The Impact of Overtime Limits on Firms and Workers: Evidence from Japan's Work Style Reform*

Gabriel Burdin
University of Siena & IZA

Ryo Kambayashi
Musashi University & IZA

Takao Kato
Colgate University & IZA

December 29, 2024

Abstract

This study provides the first analysis of Japan's 2018 Work Style Reform (WSR) and its effects on firms and workers, using payroll and survey data in a difference-in-difference design. We find that the reform's introduction of an overtime cap reduces average monthly overtime hours by 5 hours (-25%) and compresses the distribution of overtime within establishments. Total earnings decrease by 2% due to reduced overtime pay, while hourly wages remain unchanged. Notably, the reform improves life and leisure satisfaction, but these well-being gains are observed only among women. This gender difference is not explained by variations in perceived work intensification or time use. Instead, we find evidence that men (but not women) substitute paid overtime for unpaid overtime, which is consistent with the lack of well-being gains for men. Finally, we document that the reform leads to women taking more career jobs (standard employment) relative to non-career jobs (nonstandard employment) as compared to their male counterparts, highlighting the potential of working-hour regulations to promote gender equality in the labor market.

Keywords: Working Time Regulations, Overtime, Wages, Employment, Subjective Well-being, Gender, Japan, Work Style Reform

*We thank conference and seminar participants at LUBS (University of Leeds), IAFEP, AIEL, EALE, Trans-Pacific Labor Seminar (UCSB) and Workshop on Well-Being and The Future of Industrial Relations (Teiyko University) for providing valuable comments. The authors gratefully acknowledge funding from the Joint Usage and Research Center Program at the Institute of Economic Research, Hitotsubashi University. Gabriel Burdin also expresses his gratitude to the Japanese Society for the Promotion of Science (JSPS) for its generous travel funding in support of this research. This paper utilizes the micro data from the Preference Parameters Study of Osaka University's 21st Century COE Program 'Behavioral Macro-Dynamics Based on Surveys and Experiments', its Global COE project 'Human Behavior and Socioeconomic Dynamics' and JSPS KAKENHI 15H05728 'Behavioral-Economic Analysis of Long-Run Stagnation'.

1 Introduction

Work-life balance and related job amenities may be underprovided in the labor market, potentially leading to negative economic consequences.[1] Employer monopsony power, for instance, can compel individuals to work longer hours than they would prefer at prevailing wage levels (Manning, 2003). Similarly, asymmetric information may lead to a "rat race" dynamic, where workers put in inefficiently long hours to signal their ability and commitment to employers (Akerlof, 1976; Landers et al., 1996; Kato et al., 2016). Moreover, the prevalence of a gendered division of labor within households means that long working hours and rigid workloads disproportionately hinder women's career progression, thereby perpetuating persistent gender disparities in the labor market (Goldin, 2014; Frederiksen et al., 2024). Such enduring disparities present significant policy challenges for many advanced economies grappling with declining birth rates and rapidly aging populations, impacting both equity and efficiency, and underscoring the underutilization of highly educated women.

Growing concerns about workforce well-being and the need to sustain productivity have prompted many countries to implement reforms promoting work-life balance. These measures often include mandatory reductions in working hours and stricter regulations on overtime. However, robust evidence on the effectiveness of policies designed to reduce overwork across diverse contexts remains limited.[2]

This paper examines for the first time the causal impact of Japan's recently introduced Work Style Reform (WSR) on both firms and workers. The Japanese Prime Minister at the time, Shinzo Abe, described the new regulations as 'the first major reforms [to labor laws] in 70 years.' The new regulation came into force in April 2019 and introduced a maximum overtime cap of 360 hours per year (equivalent to 30 overtime hours in a typical month). Japan offers a compelling case study due to its entrenched long working-hour culture and its implication for the large and persistent gender gap in the labor market. In 2018, roughly 16% of Japanese standard workers were supplying overtime hours above the new threshold.[3] Compared to other devel-

[1] For recent evidence highlighting the adverse economic impact of work-related stress and burnout, see Nekoei et al. (2024).

[2] Another rationale for working time regulations is to reduce unemployment through "work sharing." However, the evidence generally indicates either negligible or adverse employment effects (Crépon and Kramarz, 2002; Batut et al., 2023).

[3] In Japan, a standard employee is defined as a worker termed *"seiki no jyuugyouin"* in the place of his/her employment. For Japan's labor market segmentation between standard and non-standard

oped nations, Japan ranks in the bottom fifth in work-life balance and struggles with persistent issues such as long working hours and underutilized paid leave (Jones and Seitani, 2019). Karoshi—death from overwork—remains a significant social concern (Nishiyama and Johnson, 1997). To address the adverse effects of overwork, Japan has introduced various policies over the decades, including a gradual reduction of the standard workweek from 48 to 40 hours (1980s–1990s) and increased overtime premiums for employers (Lee et al., 2012; Hamermesh et al., 2017). The 2018 WSR marked a critical milestone by implementing, for the first time, a mandatory cap on overtime hours.

The 2018 WSR is also of profound importance for Japan's policymakers who have been pursing their main policy goal of narrowing the gender gap in the labor market over the last decade or so (e.g., late Prime Minister Abe's "target of increasing the share of women in leadership positions to at least 30% by 2020 in all fields in society"). According to the OECD, the gender gap in median earnings for full-time employees in Japan was approximately 21% in 2022 (the third largest in the OECD), almost twice as high as the OECD average. The persistent large gender pay gap in Japan is particularly troublesome for Japanese policymakers, for the gender gap in educational attainment narrowed considerably in Japan. In 1980, only 12.3% of women advanced to the university-level. By 2023, it rose to 54.5% (the School Basic Survey, 2022). As the proportion of college-educated women has increased, the worker composition of full-time workers has changed dramatically. In particular, there has been a significant increase in the proportion of female university graduates among standard employees from 1.1% in 1981 to 29.0% in 2023 (Basic Survey of Wage Structure, BSWS). Despite these improvements in average female educational attainment, however, a significant gender wage gap persists in Japan. To be consistent with the sustained gender wage gap, so few women are promoted to managerial positions. According to the most recent Labor Force Survey, the proportion of female college graduates in managerial positions was still only 13% in 2022 (far short of late Prime Minister Abe's 30% target by 2020). In this paper we will show that Japan's WSR is proving to be a promising policy instrument for breaking down Japan's structural impediment for gender equality in the labor market.

For our empirical analysis, we utilize a unique array of establishment and individual-

workers, see for instance (Kambayashi and Kato, 2016).

level survey data, combining payroll and self-reported information on working hours, including paid and unpaid overtime, wages, employment and subjective well-being. Our identification strategy leverages exogenous variation in overtime hours across firms and workers resulting from the reform, which imposed an annual overtime limit of 360 hours (equivalent to 30 hours per month). Using a difference-in-differences (DiD) framework, we compare establishments with varying initial levels of affected jobs before and after the reform. We then analyze the individual-level impacts of the reform by comparing individuals whose overtime hours exceeded the new cap just before the reform compared to those who did not.

As intended, monthly overtime per worker in affected establishments decreased by 5 hours (-25%) compared to the control group. The distribution of overtime hours within these establishments also became less dispersed, with the proportion of high-overtime jobs falling. Although hourly wages remained unchanged, total monthly wages fell by 2%, primarily due to a reduction in overtime pay, which was only partially offset by an increase in base wages. Additional survey evidence supports the finding of a reduction in self-reported long workweeks and overtime at the individual level. In addition, the reduction in working hours led to improved life and leisure satisfaction, but only among female workers. These gender differences are not driven by differential changes in perceived work intensification (*jitan-harasumento*) or time use (e.g., commuting, housework). Among male workers, we find that the decline in paid overtime was partly offset by an increase in unpaid overtime. Finally, we find evidence of improvements in female job quality. Specifically, in the context of a pre-existing rising trend in standard employment among women, we observe a decline in nonstandard (precarious) female jobs in affected establishments.

This paper makes four contributions. First, it contributes to the emerging policy assessment literature on working time regulations by providing the first rigorous evidence on the effects of such a regulation introduced recently by Japan. Incorporating Japan's experience with the recent working time regulation into the literature will provide valuable new insights. First, the current literature focus mostly on European experiences where excessively long working hours are less of a concern as compared to Japan (according to the most recent Databook of International Labour Statistics 2024 by JILPT, among G7 countries, Japan has the highest share of workers working 49 or more hours per week (15.3%) in 2022, which is almost twice as high as the remaining

G7 countries. Effective working time regulations are probably most urgently needed for Japan, and Japanese policymakers have been clearly aware of the need considered and placed moderating long working hours (along with reducing the gender gap in the labor market), their top policy priority. Second, the recent literature on the gender gap in the labor market focuses on high returns to long working hours as a possible last culprit for the gender gap in wages and promotion (Bertrand et al., 2010; Kato et al., 2013; Cha and Weeden, 2014; Goldin, 2014; Cortes and Pan, 2017; Frederiksen et al., 2024). Our paper joins the literature by providing new evidence and insights on whether or not policy interventions to moderate long working hours can work as an effective means to narrow the gender gap in the labor market.

Second, most prior studies focus on understanding the employment effects of mandatory reductions in the standard workweek when firms are required to hold monthly wages constant, as implemented in some European countries. Existing research generally finds little support for work-sharing arguments, as workers and overtime hours tend to be imperfect substitutes in production (Hunt, 1999; Crépon and Kramarz, 2002; Rocheteau, 2002; Batut et al., 2023; Asai et al., 2024). Moreover, Carry (2023) documents adverse employment effects of France's minimum workweek policy, with female workers disproportionately affected. In contrast, our study explores the effects of a less common intervention: a mandatory overtime cap with no requirement for firms to maintain constant monthly earnings.

Third, we add to the body of research on overtime hours. Prior studies have examined the employment effects of overtime taxation and regulations (Trejo, 2003; Oaxaca, 2014; Cahuc and Carcillo, 2014; Martins, 2017), and, within the Japanese context, the influence of management practices and voluntarily implemented work-life balance initiatives by firms (Tanaka et al., 2022; Takahashi et al., 2024). Others have examined the determinants of unpaid overtime (Bell and Hart, 1999; Pannenberg, 2005). We add to this strand of literature by analyzing for the first time how a mandatory *paid* overtime cap causally affects the provision of *unpaid* overtime hours and the distribution of overtime hours within establishments.

Finally, we contribute to the literature on working hours and subjective well-being (SWB). Several papers have analyzed the impact of reductions in the standard workweek in countries such as Korea (Rudolf, 2014; Hamermesh et al., 2017), Portugal and France (Sánchez, 2017; Lepinteur, 2019; Berniell and Bietenbeck, 2020), and Germany

(Cygan-Rehm and Wunder, 2018). Notably, Hamermesh et al. (2017) document positive effects of shorter workweeks on life satisfaction in Japan during the 1990s. A recent study by Carcillo et al. (2023) investigates Korea's overtime limits but does not assess their effect on labour market outcomes and SWB. Our study is the first to assess the causal impact of Japan's Work Style Reform, analyzing firms' adjustment to overtime limits and changes in workers' subjective well-being within an unified framework.

The remainder of the document is organized as follows: Section 2 outlines the institutional context and details of Japan's Work Style Reform. Section 3 describes the data. Section 4 presents the identification strategy and main results regarding establishment-level responses to reform, alongside a set of robustness checks. Section 5 delves into the individual-level impacts. Finally, Section 6 concludes.

2 Institutional context and Japan's Work Style Reform

Long Working Hours and Institutions. The Japanese economy has long been known for its extensive working hours. Japan ranks in the bottom fifth in the OECD in work-life balance. In 2017, 12% of male workers worked 60 hours a week or more (Jones and Seitani, 2019). These long working hours have led to severe health issues, including the phenomenon of death from overwork, known as karoshi (Bassanini and Caroli, 2015). Following a series of institutional changes between the mid-1980s and mid-1990s, Japan's standard workweek was set at 40 hours. Employers were required to pay extra for overtime work, with rates depending on the firm's size. However, prior to the reform, there were no upper limits on overtime hours as long as employers had a written agreement with their employees (Labor Standard Act, Article 36).

The 2018 Work Style Reform. The Work Style Reform Bill, enacted in June 2018, introduced mandatory limits on overtime hours for the first time in Japan.[4] In relation to the bill, then Prime Minister Abe at that time remarked, "These are the first major reforms [to labor laws] in 70 years. We will address the problem of long working hours..." Trade unions and business organizations widely supported the legislation,

[4]The bill was reviewed by the Labor Policy Council (an advisory panel to the Minister of Health, Labour and Welfare) in September 2017 and approved by the cabinet in April 2018. Deliberations in the House of Representatives and House of Councillors (the lower and upper houses of the National Diet of Japan, respectively) took place between April and June 2018. The timing of the legislative process is significant, as firms and workers may have anticipated the reform's effects.

particularly its aim to tackle the issue of excessive overtime. [5]

The bill introduced key amendments to two central pieces of labor law in Japan: the Labor Standards Act (LSA) and the Industrial Safety and Health Act (ISHA). These laws form the backbone of the legal framework governing employment relationships. Under the new legislation, which came into force in April 2019 for large firms and April 2020 for small and medium-sized enterprises (SMEs), overtime work is capped at 45 hours in any single month and 360 hours per year (approximately 30 hours in a typical month). Non-compliant firms face fines of up to JPY 300,000 (approximately USD 2,660) per worker. In special circumstances, firms can negotiate with their employees to require up to 100 hours of overtime per month, as long as the annual limit of 720 hours is not exceeded. Certain highly specialized professionals and occupations—such as drivers, doctors, and research and development (R&D) professionals—are exempt from the overtime limits. While the overtime cap is the primary focus of our analysis, the legislation introduced additional measures aimed at improving worker health and well-being. For example, starting in April 2023, SMEs no longer benefit from a reduced overtime pay rate and are required to pay the same 50% additional wage rate for overtime hours as large firms. The new law also mandates that employers allow workers to take at least five days of annual paid leave. Finally, the bill introduces a work-interval system to ensure workers have enough rest between working days. [6][7]

3 Data

To analyze the impact of the WSR on firm and worker outcomes, we utilize two main datasets detailed in this section.

Establishment payroll data. The Basic Survey on Wage Structure (BSWS) is an employer-employee matched survey conducted annually on June 30 by the Japanese Ministry of Labor, Health, and Welfare (MLHW). It primarily replicates company payroll records, providing accurate data on *actual* working hours and wages for June, including overtime hours and pay. Additionally, the survey gathers worker charac-

[5] For further information, see here.

[6] We use the data covering 2014-2022, and hence the elimination of the overtime pay exception for SME should not affect our difference-in-difference analysis of the effects of the 2018 WSR. As shown below, our results are not influenced by variations in the number of working days.

[7] For more details, see here.

teristics such as education, gender, tenure, age, and contract type (e.g., fixed-term, nonstandard).[8] Establishment-level attributes include employment size, industry, and location (prefecture). The sample size is large, covering approximately one million workers across fifty thousand establishments each year—representing roughly 5% of all establishments in Japan. Sampling is conducted in two stages: first, establishments with five or more employees are selected using a uniform sampling method; second, employees are sampled within each selected establishment. While the dataset has a panel structure at the establishment level, it does not track individual employees over time. Consequently, we aggregate employee-level information to the establishment level for analysis.

Individual-level survey data. While establishment payroll data allows for accurate measurement of paid working hours, wages, and employment, examining other potentially important outcomes and channels requires the use of self-reported information from workers. To better understand the impact of the Work Style Reform (WSR) at the individual level, we use panel data from the Preference Parameters Study conducted by Osaka University (OPPE). This survey has been conducted annually since 2003, with the exceptions of 2014–2015 and 2019–2020, and is representative of the Japanese population aged 20 to 69 years. Several key features of this dataset are particularly relevant to our study. First, the survey collects self-reported data on wages and *usual* weekly working hours. Specifically, respondents are asked: "How many hours per week do you and your spouse usually work, including overtime?" They then report their weekly overtime hours, distinguishing between paid and unpaid overtime.[9] Second, the survey gathers data on various aspects of subjective well-being (SWB), including cognitive measures (e.g., life and job satisfaction), affective measures (e.g., happiness, stress, anxiety), and eudaimonic measures (e.g., "My daily life is fulfilling").[10] Finally, the survey collects a broad range of behavioral traits and attitudes (e.g., risk preferences, willingness to compete, conformism, social comparisons, beliefs) along with detailed demographic controls, such as gender, age, tenure, occupation, industry, and employer size.

[8] For contract types used for employment in Japan, see (Kambayashi and Kato, 2016).

[9] Previous research indicates that self-reported working hours in Japan are highly valid (Imai et al., 2016). In Section 5.5, we examine the potential impact of misreporting on our results.

[10] SWB measures have been extensively validated and correlate with neural activity and a range of behaviors (Urry et al., 2004; Clark, 2016; Liberini et al., 2017; Borga et al., 2022). However, there is no consensus on the underlying utility concept behind these measures (Benjamin et al., 2023).

Other data. We also use the Survey on Labor-Management Communication (SLMC), conducted by the Ministry of Health, Labor, and Welfare in Japan every five years. SLMC examines various aspects of labor-management communication at the workplace level. Using establishment-level data from the SLMC, we calculate the prevalence of worker voice institutions—such as joint labor-management committees, shop-floor committees, and unions—across industries.[11] This aggregated industry-level data is then applied in Section 4.5 to assess whether establishments' responses to the Work Style Reform vary according to their sectoral worker voice regime.

4 Establishment-Level Responses

In this section, we analyze the impact of the overtime cap introduced by the 2018 Work Style Reform at the establishment level, using the BSWS payroll data described earlier.

4.1 Research Design: Establishment-Level Difference-in-Differences

We construct an establishment-level panel spanning 2014–2022 by aggregating employee-level data.[12] As outlined in Section 2, the Work Style Reform, implemented in April 2019, introduced a cap on overtime hours: 45 hours in any single month and 360 hours annually (equivalent to 30 hours per typical month).

Our strategy exploits differences in establishments' structural dependence on overtime work.[13] To quantify an establishment's exposure to the regulation, we calculate the pre-reform share of full-time employees working more than 30 overtime hours per month before 2019, or *% high overtime workers*. To reduce the impact of temporary shocks, we average this exposure variable across the entire pre-reform period (2014–2019). The treatment indicator, denoted as $HighShare30hrs_j$, equals 1 if establishment j's pre-treatment average of *% high overtime workers* exceeds the median value of all establishments, and 0 otherwise.[14] This research design enables a comparison of

[11] For worker voice institutions in Japan, see for instance Kato and Morishima (2002).

[12] In Appendix TableA1, we report the cross-sectional correlates of overtime hours at the individual level. The provision of overtime hours increases with age, tenure, and is more common among individuals employed under fixed-term contracts and in large establishments. Holding constant other factors, women supplies roughly 6 overtime hours less than men. The likelihood of supplying more than 30 overtime hours per month is 10 percentage points lower among female workers.

[13] Our approach is similar to Carry (2023) and Asai et al. (2024).

[14] For the average (median) establishment in the sample, the pre-reform share of workers supplying more than 30 overtime hours is 13% (1%). In section 4.6, we assess the robustness of our main results to alternative treatment indicators.

establishments with varying initial levels of affected jobs before and after the reform. In Figure 1, we report the evolution of overtime hours per worker in treatment and control establishments, while Panel A of Table A2 in Appendix provides descriptive statistics on the estimation sample.

To obtain average post-reform treatment effects, we estimate the following difference-in-differences (DiD) specification:

$$y_{jt} = \eta(HighShare30hrs_j \times Post_t) + \mu_j + \delta_t + \psi X_{jt} + \epsilon_{jt} \quad (1)$$

where y_{jt} are the outcomes for establishment j in year t (average overtime per worker, wages, employment), $Post_t$ is a post-reform dummy equals one after 2018 and zero otherwise, $HighShare30hrs_j$ is the above-defined treatment group dummy. μ_j are establishment fixed effects accounting for time-invariant unobserved attributes, while δ_t are year dummies. We also control for time-variant establishment-level characteristics, including workforce composition (gender, age, tenure, education) and establishment size. In certain specifications, we additionally account for industry- and prefecture-specific time trends. Coefficient η captures the impact of the reform.

In our analysis, identification relies on the assumption that, absent the reform, establishments with varying structural needs for overtime work would have followed similar trajectories. Specifically, if the observed effects stem from the new overtime limits introduced by the Work Style Reform, highly exposed establishments should not display differential trends relative to the control group in the pre-reform period. To test for potential pre-trends, we estimate pre-reform effects by plotting year-specific DiD estimates for all outcome variables. These estimates correspond to the interaction between $HighShare30hrs_j$ and a full set of year dummies, with the 2018 coefficient normalized to zero.

4.2 Effects on Working Hours

Average Overtime and Total Working Hours. The primary objective of the reform was to reduce long working hours by capping overtime. A natural starting point is to assess whether the reform achieved its intended first-stage effect: did highly exposed establishments adjust overtime hours as expected in comparison to the control group? To evaluate the reform's impact over time and verify that highly exposed establishments were not on a differential pre-reform trend, Panel (a) of Figure 2 presents year-specific DiD estimates for average monthly overtime hours. During the pre-reform

period (2014–2018), the estimated effects are negligible in magnitude compared to the post-reform period, despite some statistical significance.[15] Beginning in 2019, the first year after the reform's implementation, overtime hours declined significantly, with this downward trend continuing into 2020. A slight uptick in overtime hours was observed in subsequent years.[16]

Panel (b) and (c) of Figure 2 provide estimates for base hours and total monthly working hours, respectively. These event-study graphs reveal that the reduction in overtime hours resulted in a decrease in total working hours, which was only partially offset by an increase in base hours. Table 1 reports the pooled DiD estimates, along with clustered standard errors at the establishment level in parentheses, for the entire post-reform period using Equation (1). Column (1) reports an average reduction of five monthly overtime hours per worker in affected establishments relative to the control group—a decrease of approximately 25% compared to the pre-reform mean. This reduction was partially offset by an average increase of one base hour (Column 2), resulting in a net decrease of four total working hours per worker per month (Column 3). Relative to the pre-reform average, the decline in total hours is small in magnitude (-2%).

Overtime Dispersion. We also investigate whether the reform led to a redistribution of overtime hours within establishments. To capture this, we use the intra-establishment standard deviation of overtime hours as our outcome variable. Panel (d) of Figure 2 presents the event-study graph for this measure of overtime dispersion. The results indicate that overtime hours became less dispersed following the reform. Importantly, pre-trends are minimal in comparison to the post-reform effects, underscoring the impact of the policy. The differential compression of overtime hours in affected establishments relative to the control group is further corroborated by the DiD estimates reported in Column (4) of Table 1. Additional analysis in Appendix Table A3 delves deeper into the distributional changes in overtime hours within establishments. We find that the share of workers supplying zero or only a few overtime hours per month increased, while the proportion of high-overtime jobs declined after the reform.

[15] In Section 4.6, we show that the results are robust to using alternative confidence intervals that account for potential pre-trends (Rambachan and Roth, 2023).

[16] In Section 5.2, we also document a reduction in self-reported paid overtime using individual-level survey data.

4.3 Effects on Wages

Overtime Pay and Total Monthly Earnings. We now examine how wages evolved in affected establishments compared to the control group. Notably, the Work Style Reform did not mandate firms to compensate high-overtime workers for lost income following the introduction of the overtime cap. Figure 3 presents event-study graphs illustrating changes in various components of workers' compensation packages, including overtime pay, base wages, bonuses, total monthly wages, and hourly wages. These variables are averaged at the establishment level and measured in logs.

Consistent with the reduction in overtime hours noted earlier, Panel (a) of Figure 3 depicts a sharp decline in average overtime pay among affected establishments relative to the control group. As Panel (d) demonstrates, this decline was not offset by increases in other compensation components (base wages, bonuses), resulting in a reduction in total monthly wages. Hourly wages remained stable following the reform.

Table 2 provides the corresponding DiD estimates for the pooled post-reform period. Column (2) indicates a 35% reduction in the log of overtime pay in affected establishments compared to the control group. Column (3) shows a slight increase in base wages (1%), suggesting that firms partially compensated workers for lost income, likely reflecting Japan's tight labor market conditions (Kawaguchi, 2019). Column (1) reports a net decrease of 2% in total monthly wages. As the reduction in total working hours was of a similar magnitude, hourly wages remained unchanged (see Column (4)). Unlike the European approach to working time reductions, which required firms to maintain monthly wages and thereby mechanically raised hourly wages (Crépon and Kramarz, 2002; Asai et al., 2024), the implementation of an overtime cap under Japan's Work Style Reform did not result in higher labor costs for companies.

Within-Establishment Pay Dispersion. We examine the impact on within-establishment pay dispersion. Columns (6) and (7) of Table 2 present DiD estimates for the standard deviation of total monthly wages and overtime pay. Aligned with the observed compression of overtime hours, our findings indicate a decrease in the dispersion of average overtime pay and total wages at the establishment level.

4.4 Effects on Employment.

Employment Level. Having documented a reduction in overtime hours and monthly wages, we now analyze the effect on employment. The distinction between standard and nonstandard employment is relevant in the Japanese context. Several studies have highlighted the rise of precarious nonstandard jobs relative to standard jobs and labor market dualization as significant challenges to Japan's traditional long-term employment system.[17] In our analysis, we adopt the definition of nonstandard jobs used in the BSWS, which is based on workplace titles (*seiki no jyuugyouin*). This definition encompasses the growing prevalence of nonstandard workers with open-ended contracts, offering a more comprehensive measure of the primary "good job" segment and the secondary "bad job" segment of the Japanese labour market (Kambayashi and Kato, 2016; Hijzen et al., 2015).

Figure 4 illustrates event-study evidence for the log of total, standard, and nonstandard employment. The absence of significant pre-trends differences between the treatment and the control suggests no systematic differences prior to the reform. Additionally, there is no indication of substantial employment adjustments during the post-reform period. These findings are corroborated by the DiD estimates presented in Table 3. The lack of negative employment effects is not surprising, as the reform did not increase labor costs: employers offset the changes by adjusting monthly salaries downward. As shown in the rest of the table and discussed below, however, there is a notable gender difference in the employment effects.

Workforce Composition. Firms may have responded to the reduction in overtime hours by altering the composition of their workforce. To examine this channel, Table 4 presents DiD estimates for various measures of workforce composition at the establishment level (fixed-term, nonstandard, college graduates, and part-time). Relative to the control group, the share of fixed-term and nonstandard jobs decreased by approximately 1 percentage point, representing a 4-5% reduction compared to pre-reform levels (see Columns 1-2 of Table 4). Column (3) reveals a statistically significant increase in the share of college graduates, suggesting that employers may have responded to the new regulation by enhancing worker quality. However, the magnitude of this effect is relatively small. Though the size of the effect is also small, as shown in column

[17] See, for instance, Kawaguchi and Ueno (2013).

(4), there is a statistically significant positive effect of the WSR on the share of part-time workers.

4.5 Heterogeneous Effects.

Gender. We now examine the establishment-level effects of the overtime cap separately for men and women. We replicate the analysis conducted above, but with gender-specific measures of working time, wages, and employment as outcomes. Figure 5 presents event-study plots of our main outcome variables by gender. Panels (a) and (b) of Figure 5 show reductions in overtime hours and total monthly wages for both male and female workers. Panel (c) reveals no significant changes in male employment, while there is some evidence of an increase in female employment. However, female employment in affected establishments had already been on an upward trend before the reform. As shown in Panel (d), this pre-reform trend in female employment is particularly apparent when focusing on standard jobs. Interestingly, although both female and male standard employment were on an upward trend prior to the reform, this trend diverges after the reform's implementation. The increase in standard jobs continues for female employment but abruptly halts for male employment at the time the reform takes effect. Finally, Panel (e) shows a significant reduction in nonstandard "bad" female jobs relative to the control group, with no changes observed for nonstandard male jobs. Overall, we find evidence for the positive and significant effect of the WSR on female employment in general and female standard employment in particular, but not for male employment. The evidence suggests that the WSR helps women getting and maintaining standard (good) jobs, whereas there is no similar evidence for men.[18]

Worker Voice Institutions. A well-established set of institutions facilitates labor-management communication in Japanese firms, including Joint Labor-Management Committees (JLMCs), shop-floor committees (SFCs), and unions. Notably, the WSR allows firms to exceed the new overtime limits under special circumstances, provided they negotiate a collective agreement with their workers. Unfortunately, payroll data do not contain information on the extent of labor-management communication. In-

[18]Consistent with our findings, aggregate statistics from the Labour Force Survey (Quarterly) show an upward trend in the share of standard jobs among female workers since 2014, while the share remains stable for male workers.

stead, we use the Survey on Labor Management Communication (SLMC), outlined in Section 3, to calculate the prevalence of worker voice institutions at the industry level and merge this data with our establishment-level payroll information. We then conduct DiD estimates to explore heterogeneous effects based on whether an establishment operates in an industry with a high or low incidence of worker voice institutions. Figure A1 presents event-study graphs of our main outcome variables, distinguishing establishments in industries with above- and below-median prevalence of worker voice institutions. Interestingly, we find that the reduction in overtime hours and overtime pay was less pronounced in establishments within industries characterized by a high incidence of worker voice institutions.

This finding may be consistent with the rat race/adverse selection theory of long working hours. As discussed in Frederiksen, Kato and Smith (2024), in a rat race model, the firm promotes workers with sustained low costs of working long hours. However, each worker knows his/her cost of working long hours yet the firm does not. Such asymmetric information leads to adverse selection—some workers with high costs of working long hours may have an incentive to send a false signal to the firm by working long hours in order to win promotions. To prevent such adverse selection, the firm will set its thresholds working hours inefficiently high so that nobody with high-cost of long working hours finds it optimal to send a false signal by working long hours just to win a promotion. Employee voice institutions mitigate such adverse selection by frequent and high-quality information sharing between labor and management, and thereby make it less necessary to set the threshold working hours inefficiently high. As such, the WSR is likely to be less binding and impactful for establishments with worker voice mechanisms than establishments without. It follows that the WSR will have smaller impact on overtime hours and pay for establishments with worker voice mechanisms.

4.6 Robustness

Parallel trends. Our strategy assumes that, in the absence of the WSR, establishments differently exposed to the new regulation would have evolved similarly. For all our outcomes, we report year-specific DiD estimates for the pre-reform period (2014-2018). In most cases, pre-trends are not significant or relatively small in magnitude compared to the post-reform estimates. To further assess the validity of our approach, we apply a

recent procedure proposed by Rambachan and Roth (2023) and find that the significant reduction in overtime hours remain robust to potential violations of the parallel trends assumption prior to the policy (see Appendix Figure A2).

Staggered treatment timing. We check the robustness of our main results to recently studied challenges associated with two-way fixed effect event study models with staggered treatment timing due to the presence of late-treated units. In our setting, the new regulation came into force in April 2019 for large establishments employing 300+ employees, while April 2020 was the starting application date for establishments employing not more than 300 employees (not more than 50 employees for retail businesses, not more than 100 for wholesale retail). Following Callaway and Sant'Anna (2021), we first estimate the individual cohort-time-specific treatment effects, allowing for treatment effect heterogeneity, and then aggregate these individual treatment effects to obtain overall treatment effects. The event study for overtime hours, presented in Appendix Figure A3, yields conclusions consistent with our baseline estimates.

Alternative treatment indicators. We evaluate the robustness of our findings using alternative definitions of the treatment group. First, we define treatment establishments as those with a positive share of workers exceeded 30 overtime hours per month during the pre-reform period. Second, we use a stricter threshold, considering establishments with a positive share of workers exceeding 45 overtime hours per month. Our main findings remain robust across these alternative specifications (results available upon request).

5 Additional Evidence: Individual-Level Impacts

In this section, we extend the establishment-level analysis by exploring the impact of the 2018 Work Style Reform at the individual level. To do so, we use longitudinal survey data from the Preference Parameter Study, as described in Section 3. Specifically, we leverage detailed information on self-reported working hours, including paid and unpaid overtime, perceived effort intensity, and subjective well-being (SWB).

Our analysis focuses on full-time, non-managerial employees aged 20-65 years. We exclude managers, as highly specialized professions were exempt from the new regulations. We include data from five pre-reform waves (2012, 2013, 2016, 2017, and

2018) and two post-reform waves (2021 and 2022). Panel (b) of Table A2 in Appendix summarizes descriptive statistics for the estimation sample.[19]

5.1 Individual-Level Difference-in-Differences

To examine the impact of the new overtime cap at the worker level, we classify individuals into a treatment group based on whether their paid overtime hours exceeded the new overtime cap in at least one year during the pre-reform period. The control group consists of individuals who never worked more overtime hours than the new limit in the pre-reform period. As outlined in Section 2, the reform introduced a maximum overtime threshold of 45 hours in any single month and an annual limit of 360 hours, which is equivalent to 30 hours of overtime in a typical month. Since the survey provides data on overtime hours in a typical week, our treatment group includes individuals who worked more than 30 overtime hours per month just before the reform.

We estimate the following difference-in-differences specification:

$$y_{it} = \eta(HighOvertime_i \times Post_t) + \psi X_{it} + \mu_i + \delta_t + \tau_s + \omega_p + \epsilon_{it} \quad (2)$$

where y_{it} are the outcomes for individual i in year t (e.g. long working hours, overtime hours, SWB), $Post_t$ is a post-reform dummy equals one after 2018 and zero otherwise, $HighOvertime_i$ is the above-defined treatment group dummy. μ_i are individual fixed effects, while sector τ_s and prefecture ω_p fixed effects account for time-invariant permanent differences across industries and Japanese prefectures, respectively. We also control for time-variant personal and firm-level characteristics (age, tenure, occupation and employer size). Coefficient η captures the impact of the reform. We estimate equation (1) by OLS, clustering standard errors at the individual level in order to account for serial correlation.

5.2 Self-Reported Working Time: Paid and Unpaid Overtime

First, we verify whether the reduction in overtime hours observed in the establishment-level payroll data also holds for self-reported working hours from individual survey

[19] As noted in Section 2, the reform's enforcement was staggered, impacting large firms from April 2019 and small and medium-sized firms from April 2020. However, by 2021 (the first year with post-reform data available), the legislation applied to firms of all sizes. Our results remain consistent if we restrict the analysis to individuals employed in large firms before the reform, as shown in Appendix Panel E of Table A4.

data. The results are presented in Table 5. Panel A reports unconditional DiD estimates, while Panel B includes controls for individual- and firm-level attributes (such as gender, age, tenure, firm size, occupation), along with prefecture and industry fixed effects. We also incorporate prefecture-specific time trends to account for time-varying shocks. In Panel C, we include individual fixed effects to control for unobserved time-invariant characteristics.

Our individual-level DiD estimates align with the results reported in Section 4, showing a reduction in self-reported long working hours and overtime hours. The preferred estimates in Panel C, Columns (2) and (3), indicate a 7 percentage point reduction in the incidence of long hours (i.e., individuals working more than 60 hours per week) and a reduction of 2.3 total overtime hours per week among treated individuals compared to the control group. Interestingly, as shown in Column (4) of Panel C, the decrease in total overtime is driven by a reduction in paid overtime (3.4 hours per week), partly offset by an increase in unpaid overtime (approximately 1 hour per week). Paid overtime decreased by 38% relative to the pre-reform mean.[20][21] The reduction in paid overtime holds for both men and women, though the offsetting increase in unpaid overtime is observed only for male workers.[22] Indeed, this reduction in paid overtime translates into an increase in the ratio of unpaid-to-total overtime among treated men relative to the control group (see Column (6) of Table 7). For completeness, Column (6) of Table 5 (Panel C) reports additional DiD estimates using the log of self-reported monthly wages as the dependent variable. There is a reduction in monthly wages, albeit the effect is imprecisely estimated.[23]

[20]While the results from our establishment-level and individual-level survey-based DiD estimates are qualitatively similar, the magnitude of the effect appears larger in the latter. It is important to note that survey respondents reported their own estimation of usual working hours in a typical week, while establishments reported payroll information on actual working hours in a specific month.

[21]Our preferred estimates include individual fixed effects. To further address concerns about worker sorting around the time of the reform, we conduct additional DiD estimates on a subsample of job stayers, i.e., individuals with at least 5 years of tenure with their current employer as of 2021 (the first observed post-reform year). We also perform DiD estimates using the balanced panel of individuals. The reduction in self-reported overtime hours remains robust to these modifications (see Appendix Panel (c) and (d) of Table A4).

[22]Figure 8 shows event-study plots comparing paid and unpaid overtime for both male and female workers.

[23]As mentioned in Section 2, the WSR mandates that employers provide workers with at least five days of annual paid leave. It is important to note that changes in the number of days worked may not necessarily occur within the segment of the overtime hours distribution we are analyzing. In fact, column (7) of Table 5 presents no evidence of differential changes in annual days worked. Furthermore, we do not observe significant changes in the average number of days worked in affected establishments relative to the control group (results available upon request).

5.3 Effects on Workers' Subjective Well-Being

After documenting a consistently negative first-stage effect on both payroll and self-reported working hours, we now turn to the impact of the reform on workers' subjective well-being (SWB). Specifically, we focus on measures of cognitive well-being, including life satisfaction and satisfaction with key life domains (work, spouse, leisure, and family). Due to the lack of survey data for 2019-2020, our analysis primarily captures medium-term responses to the reform. It is possible that SWB improved in the short run following the Work Style Reform, only to return to baseline levels over time due to hedonic adaptation. While anticipatory responses could be a concern in our context, such responses would likely make it more difficult to detect any genuine impact of the reform. Importantly, we find no evidence of differences in SWB during the pre-reform years (see Figure 6).[24]

Results are presented in Table 6. The dependent variables are measured on Likert scales ranging from 1 to 5, with higher values indicating greater satisfaction. In Panel C, we report results from our preferred DiD specifications, which include individual fixed effects. Overall, we find that the reduction in overtime hours did not lead to significant changes in subjective well-being (SWB).[25] However, these findings mask some heterogeneity by gender. In the female subsample (Panel D of Table 6), we observe a positive and statistically significant effect on both life satisfaction and leisure satisfaction among treated individuals relative to the control group. Specifically, the increase is 0.235 points (mean: 3.586) for life satisfaction and 0.326 points (mean: 3.353) for leisure satisfaction.

5.4 Robustness Checks and Additional Results

Parallel Trends. Similar to our analysis using establishment data, we report dynamic DiD estimates to examine whether the outcomes for treated and control individuals were on parallel trends before the reform. Figure 6 presents event-study graphs for our main individual-level outcomes—self-reported overtime hours, and measures of subjective well-being—across the years surrounding the reform. Each estimated co-

[24]There is no clear evidence of individual adaptation to reductions in working hours in other contexts (Lepinteur, 2019).

[25]Consistent with the lack of significant effects on self-reported job satisfaction, we observe no differences in the likelihood of seeking a new job, which is typically interpreted as a revealed-preference measure of job (dis)satisfaction (see Appendix Figure A5).

efficient corresponds to the interaction between T_i and a full set of year dummies, with the coefficient for 2018 normalized to zero. We find no evidence of differential pre-reform trends for any of our main variables. The reduction in overtime hours for treated individuals becomes statistically significant in 2021 and 2022. Importantly, this result holds for both unmatched and matched DiD estimates using non-parametric coarsened exact matching (Iacus et al., 2012).[26] In Figure 7, we display the results by gender and find no evidence of pre-reform trends for either male or female workers. The figures also confirm the differential increase in life and leisure satisfaction among women, as discussed earlier.

Misreporting. Our individual-level analysis uses self-reported working hours, which may be prone to measurement errors. To evaluate inconsistencies, we conduct several sensitivity checks. First, we cross-check national government statistics from both household sources (Labor Force Survey) and establishment sources (Monthly Labor Survey). This analysis indicates that the average weekly unpaid work between 2014 and 2022 was approximately 4.1 hours, closely aligning with the 3.8 hours observed in our sample. Second, we compare the sum of standard weekly hours (40 hours) and reported overtime to total usual weekly hours reported by individuals in OPPE. Figure A4 in the Appendix shows a histogram of the differences, distinguishing treated and control groups. The modal value of zero suggests generally accurate reporting, with symmetric inconsistencies around zero and some group differences. To address this, we re-estimate DiD effects excluding individuals with inconsistent reports. Panel A of Table A4 confirms robust first-stage effects on paid overtime, even with the significant reduction in sample size. Excluding extreme values (top 2%) in Panel B yields similar results.

Covid-19: Individuals' Ability to Work from Home. Our establishment-level DiD estimates show that the reduction in overtime hours is already evident in 2019, suggesting that the Covid-19 pandemic is unlikely to confound our results. However, the first post-reform survey wave used in our individual-level analysis is from 2021. While the pandemic itself should not affect our estimates unless it impacted treated

[26]This procedure improves the comparability between treated and control individuals by matching them based on observable characteristics. Specifically, we first match individuals using pre-reform characteristics measured in 2018 (such as gender, age, and firm size), and the matching weights are then used to estimate the DiD model.

and control individuals differently, we recognize that differences in the ability to work from home could influence both subjective well-being and the reporting of working hours. To address this concern, we conduct additional DiD estimates (available upon request) excluding individuals who were teleworking at least one day per week prior to the pandemic (January 2020) and one year after the outbreak (January 2021). Our main results remain robust to this sample restriction.

Other Well-Being Facets. We also examine potential impacts on other dimensions of well-being, beyond cognitive well-being. In Appendix Table A5, we report DiD estimates analyzing differential effects on a proxy of eudaimonic well-being (e.g., "My life is fulfilling") and affective well-being (measured by happiness on a 1-10 scale). We also examine other subjective indicators, such as health anxiety, feelings of stress, depression, and sleep problems. Consistent with our findings for cognitive well-being, we observe positive effects only among treated women. Specifically, for this subsample, we document a significant increase in happiness following the reform.

Perceived Effort Intensity. Employers may have responded to the new overtime cap by implementing strategies to extract more effort from workers within the reduced overtime hours. Work intensification could dampen the reform's impact on workforce well-being, while allowing firms to offset the reduction in overtime. Specifically, gender differences in work intensification may explain why the reform had heterogeneous impacts on subjective well-being between female and male workers. Anecdotal evidence suggests that firms may have placed additional pressure on workers to maintain output despite shorter overtime. In fact, the term *jitan-harasumento* (short-time harassment) was one of the most frequently used words in Japan in 2018 (Japan Times, 2018), becoming popular around the time of the Work Style Reform. The survey provides data on individuals' perceived work intensity, asking them how hard they work each day. Using our DiD framework, we examine whether these perceptions changed differentially around the time of the reform. Our outcome variable is a dummy that takes the value of one if the individual responds with either "Work hard and continuously" or "Could not work any harder than currently." We also allow for heterogeneous treatment effects by gender. As shown in Column (1) of Table 7, there is no evidence of gender differences in perceived work intensity among treated individuals relative to

the control group.[27]

Use of Freed Time. By capping overtime hours, the reform may have led to changes in time use and a reorganization of the workweek, potentially freeing up time for other activities. Changes in time use may be different for men and women, potentially explaining why the reform has gender-specific effects on SWB, as documented above. The survey provides information on daily commuting time, average housework time, and the frequency of physical activity. In Columns (3)-(5) of Table 7, we report DiD estimates with gender-specific effects. For these three dimensions, however, we find no significant changes in time use. [28]

6 Conclusions

First, using large establishment-level panel data from Japan, we have provided the first causal estimates on the effects of Japan's Work Style Reform (WSR) on firms and workers. Overtime per worker in affected establishments has been found to fall by 5 hours (-25%) compared to the control group and become less dispersed after the reform. Monthly earnings have been found to fall by 2%, primarily due to a reduction in overtime pay, while hourly wages have been found to remain unchanged.

Second, we have used unique individual worker-level survey data and have obtained collaborating evidence of a reduction in self-reported long workweeks and overtime. In addition, the reduction in working hours has been found to lead to improved life and leisure satisfaction yet the positive effect of the WSR on worker well-being has been observed only for women. We have further found that the observed gender difference in the consequences of the WSR for worker well-being had little to do with any gender-specific changes in perceived work intensification or time use. Instead, we have found intriguing evidence that among male workers, the decline in paid overtime was partly offset by an increase in unpaid overtime, which is consistent with the absence of improved life and leisure satisfaction after the reform for male workers.

[27] Firms may have modified other workplace practices in response to the reform. Unfortunately, information on workplace practices at the individual level is limited. However, by examining whether individual wages are based on work performance, we find no evidence of a differential increase in variable pay among treated individuals (see Column 2 of Table 7).

[28] Dynamic DiD estimates presented in Appendix Figure A5 further confirm the absence of differences in both commuting and housework time over the entire study period.

Finally, our analysis of the causal effects of Japan's recent WSR has yielded suggestive evidence that the reform resulted in women taking more career jobs (standard employment) relative to non-career jobs (nonstandard employment) as compared to their male counterparts. At the same time, the reform led to women improving (rather than sacrificing) their wellbeing while no such improvement found for men. To assess fully whether or not the reform will end up producing a significant increase in the proportion of female managers in the future will require many more years of data. However, our findings are certainly consistent with the growing literature stressing the importance of high returns to long working hours as the last structural impediment to gender equality in the labor market, and point to the promise of policy interventions to regulate long working hours.

References

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, 90(4):599–617.

Asai, K., Lopes, M. C., and Tondini, A. (2024). Firm-Level Effects of Reductions in Working Hours. CEPREMAP Working Papers (Docweb) 2405, CEPREMAP.

Bassanini, A. and Caroli, E. (2015). Is work bad for health? the role of constraint versus choice. *Annals of Economics and Statistics*, (119/120):13–37.

Batut, C., Garnero, A., and Tondini, A. (2023). The employment effects of working time reductions: Sector-level evidence from european reforms. *Industrial Relations: A Journal of Economy and Society*, 62(3):217–232.

Bell, D. N. F. and Hart, R. A. (1999). Unpaid work. *Economica*, 66(262):271–290.

Benjamin, D., Debnam Guzman, J., Fleurbaey, M., Heffetz, O., and Kimball, M. (2023). What do happiness data mean? theory and survey evidence. *Journal of the European Economic Association*, 21.

Berniell, I. and Bietenbeck, J. (2020). The effect of working hours on health. *Economics Human Biology*, 39:100901.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young

professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55.

Borga, L., D'Ambrosio, C., and Lepinteur, A. (2022). *Economic Perspectives on Individual Well-being*, pages 45–61. Springer Fachmedien Wiesbaden, Wiesbaden.

Cahuc, P. and Carcillo, S. (2014). The detaxation of overtime hours: Lessons from the french experiment. *Journal of Labor Economics*, 32(2):361–400.

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

Carcillo, S., Hijzen, A., and Thewissen, S. (2023). The limitations of overtime limits to reduce long working hours: Evidence from the 2018 to 2021 working time reform in korea. *British Journal of Industrial Relations*, n/a(n/a).

Carry, P. (2023). The effects of the legal minimum working time on workers, firms and the labor market. Technical report.

Cha, Y. and Weeden, K. A. (2014). Overwork and the slow convergence in the gender gap in wages. *American Sociological Review*, 79(3):457–484.

Clark, A. E. (2016). Swb as a measure of individual well-being. In Adler, M. and Fleurbaey, M., editors, *The Oxford Handbook of Well-Being and Public Policy*, pages 518–552. Oxford University Press, USA.

Cortes, P. and Pan, J. (2017). Cross-Country Evidence on the Relationship between Overwork and Skilled Women's Job Choices. *American Economic Review*, 107(5):105–109.

Crépon, B. and Kramarz, F. (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, K. and Wunder, C. (2018). Do working hours affect health? evidence from statutory workweek regulations in germany. *Labour Economics*, 53:162–171.

Frederiksen, A., Kato, T., and Smith, N. (2024). Working hours, top management appointments, and gender: Evidence from linked employer-employee data. *Journal of Labor Economics*, 0(ja):null.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4):1091–1119.

Hamermesh, D. S., Kawaguchi, D., and Lee, J. (2017). Does labor legislation benefit workers? well-being after an hours reduction. *Journal of the Japanese and International Economies*, 44:1–12.

Hijzen, A., Kambayashi, R., Teruyama, H., and Genda, Y. (2015). The japanese labour market during the global financial crisis and the role of non-standard work: A micro perspective. *Journal of the Japanese and International Economies*, 38:260–281.

Hunt, J. (1999). Has work-sharing worked in germany? *The Quarterly Journal of Economics*, 114(1):117–148.

Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20:1–24.

Imai, T., Kuwahara, K., Miyamoto, T., Okazaki, H., Nishihara, A., Kabe, I., Mizoue, T., Dohi, S., and on Occupational Health Study Group, J. E. C. (2016). Validity and reproducibility of self-reported working hours among japanese male employees. *Journal of Occupational Health*, 58(4):340–346.

Jones, R. S. and Seitani, H. (2019). Labour market reform in japan to cope with a shrinking and ageing population. OECD Working Papers 1568.

Kambayashi, R. and Kato, T. (2016). Good Jobs and Bad Jobs in Japan: 1982-2007. Working Paper Series 348, Center on Japanese Economy and Business, Columbia University.

Kato, T., Kawaguchi, D., and Owan, H. (2013). Dynamics of the gender gap in the workplace: An econometric case study of a large japanese firm. RIETI Discussion Paper Series 13-E-038.

Kato, T. and Morishima, M. (2002). The productivity effects of participatory employment practices: Evidence from new japanese panel data. *Industrial Relations: A Journal of Economy and Society*, 41(4):487–520.

Kato, T., Ogawa, H., and Owan, H. (2016). Working hours, promotion, and gender gaps in the workplace. *RIETI Discussion Paper Series*, 16(E-060):1–63.

Kawaguchi, D. and Ueno, Y. (2013). Declining long-term employment in japan. *Journal of the Japanese and International Economies*, 28:19–36.

Kawaguchi, D., M. H. (2019). The labor market in japan 2000–2018. Technical report.

Landers, R. M., Rebitzer, J. B., and Taylor, L. J. (1996). Rat race redux: Adverse selection in the determination of work hours in law firms. *The American Economic Review*, 86(3):329–348.

Lee, J., Kawaguchi, D., and Hamermesh, D. S. (2012). Aggregate impacts of a gift of time. *American Economic Review*, 102(3):612–16.

Lepinteur, A. (2019). The shorter workweek and worker wellbeing: Evidence from portugal and france. *Labour Economics*, 58:204–220.

Liberini, F., Redoano, M., and Proto, E. (2017). Happy voters. *Journal of Public Economics*, 146:41–57.

Manning, A. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Martins, P. S. (2017). Economic effects of overtime premium flexibility: Firm- and worker-level evidence from a law reform. Technical report.

Nekoei, A., Sigurdsson, J., and Wehr, D. (2024). The Economic Burden of Burnout. Technical Report 19091, C.E.P.R. Discussion Papers.

Nishiyama, K. and Johnson, J. (1997). Karoshi–death from overwork: occupational health consequences of japanese production management. *Int J Health Serv*, 27(4):625–641.

Oaxaca, R. (2014). The effect of overtime regulations on employment. Technical report.

Pannenberg, M. (2005). Long-term effects of unpaid overtime. *Scottish Journal of Political Economy*, 52(2):177–193.

Rambachan, A. and Roth, J. (2023). A More Credible Approach to Parallel Trends. *The Review of Economic Studies*, 90(5):2555–2591.

Rocheteau, G. (2002). Working time regulation in a search economy with worker moral hazard. *Journal of Public Economics*, 84(3):387–425.

Rudolf, R. (2014). Work shorter, be happier? longitudinal evidence from the korean five-day working policy. *Journal of Happiness Studies*, 15(5):1139–1163.

Sánchez, R. (2017). Does a mandatory reduction of standard working hours improve employees' health status? *Industrial Relations: A Journal of Economy and Society*, 56(1):3–39.

Takahashi, K., Kodama, N., Arita, K., Kazama, H., Sakai, S., Takeuchi, M., and Owan, H. (2024). Has japan's work style reform had the intended effect? *Applied Economics*, 0(0):1–24.

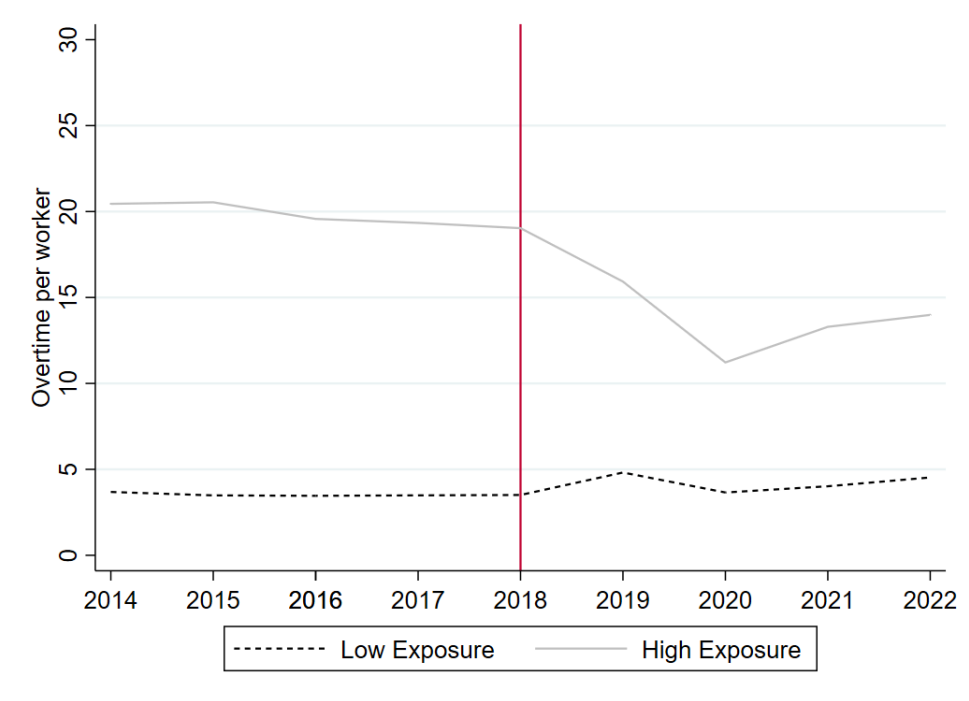
Tanaka, M., Kameda, T., Kawamoto, T., Sugihara, S., and Kambayashi, R. (2022). Managing long working hours: Evidence from a management practice survey. *Journal of Human Resources*.

Trejo, S. J. (2003). Does the statutory overtime premium discourage long workweeks? *ILR Review*, 56(3):530–551.

Urry, H., Nitschke, J., Dolski, I., Jackson, D., Dalton, K., Mueller, C., Rosenkranz, M., Ryff, C., Singer, B., and Davidson, R. (2004). Making a life worth living: neural correlates of well-being. *Psychol Sci*, Jun 15(6):367–72.

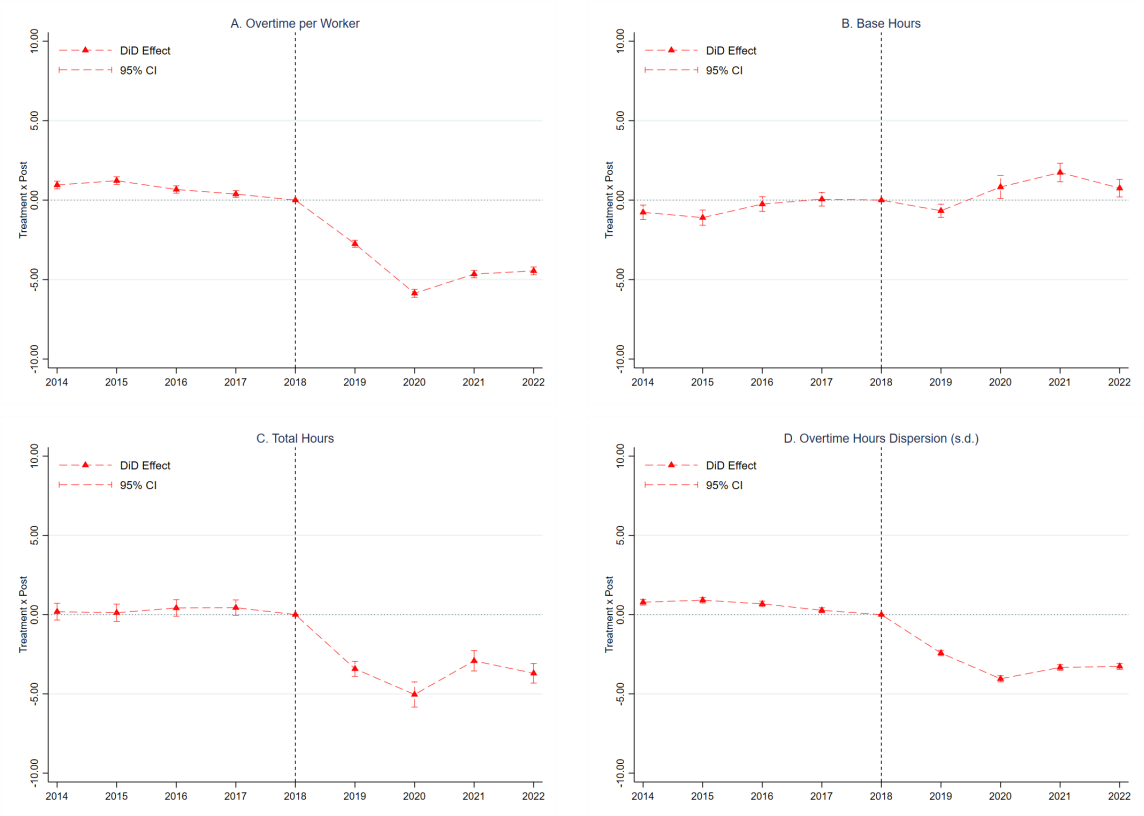
Figures and tables

Figure 1: Evolution of Overtime Hours per Worker



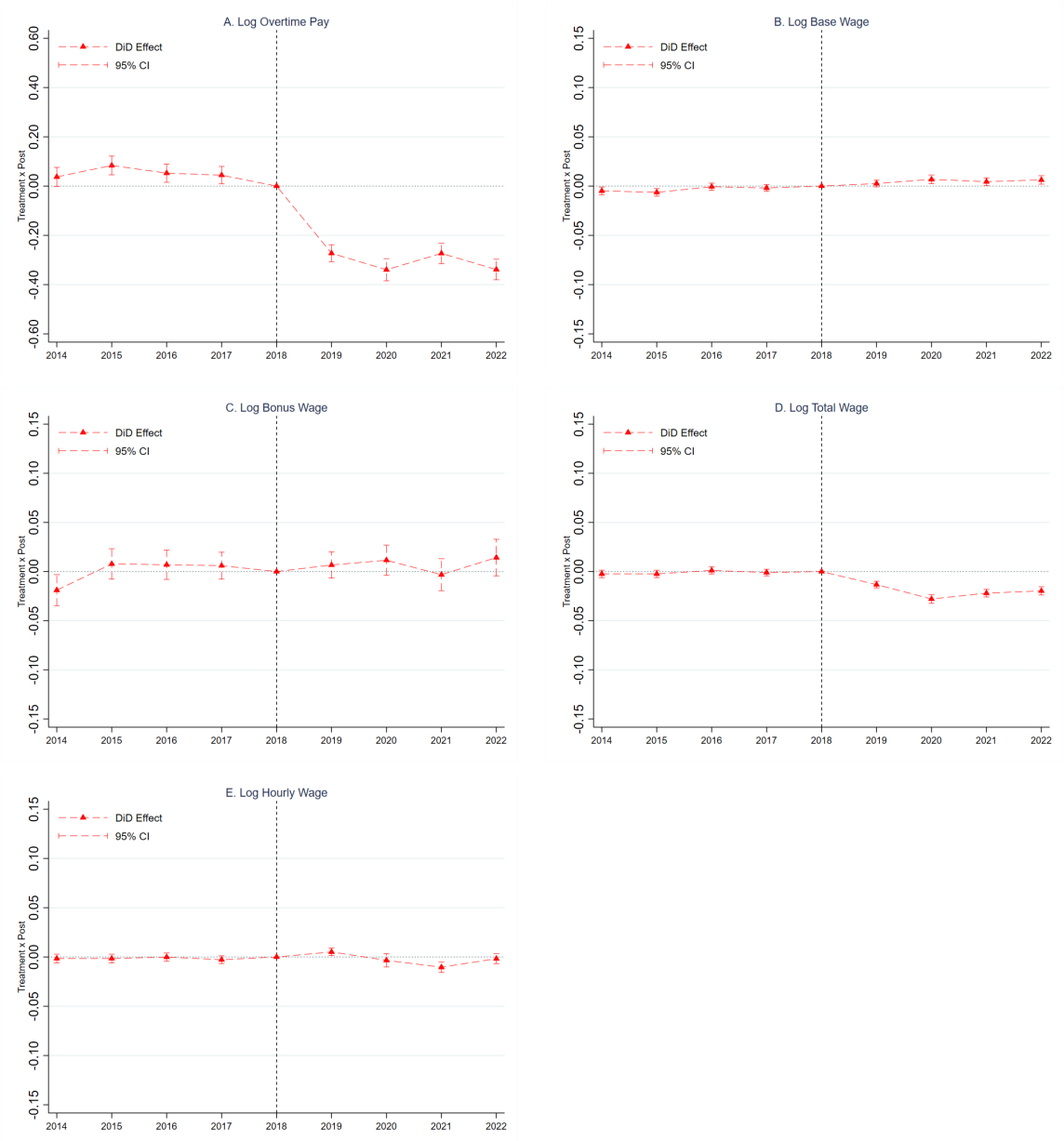
Notes: The figure plots the evolution of overtime hours per worker in ”more affected” and ”less affected” establishments. Establishments' degree of exposure considers the pre-reform share of workers supplying more than 30h overtime per month, i.e. the new WSR cap.

Figure 2: Event-Study Analysis: Hours



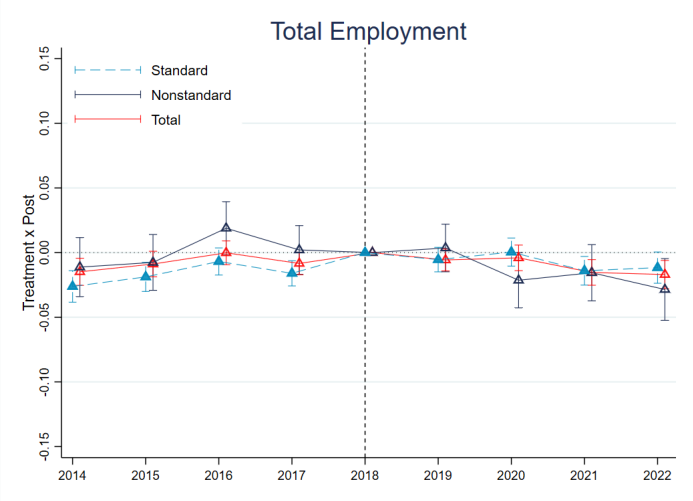
Notes: The figure year-specific DiD estimates and 95% confidence intervals. Dependent variable: Overtime per Worker (Panel A), Standard Hours (Panel B), Total Hours (Panel C) and Intra-Establishment Overtime Dispersion (sd) (Panel D). Working hours variables are measured on a monthly basis. Sample: BSWS collapsed at the establishment level. Establishment-level panel 2014-2022. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, non-standard workers), firm-size dummies. Standard errors clustered at the establishment level.

Figure 3: Event-Study Analysis: Wages



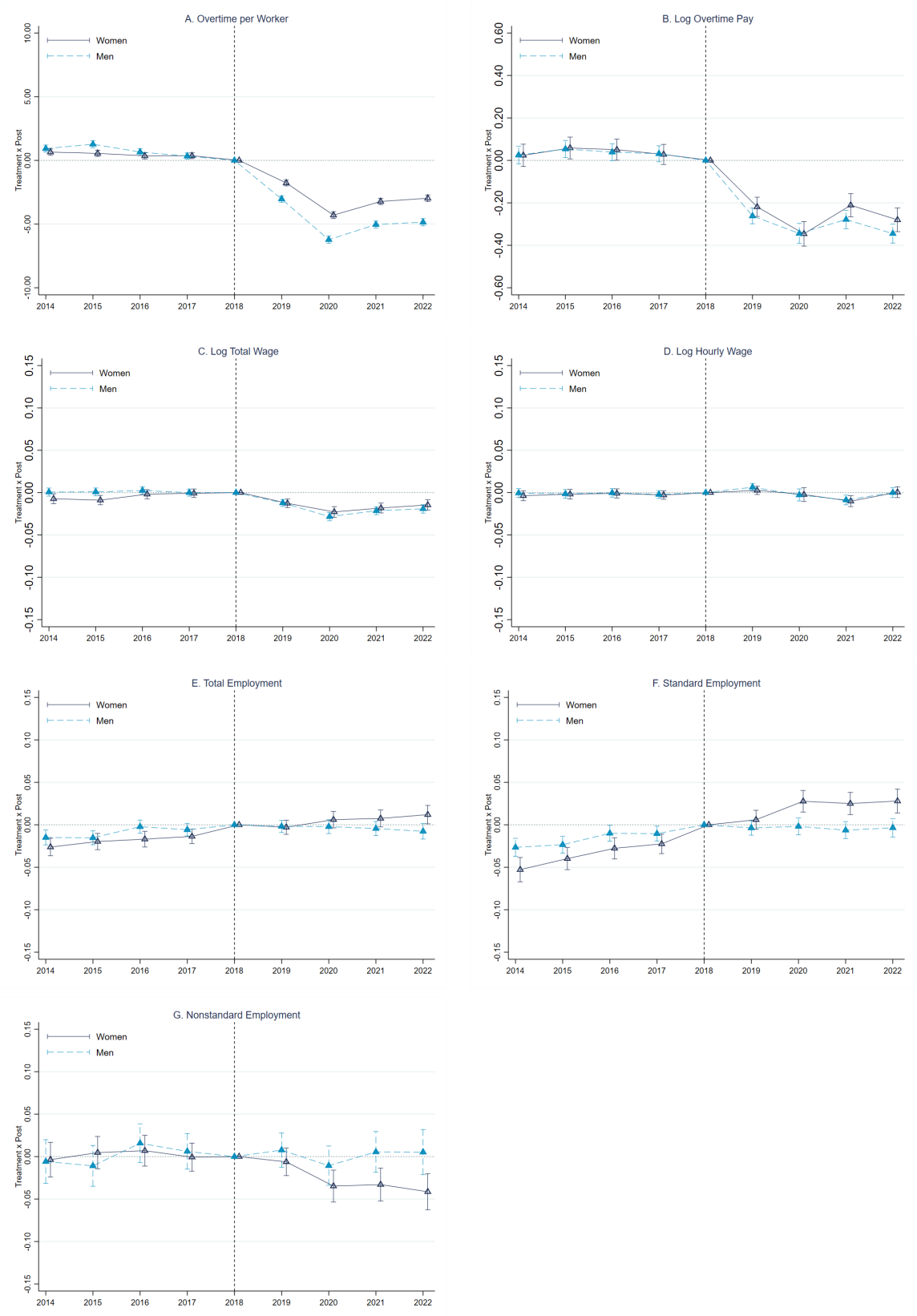
Notes: The figure year-specific DiD estimates and 95% confidence intervals. Dependent variable (in logs): Overtime Pay (Panel A), Base Wage (Panel B), Bonuses (Panel C), Total Monthly Earnings (Panel D), and Hourly Wage (Panel E). Sample: BSWS collapsed at the establishment level. Establishment-level panel 2014-2022. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies. Standard errors clustered at the establishment level.

Figure 4: Event-study Analysis: Employment



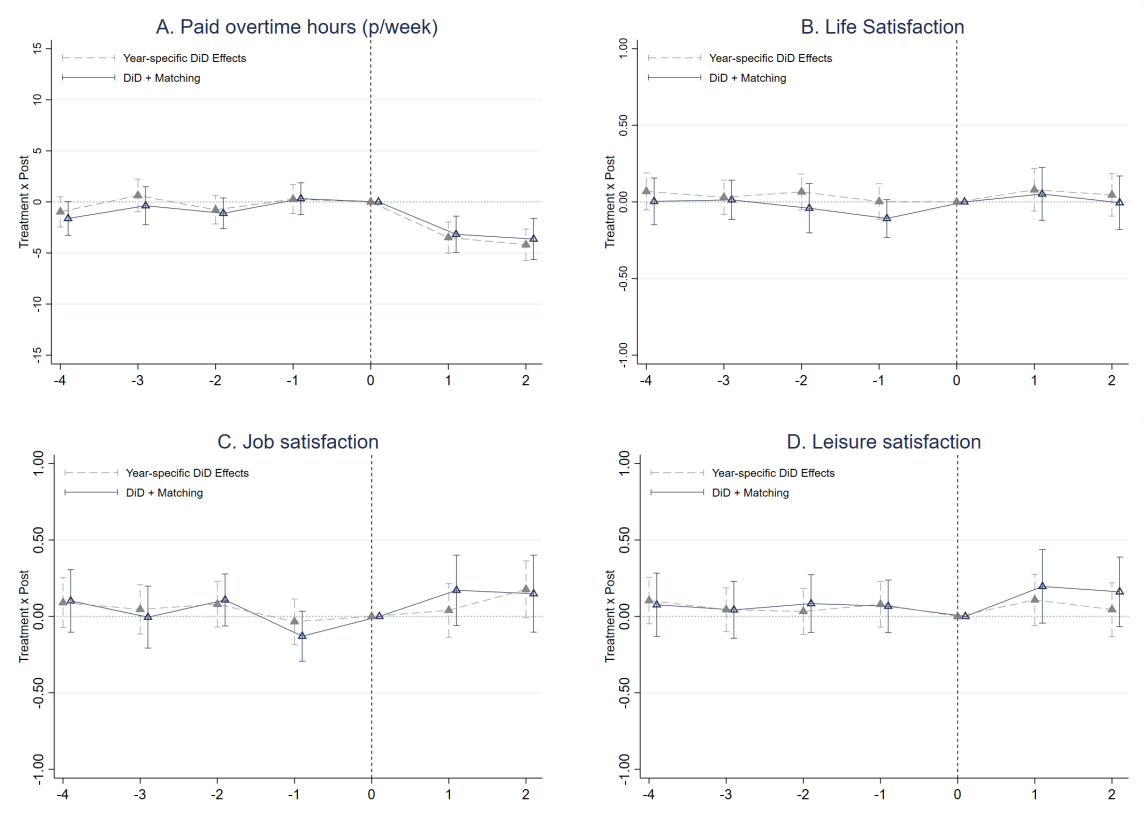
Notes: The figure year-specific DiD estimates and 95% confidence intervals. Dependent variable (in logs): Total Employment, Standard Employment, and Nonstandard Employment. Sample: BSWS collapsed at the establishment level. Establishment-level panel 2014-2022. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies. Standard errors clustered at the establishment level.

Figure 5: Heterogeneous Effects: Gender



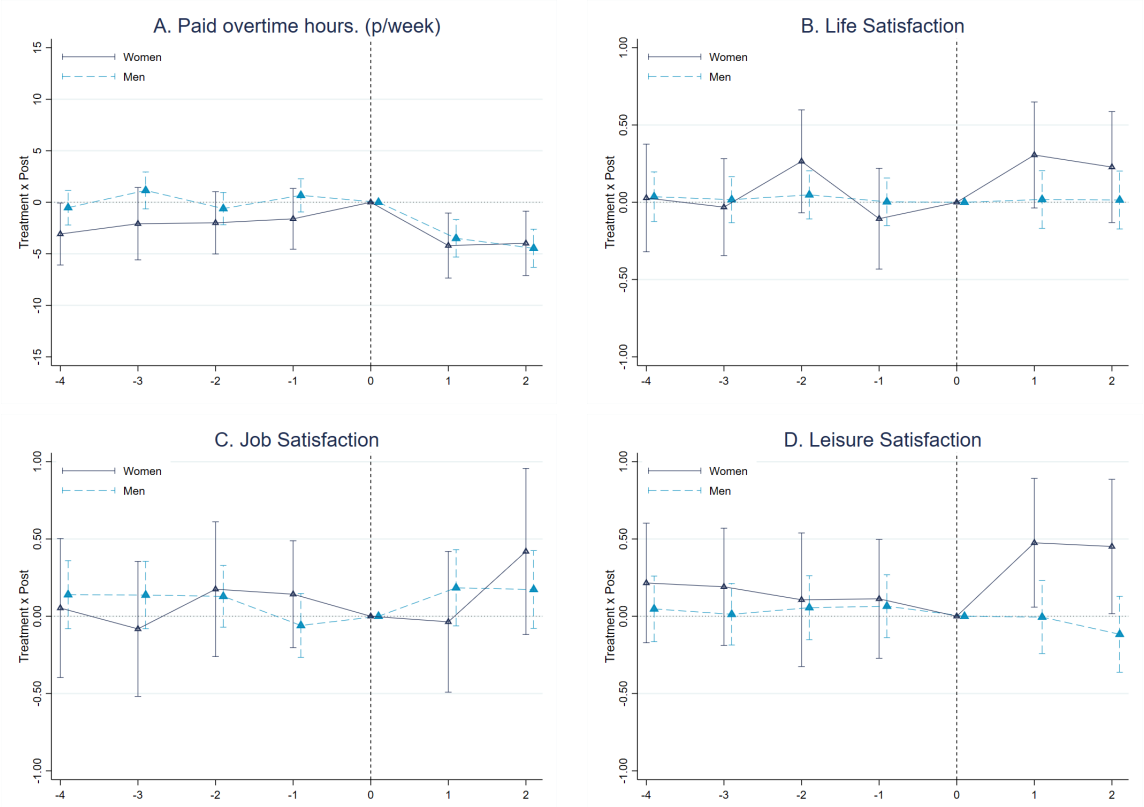
Notes: The figure year-specific DiD estimates and 95% confidence intervals. Regressions for each gender-specific outcome are estimated separately. Dependent variables are gender-specific outcomes at the establishment level: Overtime per Worker (Panel A), Overtime Pay (Panel B), Total Monthly Earnings (Panel C), Hourly Wage (Panel D), Total Employment (Panel E), Standard Employment (Panel F), and Nonstandard Employment (Panel G). Sample: BSWS collapsed at the establishment level. Establishment-level panel 2014-2022. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies. Standard errors clustered at the establishment level.

Figure 6: Event-Study Analysis: Self-Reported Overtime and Subjective Well-Being



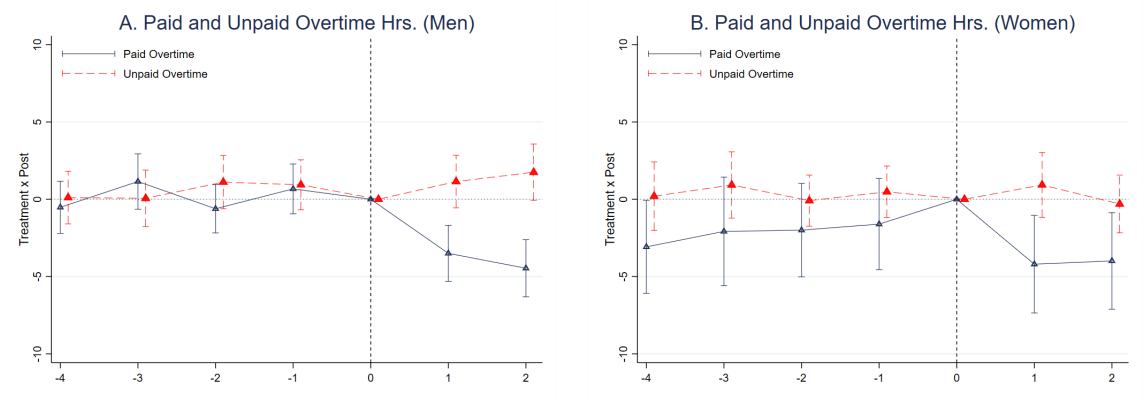
Notes: The figure plots year-specific DiD estimates and 90% confidence intervals. Dependent variables:(Weekly) Paid Overtime Hrs. (Panel A), Life Satisfaction (Panel B), Job Satisfaction (Panel C), Leisure Satisfaction (Panel D). SWB variables are 1-5 Likert scales. Sample: Individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). Reported estimates include individual and year-FE. Other controls: female, age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture effects. "DiD + Matching" refers to a re-weighted DiD estimates using a coarsened exact-matched sample of treated and control individuals. Standard errors are clustered at the individual level.

Figure 7: Event-Study Analysis: Self-Reported Overtime and Subjective Well-Being by Gender



Notes: The figure plots year-specific DiD estimates by gender and 95% confidence intervals. Dependent variables:(Weekly) Paid Overtime Hrs. (Panel A), Life Satisfaction (Panel B), Job Satisfaction (Panel C), Leisure Satisfaction (Panel D). Sample: Individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture effects. "DiD + Matching" refers to a re-weighted DiD estimates using a coarsened exact-matched sample of treated and control individuals. Standard errors are clustered at the individual level.

Figure 8: Event-Study Analysis: Paid vs. Unpaid Overtime Hours by Gender



Notes: The figure plots year-specific DiD estimates by gender and 95% confidence intervals. Sample: Individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture effects. "DiD + Matching" refers to a re-weighted DiD estimates using a coarsened exact-matched sample of treated and control individuals. Standard errors are clustered at the individual level.

Table 1: DiD estimates: Working Hours

	(1) Mean Overtime	(2) Base Hours	(3) Total Hours	(4) Overtime Dispersion (sd)
High Exposure × Post	-5.023***	1.051***	-3.972***	-3.758***
	(0.068)	(0.155)	(0.174)	(0.051)
Mean Outcome	19.804	165.167	184.971	15.906
Observations	322,572	322,572	322,572	313,193
R-squared	0.109	0.117	0.154	0.091
Establishment controls	Yes	Yes	Yes	Yes
Industry-specific time trends	Yes	Yes	Yes	Yes
Prefecture-specific time trends	Yes	Yes	Yes	Yes

Notes: DiD estimates using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies (8). Standard errors clustered at the establishment level and shown in parentheses. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 2: DiD estimates: Log Wages

	(1) Monthly Wage	(2) Overtime Pay	(3) Base Wage	(4) Bonus Wage	(5) Hourly Wage	(6) Monthly Wage (sd)	(7) Overtime Pay (sd)
High Exposure × Post	-0.020***	-0.345***	0.008***	0.006	-0.002	-54.267***	-75.167***
	(0.001)	(0.011)	(0.001)	(0.005)	(0.001)	(3.341)	(1.374)
Mean Outcome	8.032	5.673	7.904	8.723	2.819	3198.896	375.102
Observations	322,545	266,325	322,545	288,331	322,518	322,572	322,572
R-squared	0.285	0.081	0.283	0.143	0.223	0.254	0.073
Establishment controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-specific time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: DiD estimates using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies (8). Standard errors clustered at the establishment level and shown in parentheses. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 3: DiD estimates: Employment

	(1) Total	(2) Nonstandard	(3) Standard
A. Total Employment			
High Exposure × Post	-0.001	-0.004	0.001
	(0.002)	(0.006)	(0.003)
Observations	322,572	253,827	318,848
R-squared	0.640	0.334	0.441
Number of N2	180,596	145,403	177,872
B. Female Employment			
High Exposure × Post	0.012***	-0.017***	0.034***
	(0.003)	(0.006)	(0.004)
Observations	312,468	228,227	278,991
R-squared	0.455	0.285	0.361
Number of N2	174,657	131,457	152,436
C. Male Employment			
High Exposure × Post	0.002	0.006	0.002
	(0.003)	(0.008)	(0.003)
Observations	314,563	196,491	306,616
R-squared	0.530	0.257	0.406

Notes: DiD estimates using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies (8). Standard errors clustered at the establishment level and shown in parentheses. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 4: DiD estimates: Workforce Composition

	(1) % Fixed-term	(2) % Nonstandard	(3) % College Graduates	(4) % Part-time
High Exposure × Post	-0.011***	-0.009***	0.004***	0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
Mean Outcome	0.144	0.154	0.322	0.184
Observations	323,327	323,327	322,572	323,327
R-squared	0.065	0.142	0.037	0.099
Establishment controls	Yes	Yes	Yes	Yes
Industry-specific time trends	Yes	Yes	Yes	Yes
Prefecture-specific time trends	Yes	Yes	Yes	Yes

Notes: DiD estimates using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies (8). Standard errors clustered at the establishment level and shown in parentheses. Significance levels: * 0.10, ** 0.05, *** 0.01.

Table 5: DiD estimates: Self-Reported (Weekly) Working Hours.

	(1) Total Hours	(2) Long Hours (60h+)	(3) Total Overtime	(4) Paid Overtime	(5) Unpaid Overtime	(6) Log Monthly Wage	(7) Days Worked
A. Unconditional DiD							
Treatment × Post	-2.331***	-0.083***	-3.629***	-3.982***	0.044	0.033	-1.590
	(0.772)	(0.025)	(0.644)	(0.457)	(0.479)	(0.027)	(3.340)
Observations	5,120	5,120	4,850	4,950	4,994	4,652	5,042
B. Controlled DiD							
Treatment × Post	-2.648***	-0.095***	-3.811***	-3.735***	-0.294	-0.035	-2.933
	(0.792)	(0.026)	(0.652)	(0.450)	(0.491)	(0.023)	(3.545)
Observations	5,062	5,062	4,801	4,897	4,942	4,620	4,991
C. Controlled DiD - Individual FEs							
Treatment × Post	-1.493*	-0.071**	-2.272***	-3.376***	0.855*	-0.012	-0.930
	(0.858)	(0.029)	(0.660)	(0.466)	(0.499)	(0.019)	(3.696)
Mean Outcome	51.046	0.240	11.302	7.799	3.818	5.641	246.728
Observations	5,062	5,062	4,801	4,897	4,942	4,620	4,991
D. Controlled DiD - Individual FEs							
Men							
Treatment × Post	-1.099	-0.038	-2.332***	-3.997***	1.256*	-0.003	-4.349
	(1.023)	(0.035)	(0.834)	(0.610)	(0.669)	(0.020)	(4.0732)
Mean Outcome	51.727	0.267	11.853	8.033	4.138	5.698	249.496
Observations	3,537	3,537	3,343	3,416	3,445	3,242	3,497
Women							
Treatment × Post	-1.720	-0.080	-2.180	-2.552**	0.347	-0.033	4.749
	(1.895)	(0.062)	(1.342)	(1.124)	(0.819)	(0.065)	(11.800)
Mean Outcome	47.811	0.115	8.724	6.716	2.294	5.370	233.868
Observations	1,525	1,525	1,458	1,481	1,497	1,378	1,494

Notes: DiD estimates using individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). The post-reform variable equals 1 for years 2021-2022 (policy-on period), and 0 otherwise. Treatment group comprises workers who were supplying more than 30 hrs. of paid overtime in a typical month before the reform. Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture fixed effects. Standard errors clustered at the individual level and shown in parentheses. Significance levels: * 0.10, ** 0.05, ***

Table 6: DiD estimates: Subjective Well-Being.

	How satisfied are you with each of the following? (1-5 scale)				
	(1)	(2)	(3)	(4)	(5)
	Overall Life	Job	Spouse	Leisure	Family
A. Unconditional DiD					
Treatment × Post	0.034	0.069	-0.009	0.037	0.009
	(0.062)	(0.079)	(0.085)	(0.072)	(0.069)
Observations	5,134	5,175	4,074	5,180	5,024
B. Controlled DiD					
Treatment × Post	0.018	0.009	0.023	0.058	0.024
	(0.065)	(0.084)	(0.093)	(0.079)	(0.076)
Observations	5,069	5,112	4,030	5,116	4,961
C. Controlled DiD - Individual FEs					
Treatment × Post	0.025	0.036	0.064	0.034	-0.012
	(0.060)	(0.079)	(0.081)	(0.079)	(0.070)
Mean dep. var.	3.625	3.056	3.775	3.433	3.815
Observations	5,069	5,112	4,030	5,116	4,961
D. Controlled DiD - Individual FEs					
Men					
Treatment × Post	-0.022	0.072	0.058	-0.082	-0.071
	(0.067)	(0.087)	(0.088)	(0.085)	(0.077)
Mean dep. var.	3.633	3.040	3.779	3.449	3.826
Observations	3,562	3,582	3,088	3,586	3,505
Women					
Treatment × Post	0.235**	0.082	0.142	0.326**	-0.030
	(0.117)	(0.182)	(0.182)	(0.163)	(0.162)
Mean dep. var.	3.586	3.134	3.738	3.353	3.762
Observations	1,507	1,530	942	1,530	1,456

Notes: DiD estimates using individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). The post-reform variable equals 1 for years 2021-2022 (policy-on period), and 0 otherwise. Treatment group comprises workers who were supplying more than 30 hrs. of paid overtime in a typical month before the reform. Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture fixed effects. Standard errors clustered at the individual level and shown in parentheses. Significance levels: * 0.10, ** 0.05, ***

Table 7: DiD estimates: Perceived Work Intensification and Time Use

	Work Intensity	Performance Pay	Commuting time	Housework time	Freq.of Physical Activity	Unpaid-to-Tot. Overtime
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment × Post	-0.027	-0.005	-0.009	-0.095	-0.048	0.116***
	(0.041)	(0.032)	(0.056)	(0.151)	(0.113)	(0.037)
Treatment × Post× Female	0.017	0.035	0.002	0.120	0.320	-0.101
	(0.086)	(0.055)	(0.149)	(0.402)	(0.247)	(0.074)
Mean dep. var. (Men)	0.451	0.180	0.688	1.643	1.869	0.286
Mean dep. var. (Women)	0.323	0.093	0.731	3.839	1.590	0.212
Observations	5,152	5,152	5,106	4,835	5,135	3,348

Notes: DiD estimates using individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). The post-reform variable equals 1 for years 2021-2022 (policy-on period), and 0 otherwise. Treatment group comprises workers who were supplying more than 30 hrs. of paid overtime in a typical month before the reform. Commuting and housework time measured as total weekly hours (including weekends). Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture fixed effects. Standard errors clustered at the individual level and shown in parentheses. Significance levels: * 0.10, ** 0.05, ***

Online Appendix

A Supplementary Tables and Figures

Table A1: Who Supplies Overtime Hours?

	(1) Overtime hours	(2) Long overtime (30h+)
Female	-5.712***	-0.102***
	(0.017)	(0.000)
Age	0.142***	0.002***
	(0.004)	(0.000)
Age square	-0.004***	-0.000***
	(0.000)	(0.000)
Tenure	0.171***	0.003***
	(0.003)	(0.000)
Tenure square	-0.007***	-0.000***
	(0.000)	(0.000)
Junior High School	1.188***	0.019***
	(0.054)	(0.001)
Junior College	-1.427***	-0.023***
	(0.023)	(0.000)
Over 4-year college	-2.659***	-0.037***
	(0.020)	(0.000)
Fixed-term	0.216***	0.001
	(0.032)	(0.001)
Non-standard	-0.449***	-0.009***
	(0.032)	(0.001)
Large establishment	2.615***	0.040***
	(0.023)	(0.000)
Constant	19.308***	0.258***
	(0.100)	(0.002)
Observations	4,873,980	4,873,980
R-squared	0.105	0.070

Notes: Correlates of monthly overtime hours. Individual-level regressions using BSWS. Pre-reform pooled sample (2013-2018). Estimates control for industry and prefecture fixed effects (not reported).

Table A2: Estimation Samples: Descriptive Statistics

	Treated	Controls
A. Establishment-level data (BSWS)		
Average age	42.25	43.53
Average tenure	12.47	11.54
%Female	0.28	0.40
%Fixed term	0.14	0.12
%Nonstandard	0.15	0.13
%College educ.	0.33	0.30
Large establishment	0.12	0.01
Manufacturing	0.27	0.14
Big city	0.19	0.17
A. Individual-level survey (OPPS)		
Female	0.18	0.37
Age	46.15	48.89
Tenure >20y	0.42	0.48
Clerical worker	0.21	0.25
Service sector	0.17	0.21
Large firm	0.52	0.47
Big city	0.29	0.27
% Treated (pre-reform)	0.143	0.857

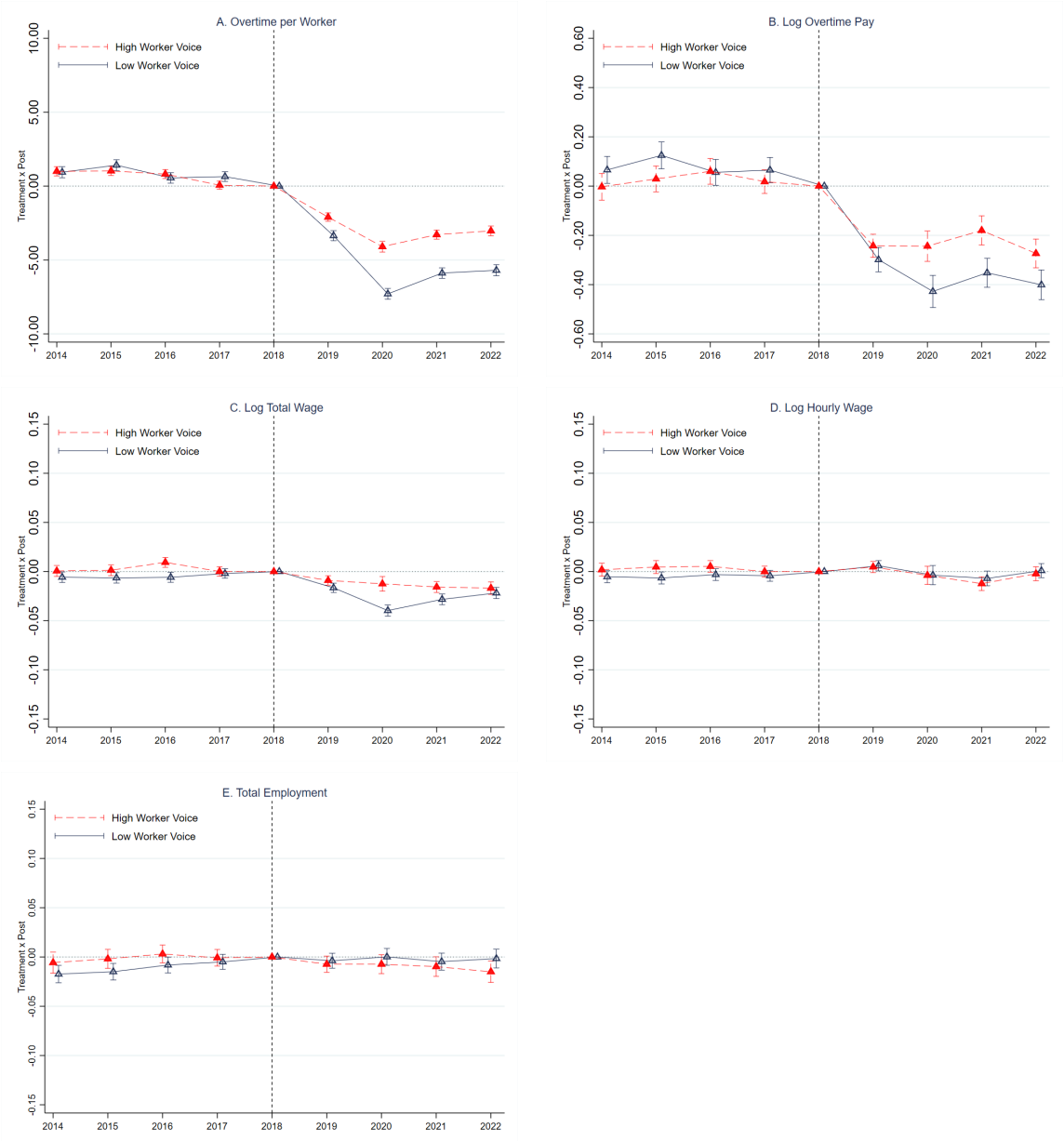
Notes: This table shows summary statistics of the estimation samples: establishments (Panel A) and individuals (Panel B). Sources: Basic Survey Wage Structure (BSWS) and Osaka Preference Parameter Study (OPPS).

Table A3: DiD estimates: Anatomy of Overtime Hours Changes Within Establishments

	(1) % Overt0hrs	(2) % Overt1-10hrs	(3) % Overt11-20hrs	(4) % Overt21-30hrs	(5) % Overt31-40hrs	(6) % Overt41-50hrs	(7) % Overt51-60hrs	(8) % Overt60hrs+
A. All workers								
High Exposure × Post	0.060***	0.044***	0.013***	-0.012***	-0.045***	-0.026***	-0.014***	-0.020***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Mean Outcome	0.294	0.166	0.132	0.124	0.119	0.075	0.039	0.051
Observations	322,572	322,572	322,572	322,572	322,572	322,572	322,572	322,572
B. Male workers								
High Exposure × Post	0.058***	0.045***	0.020***	-0.006***	-0.049***	-0.029***	-0.016***	-0.023***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean Outcome	0.304	0.133	0.121	0.125	0.130	0.084	0.045	0.059
Observations	309,501	309,501	309,501	309,501	309,501	309,501	309,501	309,501
C. Female workers								
High Exposure × Post	0.060***	0.032***	-0.008***	-0.022***	-0.032***	-0.016***	-0.006***	-0.007***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Mean Outcome	0.325	0.259	0.153	0.110	0.078	0.040	0.017	0.017
Observations	289,042	289,042	289,042	289,042	289,042	289,042	289,042	289,042

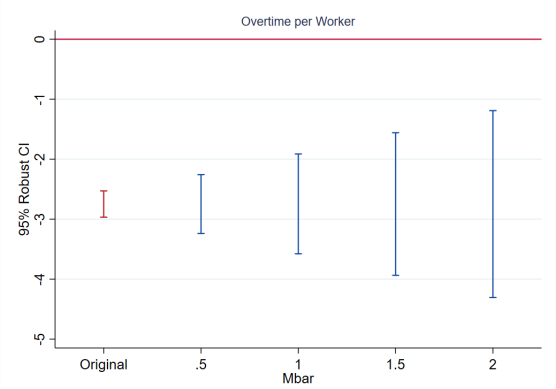
Notes: DiD estimates using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Dependent variables: share of jobs of X overtime Hrs at the establishment level. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies. Standard errors clustered at the establishment level and shown in parentheses. Significance levels: * 0.10, ** 0.05, *** 0.01.

Figure A1: Heterogeneous Effects: Worker Voice Institutions



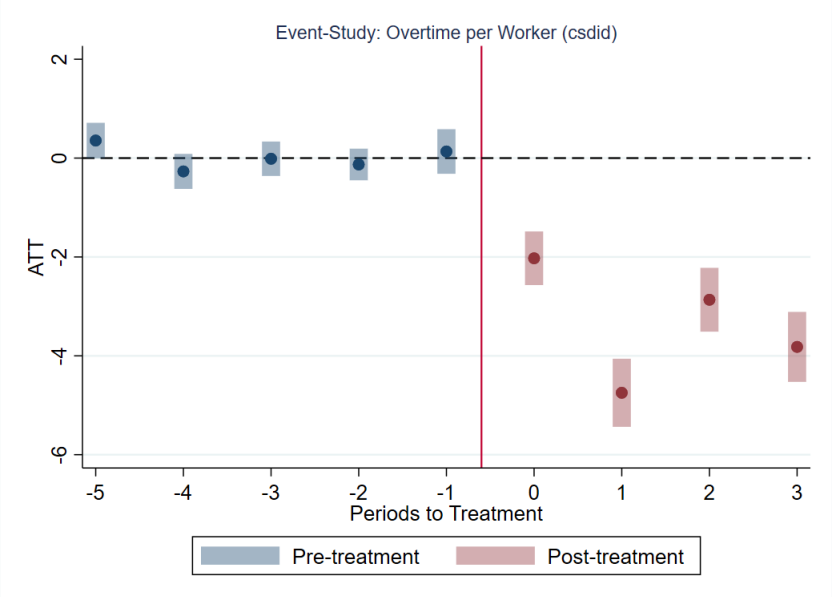
Notes: The figure year-specific DiD estimates and 95% confidence intervals. Regressions for establishments operating in high/low sectoral worker voice regimes are estimated separately. High (low) worker voice regimes are sectors with above-(below)the-median incidence of worker voice institutions at the workplace level (unions, labor-management committees, shop-floor committees) computed from SLMC. Dependent variables: Overtime per Worker (Panel A), Total Monthly Earnings (Panel B), Overtime Pay (Panel C), Total Employment (Panel D), Standard Employment (Panel E), and Nonstandard Employment (Panel F). Sample: BSWS collapsed at the establishment level. Establishment-level panel 2014-2022. Reported estimates include firm and year-FE, as well as industry- and prefecture-year effects. Other controls: workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size dummies. Standard errors clustered at the establishment level.

Figure A2: Honest Pre-Trends (Rambachan and Roth, 2023)



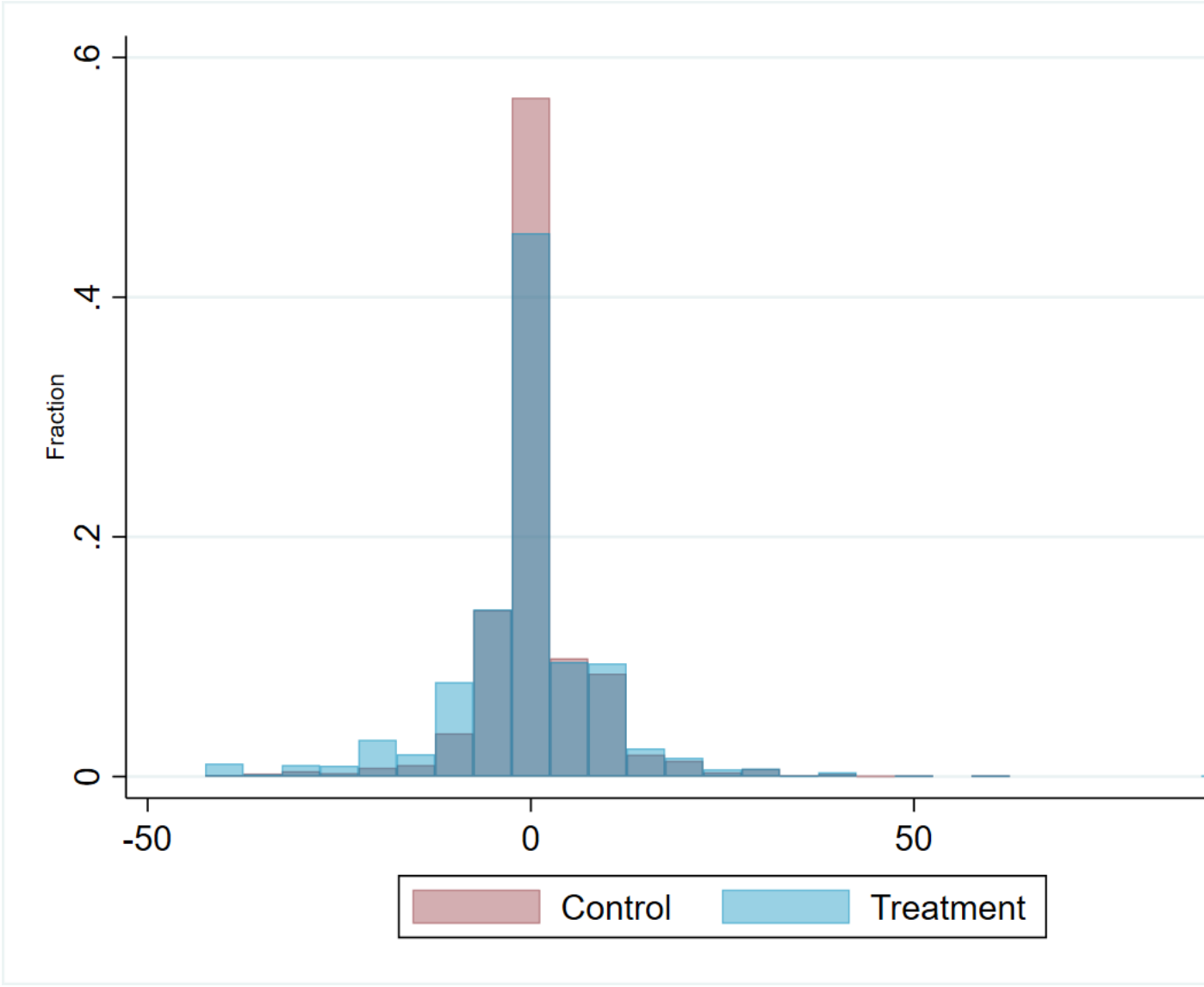
Notes: This Figure plots alternative estimated confidence intervals for η_{2019} (the first post-reform year-specific DiD effect). These 95% confidence intervals allow for deviations from parallel trend in the pre-reform period, following the method proposed by Rambachan and Roth (2023). They are calculated assuming that the post-reform violation of parallel trends is at most Mbar larger than the maximum violation of parallel trends in the pre-reform period. For example, Mbar equals to 2 means that the post-reform deviation is no more than twice that in the pre-reform period.

Figure A3: Staggered Treatment (Callaway and Sant'Anna, 2021): Overtime per Worker



Notes: Event-study analysis accounting for staggered treatment timing. Highly-exposed large firms treated from 2019 onward. Highly-exposed small-medium sized firms treated from 2020 onward. Small establishments employ not more than 300 employees (not more than 50 employees for retail busi- nesses, not more than 100 for wholesale retail). We rely on the approach and Stata routine (*csdid*) proposed by Callaway and Sant'Anna (2021) using never-treated units as controls. using establishment-level panel (2014-2022) from Basic Survey on Wage Structure (BSWS). Reported estimates include the following controls for workforce composition (age, tenure, share of female workers, share of workers with college education, share of fixed-term contracts, nonstandard workers), firm-size, industries and prefecture dummies. 95% confidence intervals, wild bootstrap-standard errors.

Figure A4: Consistency Check for Self-Reported Working Hours in OPPS



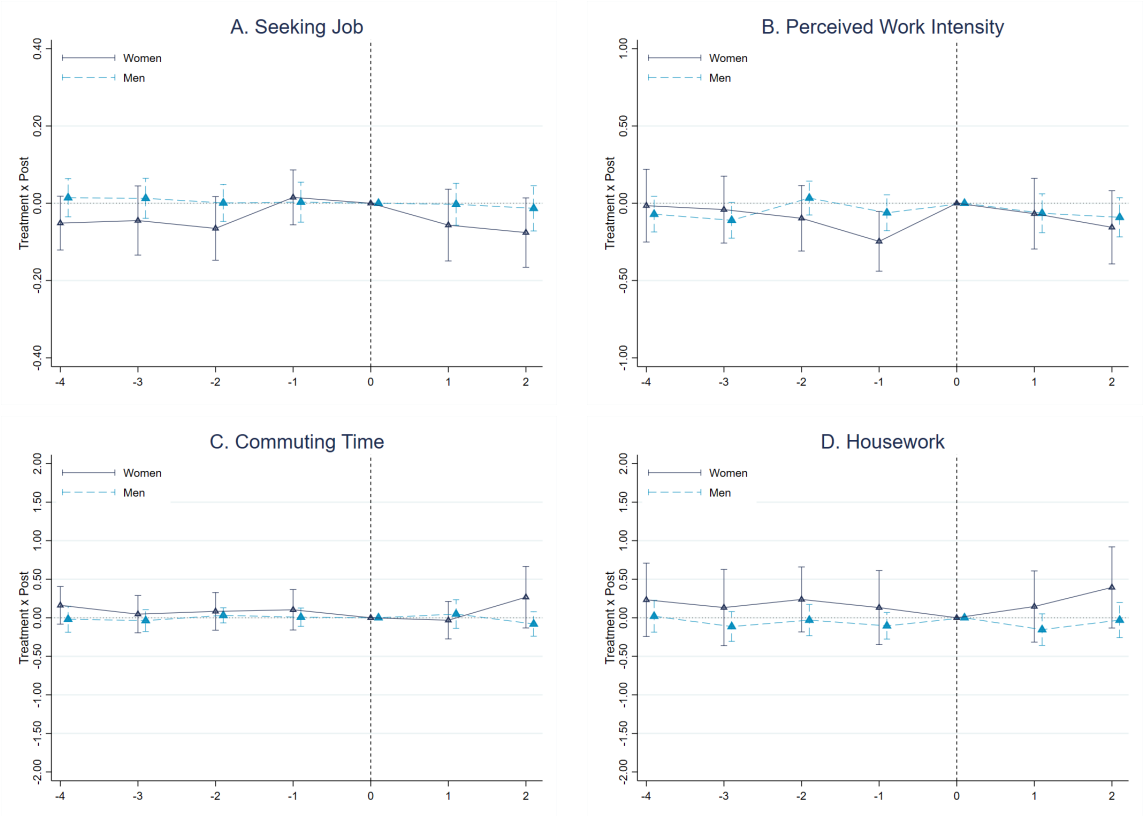
Notes: Distribution of the difference between total working hours as reported by individuals in OPPS and total working hours assuming a standard workweek (40h) plus total overtime hours reported by individuals.

Table A4: Individual-level DiD Estimates: Robustness checks

	(1) Total Hours	(2) Long Hours	(3) Paid Overtime
A. Excluding inconsistent reporting			
Treatment × Post	-1.998**	-0.049	-2.425***
	(0.939)	(0.037)	(0.600)
Observations	1,854	1,854	1,854
B. Trimming top 2% total hours			
Treatment × Post	-2.362***	-0.086***	-3.568***
	(0.686)	(0.024)	(0.430)
Observations	4,968	4,968	4,813
C. Job Stayers			
Treatment × Post	-1.297	-0.070**	-3.480***
	(0.842)	(0.028)	(0.495)
Observations	3,296	3,296	3,192
D. Balanced panel			
Treatment × Post	-1.660	-0.092***	-2.735***
	(1.029)	(0.035)	(0.585)
Observations	2,207	2,207	2,138
E. Only Individuals Employed in Large firms			
Treatment × Post	-1.620	-0.111***	-4.357***
	(1.095)	(0.038)	(0.690)
Observations	1,969	1,969	1,897
R-squared	0.066	0.039	0.063

Notes: DiD estimates using individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). The post-reform variable equals 1 for years 2021-2022 (policy-on period), and 0 otherwise. Treatment group comprises workers who were supplying more than 30 hrs. of paid overtime in a typical month before the reform. Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture fixed effects. Standard errors clustered at the individual level and shown in parentheses. Significance levels: * 0.10, ** 0.05, ***

Figure A5: Event-Study Analysis: Additional Results from OPPS)



Notes: The figure shows event studies based on a DiD model as in Equation (1). The graph displays the estimated η coefficient associated with the interaction term $T_i \times Post_t$. "Matching" refers to a re-weighted DiD estimation of a coarsened exact-matched sample of treated and control individuals. The standard errors are clustered at the individual level and the dash bars depict 95% confidence intervals.

Table A5: DiD estimates: Other Subjective Well-Being Facets

	(1) Fulfilling life	(2) Happiness	(3) Health anxiety	(4) Feeling stressed	(5) Feeling depressed	(6) Sleep problems
A. All respondents						
Treatment × Post	0.044	0.047	-0.028	-0.103	-0.055	-0.006
	(0.055)	(0.106)	(0.068)	(0.071)	(0.079)	(0.069)
Observations	5,135	5,096	5,133	5,135	5,136	5,137
R-squared	0.024	0.029	0.025	0.021	0.018	0.033
B. Men						
Treatment × Post	0.005	-0.090	-0.018	-0.099	-0.040	0.026
	(0.065)	(0.126)	(0.082)	(0.082)	(0.092)	(0.082)
Observations	3,595	3,576	3,595	3,594	3,595	3,598
R-squared	0.026	0.041	0.031	0.028	0.026	0.035
C. Women						
Treatment × Post	0.161	0.508**	-0.117	-0.116	-0.097	-0.096
	(0.115)	(0.205)	(0.129)	(0.167)	(0.171)	(0.130)
Observations	1,540	1,520	1,538	1,541	1,541	1,539
R-squared	0.055	0.044	0.052	0.042	0.028	0.058

Notes: The figure plots year-specific DiD estimates by gender and 95% confidence intervals. Sample: Individual-level panel from Osaka Preference Parameter Study (OPPS) restricted to full-time, non-managerial employees aged 20-65 years. Waves: Pre-reform (2012, 2013, 2016, 2017, 2018), Post-reform (2021, 2022). Reported estimates include individual and year-FE. Other controls: age, age squared, tenure group dummies, occupation dummies, firm size class dummies, sector and prefecture effects. "DiD + Matching" refers to a re-weighted DiD estimates using a coarsened exact-matched sample of treated and control individuals. Standard errors are clustered at the individual level.