paper broadly relates to three strands of literature. First, previous studies on labor demand-side factors that determine working hours focused on employers' responses to tax and regulations on working hours and compensation schedules (e.g., Hamermesh 1993; Trejo 1991;

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).

Crépon and Kramarz 2002).² However, empirical studies relating working hours to firm management practices are limited, with a specific focus on HRM. Among these, Bell and Freeman (2001), Gicheva (2013), and Frederiksen, Kato, and Smith (2018) show that longer working hours predict an increased probability of future promotion, higher wages, and wage growth, which they interpret as evidence supporting firms' usage of certain promotion incentive mechanisms. Our contributions to this literature are three-fold. First, we link the 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 targeting) as well as HRM practices in determining working hours. Last, we analyze changes in the distribution of working hours *within* establishments, which have rarely been examined in extant empirical studies.

Second, our study also builds on existing empirical studies using direct measures of both HRM and non-HRM practices obtained from large-scale surveys of firm managers (Bloom and Van 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 based on field experiments on certain sets of firms (e.g., Bloom et al. 2013; Bruhn, Karlan, and Schoar 2018; Gosnell, List, and Metcalfe 2020). Nevertheless, studies linking management practices to labor market outcomes are limited. A few exceptions are Bender et al. (2018) and Cornwell, Schmutte, and Scur (2021). Using cross-sectional data on management practices matched with panel data on employee earnings, they show that firms using structured management practices tend to hire and retain workers with higher abilities. In contrast, we use a panel structure of management practices matched with employees' working hours, to distinguish the effects of management practices from

² There are also studies concerning the business cycle as a determinant of hours of work (Prescott 2004; Rogerson 2006).

time-invariant firm-specific unobserved factors. This enables us to examine the relationship between changes in management practices and changes in labor outcomes of establishments, which is one of our major contributions. We show that excluding the effects of firm-fixed unobserved factors largely alters the observed relationship between management practices and working hours.

Third, our findings are consistent with the results of laboratory experiments that show that some management technologies lead to leveling workload across workers (Bewley 2000; Bartling and von Siemens 2011). This leveling may improve firm productivity, assuming that the marginal product of labor diminishes as working hours increase, as empirically supported by Brachet, David, and Drechsler (2012), Pencavel (2015, 2016), and Collewet and Sauermann (2012). This could partly explain why structured management practices improve employee productivity and firm performance, as shown in previous studies (e.g., Bloom et al. 2013; Gosnell, List, and Metcalfe 2020; Kambayashi, Ohyama, and Hori 2021).

The remainder of this paper is organized as follows. Section 2 presents the data of the study. In Section 3, we describe our analysis of working hours, management practices, and the results. Section 4 presents the results of the analysis of labor turnover and within-establishment outcomes, and Section 5 presents the conclusions.

2. Data

2.1 Data on Establishments' Management Practices

The MOPS is an internationally coordinated governmental survey that collects information on the management practices of establishments. The original survey was conducted in 2011 by the US Census Bureau for the manufacturing sector (Buffington et al. 2017), and the second wave in the US was implemented in 2016 (Buffington et al. 2017; Bloom et al. 2019). Several other countries have conducted near-identical surveys, including Pakistan in 2015 (Lemos et al. 2016) and the UK

in 2017 (Office for National Statistics 2018). In Japan, the Economic and Social Research Institute 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 practices (production monitoring and targeting) and HRM practices (bonuses, promotions, and displacement). These questions are similar to those in the World Management Survey initiated by Bloom and Van Reenen (2007). Each question in the JP-MOPS is the precise Japanese-language translation of a corresponding question in the 2015 US MOPS. The questionnaire asked about practices in the base year (2015 in the case of the JP-MOPS) and retrospectively about practices in the five years before the base year (2010 in the case of the JP-MOPS). Responses are obtained from those responsible for overall management practices in the establishment, typically presidents, executive officers, plant managers, or general or department managers (see Table A.1 in the Appendix). Less than 0.4% of the respondents were from human-resource-related departments, and approximately 90% of the respondents had tenures longer than five years, with a median tenure of about 21 years.³

The JP-MOPS sample is based on the 2014 Economic Census for Business Frame, which is a complete census of Japanese establishments in 2014, where 56,237 establishments met the criterion for being survey targets: private manufacturers with 30 or more employees.⁴ Based on a stratified random sampling by two-digit industry and size, 36,052 establishments were selected

³ Asking recall questions is a common feature of the international MOPS project. This is based on a finding from the original US-MOPS data that confirms the high response quality of such questions (Bloom et al. 2019), especially for respondents whose tenure is above five years (i.e., those who were in the same position at the time being recalled). As a robustness check, we estimate the main specification by limiting the sample to respondents who were in the same position in 2010, and confirmed that the results are qualitatively the same as the main results.

⁴ One difference in the survey design of the JP-MOPS compared to the 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 the burden on respondents; detailed information about establishments in the Japanese Census of Manufacturers is also collected only for establishments with 30 or more workers.

and sent questionnaires in the second week of January 2017. After two follow-up calls, we received 11,405 responses (31.6% response rate). In general, the establishments that participated in the JP-MOPS were larger than those that participated in the 2014 Japanese Census of Manufacturers (see Table A.2 in the Appendix).

The non-HRM section of the survey questionnaire contains questions regarding production monitoring and targeting practices, which mainly involve the collection and utilization of information about the production process by the establishment. Table A.3 in the Appendix shows the list of exact questions and response options.⁵ A question about monitoring asks “What best describes what happened at this establishment when a problem in the production process arose?”; the response options vary from “No action was taken” to “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”. The other questions on monitoring are related to the variety of key performance indicators (KPIs) collected, the frequency with which these are reviewed, and the existence of display boards showing the KPIs. Questions about targeting 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 targeting ask firms about the time frame of the production targets and who are aware of those targets.

For the HRM questions, the survey asks whether there are any performance bonuses and, if so, the determinants of the bonuses (whether individual, team, establishment, and/or firm performance) and what proportion of workers are covered by the bonuses. The questions on

⁵ For the MOPS questionnaire itself, see <https://www.census.gov/programs-surveys/mops/technical-documentation/questionnaires.html> for the US-MOPS and https://www.esri.cao.go.jp/jp/esri/prj/current_research/service/manage/chosa_seizo.pdf for the JP-MOPS.

promotion concern the primary manner in which managers and non-managers are promoted, that is, whether it is solely based on performance and ability, or other factors such as tenure or family connections also matter. Finally, the questions about displacement ask “When is an underperforming non-manager usually reassigned or dismissed?,” to which each respondent selects a response from “Within 6 months of identifying non-manager underperformance,” “After 6 months of identifying non-manager underperformance,” and “Rarely or never” The same questions were also asked about the managers.

Following the literature, we first score the response to each question by assigning a value ranging from zero to one (for details, see Table A.3 in the Appendix).⁶ For the monitoring section, a higher score is assigned when the management uses more KPIs, reviews them more frequently, shares them more widely, and continuously improves the production process. For the targeting section, a higher score is assigned when the management uses a composite of long- and short-term perspectives with a moderate level of difficulty to be achieved when designing its production targeting. The scores for the bonus section 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.

To aggregate these scores for empirical analysis, Bloom et al. (2019) employed a single score referred to as overall management score, which is the unweighted average of the scores of

⁶ After removing establishments with missing scores for at least 10 questions, we are left with 10,806 establishments in the JP-MOPS data.

the 16 questions. However, while aggregating a single score may be appropriate for the analysis of firm performance, it may not be ideal when analyzing worker outcomes. Given that HRM practices have been considered as the main channel through which management affects worker outcomes in the literature on personnel economics, we separate the responses into HRM and non-HRM practices. Furthermore, we disaggregate HRM practices into those related to wages and employment. More concretely, we used three scores by taking the unweighted average of the scores across three groups of questions: monitoring and targeting (MOPS questions 1–8), bonus and promotion (MOPS questions 9–14), and displacement (MOPS questions 15–16). We show that further disaggregation provides qualitatively similar results, although the statistical power is reduced.

2.2 Data on Employees

We match the JP-MOPS data to the Basic Survey on Wage Structure (BSWS) using the establishment identifiers provided in the establishment census. The BSWS is an employer-employee matched survey conducted annually by the Japanese Ministry of Labor, Health, and Welfare (MLHW) on June 30. The main information we use from the BSWS is employee-level overtime hours worked. The sample size is large, covering approximately one million workers across approximately fifty thousand establishments each year. The sampling is done in two stages. First, establishments are selected using a stratified sampling method from private establishments with five or more employees. Second, within each establishment selected in the first stage, employees are selected using a uniform sampling method with a predetermined sampling rate

depending on the industry and establishment size.⁷ The survey is a repeated cross-section, covering approximately 5% of establishments across Japan each year.

The survey asks human resource managers about the employees' total overtime hours worked as well as the total scheduled hours worked in June.⁸ As such, we use the total overtime hours worked in June for our main outcome variables. As a reference for the definition of scheduled and overtime hours, in Japan, the standard workweek is 40 hours per week, above which work hours are considered overtime. This implies that, for instance, the definition of long working hours in OECD statistics (OECD 2021), working for more than 50 hours per week, is the same as around 40 hours or more of overtime per month in Japan. An overtime premium of at least 25% must be paid for overtime of less than 60 hours per month and at least 50% for overtime of more than 60 hours. In addition, the survey collected worker attributes such as education, gender, tenure, and age. Throughout our analysis, we limit our attention to full-time workers who are strictly less than 60 years of age. We also exclude employees in managerial positions, anticipating varied effects on management and other employees.⁹

2.3 Data on Production in Manufacturing Establishments

In addition, we match the data to the Census of Manufacturers, which is a survey of all manufacturing establishments with four or more employees by the Japanese Ministry of Economy,

⁷ The sampling rate is 1/2 for establishments with 30–99 employees and 1/5 for manufacturing establishments with 100–499 employees, and it varies a lot across industries for larger establishments (see Ministry of Health, Labor and Welfare (2021) for details). Since the response rate of the BSWS was approximately 82% in the manufacturing sector in 2010, the sample selection bias caused by nonresponse is considered to be small.

⁸ The BSWS is a copy of firm payroll records. In Japan, there is a set of necessary information about working conditions that each labor contract must include and employers are legally obliged to retain the relevant data as the payroll record by Labor Standard Act Article 107–109.

⁹ These are employees whose title is either *Bucho* (department manager/head) or *Kacho* (section manager/head) in the BSWS. They account for approximately 3% of employees who are strictly less than 60 years of age. These managers do not include those who actually supervise other workers but not titled as above, such as line manager/head and unit manager/head (Higuchi and Kambayashi 2018).

Trade, and Industry. We use information on production, such as materials, fuel, and electricity costs, to control for establishment-level demand shocks.

2.4 Matching Data

We construct management-employee matched panel data by combining the data presented above. Note that the data have a panel structure only at the establishment and not at the employee level, because it is not possible to follow workers over time in the BSWs. The overall Venn diagram of the relationships between the data being matched is provided in Appendix Figure A.1. The JP-MOPS includes retrospective questions about management practices in 2010 and about those in 2015, allowing us to construct establishment-level panel data on 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) for June 2010–2011 and 2015–2016, respectively, assuming that management practices tend to be stable over half a year. Among the original 225,984 establishment-year observations in the BSWs in 2010, 2011, 2015, and 2016, there were 40,769 establishment-year observations in the manufacturing and private sectors. Among these, we can match 5,529 establishment-year observations to the JP-MOPS, yielding a matching ratio of 14% from the BSWs establishment-year observations.¹⁰ We then match the Census of Manufactures in 2010–2011 and 2015–2016 to use the information on the total cost of materials, fuel, and electricity for controlling establishment-level demand shocks.

Since our main empirical strategy exploits the establishment panel structure, the sample for our analysis consists of 72,180 employee-year observations in the BSWs for 827 unique establishments whose employees are observed during both the first (2010–2011) and second

¹⁰ Note that the BSWs includes establishments with less than 30 employees. Among 21,612 establishment-year observations of management practices (or 10,806 observations of unique establishments) in the JP-MOPS, 26% are matched to the BSWs.

(2015–2016) periods, whose management practice scores from the JP-MOPS are observed, and the information on the total cost of intermediate inputs is not missing.¹¹ We examine the differences in worker characteristics between the matched and unmatched samples and find that their basic statistics are similar (see Table A.2 in the Appendix). Distributions of overtime hours in matched and unmatched samples also look alike in both 2010 and 2015 (see Figure A.2 in the Appendix).¹²

2.5 Descriptive Statistics on Management Practices

Table A.5 in the Appendix provides summary statistics of the establishment-level variables for the establishments in our sample. The mean of the overall management score in 2010 was 0.507, with a standard deviation of 0.166. Table A.6 in the Appendix examines the establishment characteristics associated with higher overall management score in 2010. Overall, larger establishments with more traditional Japanese employment practices (for example long tenure was taken as proxy) tended to have higher scores on management practices in 2010. This is consistent with the fact that the original measures of management scores in Bloom and Van Reenen (2007) referred to the best practices from Japan, such as lean manufacturing, as well as practices from Western countries.

Between 2010 and 2015, the overall management score improved by about 0.044 on average, although some 33% of establishments reported no change in their overall management score. It is also notable that a substantial proportion (11%) of establishments reported a

¹¹ Among 5,529 establishment-year observations we match from the BSWs to the JP-MOPS, some establishments are observed only in the first (2010–2011) or only in the second (2015–2016) periods, in which cases these observations are useless for our panel analysis controlling for establishment fixed effects. Therefore, we excluded these observations from our analysis sample.

¹² Furthermore, we examine the differences in labor outcomes (wages and overtime hours) between the matched and unmatched samples and find few systematic differences unexplained by worker characteristics. Moreover, when we regress the labor outcomes on worker characteristics by matched and unmatched samples, we observe similar features of the coefficients. See Table A.4 in the Appendix for the results.

deterioration in their management scores (for the distribution of changes in scores, see Figure A.3 in the Appendix). Furthermore, there are variations in the directions and magnitudes of the changes in the monitoring and targeting scores, the bonus and promotion scores, and the displacement scores within establishments (see Figure A.4 in the Appendix which shows a large variation in the changes of management scores among the three components).¹³

We briefly explore what factors explain the changes in management practices. First, Table A.7 in the Appendix details the results of decomposition analysis of the changes in overall management practices by industry and region. Around 18% of the variation is explained by industry fixed effects, while prefecture fixed effects explain only 8%, suggesting the possibility of knowledge spillover within industries. Second, in the sample of establishments matched with intermediate inputs in the Census of Manufacturers, changes in the intermediate inputs (which is a proxy of demand shock) explain some of the changes in management practices, after controlling for the industry and prefecture fixed effects. A similar result is obtained when we regress changes in management practices on production shocks in the baseline years (2009–2011) measured by changes in the log of shipment values in these years (see Table A.8 in the Appendix for the results). This result implies that changes in management practices are not likely to be driven by changes in production scale or firm-specific demand changes. Third, establishments may have changed management practices to prevent employees from working long overtime hours. In this case, there is a problem of reverse causality for our main analysis. However, neither the establishment's average overtime hours nor the share of employees working more than 45 hours of overtime

¹³ 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, 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 this 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 working hours.

significantly predicts later changes in management practices (see Table A.9 in the Appendix for the results). Fourth, firms' employment practices may affect the adoption rate of new management practices. In particular, we consider traditional Japanese employment practices that have been characterized by strong employment protection and long tenure, resulting in a high share of home-grown employees (defined by regular workers who were hired directly upon graduation, or *haenuki* in Japanese) (Hashimoto and Raisian 1985). Indeed, we find that the average tenure and share of home-grown employees negatively predicts the changes in management practices. In addition, the fraction of women among the employees of the establishment in the base years positively predicts changes in management practices. Also, establishment size (measured by the log of shipment) in the base years negatively predicts changes in management practices. The detailed results are reported in Table A.9 in the Appendix. These results imply that management practices tend to improve more in small, non-traditional Japanese establishments.

2.6 Descriptive statistics of overtime hours

The basic summary statistics for the employee-level variables are given in Table A.10 in the Appendix, and more detailed summary statistics for overtime hours are provided in Table A.11 in the Appendix.¹⁴ While around 45% of workers worked less than 10 hours of overtime, 11% of workers worked for more than 45 hours of overtime per month in 2015–2016. Compared to other countries, the share of workers working long hours is high in Japan, ranking in the top five in the OECD (OECD 2021). Japanese corporate culture is often argued to be a cause of long working hours (OECD 2018). In addition, traditional Japanese employment practices may promote long working hours through the mutual expectation between employers and employees due to their

¹⁴ The minimum value of overtime hours in our data is 0. The BSWs does not contain the information on the worker's contractual hours, thus we do not analyze "negative overtime hours", that is, the hours worked less than the contractual hours.

long-term relationships (Kambayashi and Kato 2016).¹⁵ Such a large share of employees working long hours has been considered a major policy concern in Japan and a possible cause of health problems, work-life imbalance, work-related stress, and in the worst case, *karoshi* (death caused by overwork).

In the following analysis, we examine the relationships between management practices and various levels of overtime hours. In particular, to define long working hours in our setting, we consider overtime of more than 45 hours per month as a key threshold. Japanese worker's compensation rules specify that when a worker dies from a brain or heart disease and has worked more than 45 hours of overtime in the preceding month, overwork is considered a possible cause of death. This implies that 45 hours of overtime work is considered the threshold at which the risk of *karoshi* begins to increase. This threshold number is comparable to the criteria in other countries. For example, the EU Working Time Directive specifies that the maximum hours of work per week, including overtime hours, should be 48 hours. This implies that, if the legal workweek is 40 hours, as in Japan, the maximum overtime hours per month should be around 32 hours. In addition, the WHO and ILO jointly report that 3.6% of all deaths globally due to ischemic heart disease and 6.9% of those for stroke are attributable to working for more than 55 hours per week (Pega et al. 2021), which is about 60 hours of overtime per month under the Japanese legal workweek. In addition, to define short working hours, we consider overtime of more than 10 hours per month, which roughly corresponds to more than 42 hours of work per week.

3. Results: Overtime Work and Management Practices

¹⁵ Since we earlier found that management practices were likely to evolve less in establishments applying traditional Japanese employment practices, we conduct a robustness check to examine the possibility that such employment practices drive our main results. We find that our main conclusion is robust to controlling for possibly differential trends of overtime hours by proxy variables for such employment practices. See Section 3.7 for more details.

3.1 Graphical Illustration

We start by graphically illustrating the distribution of overtime hours and relating them to the overall management practices of establishments over the target years. To do this, we divide the sample of establishments into two groups according to whether the overall management score improved during 2010–2015 more than the sample median. Figure 1 depicts the distributions of overtime hours among male workers¹⁶ in 2010 (light gray) and 2015 (transparent) in the surveyed month for the two groups of establishments.¹⁷ For both groups of establishments, overtime hours increased on average between 2010 and 2015, possibly reflecting the overall economic expansion. However, there were differences in the way hours increased across the two groups. In establishments with 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 of 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 20–50 hours. Overall, Figure 1 illustrates the importance of considering the effects on the distribution of overtime hours, not only the effects on the average overtime hours.

3.2 Average Overtime Hours and Management Practices

Next, we explore this relationship in a regression framework, controlling for the observed worker and establishment characteristics. We first examine the general relationships between management

¹⁶ In this graphical illustration, we focus on the sample of male workers because the incidence of long working hours is concentrated among male workers (as shown in Table A.11 in the Appendix); hence, their figures are suitable for illustration.

¹⁷ The cases of zero overtime hours are included in Figure 1 as 0–5 overtime hours. The share of workers doing any positive hours of overtime increased by 0.048 in establishments with less improved management scores and by 0.029 in establishments with more improved management scores.

practices and average overtime hours and then move to investigate the implications for the distribution of overtime hours.

First, without controlling for establishment fixed effects, in column (1) of Table 1, overtime hours are regressed on the overall management score. We control for year fixed effects and basic worker characteristics, measured by age, age squared, tenure, tenure squared, female dummy, and three education dummies.¹⁸ In addition, noting that management practices could be correlated with establishment-level demand shocks, we control for the amount of flexible inputs used in production, as measured by the log of the total cost of intermediate inputs (materials, fuel, and electricity). The standard errors are clustered at the level of the establishment. The results show that the usage of more structured management practices is positively and significantly correlated with longer overtime hours.

However, as shown in column (2), controlling for establishment fixed effects drastically reduces the coefficient and increases the standard error, making the estimated coefficient negative and statistically insignificant. This result highlights the importance of controlling for establishment fixed effects, to understand within-establishment variation in management practices and hours. The positive and strong cross-sectional correlation between management practices and overtime hours could be driven by various other kinds of establishment characteristics that are fixed over time and the results of matching and sorting between firms and workers that have happened over a long preceding period.

It is common in the literature on management practice to use a single aggregated score of management practices (e.g., an average of all scores, as shown above), owing to the similarity of

¹⁸ To control for the worker's education attainment, we use dummy variables for middle school graduates, junior/technical college graduates, and college graduates (the reference group is high-school graduates).

the effects of these practices on production performance (Bloom and Van Reenen 2007). However, when we consider the relationship of such practices with working hours, each component of the practices should be examined separately, because of the distinct roles played by each component. In column (3), we disaggregate the overall management score into three closely related components: bonus and promotion (wage-related HRM), monitoring and targeting (non-HRM), and displacement scores (employment-related HRM). We find that the coefficient of the monitoring and targeting score is negative, while the coefficient of the bonus and promotion score is positive, and that both coefficients are relatively large. However, the standard errors remain large. This can be attributed to the heterogeneous effects of these management practices across workers, particularly by the level of overtime hours worked, as examined in the subsequent section.

3.3 Distribution of Overtime Hours and Management Practices

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, as presented in the graphical illustration. In particular, the marginal effect of the management practice score on overtime hours can be heterogeneous, depending on the level of overtime hours. One potential approach to examine this possibility is to estimate a quantile regression. However, it is not straightforward to control for establishment and year fixed effects in a quantile regression (although we do it as a robustness check in Section 4 and show that the results are similar). Therefore, as our main specification, we estimate a standard OLS using an indicator variable of working overtime for more than k hours, where $k = 5, 10, \dots, 50$, for individual i in establishment j in year t , denoted by $1(OH_{ijt} > k)$, as the dependent variable. Specifically, we estimate the following equation:

$$1(OH_{ijt} > k) = \text{Management}'_{jt}\beta^k + X'_{ijt}\gamma^k + \delta^k \text{Demand}_{jt} + \phi_j^k + \phi_t^k + \mu_{ijt}^k \quad (1)$$

for each $k = 5, 10, \dots, 50$. $Management_{jt}$ is a vector of establishment j 's management scores in year t that contains the bonus and promotion score, the monitoring and targeting score, and the displacement score. The subscript t indicates one of the four years, 2010, 2011, 2015, or 2016. 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 of materials, fuel, and electricity as a proxy for 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. Standard errors are clustered at the level of the establishment.

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 from 2010 to 2015¹⁹) is associated with a 4.7% increase in the share of workers working more than 10 hours of overtime. The magnitudes of the coefficients peak around overtime above 10–20 hours and become smaller for longer hours. Conversely, the coefficients of the monitoring and targeting score are negative for all ranges, and they are small in magnitude for overtime above 5–10 hours and become larger for longer hours. In particular, for overtime above 40–50 hours, the coefficients are statistically significant at the 10% level. The coefficients are large; for example, an improvement in the monitoring and targeting score by 0.1 (roughly equivalent to a one-standard-deviation of the changes in the score from 2010

¹⁹ For example, a change in the bonus and promotion score by 0.1 can be roughly achieved if an establishment changes the bonus system of the non-managers to be considered on the basis of team performance rather than establishment performance, and modifies their promotion system to be based solely on individual performance rather than both individual performance and tenure.

to 2015²⁰) is associated with a 7.7% reduction in the share of workers working more than 45 hours.²¹

The striking differences in the results between HRM and non-HRM practices imply that they play different roles in managing working hours. We discuss possible interpretations of each of their distinct roles below.

3.4 Roles of HRM Practices

We start by exploring the role of wage-related HRM practices in more detail. An improvement in the bonus and promotion score here means either a change from a fixed wage scheme to a group-based incentive scheme or a change from a group-based to an individual-based incentive scheme. Hence, a natural interpretation of the results for the bonus and promotion score is that such changes prevent moral hazard (hidden action problems), inducing some workers to exert greater effort, as widely shown in the literature on performance pay and career concern (Bloom and Van Reenen 2011; Holmström 1999), in the form of increasing working hours. Interestingly, our results show that the coefficients of bonus and promotion score are large only for short overtime hours, and small and insignificant for excessively long hours. One explanation is that the marginal product of labor decreases with additional overtime hours (as per evidence by Brachet, David, and Drechsler (2012), Pencavel (2015, 2016), and Collewet and Sauermann (2012)), while the marginal cost of effort possibly increases. Therefore, payment schemes based on individual output would not encourage workers to work beyond the level at which the marginal effort cost exceeds the marginal increase in (expected) payment. In other words, the variation in coefficient sizes may reflect the

²⁰ For example, a change in the monitoring and targeting score by 0.1 can be achieved if an establishment increases the number of KPIs collected from 1 to 2 to 3–9 and starts to show the KPIs on not just one but multiple display boards.

²¹ We obtain a similar sizable result for monthly overtime of over 60 hours, although the coefficient is estimated less precisely. The coefficient of the monitoring and targeting score for overtime of over 60 hours is -0.033 with a standard error of 0.026, and the mean of the dependent variable is 0.045.

balance between the marginal cost and the marginal benefit of working an additional hour at each level of overtime hours.

Alternatively, this result may also be interpreted by an adverse selection (hidden information) model. In theory, the introduction of a promotion scheme based on individual performance can alleviate the adverse selection problem, which induces some workers to signal their ability in the form of long working hours (Sousa-Poza and Ziegler 2003; Landers, Rebitzer, and Taylo 1996). More specifically, the bonus and promotion score in the JP-MOPS improves when promotion schemes change from reflecting factors other than individual performance to evaluation solely based on individual performance. Such a change potentially comes together with improved measures of individual performance or ability (e.g., individual production output) to be used for promotion decisions. According to signaling model (Sousa-Poza and Ziegler 2003; Landers, Rebitzer, and Taylo 1996), this could reduce the need for workers to signal ability by working long hours. This mechanism would reduce workers who work long overtime hours and increase those who work short overtime hours instead. While this explains the positive and decreasing coefficients of the bonus promotion score for short and medium hours, it does not explain the non-negative coefficients for long hours. However, we should note that our data on bonus and promotion schemes do not explicitly indicate whether the measures for individual performance evaluation still include working hours. If the measures still include working hours, the non-negative coefficients of the bonus and promotion score for long overtime hours may reflect this reason. A better way to examine the implication of the signaling hypothesis is to look at the interaction with non-HRM practices (monitoring and targeting scores), which potentially convey information about the degree of data collection regarding individual performance. This is discussed in Section 3.6.

Next, we explore whether the results of wage-related HRM practices in Table 2 are mainly driven by changes in payment schemes (short-term incentives) or changes in promotion schemes (long-term incentives). One way to examine this question is to divide the sample by tenure. The theory of career concern predicts that the incentive for promotion exists only during the earlier career stages and, therefore, should not affect long-tenured workers. However, there is no strong reason to believe that performance pay differentially affects long- and short-tenured workers. To examine this possibility, in Panel A of Table 3, we estimate the equations for overtime above 10 (“short”) and 45 (“long”) hours using the sample divided by tenure group, with tenure being less than the median (11 years) (“junior”) or not (“senior”). On average, junior workers tend to work longer overtime hours than senior workers do. The coefficients of the bonus and promotion score are larger for junior workers than senior workers for all levels of overtime. Notably, although the coefficient for long hours is large, positive and significant among junior workers, this coefficient is insignificant and small among senior workers. This supports that the results of wage-related HRM practices are mainly driven by promotion schemes. 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 positive for junior workers and negative for senior workers for short and long overtime, although these differences are insignificant (see Table A.12 in the Appendix).²²

Another interesting aspect of HRM practices and working hours is gender. Female workers tend to work fewer overtime hours than male workers on average, and this is considered the source of the gender wage gap, through promotion (Gicheva 2013). Panel B in Table 3 divides the sample

²² The insignificant results for the promotion score can be attributed to the high correlation between 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.

by gender. The bonus and promotion coefficients are larger for female workers on both short and long hours. The coefficients for female workers are sizable; for example, an increase in the bonus and promotion score of 0.1 is associated with a 30.4% increase in the share of workers working longer hours. As a result, the improvements in bonus promotion scores are associated with a narrower gender gap in the share of workers working long overtime. One potential explanation for this result is that performance-based promotion policies help break the gender glass ceiling, which may otherwise prevail, for example, under tenure-based policies. An alternative interpretation is that working hours signal the workers' commitment to the firm, and such a signal is stronger for female workers, consistent with the model and evidence provided by Kato, Ogawa, and Owan (2017). As before, the results suggest that promotion schemes rather than payment schemes (bonuses) seem to be the main driver of the results for HRM practices. This is consistent with the study by Bandiera et al. (2021), which showed that men and women are equally affected by performance pay.

3.5 Roles of Non-HRM Practices

Next, for the interpretation of the results on non-HRM practices (monitoring and targeting), one of the key features to note is that all of their coefficients in Table 2 are negative. One explanation is that these practices eliminate waste in production, reducing the total labor hours needed to produce a given amount of output. For instance, collecting and reviewing detailed data on the production progress may alert managers to problems in the production system, and continuously improving a production system could reduce problems that lead to a waste of labor hours. Another key feature of the results is that the coefficient sizes for shorter hours are small and that they become larger for longer hours. One interpretation of this feature is that monitoring and targeting practices also allow leveling of workload across workers and time. In theory, 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 every day. However, the actual workload can be heterogeneous across workers and days for various reasons, such as miscommunication (bad monitoring practices), planning mistakes (bad targeting practices), and workers' preference for working long hours (Hamermesh and Slemrod 2008). In such a case, leveling of the workload across workers and time through more frequent and closer production monitoring improves allocative efficiency. In addition, setting and sharing targets at a moderate level of difficulty would help adjust workloads across time, thereby reducing excessively long overtime and possibly increasing short overtime. The negative coefficients, even at short hours in our results, can be explained if we combine this hypothesis with the waste elimination hypothesis. Another explanation is that workload reallocation does not necessarily result in an increase in short-medium overtime if the work can be done by other workers or other times within standard work hours. In such a case, it implies an increase in efficiency during standard work hours, which we cannot measure in our data.

The abovementioned hypothesis suggests that structured monitoring and targeting practices control excessive overtime, and the magnitude of this effect should be larger for workers who tend to do more of it. A typical example of such workers would be junior workers, who tend to have stronger career concerns and may have insufficient experience in managing their own time and workload. Consistent with this result, in Panel A of Table 3, we find that the negative coefficient of the monitoring and targeting score is large and statistically significant at the 5% level

for overtime of above 45 hours for junior workers but smaller and statistically insignificant for senior workers.²³

We further explore this possibility by dividing the sample of workers by their propensities for working overtime over 10 and 45 hours. By estimating the probability of working overtime above 10 or 45 hours as a function of gender, age, tenure, and education level, we exploit the full variation in workers' observed characteristics. We use the 2009 BSWs data, as it is the year before 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 tended to work longer overtime. In addition, the probability of working overtime increases but is strongly concave in age, and college graduates tend to work overtime less than high school graduates. We use these estimated probit models to predict the propensity of working overtime for all workers in our analysis sample in 2010–2011 and 2015–2016. We then divide the sample by the median of the estimated propensity scores. The results are reported in Panel C of Table 3. The monitoring and targeting score coefficients are large, negative, and statistically significant for workers with a greater propensity to work long hours and insignificant for the other workers. The coefficients are sizable, suggesting that a 0.1 increase in the monitoring and targeting score is associated with a 7.7% reduction in the share of workers working long hours. This result supports our conjecture that structured monitoring and targeting practices control excessive amounts of overtime by workers who are likely to do more of it.

3.6 Interaction between HRM and non-HRM Practices

²³ When we alternatively divide the sample by age, we find that younger workers are more likely to work long hours and the negative coefficient of the monitoring and targeting score for long hours is larger for them than for older workers (see Table A.13 in the Appendix).

In the analysis so far, we implicitly assumed that each management practice is independently related to overtime hours. However, HRM and non-HRM practices may work as substitutes or complements. In particular, if better monitoring and targeting practices provide more accurate information about individual performance to the firm, they may also influence the way that bonuses and promotions affect overtime hours. Whether these practices work as substitutes or complements depends on the underlying features of asymmetric information. Recall that we have two interpretations of the effects of the bonus and promotion score: hidden action and hidden information hypotheses. Under the hidden action hypothesis, the two practices are likely to be complements, because a shift to individual-level performance-based pay/promotion would induce greater effort when accurate information about individual performance is measured and reported to the employer, compared to the case where such information is not measured or reported. In contrast, under the hidden information hypothesis, the practices are likely to be substitutes, since workers have less need to signal their ability by working overtime if individual performance is accurately measured and reported to the firm (Sousa-Poza and Ziegler 2003; Landers, Rebitzer, and Taylo 1996).

To examine these hypotheses, we estimate the following equation, which adds an interaction term between the bonus and promotion, and the monitoring and targeting practices to equation 1.²⁴

$$1(OH_{ijt} > k) = \beta_1^k BP_{jt} + \beta_2^k MT_{jt} + \beta_3^k DP_{jt} + \alpha^k BP_{jt} MT_{jt} + X'_{ijt} \gamma^k + \delta^k Demand_{jt} + \phi_j^k + \phi_t^k + \mu_{ijt}^k, \quad (2)$$

²⁴ The use of interaction term for testing complementarity of factors in production function is common in the literature. See, for instance, Bresnahan, Brynjolfsson, and Hitt (2002).

where BP_{jt} , MT_{jt} , and DP_{jt} denote the bonus and promotion score, the monitoring and targeting score, and the displacement score, respectively. The result is presented in Table 4. The estimates of α^k are positive and statistically significant for overtime over 10 hours, implying complementarity of the two practices. This result supports the hidden action hypothesis that the two practices are net complements. Note that this result does not reject the signaling hypothesis. While the two practices can have the elements of both complements and substitutes, the evidence suggests that the complementary effects are stronger.

To better illustrate this result, Panel (a) of Figure 2 plots the estimates of the “marginal contribution” of the bonus and promotion score (defined by $\widehat{\beta}_1^k \Delta BP_{jt} + \widehat{\alpha}^k MT_{jt} \Delta BP_{jt}$) in percentage terms of the mean of the dependent variable, where ΔBP_{jt} is one standard deviation of the change of the bonus and promotion score from 2010 to 2015. The panel depicts these values at three different levels of the monitoring and targeting score: MT_{jt} is at the 10th percentile, at the sample average, and at the 90th percentile in 2010. At the average level of MT_{jt} , the marginal contributions of BP_{jt} are positive and around 5-10% of the dependent variable for all levels of overtime hours. The marginal contributions are larger when MT_{jt} is at the 90th percentile. This is consistent with the hidden action hypothesis, which implies that the introduction of pay/promotion tied to individual performance induces greater effort when the employer collects more accurate information about individual performance. Interestingly, when MT_{jt} is at the 10th percentile, the marginal contribution of the bonus and promotion score is small for short overtime hours, and it is even negative for long overtime hours. This result could imply that the bonuses and promotions tied to individual performance induce only little additional effort under poor performance monitoring and targeting practices. Also, in such a case, the bonuses and promotions tied to

group/firm performance could induce relatively more effort than those tied to individual performance.

Panel (b) of Figure 2 similarly plots the estimates of the “marginal contribution” of the monitoring and targeting score (defined by $\widehat{\beta}_2^k \Delta MT_{jt} + \widehat{\alpha}^k BP_{jt} \Delta MT_{jt}$) in percentage terms of the mean of the dependent variable, where ΔMT_{jt} is one standard deviation of the change of the monitoring and targeting score from 2010 to 2015. At the average level of BP_{jt} , the marginal contributions of MT_{jt} are negative and decreasing in the level of overtime hours. The marginal contributions of MT_{jt} are negative and large at the 10th percentile of BP_{jt} , while they are close to zero at the 90th percentile of BP_{jt} . This is possibly because two countervailing effects of MT_{jt} are offsetting each other. On one hand, an increase in MT_{jt} could imply collecting more accurate information about individual performance, and this would induce greater effort of workers, especially when their pay/promotions are more strongly tied to the individual performance (as implied by the positive interaction term). On the other hand, an increase in MT_{jt} could also represent improving efficiency of workload allocation across workers and time, which would lead to a reduction of long overtime (as discussed in section 3.5), irrespective of the level of BP_{jt} . Therefore, at a low level of BP_{jt} , the second effect dominates the first effect, while the two effects offset each other at a high level of BP_{jt} .

Overall, the result above suggests that the relationship between the management practices and disparities of overtime hours across workers can depend on the baseline level of the management practice scores. In particular, the adoption of more structured bonus and promotion practices is associated with increasing disparities at a high level of monitoring and targeting score, while it is associated with decreasing disparities at a low level of monitoring and targeting score. However, the introduction of more structured monitoring and targeting practices is generally

associated with a reduction in disparities of overtime hours across workers at almost any level of bonus and promotion score.

3.7 Other Channels

We have explored the explanations for the changes in *actual* overtime hours in the previous sections. An alternative interpretation of the result on non-HRM practices is that a higher monitoring score is associated with changes in *reported* overtime hours (not necessarily with actual overtime hours) through better monitoring of overtime hours.²⁵ If some workers over-report hours, better monitoring of hours may reduce the reported hours to the level of actual hours. Since our data on working hours is in principle a payroll record of the workers, it indicates the working hours for which the establishment paid the workers. Therefore, under this alternative explanation, the results imply that an improvement in the monitoring score eliminates a part of the establishment's wage bill that was used to pay workers to produce nothing. However, we argue that this alternative does not explain our results. This is because under this assumption, the coefficients of monitoring and targeting should be universal across all levels of overtime hours (because workers have an incentive to misreport their overtime hours), while we observe large and significant coefficients only for long overtime hours.

Finally, the results are robust to the alternative specifications. First, we find similar results when we change the variable controlling for demand shock from the intermediate input cost to the log of the establishment's sales (see Table A.15 in the Appendix). Second, our main results are robust to controlling for possible differential trends of overtime hours by the establishment's

²⁵ A 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 were mostly top-level managers of the establishments, and there were only a few respondents from HRM departments. Even so, there is a possibility that monitoring practices regarding production and hours are correlated.

baseline characteristics (see Table A.15 in the Appendix for the results). As discussed in Section 2.5, we found that small non-Japanese establishments tend to improve management practices more than other types of establishments. If these establishments have differential time trends of overtime hours for reasons that are not related to changes in management practices, our main results are biased. We examined this possibility by additionally controlling for the establishment's baseline characteristics (including the proxy variables for the traditional Japanese employment practices) interacting with the second period (years 2015–2016) dummy to equation 1 and found similar patterns to the baseline results.²⁶ Third, we test whether the results are driven by recall errors in retrospective questions on practices in 2010.²⁷ We found that our results remain qualitatively same when we restrict the sample to JP-MOPS respondents whose tenure is longer than seven years (see Table A.16 in the Appendix). Fourth, the results are also robust to redefining categorical management scores by dropping MOPS questions regarding managers (questions 11, 12, 14, and 16 in Table A.3 in the Appendix), as these questions may be less relevant for workers in our sample.²⁸ As Table A.17 in the Appendix shows, using the redefined scores gives a similar result, although the coefficient estimates become noisier.

4. Discussions

²⁶ In this analysis, we assume that the time trend (that is, the conditional expectation of the change in the dependent variable from 2010–2011 to 2015–2016) is linear in the establishment's baseline characteristics. An alternative specification that avoids such a linearity assumption is to control for some categorical variables constructed from the establishment's baseline characteristics interacting with second period dummy. We also conducted this exercise and found qualitatively similar results. The results are provided upon request.

²⁷ According to Bloom et al. (2019), a comparison of management practice scores in 2010 using the 2010 and 2015 US-MOPS, which asked the same retrospective questions as ours, shows that the tenure of the manager responding to the survey at the establishment is an important determinant of recall errors. In particular, they show that the response quality is high if the respondent had been in the establishment at least one year before the period of recall.

²⁸ Although we dropped some managers in the BSWs from our analysis sample, the definition of these managers is solely based on job titles and does not include some workers who actually supervise other workers (Higuchi and Kambayashi 2018). On the other hand, the JP-MOPS defines managers more broadly within its questionnaire based on task content. Therefore, questions regarding managers in the JP-MOPS would still apply to managers who remain in our BSWs analysis sample. For this reason, we use all MOPS questions, including those regarding managers in our main specification.

4.1 Labor Turnover and Management Practices

The results in the previous sections could be potentially affected by changes in worker composition through hiring and separation, which in turn, may be influenced by changes in displacement policy or other management practices. To examine this possibility, we estimate the association between changes in management practices and changes in labor turnover measured using three types of data.

First, we use the data on hiring and separation at establishments from the Survey of Employment Trends (SET) conducted by the MLHW (see the Data Appendix A.1 for details about the SET). From the data, we define the “hiring rate” (or “separation rate”) by the total number of hires (or separations) of full-time workers in a calendar year divided by the number of full-time workers at the beginning of the year at each establishment. We then match the hiring and separation rates in 2010 and 2011 to the management data for 2010 from the JP-MOPS, and those in 2014 and 2015 to the management data for 2015. Table 5 shows the results of regressing the hiring and separation rates on the management scores, year fixed effects, establishment fixed effects, controlling for demand shocks by either the log of intermediate input cost or the log of shipment value in the year.

Overall, the results indicate only a weak association between changes in management practices and changes in labor turnover. In particular, we may expect that modification of management policies toward more frequent displacements results at a higher rate of separation, in which case there may be negative or positive selection of the remaining workers (Jovanovic 1979). However, the coefficients of the displacement score for the separation rate are small and

statistically insignificant.²⁹ This result is not surprising, however, given the persistent prevalence of the Japanese long-term employment system in which full-time workers are rarely fired in normal times. The result suggests that, although changes in displacement score might have resulted in a few cases of actual displaced full-time workers, it did not have a large effect on worker composition. The coefficients of the other management scores were also small and statistically insignificant. One limitation of this analysis, however, is that the sample size of the SET that matched the JP-MOPS was small, with around 400 establishment-year observations. Thus, we complemented our analysis using other types of data.

The second set of data used was the employee sample of the BSWS. We aggregate the employee data at the establishment-year level by computing the share of workers hired in the past five years, using the information on tenure. We then decompose this share into new graduate and mid-career hires using information on age. The remainder, those whose tenure is more than five years, are those who have stayed in the establishment for the preceding five years. In Table A.18 in the Appendix, we regress the share of newly hired workers and that of staying workers on management practices, controlling for year fixed effects, establishment fixed effects, and the log of intermediate input cost. The estimated coefficients of the displacement score are small and insignificant, which is consistent with earlier results. Furthermore, we find only a weak and insignificant association between changes in other management scores and changes in the share of newly hired workers.³⁰

Lastly, we can also use the establishment survey of the BSWS, which asks establishments about the number of new graduate hires. Using these data matched with the JP-MOPS, we find

²⁹ The coefficients imply that a one-standard-deviation increase in the displacement score is associated with just 0.70 percentage points ($=0.162*0.043*100$) increase in the separation rate, while the average separation rate in the sample is 10%.

³⁰ For example, an improvement in the monitoring and targeting score by 0.1 (which is roughly equivalent to the standard deviation of changes in this variable) is associated with only a 1.8% reduction in the share of hired workers.

only an insignificant association between changes in hiring of new graduates and changes in management practice scores (see Table A.19 in the Appendix). The results are similar when we decompose the number of new graduate hires by gender.

In sum, these results indicate only a weak and insignificant relationship between changes in management practices and changes in labor turnover. This result may reflect two features of our empirical study. First, the Japanese labor market is known for its slow and inflexible labor adjustment for full-time workers relative to other OECD countries. For instance, comparing the average rate at which workers flow into and out of unemployment³¹ during 1968–2009 across countries, Barnichon and Garda (2016) show that Japanese inflow and outflow rates are approximately 0.5% and 20%, respectively, while those in the US are approximately 3.5% and 57%, respectively. Additionally, we examine changes in management and outcomes over only five years, which may be too short for the adoption of new management practices to affect the composition of workers. Therefore, our results are best interpreted as short-run relationships between management practices and labor outcomes before substantial compositional adjustments occur. Although the long-run analysis is out of the scope of this paper, more structured management may gradually attract workers who prefer these practices, in which case, the long-run relationships between management practices and overtime hours are also possibly affected by these newly hired workers' preferences for overtime hours.

4.2 Inequality of Overtime Hours within Establishments

The results of our analysis suggest that improving some of the management scores – in particular, the monitoring and targeting score – is associated with declining inequality in overtime hours

³¹ Given the lack of internationally comparable data on recent hiring rates across countries, labor turnover into and out of unemployment is a useful approximation of the frequency of hiring and separation.

within establishments. Next, we directly test this hypothesis in two ways. The first is the use of quantile regression. Note that, while our baseline specification allows us to examine changes in the likelihood of working above certain hours, it does not fully explain changes in the shape and location of within-establishment distributions of overtime hours. Quantile regressions shed light on these aspects. Following the quantile regression method with panel fixed effects proposed by Canay (2011), we use the residuals after the 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 the overall management scores from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term by quantile regression. The estimated coefficients of the interaction term, summarized in Panel (a) of Figure 3, indicate that an improvement in the overall management scores is associated with small increases in overtime hours at around 45th and 50th percentiles of the distributions and large decreases around 80th to 95th percentiles.³² When we repeat a similar process for estimating the quantile regression for each of the three disaggregated scores (Figure 3 Panel (b)–(d)), we find that the former is driven by changes in the bonus and promotion scores, and the latter is mostly driven by changes in the monitoring and targeting scores, which is consistent with our baseline results.

The second way to test the implications of within-establishment inequality is to regress establishment-year-level inequality indicators on the management scores, controlling for year and 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) and on the three disaggregated management scores in column (2). Consistent with the results in Table 1, the

³² Since the distributions of overtime hours are truncated at 0, it is noteworthy that nonzero changes at lower percentiles of the distributions are observed only for a limited fraction of the establishments. Therefore, the results for lower quantiles may not be informative for selection issues (see Figure A.5 in the Appendix for more details).

coefficient of the overall management score is positive and insignificant, and the coefficient on the bonus and promotion score is positive and significant. In columns (3)–(4) and columns (5)–(6), the dependent variables are the difference in overtime hours between the maximum and the minimum in the establishment and the difference between the 95th and 5th percentiles in the establishment, respectively. The coefficients of the overall management practice are negative in columns (3) and (4), which seem to be mainly driven by the negative coefficients of the monitoring and targeting score. Although these coefficient estimates are noisy due to the low statistical power, their sizes and directions are consistent with the earlier results, suggesting that establishments that improved their monitoring and targeting scores narrowed the gap between the high and low tails of the distribution of overtime hours.

5. Concluding Remarks

Since overwork can cause problems for both productivity and workers' lives, firms' management of long working hours is an important issue for society. Nevertheless, how firms' management practices are related to working hours is still largely unknown. For economists, how working hours are affected by both labor demand and supply factors has been a topic of great interest. However, to date, despite recent empirical studies (using large-scale data obtained from surveys of firm managers) showing that HRM and non-HRM factors strongly affect firm performance, empirical evidence linking working hours and management practices has been limited to a few studies showing the link between promotion and long working hours relying on worker-side data. Therefore, in this study, we obtained large panel data concerning management practices of Japanese manufacturing establishments and matched them with the overtime hours data of their employees.

We obtained the following results. The adoption of more structured bonus and promotion practices (wage-related HRM) is associated with increasing share of employees working short-to-medium overtime hours. This result is consistent with the standard theoretical result which predicts the compensation and promotion schemes tied to individual performance induce worker effort. However, we also find that the introduction of more structured monitoring and targeting practices (non-HRM) is associated with reductions in long overtime hours. One explanation is that when firms collect detailed production processes and systems data and review them regularly, they can make continuous improvements to their production processes and systems, which in turn allows leveling of workload and time across workers and reducing the problems that lead to unnecessary overtime. In addition, setting production targets at a moderate level of difficulty and sharing them across workers helps adjust workload and time across workers, which in turn, also reduces the problems that cause unnecessary overtime. As a result, using more structured non-HRM practices seems to play a key role in narrowing the disparity in hours across workers within establishments.

Finally, we find only a weak and insignificant association between changes in management practices and changes in hiring and separation rates. This implies that changes in worker composition induced by changes in management practices do not significantly affect the above results. Although this result does not exclude the possibility that changes in management practices lead to significant compositional adjustments in the long run, we interpret our results as short-run relationships between the changes in management practices and those in the labor market.

References

- Bandiera, Oriana, Greg Fischer, Andrea Prat, and Erina Ytsma. 2021. “Do Women Respond Less to Performance Pay? Building Evidence from Multiple Experiments.” *American Economic Review: Insights*, 3(4): 435-54.
- Barnichon, Regis, and Paula Garda. 2016. “Forecasting Unemployment across Countries: The Ins and Outs.” *European Economic Review*, 84: 165–183.
- Bartling, Björn, and Ferdinand A. von Siemens. 2011. “Wage Inequality and Team Production: An Experimental Analysis.” *Journal of Economic Psychology*, 32(1): 1–16.
- Bell, Linda A., and Richard B. Freeman. 2001. “The Incentive for Working Hard: Explaining Hours Worked Differences in the US and Germany.” *Labor Economics*, 8(2): 181–202.
- 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.
- Berniell, Inés, and Jan Bietenbeck. 2020. “The Effect of Working Hours on Health.” *Economics & Human Biology*, 39: 100901.
- Bewley, Truman 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*, 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.
- Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence.” *The Quarterly Journal of Economics*, 117(1): 339–376.
- 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(3): 368–386.
- Collewet, Marion, and Jan Sauermann. 2012. “Working Hours and Productivity.” *Labour Economics*, 47: 96–106.
- Cornwell, Christopher, Ian M. Schmutte, and Daniela Scur. 2021. “Building a Productive Workforce: The Role of Structured Management Practices.” *Management Science*, 67(12): 7308–7321.

- 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.
- Cygan-Rehm, Kamila, and Christoph Wunder. 2018. “Do Working Hours Affect Health? Evidence from Statutory Workweek Regulations in Germany.” *Labour Economics*, 53: 162–171.
- 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.
- Gosnell, Greer K., John A. List, and Robert D. Metcalfe. 2020. “The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains.” *Journal of Political Economy*, 128(4): 1195–1233.
- 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.
- Hamermesh, Daniel S., and Joel B Slemrod. 2008. “The Economics of Workaholism: We Should Not Have Worked on This Paper.” *The B.E. Journal of Economic Analysis & Policy*, 8(1): 1–30.
- Hamermesh, Daniel S., Daiji Kawaguchi, and Jungmin Lee. 2017. “Does Labor Legislation Benefit Workers? Well-being after an Hours Reduction.” *Journal of the Japanese and International Economies*, 44: 1–12.

- Hashimoto, Masanori, and John Raisian. 1985. "Employment Tenure and Earnings Profiles in Japan and the United States." *American Economic Review*, 75(4): 721–735.
- Higuchi, Yoshio, and Ryo Kambayashi. 2018. "One Aspect of Japanese Mangers." RIETI Discussion Paper 18-J-013.
- Holmstrom, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies*, 66(1): 169–182.
- Jovanovic, Boyan. 1979. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(5): 972–990.
- Kambayashi, Ryo, and Takao Kato. 2016. "Long-Term Employment and Job Security over the Past 25 Years: A Comparative Study of Japan and the United States." *Industrial Labor Relations Review*, 70(2): 359–394.
- Kambayashi, Ryo, Atsushi Ohyama, and Nobuko Hori. 2021. "Management practices and productivity in Japan: Evidence from six industries in JP MOPS." *Journal of the Japanese and International Economies*, 61: 101152.
- Kato, Takao, Hiromasa Ogawa, and Hideo Owan. 2017. "Working Hours, Promotion, and the Gender Gap in the Workplace." IZA DP No. 10454.
- Landers, René M., James B. Rebitzer, and Lowell J. Taylo. 1996. "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms." *The American Economic Review*, 86(3): 329–348.
- Lemos, Renata, Ali Choudhary, John Van Reenen, and Nicholas Bloom. 2016. "Management in Pakistan: First Evidence from Punjab." IGC Working Paper.
- Ministry of Health, Labor and Welfare. 2021. "Guidance for Completing the Basic Survey on Wage Structure." Ministry of Health, Labor and Welfare.

OECD. 2018. Working Better with Age: Japan, Ageing and Employment Policies. OECD Publishing.

OECD. 2021. “Better Life Index.” <https://www.oecdbetterlifeindex.org/topics/work-life-balance/>, Accessed: 2021-11-03.

Office for National Statistics. 2018. “Management Practices and Productivity in British Production and Services Industries – Initial Results from the Management and Expectations Survey: 2016.” Office for National Statistics.

Pega, Frank, Bálint Náfrádi, Natalie C. Momen, Yuka Ujita, Kai N. Streicher, Annette M. Prüss-Üstün, Alexis Descatha, Tim Driscoll, Frida M. Fischer, Lode Godderis, Hannah M. Kiiver, Jian Li, Linda L. Magnusson Hanson, Reiner Rugulies, Kathrine Sørensen, and Tracey J. Woodruff. 2021. “Global, Regional, and National Burdens of Ischemic Heart Disease and Stroke Attributable to Exposure to Long Working Hours for 194 Countries, 2000–2016: A Systematic Analysis from the WHO/ILO Joint Estimates of the Work- related Burden of Disease and Injury.” *Environment International*, 154: 106595.

Pencavel, John. 2015. “The Productivity of Working Hours.” *Economic Journal*, 125(589): 2052–2076.

Pencavel, John. 2016. “Recovery from Work and the Productivity of Working Hours.” *Economica*, 83(332): 545–563.

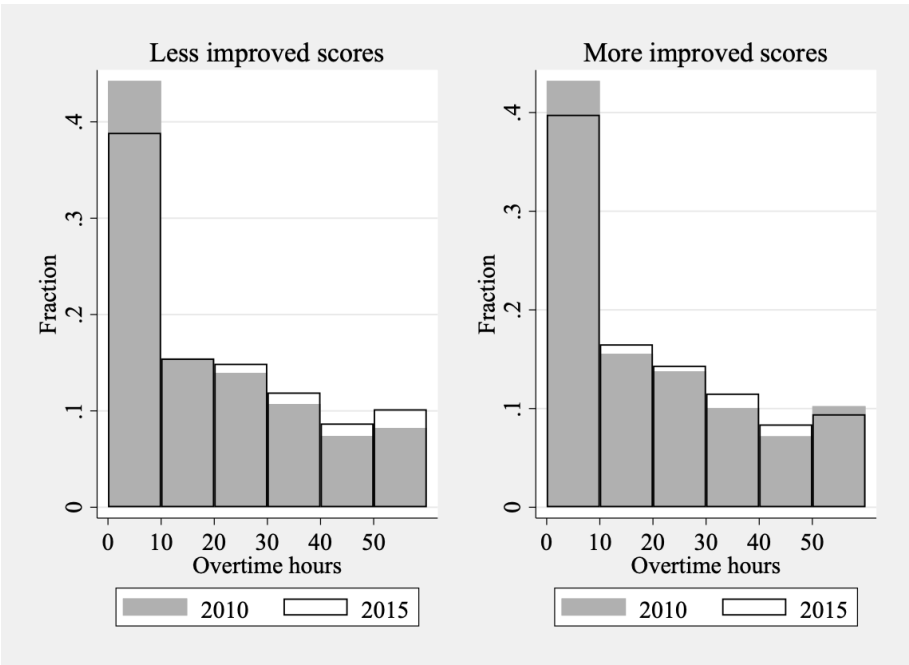
Prescott, Edward C. 2004. “Why Do Americans Work So Much More Than Europeans?” *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1): 2–12.

Rogerson, Richard. 2006. “Understanding Differences in Hours Worked.” *Review of Economic Dynamics*, 9(3): 365–409.

Sousa-Poza, Alfonso, and Alexandre Ziegler. 2003. "Asymmetric Information about Workers' Productivity as a Cause for Inefficient Long Working Hours." *Labour Economics*, 10(6): 727–747.

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

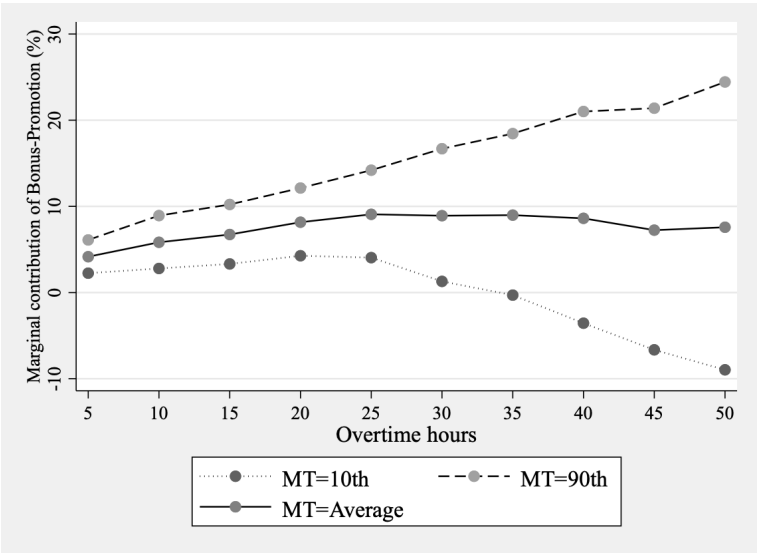
Figure 1. Distribution of Overtime Hours by Changes in Management Scores



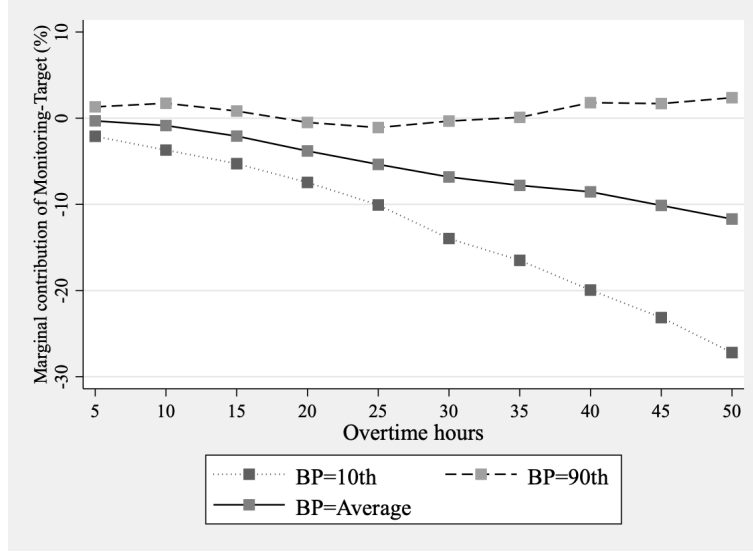
Notes: We divide the sample of establishments into two groups according to whether the change of the overall management scores from 2010 to 2015 is below (left panel) or above (right panel) the median of the changes in the sample. The figures show the histograms of overtime hours among male workers by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

Figure 2. Complementarity of Bonus-Promotion and Monitoring-Targeting: Graphical Representation

(a) Marginal Contribution of Bonus–Promotion, holding Monitoring–Targeting fixed

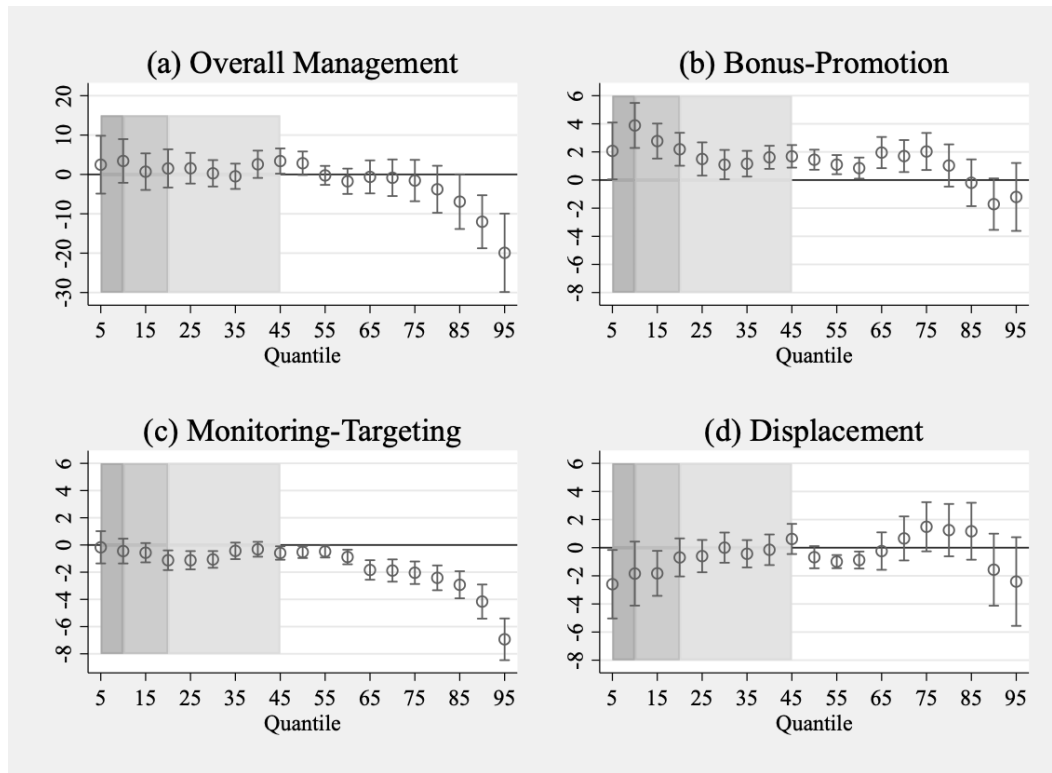


(b) Marginal Contribution of Monitoring–Targeting, holding Bonus–Promotion fixed



Notes: The figure visualizes the results of Table 4. In panel (a), we plot the estimates of the marginal contribution of bonus and promotion score, defined by $\widehat{\beta}_1^k \Delta BP_{jt} + \widehat{\alpha}^k MT_{jt} \Delta BP_{jt}$, where ΔBP_{jt} is one standard deviation of the change of the bonus and promotion score from 2010 to 2015. The panel depicts this value at three levels of the monitoring and targeting score (MT_{jt}): the 10th percentile, the sample average, and the 90th percentile in 2010. In panel (b), we plot the estimates of the marginal contribution of monitoring and targeting score, defined by $\widehat{\beta}_2^k \Delta MT_{jt} + \widehat{\alpha}^k BP_{jt} \Delta MT_{jt}$, where ΔMT_{jt} is one standard deviation of the change of the monitoring and targeting score from 2010 to 2015. The panel depicts this value at three levels of the bonus and promotion score (BP_{jt}): the 10th percentile, the sample average, and the 90th percentile in 2010. The marginal contributions in the y-axis are shown in percentage terms of the mean of the dependent variable (that is, an indicator of working overtime above the hours shown on the x-axis).

Figure 3. Quantile Regressions for Overtime Hours



Notes: The figure summarizes the results of quantile regressions for overtime hours. We first obtain the residuals after the 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 management scores from 2010 to 2015, an indicator for observations in 2015 or 2016, and their interaction term using quantile regressions. As the management score, Panel (a) uses the overall management score, Panel (b) uses the bonuses and promotion score, Panel (c) uses the monitoring and targeting score, and Panel (d) uses the displacement score, respectively. The estimated coefficients of the interaction term are presented in the figure. The 95% confidence intervals were calculated based on standard errors estimated by bootstrapping with 500 replications. As the distributions of overtime hours are truncated at 0, non-zero changes at the lower percentiles of the distributions are observed only for a limited fraction of establishments. The gray areas indicate the magnitude of this issue by color. As shown in Figure A.5 in the Appendix, 1) about three-quarters of the establishments have 0 as the 10th percentile of the establishment's distribution of overtime hours, 2) about half of the establishments have 0 as the 20th percentile of the distribution of overtime hours, and 3) about a quarter of the establishments have 0 as the 45th percentile of the distribution of overtime hours.

Table 1. Average Overtime Hours and Management Practices

	(1)	(2)	(3)
	Overtime Hour		
Overall Management Score	2.553*** (0.495)	-2.436 (4.747)	
Bonus–Promotion			8.623** (4.053)
Monitoring–Targeting			-5.579 (3.713)
Displacement			-1.228 (2.323)
Observations	72,180	72,180	72,180
Worker Attributes	Yes	Yes	Yes
Demand Control	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Establishment FE	No	Yes	Yes
Mean dep. var.	18.07	18.07	18.07

Notes: Column (1) controls only for year fixed effects, worker attributes (age, age squared, tenure, tenure squared, female dummy, and three education dummies), and the log of intermediate input cost. Columns (2)–(3) additionally control for the establishment of fixed effects. Standard errors are clustered at the level of the establishment and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. Distribution of Overtime Hours and Management Practices

	(1) Over 5h	(2) Over 10h	(3) Over 15h	(4) Over 20h	(5) Over 25h	(6) Over 30h	(7) Over 35h	(8) Over 40h	(9) Over 45h	(10) Over 50h
Bonus–Promotion	0.216*** (0.083)	0.253*** (0.083)	0.250*** (0.093)	0.253** (0.101)	0.227** (0.103)	0.171* (0.093)	0.132* (0.073)	0.091 (0.062)	0.050 (0.050)	0.038 (0.044)
Monitoring–Targeting	-0.007 (0.071)	-0.027 (0.078)	-0.070 (0.089)	-0.115 (0.096)	-0.133 (0.095)	-0.133 (0.085)	-0.118* (0.068)	-0.097* (0.058)	-0.082* (0.048)	-0.071* (0.040)
Displacement	-0.073 (0.057)	-0.042 (0.055)	-0.004 (0.055)	-0.015 (0.051)	-0.012 (0.050)	-0.016 (0.046)	-0.034 (0.040)	-0.028 (0.036)	-0.034 (0.032)	-0.021 (0.030)
Observations	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180
Mean dep. var.	0.634	0.536	0.457	0.380	0.309	0.249	0.195	0.149	0.106	0.0797

Notes: “Over a H” is defined by an indicator variable that takes one if the worker works more than a hours of overtime in the week, and zero otherwise. 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 three education dummies). Standard errors are clustered at the level of the establishment and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3. Distribution of Overtime Hours and Management Practices by Workers'

Characteristics

Panel A	(1)	(2)	(3)	(4)
	Over 10h	Over 10h	Over 45h	Over 45h
Sample: By Tenure	Junior	Senior	Junior	Senior
Bonus–Promotion	0.313*** (0.093)	0.179* (0.096)	0.110** (0.056)	-0.007 (0.067)
Monitoring–Targeting	-0.025 (0.093)	-0.051 (0.085)	-0.106** (0.054)	-0.058 (0.059)
Displacement	-0.065 (0.067)	-0.022 (0.061)	-0.055 (0.044)	-0.021 (0.030)
Observations	35,268	36,906	35,268	36,906
Mean dep. var.	0.563	0.510	0.119	0.0926

Panel B	(1)	(2)	(3)	(4)
	Over 10h	Over 10h	Over 45h	Over 45h
Sample: By Gender	Male	Female	Male	Female
Bonus–Promotion	0.234*** (0.086)	0.342*** (0.118)	0.018 (0.057)	0.124* (0.066)
Monitoring–Targeting	-0.027 (0.079)	-0.056 (0.128)	-0.090 (0.056)	-0.058 (0.058)
Displacement	-0.049 (0.056)	-0.009 (0.081)	-0.038 (0.039)	-0.022 (0.035)
Observations	56,322	15,842	56,322	15,842
Mean dep. var.	0.587	0.356	0.124	0.0408

Panel C	(1)	(2)	(3)	(4)
	Over 10h	Over 10h	Over 45h	Over 45h
Sample: By Propensity	High P(10h)	Low P(10h)	High P(45h)	Low P(45h)
Bonus–Promotion	0.288*** (0.095)	0.251*** (0.095)	0.054 (0.060)	0.049 (0.052)
Monitoring–Targeting	0.013 (0.089)	-0.079 (0.089)	-0.111* (0.065)	-0.046 (0.043)
Displacement	-0.084 (0.064)	-0.001 (0.063)	-0.047 (0.045)	-0.024 (0.026)
Observations	36,383	35,791	36,243	35,931
Mean dep. var.	0.639	0.431	0.144	0.0672

Notes: In Panel A, columns labeled “Junior” use the sample of workers whose tenure is less than 11 years (median of tenure in the sample), and columns labeled “Senior” use the rest. In Panel B, columns labeled “Male” use the sample of male workers, and columns labeled “Female” use the female sample. In Panel C, the propensity function $P(ah)$ is defined for each worker by $\text{Prob}(\text{Overtime hours} \geq a | X)$ for each a ($=10$ and 45) and estimated using a probit model using the BSWs data in 2009 (see Appendix Table A.3 for the estimation result), where X is the set of workers’ attributes. Columns labeled “High $P(ah)$ ” use the sample of workers having above the median level of the estimated $P(ah)$, and columns labeled “Low $P(ah)$ ” use the rest of the sample. 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 three education dummies. Standard errors are clustered at the level of the establishment and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. Testing Complementarity of Bonus–Promotion and Monitoring–Targeting

	(1) Over 5h	(2) Over 10h	(3) Over 15h	(4) Over 20h	(5) Over 25h	(6) Over 30h	(7) Over 35h	(8) Over 40h	(9) Over 45h	(10) Over 50h
Bonus–Promotion	0.051 (0.157)	0.032 (0.131)	0.038 (0.137)	0.052 (0.142)	0.015 (0.142)	-0.086 (0.133)	-0.114 (0.110)	-0.156* (0.094)	-0.149* (0.081)	-0.142* (0.076)
Monitoring–Targeting	-0.235 (0.193)	-0.333* (0.178)	-0.363* (0.191)	-0.392* (0.201)	-0.424** (0.203)	-0.489** (0.190)	-0.458*** (0.168)	-0.438*** (0.147)	-0.357*** (0.132)	-0.318*** (0.121)
Displacement	-0.070 (0.055)	-0.039 (0.053)	-0.000 (0.053)	-0.011 (0.049)	-0.009 (0.049)	-0.012 (0.045)	-0.030 (0.039)	-0.024 (0.034)	-0.031 (0.031)	-0.018 (0.029)
Monitoring–Targeting × Bonus–Promotion	0.352 (0.282)	0.473* (0.251)	0.453* (0.251)	0.430* (0.254)	0.451* (0.253)	0.551** (0.246)	0.526** (0.219)	0.528*** (0.197)	0.426** (0.176)	0.383** (0.163)
Observations	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180	72,180
Mean dep. var.	0.634	0.536	0.457	0.380	0.309	0.249	0.195	0.149	0.106	0.0797

Notes: This table shows the estimates of equation 2 that adds the interaction term of the bonus–promotion and monitoring–targeting scores to the baseline specification in Table 2. Standard errors are clustered at the level of the establishment and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5. Labor Turnover

	(1) Hiring rate	(2)	(3) Separation rate	(4)
Bonus–Promotion	-0.225 (0.155)	-0.205 (0.156)	0.080 (0.158)	0.095 (0.163)
Monitoring–Targeting	-0.068 (0.079)	-0.055 (0.081)	0.030 (0.065)	0.041 (0.064)
Displacement	-0.040* (0.022)	-0.033 (0.023)	0.043 (0.027)	0.042 (0.027)
Ln(Cost of intermediate inputs)	0.008 (0.008)		-0.008 (0.006)	
Ln(Shipment)		-0.002 (0.008)		-0.017** (0.009)
Observations	398	401	398	401
Mean dep. var.	0.110	0.111	0.101	0.102

Notes: This table uses the sample of the Survey of Employment Trends (2010, 2011, 2014, 2015) in the balanced panel of establishments that are matched to the JP-MOPS for at least one year in each period of 2010–2011 and 2014–2015 and have non-missing observations for the establishment’s total cost of inputs. All regressions control for year fixed effects, establishment fixed effects, and either the log of intermediate input cost or the log of the shipment value of the year. “Hiring rate” is the number of full-time workers hired during the year divided by the number of full-time workers at the beginning of the year. “Separation rate” is the number of full-time workers that left their jobs during the year divided by the number of full-time workers at the beginning of the year. Standard errors are clustered at the level of the establishment and are reported in parentheses.

Table 6. Distribution of Overtime Hours within Establishments

	(1) Average overtime hours	(2) Average overtime hours	(3) Max - min of overtime hours	(4) Max - min of overtime hours	(5) 95th - 5th percentile of overtime hours	(6) 95th - 5th percentile of overtime hours
Overall Management Score	1.308 (4.456)		-9.955 (9.261)		-5.469 (8.816)	
Bonus–Promotion		6.370* (3.613)		1.093 (7.891)		3.568 (7.525)
Monitoring–Targeting		-1.447 (3.518)		-7.199 (7.037)		-3.353 (6.964)
Displacement		-1.667 (2.342)		-0.614 (5.508)		-2.110 (4.816)
Observations	2,307	2,307	2,307	2,307	2,307	2,307
Mean dep. var.	17.26	17.26	49.05	49.05	41.29	41.29

Notes: This table uses the sample of the BSWS (in 2010–2011 and 2015–2016) aggregated at the level of establishment-year and matched to the JP-MOPS data on management practices in 2010 and 2015. All regressions control for year fixed effects, establishment fixed effects, and the log of intermediate input cost. “Average overtime hours” is the average of overtime hours in the establishment in the year. “Max - min of overtime hours” denotes the difference between the maximum and the minimum overtime hours in the establishment in the year. “95th - 5th percentile of overtime hours” denotes the difference between the 95th percentile and 5th percentile in the year. Standard errors are clustered at the level of establishments and are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.