International Archives of Occupational and Environmental Health (2025) 98:79–98
https://doi.org/10.1007/s00420-024-02113-8

ORIGINAL ARTICLE

Causal effects of promotion to managerial positions on mental health and satisfaction in Japanese male workers

Ryohei Kashima[1] · **Masaya Takahashi**[1]

Received: 20 February 2024 / Accepted: 21 November 2024 / Published online: 27 December 2024

Abstract

Purpose We estimated the static and dynamic effects of promotion to managerial positions on mental health and satisfaction using the recent estimation method; we addressed the gap in extant research by considering the heterogeneity in treatment timing to reconfirm evidence regarding the causal effects of promotion.

Methods We used a modified estimation method of two-way fixed effects recommended by (Callaway and Sant'Anna, J Econom 225:200–230, 2021). To check for robustness, we compared the analyses using propensity sample matching, an alternative treatment timing heterogeneity-robust analysis, and a normal two-way fixed effects event study. We used four years of complete panel data on a sample of 1454 Japanese males (5816 overall).

Results The results indicated that becoming a manager worsened mental health with a two-year lag. We tested the effects of promotion to managerial positions on variables shown in previous studies to have negative effects on mental health (e.g., increased working hours) and positive effects (e.g., increased job role clarity) as outcomes to explore the mechanisms. Some of both variables indicate statistical significance. These results were robust.

Conclusion The promotion to managerial positions deteriorated mental health with a two-year lag. Promotion to managerial positions exhibited both positive and negative effects. The findings have potential implications for firms' human resources allocation.

Keywords Promotion to managerial positions · Mental health · Staggered difference-in-difference · Two-way fixed effects

Introduction

In the fields of public health and health economics, scholars have focused on the impact of being promoted to managerial positions on mental and physical health. Similar to other positive correlations between high socio-economic status and health[1] (Lantz et al. 1998; Johnson et al. 1999; Smith 1999; Marmot 2005; Fukuda et al. 2005; Mackenbach et al. 2008), simple analyses have shown that job title and position are positively associated with mental and physical health (Marmot et al. 1997; Kunst et al. 1999; Rahkonen et al. 2006; Green 2011; Stringhini et al. 2011; Kokkinen et al. 2020). Then, does promotion to managerial positions exhibit a positive causal effect as well as correlation? Previous studies reported negative aspects of health, such as high morbidity and burnout among managers (Cordes et al. 1997; Kubo and Tao 1994). Individual attributes differ greatly between potential managers and others (Fuchs 1980; Adams et al. 2003). Therefore, correlation and causality do not necessarily coincide. It is important to accumulate results using a causal inference approach to confirm the effects of promotions on managers and avoid misinterpreting the direction of causality.

The literature on the causal effects of promotions indicates that conducting experiments using random assignment or natural experiments, such as institutional changes, is difficult because the issue is one of personnel allocation within a company. Therefore, studies using panel data to test the causal effects of promotions use two-way fixed effects

✉ Ryohei Kashima
kashima-r@h.jniosh.johas.go.jp

[1] Social Science and Occupational Health Research Group, National Institute of Occupational Safety and Health, Nagao 6-21-1, Tama-Ku, Kawasaki 214-8585, Japan

[1] Many studies have been conducted in the public health and medical fields, particularly regarding mortality.

(TWFE), which can control for unobservable individual and time confounding factors.[2] Previous studies using TWFE reported that promotion to managerial positions could deteriorate mental and physical health (Anderson and Marmot 2012; Boyce and Oswald 2012; Johnston and Lee 2013; Sato 2015; Nyberg et al. 2017; Ikesu et al. 2021).

TWFE is understood as a causal inference method because it can control for unobservable individual and time confounding factors. However, the negative aspects of TWFE have recently begun to be understood. In cases where treatment timing is staggered, TWFE fails with regard to proper comparison and weighting, and it becomes problematic to claim the point estimate as a causal effect (De Chaisemartin and d'Haultfoeuille 2022; Borusyak et al. 2021; Goodman-Bacon 2021; Baker et al. 2022). Considering the treatment of promotion to managerial positions, the treatment timing is usually staggered. No study has seriously considered and analyzed the heterogeneity of treatment timing in the causal effects of promotions on mental health.

In this context, we aimed to re-examine the causal effects of promotions to managerial positions on workers' mental health by using the staggard difference-in-difference (SDID) estimation method proposed by Callaway and Sant'Anna (2021) on individual worker panel data from eight Japanese workplaces that were available for the four-year period between 2010 and 2013.

Methods

Our data

We used Japanese longitudinal data from the Japanese Study of Health, Occupation and Psychosocial Factors Related Equity (J-HOPE) (Eguchi et al. 2015; Inoue et al. 2016, 2018; Miyaki 2013; Tsutsumi 2011). Appendix Table presented the J-HOPE raw data sample status by company (anonymized by industry), male sample status in the raw data, and sample status in our analysis data. J-HOPE surveyed 6400 employees at 10 arbitrary companies in 2010 and at 13 arbitrary companies between 2011 and 2013 (10,228 employees in 2013). In this dataset, we restricted our study sample to nine companies with individuals who were promoted to managerial positions, to men, and to individuals who could be tracked for four years; further, we excluded from the analysis individuals who were in managerial positions at the first period in 2010, individuals who transferred from managerial to other positions, individuals whose working hours are less than 30, and individuals who indicated that their education had decreased. Individuals with missing data for each variable used in the analysis were also excluded. Finally, four waves of complete panel data for 1454 individuals (5816 in total for the four waves) were used in the analysis.

Treatment variable: promotion to managerial positions

J-HOPE does not contain a direct question regarding promotions. Among the eight occupational categories, managerial positions are included in one of these occupational categories. The analysis was conducted by considering promotion to managerial occupations in 2011 and later as the promotion to managerial positions," which is the focus of our study. The data exclude individuals who have been in managerial positions since 2010, the first period, and those who have changed from managerial positions to other types of jobs. It is difficult to assume that the effects of promotion from other occupations to management, and from management to other occupations, are identical. Therefore, we focused on the former to avoid contaminating the effects we desired to estimate.

Step 1's outcome variables: mental health and satisfaction

Kessler et al.'s (2002) Kessler-s (K6) score, which has been validated for Japan (Furukawa et al. 2008), was used in our analysis as both a continuous and binary variable that takes the value of 1 when the score is 5 or higher. Scores of 5 or higher indicate that individuals have depressive and anxiety tendencies. Higher scores indicate worse mental health.

We also used the results of a study employing a four-point scale on satisfaction with work and family life (1 = unsatisfied, 2 = somewhat unsatisfied, 3 = fairly satisfied, and 4 = satisfied; Shimomitsu 2000). We used the index as a continuous variable in our analysis.

Step 2–1's outcome variables: potential negative mechanisms

The increase in working hours negatively affects mental health (Virtanen et al. 2011; Amagasa and Nakayama 2013; Kim et al. 2016; Afonso et al. 2017; Ogawa et al. 2018). J-HOPE asked about weekly working hours using the following option: 30 h or less, 31–40 h, 41–50 h, 51–60 h, and 61 h or more. We used these responses in our analysis as continuous values from 1 to 5, beginning with less than 30 h as 1.

J-HOPE used the Japanese version index of job demand in the job-demand control model (Karasek 1979; Kawakami et al. 1995), effort-reward-imbalance model

[2] Anderson and Marmot (2012) performed IV estimation using the initial department of affiliation as IV variables.

(ERI; Siegrist 1996; Tsutumi et al. 2011), and job-insecurity model (Karasek et al. 1998; Kawakami et al. 2004). Higher scores indicate worse mental health. These scores were used in our analysis as continuous variables.

Step 2–2's outcome variables: potential positive mechanisms

Apouey and Clark (2015) and Kim and Koh (2021) used lottery winnings as the context for exogenous income and indicated that increases in exogenous income positively impact individuals' mental health in the short run, besides having long-term effects. J-HOPE does not ask about individual or annual household income. Although the movements in individual and household incomes do not necessarily coincide depending on the trends of people living together, the household income response was used as the outcome variable because it is expected to be generally consistent. J-HOPE asked about annual household income using the following choices: less than ¥2.99 million, ¥3 million to ¥4.99 million, ¥5 million to ¥7.99 million, ¥8 million to ¥9.99 million, ¥10 million to ¥14.99 million, and ¥15 million or more. As there was a marked difference in the range of each item, the median value for each option was log-transformed (¥15 million was used for ¥15 million and above) and used in the analysis as a continuous value.

J-HOPE utilized the Japanese version of the job control index in the job-demand control model (Karasek 1979; Kawakami et al. 1995), work engagement (Schaufeli et al. 2006; Shimazu et al. 2008), and job role clearness (Coverman 1989). These scores were used in our analysis as continuous variables. Increasing these scores is thought to have a positive effect on mental health. These scores were used in our analysis as continuous variables.

Controlling time variant confounders

The birth of a child may have a positive or negative effect on mental health. Since we do not have access to data on the number of children, we use the total number of family members, which is a proxy variable in the survey's acquisition information, as a control variable. Since the individual and time confounder are considered, we do not use any other control variables.

Promotions to managerial positions in Japan

The Global Employment Status and Growth Attitudes Survey (Persol Research and Consulting Co. Ltd. 2022) was conducted in 18 countries (13 Asia-Oceania countries plus the United States, the United Kingdom, Germany, France, and Sweden). Respondents were asked, "Do you feel you would like to be a manager in your current company?" and their responses were rated on a five-point scale. The results showed that Japan had the lowest percentage of positive responses (i.e., "agree" and "somewhat agree"), at 19.8%. As Australia's next lowest rate was 38.0% and all countries' average rate was 58.6%, Japan could be a negative country in terms of promotion to managerial positions.

Estimation method

Previous studies' method: conventional TWFE and event study

Previous studies typically performed conventional TWFE estimation using panel data as follows:

$$y_{it} = \delta_i + \delta_t + bPostTreat_{it} + error_{it} \quad (1)$$

PostTreat denotes a binary indicator that continues to take the value of 1 after an individual has been treated, and δ_j (j = i,t) denotes the fixed-effect term of the j factor. Instead of ordinary least squares estimation, TWFE has been widely used in empirical studies to address endogeneity issues successfully because it controls for individual and time-derived confounding factors that are not observable in the data. In the context of promotion effects, studies have used mental and physical health for Y in Eq. (1) (Anderson and Marmot 2012; Boyce and Oswald 2012; Johnston and Lee 2013; Sato 2015; Nyberg et al. 2017; Ikesu et al. 2021). Since TWFE was considered an extension of difference-in-difference (DID) estimation, the coefficient value b in (1) can be considered as average treatment effect on treated (ATT) as well as DID estimation.

Along with TWFE, a dynamic decomposition of the onset of the TWFE effect, called an event study, was estimated as follows:

$$y_{it} = \delta_i + \delta_t + \sum_{g=-2}^{-G} \eta_g Lag_{it}^{g} + \sum_{p=0}^{P} \zeta_p Lead_{it}^{p} + error_{it} \quad (2)$$

Lag_{it}^{g} denotes a binary indicator that takes the value of 1 if individual i is in period t before the gth period of treatment, and $Lead_{it}^{p}$ denotes a binary indicator that takes the value of 1 if individual i is in period t after the pth period of treatment. The baseline is generally measured one period before treatment. This dynamic estimation of (2) allows us to confirm the prior parallel trend and determine how long the effect lasts and whether it occurs with a lag. The presence or absence of a statistical significance in η_g suggests whether a conditional parallel trend could hold in g prior period from treatment (if it could be satisfied, the results could be interpreted as a causal relationship). ζ_p was interpreted as

Table 1 Promotions to managerial positions

Category	Survey years				
	2010	2011	2012	2013	Four years total
Managers	0	53	94	133	280
Non-managers	1454	1401	1360	1321	5536
Total	1454	1454	1454	1454	5816
Promotion (rate)	0	53 (3.6%)	41 (2.8%)	39 (2.7%)	133 (2.3%)

The percentage of promotions to managerial positions represents the number of promotion occurrences in a single year as a percentage of the total sample of 1454 individuals

dynamic ATT after treatment period p. The promotion effect literature has been used as an indicator of the number of terms before or after a promotion. For example, Johnston and Lee (2013) used the Household, Income and Labour Dynamics (HILDA) survey data for Australia to show that mental health deteriorates two years after a promotion.

Recent modified estimation method: Callaway and Sant'Anna's (2021) estimation

However, the recent advances in the methodological and theoretical understanding of TWFE have revealed the difficulties in considering the TWFE coefficient (1)'s *b* to be ATT (De Chaisemartin and d'Haultfoeuille 2022; Borusyak et al. 2021; Goodman-Bacon 2021; Baker et al. 2022). Goodman-Bacon (2021) showed that if the sample is divided into three groups—never treated, early treated, and later treated—the TWFE can be broken down into a 2×2 DID comparison for each group. It pointed out that TWFE contains a DID estimation in which the early treated group's post-treatment status is considered the control group, as a so-called "forbidden comparison." Therefore, if there were some heterogenous treatment timing, TWFE could be biased. Borusyak et al. (2021) point out that the occurrence of such bias remains even with event study specification.

Callaway and Sant'Anna's (2021) method was developed for treatments with staggered timing. They modified TWFE not to conduct forbidden comparisons and their method estimates ATT based only on appropriate comparisons. Our Appendix presents a detailed discussion of the analysis method (TWFE's bias), Callaway and Sant'Anna's (2021) method, and their definition of ATT.

Our estimation in this research

The timing of the promotion to managerial positions was heterogeneous for individuals within the data period. In this context, we confirmed dynamic and static treatment effects as our main result by considering the staggeredness in treatment timing and using Callaway and Sant'Anna's (2021) specification. We also attached static TWFE results of Eq. (1) to compare the findings. We favored dynamic over static results. We did not judge the presence of a causal effect based only on the presence of statistical significance. We display the dynamic effects as figures. We argue that causal effects exist when the pre-treatment motion is stable and post-treatment coefficient deviates from pre-treatment coefficients[3].

The analysis was conducted in two steps. In Step 1, we estimated the dynamic and static treatment effects without controlling for covariates, using mental health status and job/life satisfaction as outcome variables. In Step 2, we estimated the dynamic treatment effect using potential factors underlying the mechanism for mental health status as the outcome variable y, for which previous studies have suggested a positive or negative relationship. Previous studies have shown that income, job control, work engagement, and high job clarity have positive effects on mental health (Apouey and Clark 2015; Kim and Koh 2021; Karasek 1979; Kawakami et al. 1995; Schaufeli et al. 2006; Shimazu et al. 2008; Coverman 1989). Since personal income was not included in our data, we substituted it for household income, which was included in our data. Previous studies have shown that working hours, job demands, effort-reward imbalance (ERI), and high job insecurity negatively affect mental health (Virtanen et al. 2011; Amagasa and Nakayama 2013; Kim et al. 2016; Afonso et al. 2017; Ogawa et al. 2018; Karasek 1979; Kawakami et al. 1995, 2004; Siegrist 1996; Tsutsumi et al. 2001; Karasek Jr et al. 1998).

In Step 1, we estimated the aggregated overall effect of promotion on mental health and satisfaction, including the effects of various pathways. Step 2 explored the mechanisms responsible for the total effect obtained in Step 1. Because of data limitations, we could check the results only for the four positive and four negative elements.

We made the following three estimations to check the robustness of the results: First, we estimated the probability of being a worker promoted during 2011–2013. We used the 2010 information of the K6 total score, working hour, ERI score, log household income, work engagement score, age, years of experience, number of household members, education level, and firm. We constructed a 1:1 matched sample settled caliper to 0.001 and used Callaway and Sant'Anna's (2021) estimation.[4] Second, we used Sun and

[3] They performed the Hausman test and used a random effects model if the Hausman test was not rejected.

[4] Several outcome variables were used here to estimate propensity scores. While this is not valid for cross-sectional data, the use of pre-treatment outcome variables to calculate propensity scores is valid for panel data.

Table 2 Summary statistics for the treatment and control groups

		Control group		Treatment group		Control—Treatment	
		Mean	SD	Mean	SD	Difference	*t* test statistics
Outcomes	K6 total score	5.53	5.04	4.83	4.30	− 0.708**	− 3.13
	K6 score ≧ 5	0.49	0.50	0.44	0.50	− 0.048*	− 2.11
	Job satisfaction	2.61	0.79	2.78	0.79	0.166***	4.61
	Life satisfaction	3.02	0.79	3.07	0.78	0.052	1.47
	Working hours	2.93	0.71	3.32	0.78	0.382***	11.66
Potential negative mechanisms	Job demands	32.79	5.30	33.94	5.27	1.154***	4.78
	Effort-reward imbalance (ERI)	1.08	0.37	1.06	0.33	− 0.023	− 1.43
	Job insecurity	6.66	1.40	6.45	1.17	− 0.209***	− 3.32
	log Annual household income	6.44	0.35	6.70	0.31	0.262***	16.46
Potential positive mechanisms	Job control	66.62	9.38	71.95	7.74	5.327***	12.67
	Work engagement	2.76	0.95	3.11	0.84	0.345***	8.10
	Clarity regarding job role	29.90	5.83	31.82	5.58	1.928***	7.30
Individual characters	Age	40.66	9.30	42.37	5.48	1.709***	4.16
	Experience	17.96	10.73	18.74	6.70	0.784	1.65
	Number of family members	2.80	1.42	3.25	1.25	0.443***	6.94
	Years of education	14.61	2.38	15.71	1.72	1.100***	10.40
	Observations	5284		532		5816	

Summary statistics were generated by pooling data from 2010 to 2014 using non-promotion as the control group and promotion to managers as the treatment group. In "Control—Treatment," we performed t-tests on the difference between the mean of the control and treatment groups without adjusting for covariates for each variable and compared the results. Statistical significance was set as follows: $*p<0.05, **p<0.01, ***p<0.001$

Abraham's (2021) approach as a robust method to address the staggeredness of treatment timing other than Callaway and Sant'Anna's (2021) method. Third, we performed conventional TWFE event studies.

Results

Summary statistics

Table 1 shows the number of promotions to managerial positions that occurred during the data period in the first place. Because individuals who move in a demotion-like manner from managerial to non-managerial positions are excluded from the data refinement process, the binary indicator for managerial positions continues to take the value of 1 after conversion to managerial positions in this dataset. Out of the 1454 individuals in our sample, an average of 3.0% (53 to 39) converted to management positions in a single year, and 133 individuals (9.1% of the analytical individuals) were promoted to managers within three promotional chance years.

Table 2 presents the summary statistics on the outcome variables and underlying individual attributes for the two groups—promoted to managerial and not promoted. Even when individuals originally in managerial positions were excluded from the sample, the correlations indicated that those who were promoted to managerial positions have lower K6 scores and higher job satisfaction than other workers. Checking each latent mechanism for mental health revealed that all positive mechanism variables are better for those who were promoted to managerial positions. As for the negative latent mechanisms, managers tended to have higher working hours and job demands but lower job instability.

These are only correlations, and whether promotion to managerial positions had a causal effect on these outcomes needs further investigation and is the focus of this study. Moreover, those who changed to managerial positions had higher age and a higher number of household members and years of education.

Step 1: total effect on mental health and satisfaction

Table 3 reports the dynamic and static total effects of promotion on mental health and satisfaction estimated by Callaway and Sant'Anna (2021) and ordinal static TWFE. Figure 1 plots the effects of the dynamics presented in Table 3, showing promotion to managerial positions occurring in normalized period 0, indicating how many periods off from the timing of promotion, and plotting

Table 3 Dynamic and static total effects of promotion to managerial positions on mental health and satisfaction

	(1) K6 total score	(2) K6 score≧ 5	(3) Job satisfaction	(4) Life satisfaction
lead2	− 0.959	− 0.151	0.135	− 0.036
	(0.580)	(0.084)	(0.125)	(0.116)
lead1	− 0.263	0.022	0.113	0.015
	(0.347)	(0.047)	(0.076)	(0.072)
lag0	− 0.260	0.002	0.041	− 0.089
(Treatment timing)	(0.284)	(0.038)	(0.065)	(0.057)
lag1	− 0.039	0.080	0.047	− 0.009
	(0.374)	(0.046)	(0.074)	(0.062)
lag2	0.373	0.174*	0.108	− 0.052
	(0.488)	(0.075)	(0.101)	(0.079)
N	5816	5816	5816	5816
Static ATT	− 0.066	0.061	0.055	− 0.055
Ordinal TWFE ATT	− 0.399	0.017	0.111*	− 0.053
p value of pre trend test (H0: All Pre-treatment are equal to 0)	0.234	0.340	0.144	0.983

After estimating Eq. (10) using Callaway and Sant'Anna's (2021) method, the estimated ATTg,t is the weighted average as the event study form above N, and as Static ATT form below N. Control group is never treated individuals. Ordinal TWFE ATT indicates the results of TWFE. All results are adjusted by the number of family members. Statistical significance was set as follows: $*p<0.05, **p<0.01, ***p<0.001$

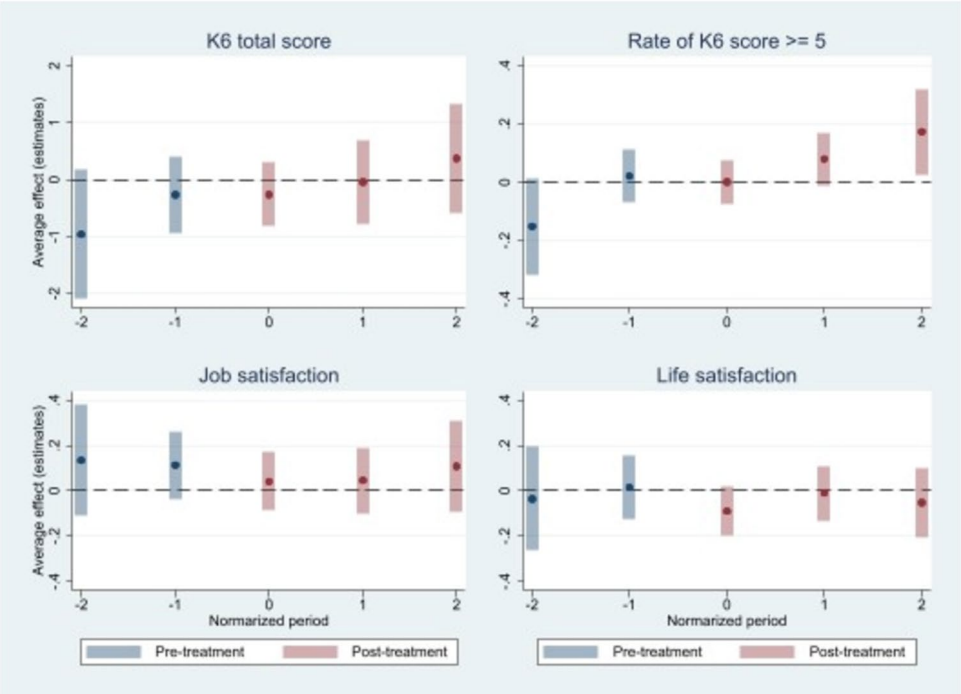

Fig. 1 Dynamic total effects of promotion to managerial positions on mental health and satisfaction

the ATT estimates and 95% confidence intervals. We could not reject the null hypothesis that all pre-treatments are equal to 0, indicating that conditional parallel trend assumptions could hold. Therefore, we could interpret all estimated coefficients as causal effects. Results of the static ATT indicated that promotion to managerial positions had no statistically significant effect. Conversely, ordinal TWFE results indicated that promotion to managerial positions significantly increased job satisfactions ($p<0.05$). The results of the effect dynamics

demonstrated that promotion to managerial positions increased the percentage of individuals with K6 ≧ 5 to 17.4% ($p < 0.05$) after two years from the promotion.

Step 2: potential mechanisms of total effects

Effect on negative latent mechanisms

Table 4 reports the dynamic and static effects of promotion to managerial positions on the latent mechanisms negatively

Table 4 Dynamic and static effects of promotion to managerial positions on the latent mechanisms negatively associated with mental health and satisfaction

	(1) Working hour	(2) Job demand	(3) ERI	(4) Job insecurity
Lead2	0.099	− 0.884	0.030	− 0.025
	(0.100)	(0.798)	(0.066)	(0.241)
Lead1	− 0.034	− 0.385	− 0.092*	0.186
	(0.098)	(0.438)	(0.036)	(0.159)
Lag0	0.255***	0.936*	− 0.002	− 0.216
(Treatment timing)	(0.074)	(0.413)	(0.022)	(0.120)
Lag1	0.179*	0.489	− 0.019	− 0.168
	(0.082)	(0.542)	(0.032)	(0.141)
Lag2	0.167	0.637	0.012	0.281
	(0.103)	(0.877)	(0.044)	(0.175)
N	5816	5816	5816	5816
Static ATT	0.213***	0.729	− 0.005	− 0.106
Ordinal TWFE ATT	0.233***	0.419	− 0.044	− 0.076
p value of pre trend test (H0: All Pre-treatment are equal to 0)	0.233	0.342	0.048*	0.502

After estimating Eq. (10) using Callaway and Sant'Anna's (2021) method, the estimated ATTg,t is the weighted average as the event study form above N, and as Static ATT form below N. Ordinal TWFE ATT indicates the results of TWFE. All results are adjusted by the number of family members. Statistical significance was set as follows: $*p < 0.05, **p < 0.01, ***p < 0.001$

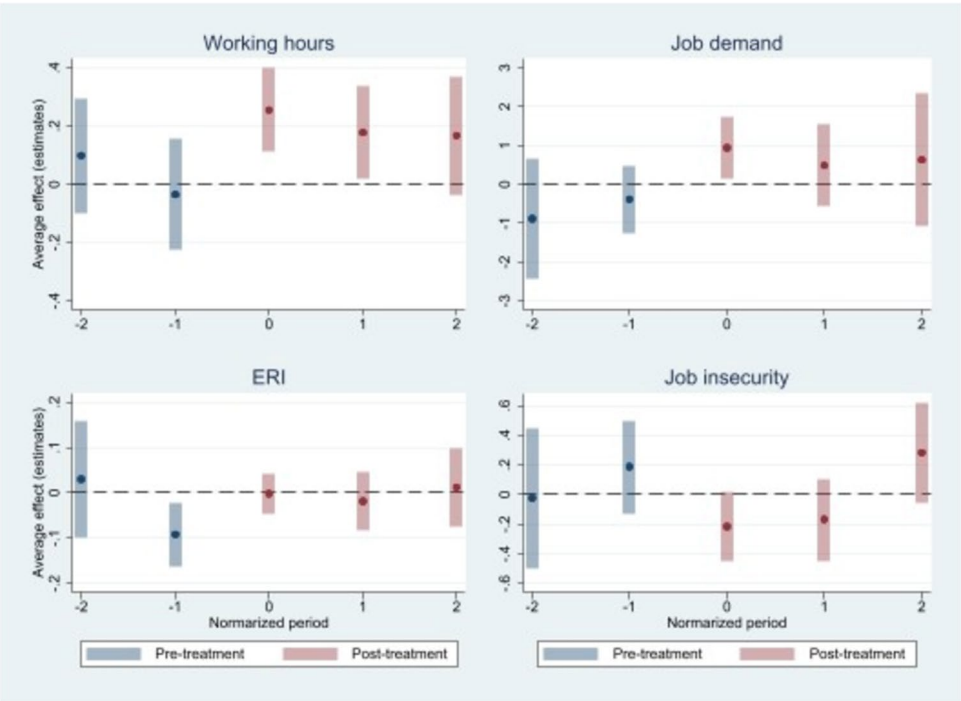

Fig. 2 Dynamic effects of promotion to managerial positions on the latent mechanisms negatively associated with mental health and satisfaction

associated with mental health and satisfaction. Figure 2 plots the effects of the dynamics presented in Table 4. Except ERI estimations ($p < 0.05$), conditional parallel trends held. The statistical significance of the static ATT indicated that promotion to managerial positions significantly increased working hours by 0.213 ($p < 0.001$). The results of the effect dynamics demonstrated that promotion to managerial positions increased working hours by 0.255 ($p < 0.01$) immediately after promotion, by 0.179 ($p < 0.05$) after one year, and insignificantly after two years. Moreover, job demand score significantly increased to 0.936 ($p < 0.05$) immediately after promotion to managerial positions.

Effect on positive latent mechanisms

Table 5 reports the dynamic and static effects of promotion to managerial positions on the latent mechanisms positively associated with mental health and satisfaction. Figure 3 plots the effects of the dynamics presented in Table 5 for all the estimations of positive latent mechanisms, with conditional parallel trends held. Judging by statistical significance of the static ATT, promotion to managerial positions significantly increased job role clarity score to 0.807 ($p < 0.05$), which was also significant with the ordinal TWFE (0.807, $p < 0.01$). Although ordinal TWFE results indicated that promotion to managerial positions significantly increased work engagement score ($p < 0.05$), static ATT indicated statistical insignificance. The results of the effect dynamics demonstrated that promotion to managerial positions significantly increased work engagement to 0.202 ($p < 0.05$) and job role clarity score to 1.644 ($p < 0.01$) two years after the promotion. There were no significant effects immediately after and one year after the promotion. Although job control was insignificant, the shape was close to work engagement and job role clarity dynamics.

Robustness check

We performed three robustness checks as we explained in the method section: first, estimation of Callaway and Sant'Anna's (2021) method using matched data; second, estimation using Sun and Abraham's (2021) approach; and third, a conventional event study using TWFE. All the tables and figures for the robustness checks are exhibited in the Appendix.

Appendix Tables 6, 7 indicates the changes in the descriptive statistics of individual attributes in the treatment and control groups after matching. While the t-test for the pre-matching data in Table 2 indicated mostly significant difference between the two groups, the t test for the post-matching data no longer confirmed significant differences.

Appendix Figs. 4, 5, and 6 verify the robustness of Figs. 1, 2, and 3, respectively. We mainly checked the shape of event plots. The results revealed three points. First, the

Table 5 Dynamic and static effects of promotion to managerial positions on the latent mechanisms positively associated with mental health and satisfaction

	(1) Log annual household income	(2) Job control	(3) Work engagement	(4) Clarity regarding job role
Lead2	− 0.014	− 1.332	0.190	0.495
	(0.029)	(1.105)	(0.121)	(0.876)
Lead1	0.021	0.599	0.056	0.689
	(0.026)	(0.584)	(0.068)	(0.506)
Lag0	− 0.004	0.862	0.057	0.437
(Treatment timing)	(0.020)	(0.558)	(0.052)	(0.392)
Lag1	0.003	− 0.065	0.060	0.858
	(0.025)	(0.897)	(0.072)	(0.506)
Lag2	0.024	2.091	0.202*	1.644*
	(0.031)	(1.183)	(0.089)	(0.657)
N	5816	5816	5816	5816
Static ATT	0.004	0.783	0.085	0.807*
Ordinal TWFE ATT	0.002	0.662	0.101*	0.887**
p value of pre trend test	0.602	0.259	0.060	0.241
(H0: All Pre-treatment are equal to 0)				

After estimating Eq. (10) using Callaway and Sant'Anna's (2021) method, the estimated ATTg,t is the weighted average as the event study form above *N*, and as Static ATT form below N. Ordinal TWFE ATT indicates the results of TWFE. All results are adjusted by the number of family members. Statistical significance was set as follows: $*p < 0.05, **p < 0.01, ***p < 0.001$

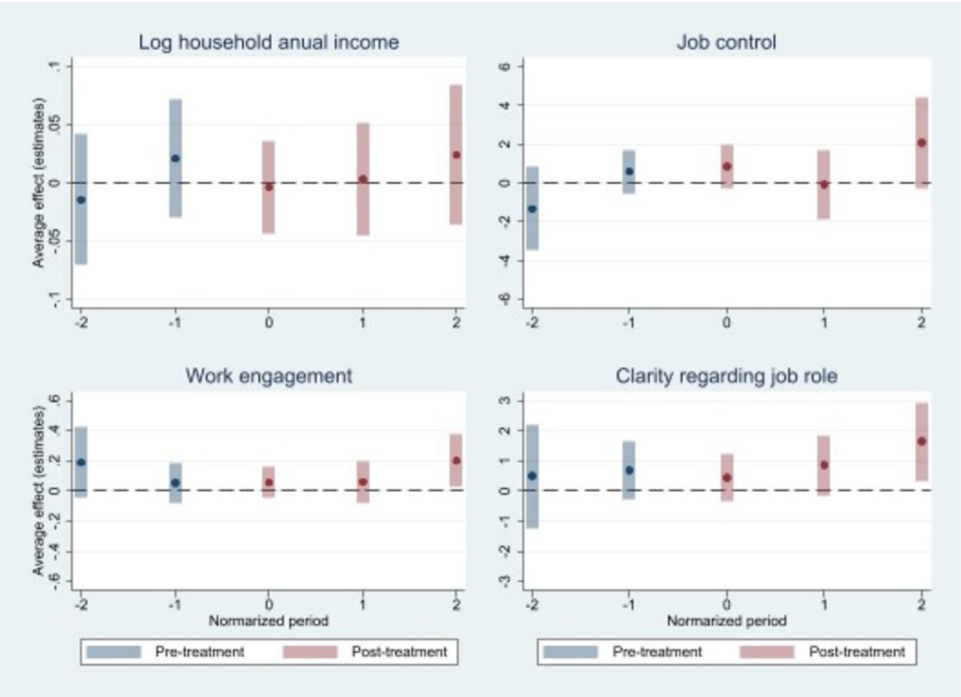

Fig. 3 Dynamic effects of promotion to managerial positions on the latent mechanisms positively associated with mental health and satisfaction

treatment and control groups may be an inappropriate comparison because of concerns that the difference between them could be too large for those promoted to managerial positions. Our matched data results showed similar shape of plots, which revealed no such concerns. Second, estimation results could be changed by the SDID estimation method. Our results showed that the shape of SDID did not change substantially after adopting Sun and Abraham's (2021) approach (other than base period difference). Third, comparing the estimates from the event study using the TWFE confirmed that the shape of dynamics did not change significantly. As the main results, K6 score ≥ 5, working hours, job demand, and clearness of job role showed similar dynamics.

Regarding the detailed statistical significance, in some cases, the results deviated between TWFE and SDID. For example, Callaway and Sant'Anna's (2021) estimation with matching data suggests a significant increase in households' annual income; other estimation did not indicate significant increases. However, our results showed that major concerns raised in the SDID literature, such as diametrically opposite coefficient values, did not arise in our analyses.

Discussion

Interpretation of results

The results indicated that, while promotions to managerial positions have positive effects (through increased work engagement and job role clarity), they also have negative effects (owing to increased work hours and job demands) that potentially link to mental health deterioration with a two-year lag.

As an underlying mechanism of the K6 result, we interpreted that promotions to managers would have both positive and negative effects on mental health, including factors not observed in our data. While both changes would be balanced and cancel each other out immediately after the promotion and one year later, the positive effect would disappear faster than the negative effect after two years, and the negative factors would continue to have more influence two years later, resulting in an imbalanced result with a two-year lag. These results could be related to the hedonic treadmill model (Diener et al. 2006); negative emotions may have a more prolonged impact than positive emotions and may have a slower return speed to individuals' original mental health level. Our results showed that promotion increases the factors that negatively affect mental health (working hours and job demand) and increases the factors that positively affect it (work engagement and clarity of job role). We explored only four factors due to data limitations, but we expect the presence of other relevant factors, such as increased self-esteem and higher wages as potential positive factors, and increased responsibility, pressure for results, and deteriorating work life balance as potential negative factors. Examining the effects of more potential positive and negative factors would provide deeper insight into why mental health deteriorates with a lag. Therefore, a further large-scale panel study asking potential factors such as individual self-esteem and responsibility should be conducted to reveal the mechanism.

We expected wages to increase after a promotion, but our results showed that promotions do not significantly increase household income. This result may be due to measurement error originating from the fact that the income responses had a wide range or asked about household income rather than individual income. The median range of responses was 1–3.5 million yen. An increase in income of, say, 30–50 thousand yen per month due to promotion could be considered an increase in income that does not reach this range, possibly underestimating the effect of income growth. In addition, if the partner quits or starts a job because of household income, the change in household income may not indicate an increase in personal income due to promotion. As shown in Table 2, the average age of the analyzed sample is 40 years, which suggests that children and other factors may affect married households. Since Johnston and Lee (2013) and our matching estimation depicted in Appendix Fig. 6 have shown a significant positive increase in income, an increase in income could be a factor that positively affects mental health, and we should confirm it by an analysis based on appropriate data.

Although the finding regarding mental health effect with a lag of two years seems unusual at first glance, Johnston and Lee (2013) showed similar results using Australian panel data (HILDA). Our results showed the same possibility in Japan using this modified estimation technique. However, while Johnston and Lee (2013) showed significant effects for many other factors, such as job control, we had considerably fewer outcome variables with confirmed significant effects by comparison. Three possible explanations for this difference are that our sample size was approximately a quarter smaller, we handled Japanese workers, and we considered the staggeredness of treatment timing. Differences in attitude about promotion between Japan and Australia may have affected the less significant outcomes. However, our interpretation coincided with that of Johnston and Lee (2013), indicating the presence of both positive and negative effects in mental health.

A comparison of the main results with the event study format based on SDID modification suggests that TWFE did not have major problems, such as reversal of the sign of the point estimates (Goodman-Bacon 2021). However, while no considerable problems were found with the shape of the dynamics, our robustness check results suggested that some statistical significance of static and dynamic TWFE estimates could change. Interpretation based solely on the statistical significance of the TWFE coefficients could lead to misunderstanding conclusions.

Sample restrictions

We focused on four-year consecutive male participants with no missing values for our analytical sample. We focused on men since only 10 female workers were promoted within the complete four-period data (see Appendix Tables 7, 8), and Japan's employment situation largely differs between men and women (Kikuchi et al. 2020; Shimazu et al. 2011). If the employment environment and the impact of promotions could be considered the same for men and women, it would be desirable to aggregate the two to increase the sample size. However, in Japan, such a situation is not suitable. We believe that more than 10 times the additional treated (promotion) sample size is better to analyze the effect of promotion on women and expect the examination of gender differences to be an interesting endeavor. However, this can only be achieved through other data.

We restricted the analysis to complete data with no missing values since the current main estimation model by Callaway and Sant'Anna (2021) can only define ATT with balanced complete panel data. The SDID's main concept is the appropriate comparison. Certainly, we can perform TWFE with unbalanced panel data, but this is largely dependent on the staggeredness of the treatment and the fact that the sample with data periods after the occurrence of the treatment may be large or small; thus, this estimation lacks assurance of the adequacy of the comparison. While the analysis of unbalanced panels may be thought to increase the external validity of data, the high degree of results contamination can undermine the reliability of the results, rendering them less reliable compared to those of clean complete data. Against this background, balanced samples are used in the analysis to pursue internal validity.

Based on the above reasons, we focused on our sample of male workers' complete and balanced data with no missing values. Having a sample with complete data indicates that the workers surveyed in this study may be more diligent than the typical Japanese male worker. Moreover, because we are focusing on currently working individuals, we are not able to include those who quit their jobs due to depression in our data. This could be a source of underestimation in the effect on mental health. Umeda et al. (2015) reported that statistics of poor mental health are higher among Japanese female managers than among those in other occupations at the cross-sectional level. Including female workers in the analysis of job promotion effects in Japan could show larger effects than those revealed here.

Limitations

Our study has three main limitations rooted in the data. First, as the data were only available for four periods, we could only track effects up to two years after the treatment occurs. The J-HOPE was limited to a four-period analysis. For example, the two-year effect in this study was estimated using only the group that was promoted to managerial positions in 2011—the second period of the data. One possible

explanation is that this group may have been unique. It would be more desirable to conduct an analysis that includes groups whose timings of promotion to managerial positions are varied and to adequately consider the weighted average results. The availability of more extended panel data for time-series analyses could address this issue.

Second, external validity was weak. We sacrificed substantial data to improve the internal validity of the data on approximately 6400 individual workers present in a single year at J-HOPE. Consequently, the internal validity of the current analysis improved based on 1454 complete J-HOPE male individuals. The sample used in this study consists of worker data from nine arbitrary Japanese companies that agreed to be surveyed. Therefore, generalizations of this study's results should be handled with caution.

Third, we considered only variations in job title codes and did not specify promotion positions; we analyzed the general effect owing to our data constraints. Adding an event or a job title question would have helped identify the job title for which the effect was more detailed. However, this was challenging owing to our data constraints. As the average age of the respondents was approximately 40 years, the present study may have substantially captured promotion to the first managerial position as the primary treatment. However, such an interpretation requires further clarification. The effect of treatment may be heterogeneous depending on the type of the promotion, and this would be an exciting topic for future studies.

Furthermore, our outcomes were evaluated based on subjective symptoms and not on clinical diagnoses, which is likely to be a source of recall bias. We did not sufficiently unpack the main result mechanisms or why mental health worsens after two years of promotion.

Conclusion

The results indicated that promotions to managerial positions have both positive (through increased work engagement and job role clarity) and negative (through increased work hours and job demands) effects on mental health. The negative effect points to such promotions causing mental health deterioration with a two-year lag.

Again, the sample used in this study consisted of worker data from nine arbitrary Japanese companies that agreed to be surveyed. Therefore, caution is needed when generalizing this study's results. However, our results have the following potential implications for firms' allocation of human resources to managerial positions: While being promoted to a managerial position can positively affect Japanese workers' mental health, it may also be an adverse event with negative effects on the same. If promotions can be implemented such that factors adversely affecting mental health were minimized, promotions to managerial positions could have a less damaging effect on individuals' mental health. For example, avoiding increased working hours for recently promoted workers could reduce negative mental effects, and maintaining good mental health status among workers may lead to increased productivity (Bubonya et al. 2017).

We sought to achieve the highest possible internal validity in our analysis of J-HOPE data by combining the recent developments in the literature on the modified methods of estimating TWFE with the literature on the causal effect of promotions to managerial positions; the latter has witnessed progress primarily through TWFE. Our main results, based on Callaway and Sant'Anna's (2021) estimation approach, were similar in the shape of the event plots compared to the event study form of TWFE. Despite the limitations, our results weakly support the certainty of previous studies that conducted TWFE and event study analyses. Our study also provided an example of dealing with treatment timing heterogeneity for future research on promotion.

Table 6 J-HOPE populations' and analytical complete male sample firms' background

		Survey years				
Anonymous firms' industry		2010	2011	2012	2013	Total
Population of J-HOPE	IT	1080	861	741	586	3268
	Information service	133	751	768	736	2388
	Pharmaceutica	343	335	339	355	1372
	Veterinarian	8	7	8	7	30
	Shipping	1406	1441	1612	1301	5760
	Service 1	46	45	43	42	176
	Service 2	937	976	1111	917	3941
	Medical 1	146	643	503	453	1745
	Medical 2	36	35	32	31	134
	Manufacturing 1	2265	2397	2195	2175	9032
	Manufacturing 2	0	3462	3344	3179	9985
	House-maker 1	0	765	451	349	1565
	House-maker 2	0	119	109	97	325
	Total	6400	11,837	11,256	10,228	39,721
	Manager	1308 (25.6%)	1914 (21.6%)	1823 (21.4%)	1758 (22.8%)	6803 (22.5%)
Male (in J-HOPE population)	IT	839	674	589	471	2573
	Information service	85	505	517	494	1601
	Pharmaceutica	160	156	159	170	645
	Veterinarian	1	1	1	1	4
	Shipping	1356	1383	1554	1259	5552
	Service 1	20	19	15	12	66
	Service 2	572	625	746	604	2547
	Medical 1	45	152	121	121	439
	Medical 2	14	15	13	12	54
	Manufacturing 1	2025	2138	1948	1943	8054
	Manufacturing 2	0	2498	2428	2296	7222
	House-maker 1	0	611	352	282	1245
	House-maker 2	0	74	69	62	205
	Total	5117	8851	8512	7727	30,207
	Manager	1263 (24.7%)	1853 (20.9%)	1763 (20.7%)	1686 (21.8%)	6565 (21.7%)
Complete Male (analytical sample)	IT	263	263	263	263	1052
	Information service	20	20	20	20	80
	Pharmaceutica	43	43	43	43	172
	Veterinarian	0	0	0	0	0
	Shipping	164	164	164	164	656
	Service 1	9	9	9	9	36
	Service 2	63	63	63	63	252
	Medical 1	24	24	24	24	96
	Medical 2	5	5	5	5	20
	Manufacturing 1	863	863	863	863	3452
	Manufacturing 2	0	0	0	0	0
	House-maker 1	0	0	0	0	0
	House-maker 2	0	0	0	0	0
	Total	1454	1454	1454	1454	5816
	Manager	0 (0%)	53 (3.6%)	94 (6.5%)	133 (9.1%)	280 (4.8%)

"Population" indicates J-HOPE raw data, "Male" indicates J-HOPE raw male data, and "Complete Male" indicates analytical sample. Anonymous companies were not listed by name, only by industry. Since "Complete Male" focuses on promotion to management positions, in 2010, managers became 0. "Complete Male" focused on a sample with no missing outcomes (K6, satisfactions, etc., and positive and negative potential mechanisms), control (number of family members), summary statistics variables (tenure, education), and occupation reporting. We also focused entirely on 4-period tractable individuals who were working more than 30 h, not reporting a decreased year of education, non-managers at Period 1, and those with no experience in demotion (manager to the other). Answering J-HOPE was not mandatory; some company's sample sizes in 2011 became larger than in 2010. That is one of the features of J-HOPE

Table 7 Summary statistics for the treatment and control groups with matched data

Individual Characters	Control group		Treatment group		Control—treatment	
	Mean	SD	Mean	SD	Difference	*t* test statistics
Age	41.54	7.93	40.86	5.45	− 0.678	− 0.76
Experience	18.19	9.76	17.14	6.69	− 1.045	− 0.95
Family member	3.15	1.26	3.06	1.27	− 0.087	− 0.52
Years of education	15.44	2.06	15.57	1.75	0.122	0.48
Observations	115		115		230	

Summary statistics were generated by pooling data from 2010 using non-promotion as the control group and promotion to managers as the treatment group with 1:1 matched data. In 'Control—Treatment' we performed *t* tests on the difference between the mean of the control and treatment groups without adjusting for covariates for each variable and compared the results. Statistical significance was set as follows: $*p<0.05, **p<0.01, ***p<0.001$

Table 8 Female workers' promotion in J-HOPE as Table 2

Category	Survey years				
	2010	2011	2012	2013	Four years total
Population (Female)	1283	2985	2744	2498	9510
Complete managers	0	2	4	10	16
Complete non-managers	384	382	380	374	1520
Total	384	384	384	384	1536
Promotion (rate)	0	2 (0.5%)	2 (0.5%)	6 (1.6%)	10 (0.7%)

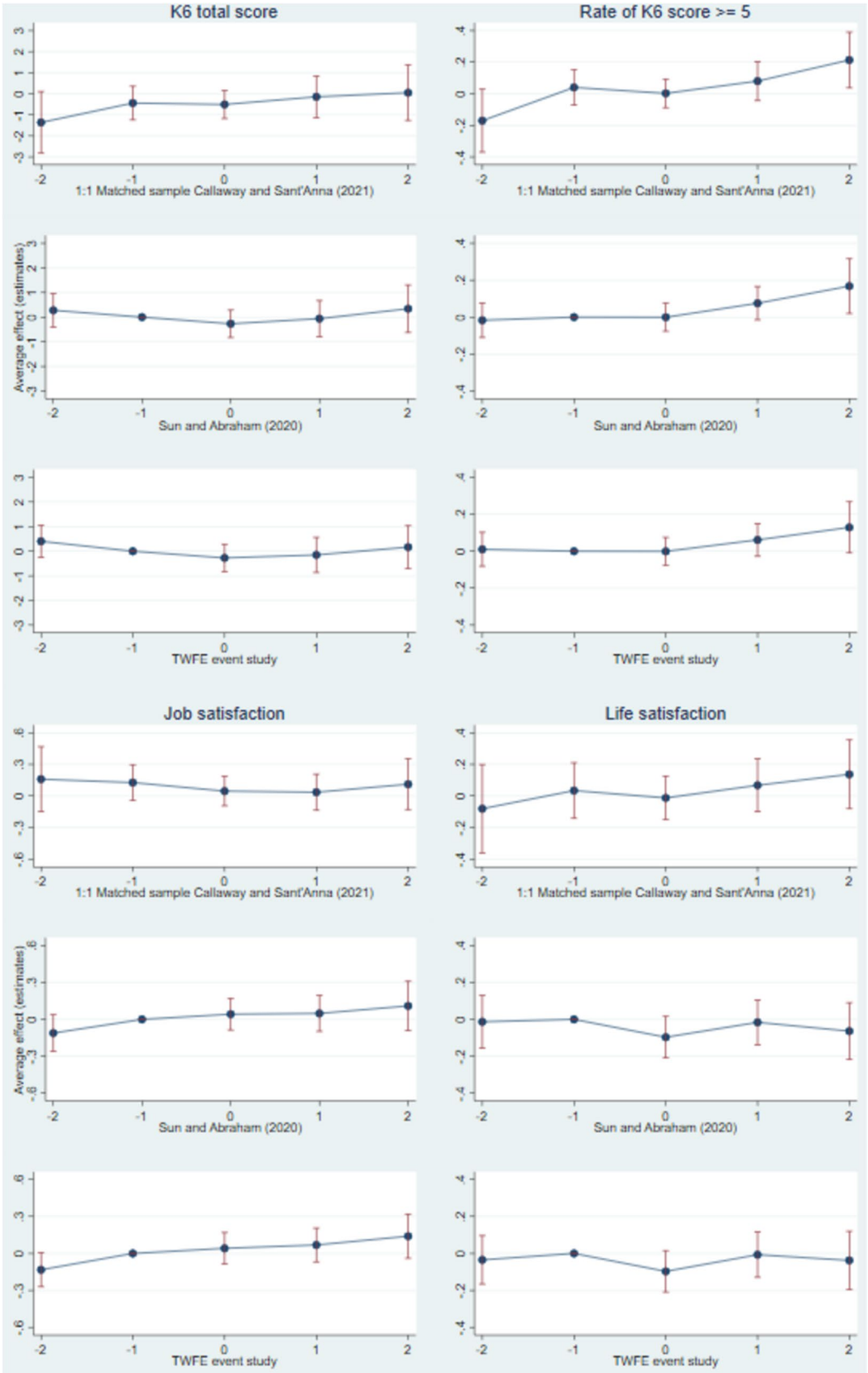

Fig. 4 Robustness check for dynamic total effects

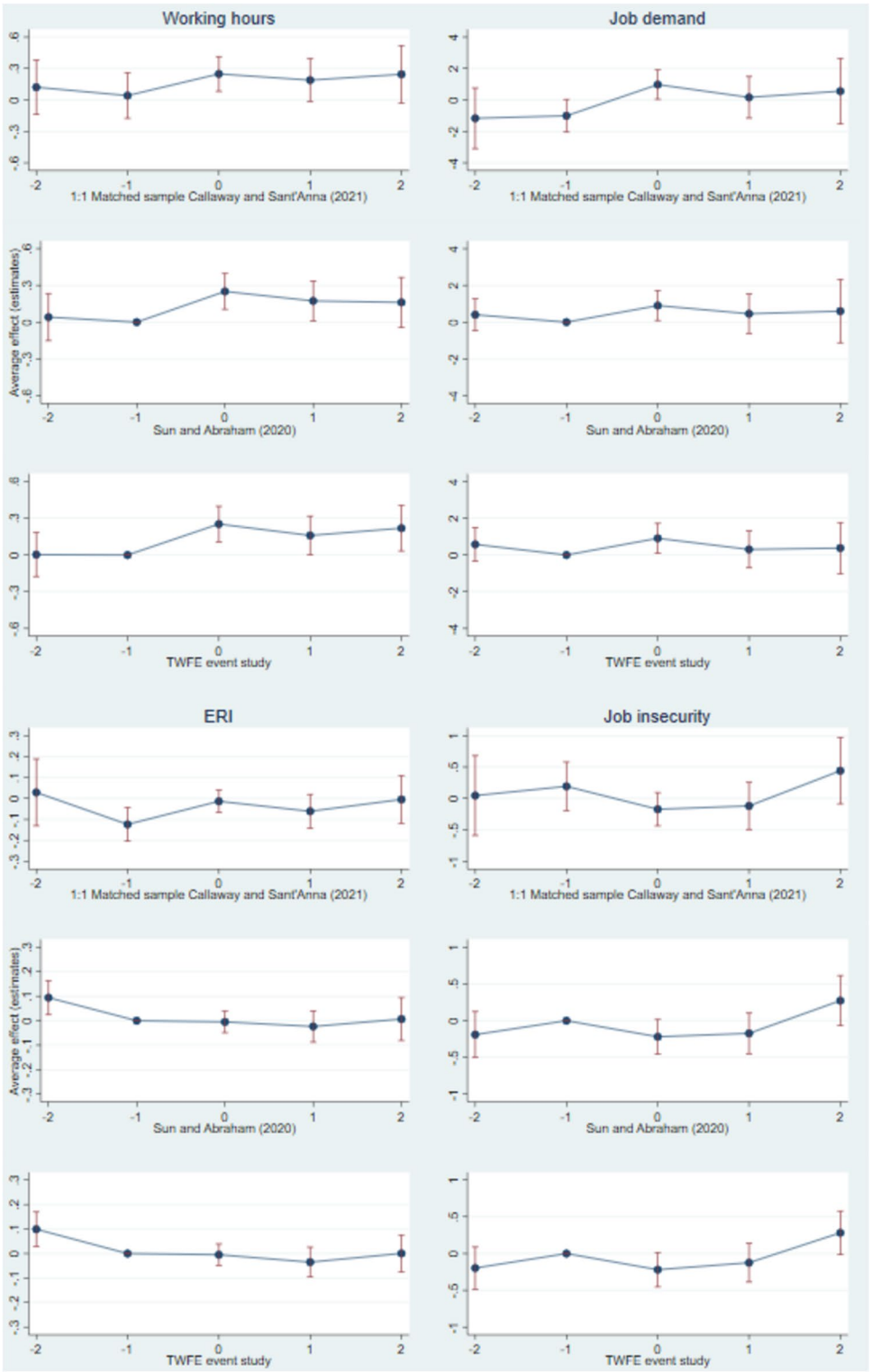

Fig. 5 Robustness check for negative latent mechanisms

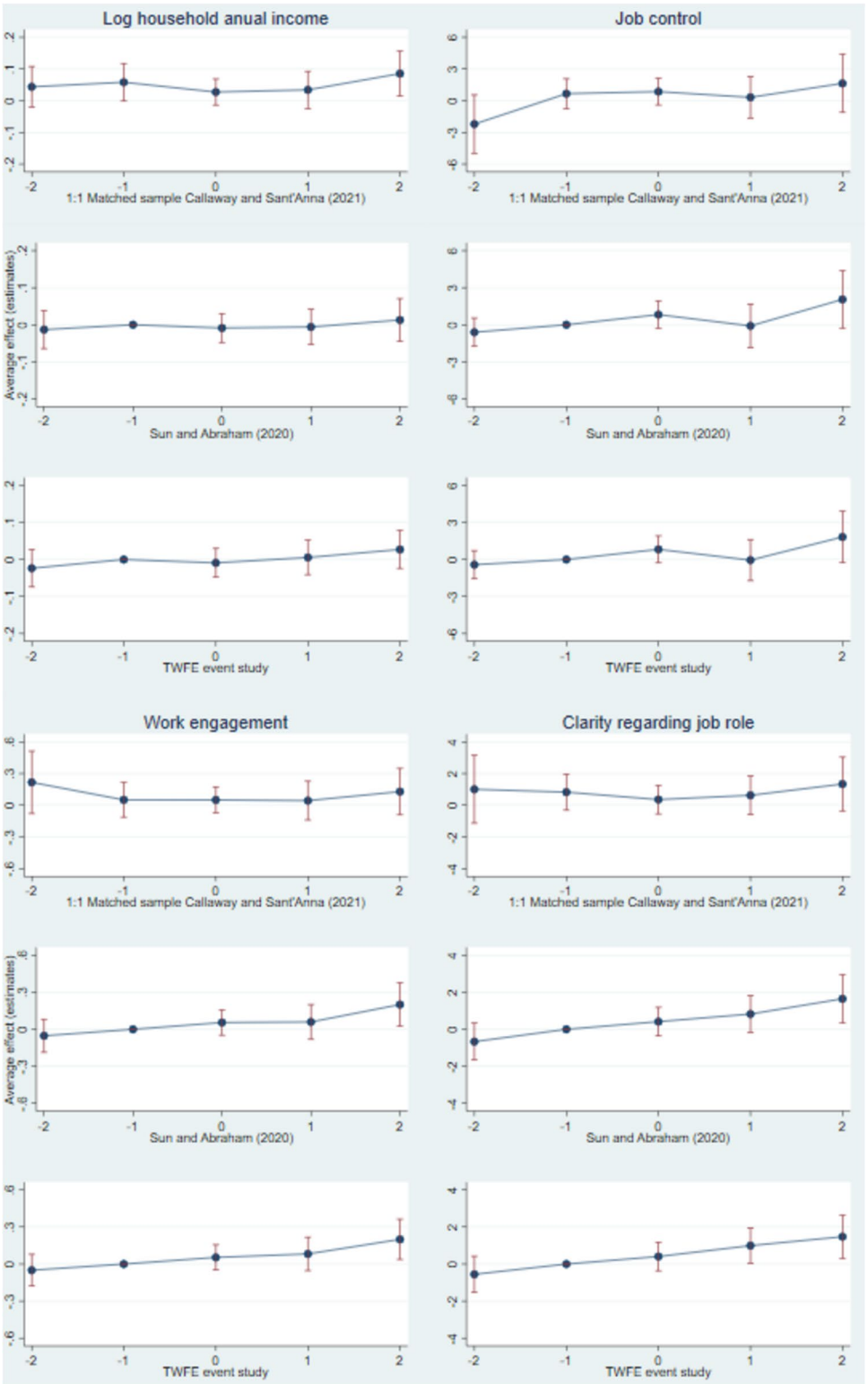

Fig. 6 Robustness check for positive latent mechanisms

Appendix

Detailed description about staggered DID method, TWFE, and ATT

We defined ATT based on Rubin's (1974) causal model. First, we showed that the usual TWFE estimation is a type of DID estimation, which is a representative method for estimating ATT. Next, we demonstrated that TWFE cannot correctly estimate ATT given heterogeneity in treatment timing, such as promotion to managerial positions. We then explained Callaway and Sant'Anna's (2021) estimation, which is robust to treatment timing heterogeneity and is our primary estimation method. Finally, we described the details of the actual estimation.

ATT and DID estimation

First, the outcome variable was defined as Y, and D was defined as a binary indicator variable indicating whether an individual received treatment (treatment group) or not (control group). Y_1 indicates the outcome after treatment, and Y_0 indicates the untreated outcome. The subscript i denotes individuals from 1 to n, t denotes time, and treatment occurs between periods t (ex-ante) and t + 1 (ex-post). The ATT of interest can be described as follows:

$$ATT \equiv E(Y^1_{it+1} - Y^0_{it+1}|D_i = 1) = E\left(Y^1_{it+1} - Y^0_{it}|D_i = 1\right) - E\left(Y^0_{it+1} - Y^0_{it}|D_i = 1\right) \quad (3)$$

If the treatment group is not treated in the t + 1 period after treatment, $E\left(Y^0_{it+1}|D_i = 1\right)$, is the unobservable outcome, which is required in the definition of ATT, as in (3). The movement of the control group ($D_i = 0$) did not receive any treatment. If the parallel trend assumption is satisfied, then we can consider $E\left(Y^0_{it+1} - Y^0_{it}|D_i = 1\right) \approx E\left(Y^0_{it+1} - Y^0_{it}|D_i = 0\right)$ and can replace the unobservable potential outcome in (3) with an approximation, defining ATT as follows:

$$ATT \approx E\left(Y^1_{it+1} - Y^0_{it}|D_i = 1\right) - E\left(Y^0_{it+1} - Y^0_{it}|D_i = 0\right) \quad (4)$$

ATT can be estimated from actual data using the following:

$$y_{it} = a_0 + a_1 D_i + a_2 After_t + a_3 D_i After_t + error_{it} \quad (5)$$

The variables y and D are defined as in (3); "After" is a binary indicator that takes 1 if the time becomes t + 1 period, and "error" is the error term. We can estimate ATT as ATT $\approx a_3$ by using (5).

Estimating ATT using TWFE and event study

We performed TWFE estimation using panel data, as follows:

$$y_{it} = \delta_i + \delta_t + b\, Post\ Treat_{it} + error_{it} \quad (6)$$

PostTreat denotes a binary indicator that continues to take the value of 1 after an individual has been treated, and δ_j (j = i,t) denotes the fixed-effect term of the j factor. The DID estimation in (5) can be considered as a two-period case in (6). This is because the fixed-effect term includes all the attributes of the subscript; thus, we can think of $D_i \in \delta_i$ and $After_t \in \delta_t$. Thus, if the common trend assumption is satisfied, the coefficient value b can be considered as ATT as well as a_3 for the DID estimate.

Along with TWFE, a dynamic decomposition of the onset of the TWFE effect, called an event study, was estimated as follows:

$$y_{it} = \delta_i + \delta_t + \sum_{g=-2}^{-G} \eta_g Lag^g_{it} + \sum_{p=0}^{P} \zeta_p Lead^p_{it} + error_{it} \quad (7)$$

Lag^g_{it} denotes a binary indicator that takes the value of 1 if individual i is in period t before the gth period of treatment, and $Lead^p_{it}$ denotes a binary indicator that takes the value of 1 if individual i is in period t after the pth period of treatment. The baseline is generally measured one period before treatment. This dynamic estimation of (7) allows us to confirm the prior parallel trend (η_g suggests whether a conditional parallel trend could hold in g prior period before the treatment) and determine how long the effect lasts and whether it occurs with a lag. ζ_p was interpreted as ATT after treatment period p.

Situations under which TWFE coefficients do not coincide with ATT

The recent advances in the metrological and theoretical understanding of TWFE have revealed the difficulties in considering the TWFE coefficient b to be an estimate of ATT (De Chaisemartin and d'Haultfoeuille 2020; Borusyak et al. 2021; Goodman-Bacon 2021). Goodman-Bacon (2021) showed that if the sample is divided into three groups—never treated, early treated, and later treated—the TWFE can be broken down into a 2 × 2 DID comparison for each group. This implies that TWFE contains a DID estimation in which the early-treated group's post-treatment status is

considered the control group, as a so-called "forbidden comparison." Therefore, in the context of TWFE and SDID, in some situations, negative weights can cause problems with the estimates, including sign reversal.

Goodman-Bacon (2021) showed that in the presence of treatment timing heterogeneity, the estimated coefficient value b of TWFE estimated in (6) can be decomposed as follows:

$$b = VWATT + VWCT - \Delta ATT \tag{8}$$

VWATT indicates variance-weighted average ATT, VWCT indicates variance-weighted average common trends, and ΔATT indicates the difference in magnitude of ATT for each treatment timing. VWATT is closest to the ordinal ATT and can be considered a weighted average causal effect. Thus, when $VWCT = 0 \wedge \Delta ATT = 0$, we can estimate the causal effect as $b = VWATT$. $VWCT = 0$ when the common trend assumption is satisfied, as in an ordinary DID. $\Delta ATT = 0$, when the effect size is constant regardless of the treatment at any point in time. Conversely, TWFE is biased when these two conditions cannot be satisfied. Borusyak et al. (2021) point out that the occurrence of such bias remains even with event study specification.

Callaway and Sant'Anna's (2021) estimation method

Callaway and Sant'Anna's (2021) estimation method was developed for treatments with staggered timing. The timing at which the treatment occurs for the first time is period g, and G_g is a binary indicator that takes the value 1 for treatment group members who receive the treatment in period g. C is a binary indicator that takes the value of 1 for the never-treated group. They defined the target ATT for each treatment timing group as follows:

$$ATT_{g,t} = E\left(Y_t^1 - Y_t^0 | G_g = 1\right) \tag{9}$$

At $(g,t) = (2,2)$, $ATT_{2,2}$ is equal to the ATT estimated from the DID at 2×2. At $t \geq g$, $E\left(Y_t^0 | G_g = 1\right)$ becomes an unobservable potential outcome. Estimating $ATT_{g,t}$ requires an approximation using observable data. The outcome in period t can be decomposed into the outcome at the level of period g − 1 and the variation from period g − 1 to period t when the period immediately before treatment is used as the reference point. Hence, when the parallel trend assumption holds, we can rewrite as follows:

$$E\left(Y_t^0 | G_g = 1\right) = E\left(Y_{g-1}^0 | G_g = 1\right) + E\left(Y_t^1 - Y_{g-1}^0 | C = 1\right) \tag{10}$$

Therefore, we can approximate this as follows:

$$ATT_{g,t} = E\left(Y_t^1 - Y_t^0 | G_g = 1\right) \approx E\left(Y_t^1 - Y_{g-1}^0 | G_g = 1\right) - E\left(Y_t^0 - Y_{g-1}^0 | C = 1\right) \tag{11}$$

By appropriately weighting these estimated $ATT_{g,t}$, we can capture causal effects as a modified version of TWFE's b in (6) and ζ_p in the event study in (7). Estimation results could be aggregated as follows (e denotes effect timing, T denotes terminal period T):

$$\theta_{\text{event}}(e) = \sum_{g=2}^{T} 1\left[g + e \leq T\right] ATT_{g,g+e} P(G = g | G + e \leq T, C \neq 1) \tag{12}$$

Acknowledgements The authors would like to thank participants at the Hias Lunch Seminar held at Hitotsubashi University and the Health Economics Conference at Chiba University in 2023 for their discussion and comments on the estimation method. We also thank Mr. Takeshi Sasaki for his comments on sample selection and other support. All remaining errors are the authors' responsibility.

Authors' contributions Conceptualization: Ryohei Kashima, Masaya Takahashi; Methodology: Ryohei Kashima, Masaya Takahashi; Formal analysis and investigation: Ryohei Kashima; Writing—original draft preparation: Ryohei Kashima; Writing—review and editing: Masaya Takahashi, Ryohei Kashima; Funding acquisition: Ryohei Kashima; Resources: Masaya Takahashi; Supervision: Masaya Takahashi.

Funding The present work was supported by an internal grant of National Institute of Occupational Safety and Health, Japan (JNIOSH) (Grant No. N-S03-03).

Data availability Data are available upon reasonable request to Professor Akizumi Tsutsumi (Department of Public Health, Kitasato University School of Medicine, Sagamihara, Kanagawa, Japan) as the Principal Investigator of J-HOPE. Stata code for data constructions and analysis are provided by emailing the corresponding author. Please see the following link for access to J-HOPE: https://www.med.kitasato-u.ac.jp/lab/publichealth/eng/jhope.html.

Declarations

Conflict of interests The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval The Ethics Committee of the Graduate School of Medicine/Faculty of Medicine, The University of Tokyo, and the Ethics Committee of the National Istitute of Occupational Safety and Health, Japan (JNIOSH) reviewed and approved the study protocol (No. 2772 and H22010, respectively).

Consent to participate Participants for J-HOPE provided their informed consent before the study.

Consent to publish The authors hereby consent to publication of the current work in the International Archives of Occupational and Environmental Health.

References

Adams P, Hurd MD, McFadden D, Merrill A, Ribeiro T (2003) Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. J Econom 112:3–56. https://doi.org/10.1016/S0304-4076(02)00145-8

Afonso P, Fonseca M, Pires JF (2017) Impact of working hours on sleep and mental health. Occup Med 67:377–382. https://doi.org/10.1093/occmed/kqx054

Amagasa T, Nakayama T (2013) Relationship between long working hours and depression. J Occup Environ Med 55:863–872

Anderson M, Marmot M (2012) The effects of promotions on heart disease: Evidence from Whitehall. Econ J 122:555–589. https://doi.org/10.1111/j.1468-0297.2011.02472.x

Apouey B, Clark AE (2015) Winning big but feeling no better? The effect of lottery prizes on physical and mental health. Health Econ 24:516–538. https://doi.org/10.1002/hec.3035

Baker AC, Larcker DF, Wang CCY (2022) How much should we trust staggered difference-in-differences estimates? J Financ Econ 144:370–395. https://doi.org/10.1016/j.jfineco.2022.01.004

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

Boyce CJ, Oswald AJ (2012) Do people become healthier after being promoted? Health Econ 21:580–596. https://doi.org/10.1002/hec.1734

Bubonya M, Cobb-Clark DA, Wooden M (2017) Mental health and productivity at work: Does what you do matter? Labour Econ 46:150–165. https://doi.org/10.1016/j.labeco.2017.05.001

Callaway B, Sant'Anna PHC (2021) Difference-in-differences with multiple time periods. J Econom 225:200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cordes CL, Dougherty TW, Blum M (1997) Patterns of burnout among managers and professionals: A comparison of models. J Organ Behav 18:685–701. https://doi.org/10.1002/(SICI)1099-1379(199711)18:6%3c685::AID-JOB817%3e3.0.CO;2-U

Coverman S (1989) Role overload, role conflict, and stress: addressing consequences of multiple role demands. Soc Forces 67:965–982. https://doi.org/10.1093/sf/67.4.965

De Chaisemartin C, d'Haultfoeuille X (2020) Two-way fixed effects estimators with heterogeneous treatment effects. Am Econ Rev 110:2964–2996. https://doi.org/10.1257/aer.20181169

De Chaisemartin C, d'Haultfoeuille X (2022) Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical Report

Diener E, Lucas RE, Scollon CN (2006) Beyond the hedonic treadmill: Revising the adaptation theory of well-being. Am Psychol 61(4):305–314. https://doi.org/10.1037/0003-066X.61.4.305

Eguchi H, Shimazu A, Kawakami N, Inoue A, Nakata A, Tsutsumi A (2015) Work engagement and high-sensitivity C-reactive protein levels among Japanese workers: A 1-year prospective cohort study. Int Arch Occup Environ Health 88:651–658. https://doi.org/10.1007/s00420-014-0995-9

Fuchs VR (1980) Time preference and health: An exploratory study. Technical Report

Fukuda Y, Nakamura K, Takano T (2005) Cause-specific mortality differences across socioeconomic position of municipalities in Japan, 1973–1977 and 1993–1998: Increased importance of injury and suicide in inequality for ages under 75. Int J Epidemiol 34:100–109. https://doi.org/10.1093/ije/dyh283

Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, Tachimori H, Iwata N, Uda H, Nakane H, Watanabe M (2008) The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 17:152–158. https://doi.org/10.1002/mpr.257

Goodman-Bacon A (2021) Difference-in-differences with variation in treatment timing. J Econom 225:254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Green F (2011) Unpacking the misery multiplier: How employability modifies the impacts of unemployment and job insecurity on life satisfaction and mental health. J Health Econ 30:265–276. https://doi.org/10.1016/j.jhealeco.2010.12.005

Ikesu R, Miyawaki A, Svensson AK, Svensson T, Kobayashi Y, Chung U-I (2021) Association of managerial position with cardiovascular risk factors: a fixed-effects analysis for Japanese employees. Scand J Work Environ Health 47:425–434. https://doi.org/10.5271/sjweh.3966

Inoue A, Kawakami N, Eguchi H, Tsutsumi A (2016) Modifying effect of cigarette smoking on the association of organizational justice with serious psychological distress in Japanese employees: a cross-sectional study. Int Arch Occup Environ Health 89:901–910. https://doi.org/10.1007/s00420-016-1128-4

Johnson NJ, Sorlie PD, Backlund E (1999) The impact of specific occupation on mortality in the US National Longitudinal Mortality Study. Demography 36:355–367. https://doi.org/10.2307/2648058

Johnston DW, Lee W-S (2013) Extra status and extra stress: are promotions good for us? ILR Rev 66:32–54. https://doi.org/10.1177/001979391306600102

Karasek RA Jr, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B (1998) The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol 3:322. https://doi.org/10.1037/1076-8998.3.4.322

Karasek Jr RA (1979) Job demands, job decision latitude, and mental strain: Implications for job redesign. Adm Sci Q:285–308

Kawakami N, Kobayashi F, Araki S, Haratani T, Furui H (1995) Assessment of job stress 315 dimensions based on the job demands-control model of employees of 316 telecommunication and electric power companies in Japan: reliability and validity 317 of the Japanese version of the Job Content Questionnaire. Int J Behav Med 2:358–375. https://doi.org/10.1207/s15327558ijbm0204_5

Kawakami N, Haratani T, Kobayashi F, Ishizaki M, Hayashi T, Fujita O, Aizawa Y, Miyazaki S, Hiro H, Masumoto T, Hashimoto S (2004) Occupational class and exposure to job stressors among employed men and women in Japan. J Epidemiol 14:204–211. https://doi.org/10.2188/jea.14.204

Kessler D, Bennewith O, Lewis G, Sharp D (2002) Detection of depression and anxiety in primary care: follow up study. Br Med J 325:1016–1017. https://doi.org/10.1136/bmj.325.7371.1016

Kikuchi H, Odagiri Y, Ohya Y, Nakanishi Y, Shimomitsu T, Theorell T, Inoue S (2020) Association of overtime work hours with various stress responses in 59,021 Japanese workers: retrospective cross-sectional study. PLoS ONE. https://doi.org/10.1371/journal.pone.0229506

Kim S, Koh K (2021) The effects of income on health: Evidence from lottery wins in Singapore. J Health Econ 76:102414. https://doi.org/10.1016/j.jhealeco.2020.102414

Kim W, Park E-C, Lee T-H, Kim TH (2016) Effect of working hours and precarious employment on depressive symptoms in South Korean employees: a longitudinal study. Occup Environ Med 73:816–822. https://doi.org/10.1136/oemed-2016-103553

Kokkinen L, Muntaner C, Koskinen A, Väänänen A (2020) Occupational class, capitalist class advantage and mortality among working-age men. J Epidemiol Community Health 74:3–6. https://doi.org/10.1136/jech-2019-212952

Kubo M, Tao M (1994) Burnout among nurses the relationship between stresses and burnout. Jpn J Exp Soc Psychol 34:33–43

Kunst AE, Groenhof F, Andersen O, Borgan JK, Costa G, Desplanques G, Filakti H, Giraldes MD, Faggiano F, Harding S, Junker C (1999) Occupational class and ischemic heart disease mortality in the United States and 11 European countries. Am J Public Health 89:47–53

Lantz PM, House JS, Lepkowski JM, Williams DR, Mero RP, Chen J (1998) Socioeconomic factors, health behaviors, and mortality: Results from a nationally representative prospective study of US adults. JAMA 279:1703–1708. https://doi.org/10.1001/jama.279.21.1703

Mackenbach JP, Stirbu I, Roskam A-JR, Schaap MM, Menvielle G, Leinsalu M, Kunst AE (2008) Socioeconomic inequalities in health in 22 European countries. N Engl J Med 358:2468–2481. https://doi.org/10.1056/NEJMsa0707519

Marmot M (2005) Social determinants of health inequalities. Lancet 365:1099–1104. https://doi.org/10.1016/S0140-6736(05)71146-6

Marmot MG, Bosma H, Hemingway H, Brunner E, Stansfeld S (1997) Contribution of job control and other risk factors to social variations in coronary heart disease incidence. Lancet 350:235–239. https://doi.org/10.1016/S0140-6736(97)04244-X

Miyaki K, Song Y, Taneichi S, Tsutsumi A, Hashimoto H, Kawakami N, Takahashi M, Shimazu A, Inoue A, Kurioka S et al (2013) Socioeconomic status is significantly associated with the dietary intakes of folate and depression scales in Japanese workers (J-HOPE Study). Nutrients 5(2):565–578. https://doi.org/10.3390/nu5020565

Nyberg A, Peristera P, Westerlund H, Johansson G, Hanson LLM (2017) Does job promotion affect men's and women's health differently? Dynamic panel models with fixed effects. Int J Epidemiol 46:1137–1146. https://doi.org/10.1093/ije/dyw310

Ogawa R, Seo E, Maeno T, Ito M, Sanuki M, Maeno T (2018) The relationship between long working hours and depression among first-year residents in Japan. BMC Med Educ 18:50. https://doi.org/10.1186/s12909-018-1171-9

Persol Resarch and Consulting Co. Ltd. (2022) Global employment status and growth attitude survey. [Translated from Japanese: グローバル就業実態•成長意識調査] https://rc.persol-group.co.jp/thinktank/assets/global-2022.pdf

Rahkonen O, Laaksonen M, Martikainen P, Roos E, Lahelma E (2006) Job control, job demands, or social class? The impact of working conditions on the relation between social class and health. J Epidemiol Community Health 60:50–54. https://doi.org/10.1136/jech.2005.035758

Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 66:688. https://doi.org/10.1037/h0037350

Sato K (2015) How does promotion to a managerial position affect mental health. [Translated from Japanese: 管理職への昇進はメンタルヘルスにどのような影響を及ぼすのか] https://www.rieti.go.jp/jp/publications/dp/15j062.pdf

Schaufeli WB, Bakker AB, Salanova M (2006) The measurement of work engagement with a short questionnaire: a cross-national study. Educ Psychol Meas 66:701–716. https://doi.org/10.1177/0013164405282471

Shimazu A, Schaufeli WB, Kosugi S, Suzuki A, Nashiwa H, Kato A, Sakamoto M, Irimajiri H, Amano S, Hirohata K, Goto R (2008) Work engagement in Japan: Validation of the Japanese version of the Utrecht Work Engagement Scale. Appl Psychol 57:510–523. https://doi.org/10.1111/j.1464-0597.2008.00333.x

Shimazu A, Demerouti E, Bakker AB, Shimada K, Kawakami N (2011) Workaholism and well-being among Japanese dual-earner couples: a spillover-crossover perspective. Soc Sci Med 73(3):399–409

Shimomitsu T (2000) The final development of the Brief Job Stress Questionnaire mainly used for assessment of the individuals. Ministry of Labour sponsored grant for the prevention of work-related illness: The 1999 report:126–164

Siegrist J (1996) Adverse health effects of high-effort/low-reward conditions. J Occup Health Psychol 1:27–41. https://doi.org/10.1037/1076-8998.1.1.27

Smith JP (1999) Healthy bodies and thick wallets: The dual relation between health and economic status. J Econ Perspect 13:145–166. https://doi.org/10.1257/jep.13.2.145

Stringhini S, Dugravot A, Shipley M, Goldberg M, Zins M, Kivimäki M, Marmot M, Sabia S, Singh-Manoux A (2011) Health behaviours, socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. PLoS Med. https://doi.org/10.1371/journal.pmed.1000419

Sun L, Abraham S (2021) Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J Econom 225:175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Tsutsumi A, Ishitake T, Peter R, Siegrist J, Matoba T (2001) The Japanese version of the Effort-Reward Imbalance Questionnaire: a study in dental technicians. Work Stress 15:86–96. https://doi.org/10.1080/02678370118173

Tsutsumi A, Kayaba K, Ishikawa S (2011) Impact of occupational stress on stroke across occupational classes and genders. Soc Sci Med 72(10):1652–1658. https://doi.org/10.1016/j.socscimed.2011.03.026

Umeda M, McMunn A, Cable N, Hashimoto H, Kawakami N, Marmot M (2015) Does an advantageous occupational position make women happier in contemporary Japan? Findings from the Japanese Study of Health, Occupation, and Psychosocial Factors Related Equity (J-HOPE). SSM-Population Health 1:8–15. https://doi.org/10.1016/j.ssmph.2015.09.002

Virtanen M, Ferrie JE, Singh-Manoux A, Shipley MJ, Stansfeld SA, Marmot MG, Ahola K, Vahtera J, Kivimäki M (2011) Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. Psychol Med 41:2485–2494. https://doi.org/10.1017/S0033291711000171