Happiness Makes Workers More Productive:
Evidence from Large-Scaled Experiments

Keiko Iwasaki *

NLI Research Institute

Abstract

There is an increasing interest among firms in investing in the happiness of their employees. However, the empirical evidence of the causal relationship between happiness and productivity is limited (Oswald et al., 2015; Bellet et al., 2020). Therefore, we conducted two different styles of large-scaled experiments which exogenously provide the variation in the level of happiness among employees and civil servants in Japan (n=6,201) to test the causal relationship. The first experiment is a Randomized Controlled Trial (RCT) showing a comedy clip to the treatment group while showing a control clip of moving shapes to the control group to test if the raised happiness induced by the comedy clip makes participants more productive. The second experiment is a natural experiment exploiting exogenous real-life negative shocks on happiness (death or serious illness of the spouse) to test if the lowered happiness caused by the shocks decreases productivity. The productivity of each participant is measured by the number of correct answers of timed mathematical additions that participants solve for monetary incentives after watching a comedy/control clip. Both experiment results support the causal relationship of happiness raising the productivity of workers.

Keywords: Happiness; Productivity; Randomized Controlled Trial

Journal of Economic Literature (JEL) Classification Codes: J24, D90, C90

***** Researcher, NLI Research Institute. 4-1-7 Kudan-kita, Chiyoda-ku, Tokyo 102-0073, Japan.
E-mail: kiwasaki@nli-research.co.jp

1. Introduction

There is an increasing interest among firms in investing in the happiness of their employees. However, most of the empirical studies on the relationship between happiness and productivity are correlational studies and the causal evidence is scarce. Exceptions are Oswald et al. (2015) and Bellet et al. (2020). Oswald et al. (2015) conducted different types of laboratory experiments inducing the variety in the happiness among elite university students and provided a first causal evidence that happiness makes the students more productive. Bellet et al. (2020) provided the first field evidence by studying call sales employees of one of the largest firms in the United Kingdom exploiting the changes in weather as exogenous shocks to influence employees' daily happiness levels. However, since the sample characteristics are limited in both of the studies, further empirical studies are necessary to provide externality to the causal relationship between happiness and productivity.

Therefore, we conduct two different styles of large-scaled experiments which exogenously provide the variation in the level of happiness among employees of firms and civil servants in Japan (n=6,201) to test the causal relationship. Applying the laboratory experiments conducted by Oswald et al. (2015), we include two types of experiments in a carefully designed web survey, which enables us to have large participants of workers with the variety of characteristics.

The first experiment is a Randomized Controlled Trial (RCT) showing a comedy clip to the treatment group while showing a control clip of moving shapes to the control group to test if the raised happiness induced by the comedy clip makes participants more productive. The second experiment exploits exogenous real-life negative shocks on happiness (death or serious illness of the spouse) to test if the lowered happiness caused by the shocks decreases productivity. The productivity of each participant is measured by the number of correct answers of timed mathematical additions that participants solve for monetary incentives after watching a comedy/control clip.

As to the first experiment, the intervention of the comedy clip to raise happiness was overall not successful but successful only for those who live in Tokyo probably because of regional differences in the sense of humor. Therefore, we further study participants from Tokyo and find that those who are in the treatment group solve more additions correctly than those who are in the control group. As a result of the second experiment, we find that those who have experienced spousal bereavement or the serious illness of the spouse show lower happiness levels as well as lower productivity compared to those who have not. Both experiment results support the causal relationship of happiness raising the productivity of workers.

The rest of the paper is organized as follows. Section 2 explains our research design to examine the impact of happiness on productivity. Section 3 present the empirical results of two types of the experiments and Section 4 contains concluding remarks on our findings.

2. Research Design

A. General Information of the Data

Our original data is obtained by a carefully designed web survey containing Randomized Controlled Trial (RCT) conducted in March 2019. Sample size is 6,201. All the participants are employees of firms or civil servants among the monitor members of Cross Marketing Inc. Gender, age and regional distributions of the sample are adjusted to their distributions of the Japan Census 2015. Out of 6,201 participants, information on 4,568 are used for analysis to avoid rough answers attributing from the nature of the web survey. To explain further, we dropped those who made a wrong choice in a trap question included in the survey and those who did not make any correct answers in the mathematical additions trial which is used as the productivity measure (details of the variable will be explained as follows) from the sample for analysis.

The survey consists of three sections. The first section includes questions about the basic backgrounds of the participants such as age, gender, subjective health status and life events as well as happiness and emotional status. In the second section, the intervention of the RCT (watching a minute video clip) is performed. In the final section, participants are asked to solve timed (three minutes) mathematical additions for monetary incentives.[1]

B. Experiment 1

The first experiment is an RCT testing the impact of positive emotion using a mood induction by a one-minute comedy clip. The experiment consists of five steps.

Step 1. All the participants answer questions to measure the original level of positive emotion.

Step 2. Participants are randomly divided into a treatment group or a control group and those who are in the treatment group watch a one-minute comedy clip while those who are in the control group watch a one-minute control clip showing moving shapes.

Step 3. Participants answer a simple multiple-choice question asking about the content of the clip.

---

[1] In the explanation before the trial, participants are told that top 800 participants will receive bonus100 point (≒100 JPY). Considering that each participant receives about 150 point as a base contribution point to the survey, the magnitude of the point bonus cannot be considered a subtle.

Step 4. All the participants answer questions to measure the level of positive emotion again to check the impact of the clip on their levels of positive emotion.

Step 5. All the participants solve mathematical additions for three minutes for monetary incentives.

To measure the level of positive emotion in Step 1(before watching the clip) and Step 4(after watching the clip), we employ the Japanese version of “Brief, Momentary Mood Checklists (BMMC)” (Thomas, D. L., & Diener, E., 1990 ; Tanaka, 2008) following (Kurokawa et al., 2014). BMMC uses four positive words (happy, joyful, pleased, enjoyment/fun) and five negative words (depressed/blue, unhappy, frustrated, angry/hostile, worried). Participants rate the degree to which he/she is experiencing with a 7-point scale ranging from 0 (not at all) to 6 extremely much) to each of nine words. We define the sum of the four positive words’ scores before watching the clip as *PE (before)* and the sum after watching the clip as *PE (after)*.

Since we use the web survey method, we cannot actually observe participants watching the provided clip like laboratory experiments. Therefore, to check if a participant has actually watched the clip or not, we include one simple multiple-choice question asking about the content of the clip after the treatment as Step 3. Among, 4,568 participants, 3,932 participants answered the question correctly so that we use information of those 3,932 participants for analyzing the impact of the mood induction on productivity.

To capture the productivity (*PR*) of each participant, we use the number of correct answers of the mathematical additions within 3 minutes (Step 5). Before starting the trial, participants are informed that top scored 800 participants will get bonus 100 points (values100 JPY) in addition to the base contribution points which is about 150 points.

For analysis, we first check if the random assignment to the treatment/control group was successful by a balancing test. Second, we test if the comedy clip successfully raised the *PE* of the participants with OLS regression. After making sure about the impact of the treatment on participants’ emotional status, we test if the treatment also raised productivity or not using both the reduced form model and the two-stage leased squares (2SLS) model using the treatment dummy (*Treatment*) as an exogenous variable affecting *PE (after)*. The summary statistics of all variables used for analysis of both experiment 1 and 2 are available in Table 1.

C. Experiment 2

While the first experiment tests the impacts of the short-term mood induction on productivity, the second experiment tests the impact of relatively long-term happiness status on productivity exploiting the real-life negative shocks on happiness. We define the negative shocks as the death of the spouse within a past year and the current serious illness of the spouse.

We define a variable *Sad event* taking one when one has experienced either of the negative shocks. The level of happiness (*Happiness*) is measured with 11-point scale general happiness question (11: Very happy…1: Miserable). We use *PR* as a productivity variable same as the analysis of the first experiment.

Though the second experiment is not an RCT, we use *Sad event* as an exogenous variable to affect participants' level of happiness since the shocks are unintended natural events in most of the cases as Oswald et al. (2015) explains. Therefore, it can be considered as a natural experiment and the exogeneity of *Sad event* would be more plausible after controlling for age, sex, one's health status, income, and ability. To control for one's ability, we use a measurement of numeracy and language and math grades of the junior high school. As to numeracy measurement, we use the four mathematical questions from the Berlin Numeracy Test (Cokely et al., 2012) following Ikawa and Kusumi (2018).[2] The number of correct answers out of the four mathematical questions is used as a *Numeracy* variable. As to the language and math grade of the junior high school, we ask one's own grades when they were in junior high school retrospectively.

In same way as the first experiment, after making sure about the impact of the *Sad even*t on the participants' level of happiness, we test if those who have experienced *Sad event* show lower productivity compared to those who have not, with both reduced form model and 2SLS model using the *Sad event* as an exogenous variable affecting the level of *Happiness*.

3. Results

A. Experiment 1

First, we check if the randomization of the treatment is successful with a balancing test (See Table A1). With $p < 0.05$ significance level, only one variable (*Age*) is significant. Considering that there are 21 variables included in the model (Column 2 of Table A1) and 5

[2] The four questions used in the Berlin Numeracy Test by Cokely et al. 2012 are following:
1. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? ________
2. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (please indicate the probability in percent).
3. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6?
4. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red?
The correct answers are 1 = 30; 2 = 25; 3 = 20; 4 = 50.

percent of them can mistakenly reject the null, our randomization can be considered quite successful. For analysis, not only we control for *Age*, we use other insignificant variables as controls for robustness check as well though our randomization looks successful.

Second, we test if the treatment (the comedy clip) raised the *PE* of the participants. As shown in Figure A1 and Table A2, unfortunately, the treatment was not successful in raising *PE*. We attribute this from the well-known cultural difference in the sense of humor (Ura, 2018; Senuma, 2015) and further searched for the areas where the treatment was affective. We first use the regional categorization (10 categories) to find regions where the treatment was affective. However, as shown in Table A3, we find no regions where the treatment was affective. Therefore, we further investigated in the prefectural level (47 prefectures) and found that the treatment successfully raised happiness only among those who live in Tokyo as shown in Table A4. The comedy clip was selected by the group of researchers in NLI research institute, which is located in Tokyo. Accordingly, it is possible that the selection of the comedy clip reflects the cultural sense of humor of those who live in Tokyo, which could be a reason why the comedy clip raised *PE* of only those who live in Tokyo. Hence, we further investigate the impact of the treatment on productivity only using the information of those who live in Tokyo (n=494).

Since we use the sub sample (participants from Tokyo) for analysis because of the heterogenous treatment effect depending on prefectures, we conduct a balancing test only using the information of participants from Tokyo. As shown in Table A5, with $p < 0.05$ significance level, only one variable (*Language grade: Don't know*) is significant, which shows that the randomization is quite successful in the sub sample as well.

Figure 1 indicates the difference of *PE (after)* between the treatment group and the control group among participants from Tokyo. As found in Table A4, we can see that *PE (after)* of the treatment group is higher than that of the control group. OLS regression results regressing *PE (after)* on *Treatment* is presented in Table 2, column (1) and (2). Column (1) shows the estimation result without any control variables and (2) is with control variables. Both results verify that the treatment successfully raised *PE* of participants from Tokyo.

As to the impact of the treatment on productivity (*PR*), we can see in Figure 2 that *PR* of the treatment group is higher than that of the control group. OLS regression results of regressing *PR* on *Treatment* is presented in Table 2, column (3) and (4). Column (3) presents the estimation result without any control variables and (4) is with control variables. Both results confirm that *PR* of the treatment group is higher than that of the control group. According to the estimation results of column (3) and (4), the treatment raised *PR* 4.5 to 6.0 point. Considering that the mean of *PR* of the control group among participants from Tokyo is 52, our results show that the treatment raised productivity about 9 to 12 percent.

Furthermore, using *Treatment* as an instrument variable, we estimate the 2SLS model.

Column (5) in Table 2 present the result of the second stage estimation and column (2) in Table 2 presents the result of first stage estimation. These 2SLS model estimation results confirm that *PE* raises *PR*. Though weak instrument is a suspect because maximal IV size is not less than 15 percent but 25 percent, the concern of resulting bias is not a serious problem here since our model is just-identified 2SLS model which can be considered approximately unbiased (Angrist, J. & Pischke, J. S., 2009).

B. Experiment 2

First, using all the available sample (n=4,578), we test if those who have experienced the sad events (spousal bereavement within a past year or serious illness of the spouse) report lower level of happiness or not. As shown in Figure 3, though those who are not married show the lowest level of happiness, those who have experienced the sad events report lower level of happiness than those who are married and have not experienced the sad events. Figure 4 incorporates the gender difference of the impact of the sad events. It shows that women report relatively higher level of happiness, but the sad events lower the level of happiness among both men and women. OLS regression results regressing *Happiness* on *Sad event* are presented in Table 3, column (1) and (2). Column (1) shows the estimation result without any control variables and column (2) shows the estimation result with control variables. Both results confirm that those who have experienced the sad events report lower level of happiness.

As to the impact of the sad events on *PR*, we can see in Figure 5 that *PR* of those who have experienced the sad events is lower than those who are not married or those who are married and have not experienced the sad events. Figure 6 incorporates the gender difference of the impact of the sad events, which shows that productivity of those who have experienced sad events is lower than that of others both among men and women.

OLS regression results of regressing *PR* on *Sad event* is presented in Table 2, column (3) and (4). Column (3) presents the estimation result without any control variables and (4) is with control variables. Both results confirm that *PR* of those who have experienced the sad events is lower than that of other people. According to the estimation results of column (3) and (4), experience of the sad events lowers 6 to 9 point of *PR*. Considering that the mean of *PR* among those who are married and have not experienced the sad events is 53, our results show that the sad events lower productivity about 11 to 17 percent.

Furthermore, using *Sad event* as an instrument variable, we estimate the 2SLS model. Column (5) in Table 3 presents the result of the second stage estimation and column (2) in Table 3 presents the result of the first stage estimation. The estimation results confirm that *Happiness* raises *PR*. Maximal IV size is less than 15 percent and supports the validity of our model.

As a robustness check, we conduct estimations using two different variables as treatment variables instead of *Sad event*, which are a dummy for those who have experienced spousal bereavement within a past year (*Spouse death*) and a dummy for those whose spouses are ill (*Spouse ill*). Figure A2 presents the comparison of the mean score of *Happiness* depending on those who are not married, those who are married, and those whose spouse have died within a past year. Since there are only 5 people in our sample whose spouse have died within a past year, the standard deviation is large for the *Happiness* of those whose spouse have died within a past year. However, we can still observe that mean score of *Happiness* among those who have experienced spousal bereavement within a past year is relatively lower than that of those who are married. Furthermore, the mean score incorporating the gender difference presented in Figure A3 shows that the negative impact of the spousal bereavement on *Happiness* is especially for men but not for women which is consistent with a previous findings such as Spahni et al., 2015, which shows that more women tend to have resilience from the mental shock of the spousal bereavement than men.

Figure A4 presents the comparison of the mean score of *Happiness* of those who are not married, those who are married and their spouses are not ill and those whose spouses are ill. We can observe that the mean score of *Happiness* among those whose spouses are ill is lower than that of those who are married and their spouses are not ill. Unlike the impact of the death of the spouse which shows heterogenous effect on happiness depending on gender difference, illness of spouses has a negative impact on happiness for both men and women as shown in Figure A5.

OLS regression results regressing *Happiness* on *Spouse death* and *Spouse ill* are presented in Table A6, column (1) and (2). Column (1) shows the estimation result without any control variables and column (2) shows the estimation with control variables. Both results confirm that death of the spouse and illness of the spouse have negative impact on the level of *Happiness*, and the death of the spouse have heterogenous impact on *Happiness* between men and women.

As to the impact of the death of the spouse and the illness of the spouse on *PR*, we observe that either impact have negative affect on *PR* among both men and women (See Figure A6 to A9). OLS regression results of regressing *PR* on *Spouse death* and *Spouse ill* is presented in Table A6, column (3) and (4). Column (3) presents the estimation result without any control variables and (4) is with control variables. Both results confirm that both *Spouse death* and *Spouse ill* have negative impact on *PR*. According to the estimation result of column (3) and (4), point estimates of *Spouse death* and *Spouse ill* are -34 to -20 and -9 to -6, respectively. Considering that the mean of *PR* among those who are married and have not experienced either sad events is about 53, our results show that death of the spouse can lower *PR* 38 to 64 percent and the illness of the spouse can lower *PR* 11 to 17 percent.

Furthermore, using *Spouse death* and *Spouse ill* as instrument variables, we estimate the

2SLS model. Column (5) in Table A6 presents the result of the second stage estimation and column (2) in Table A6 presents the result first stage estimation. The endogenous variable, *Happiness*, is positive and significant in column (5), which confirms the positive impact of *Happiness* on *PR*. Both the overidentification test and the weak identification test support the validity of our model.

4. Conclusion

Our study provides a causal evidence that happiness raises productivity of workers. We conducted two types of experiments which are an RCT and a natural experiment using the real-life negative shocks following Oswald et al. (2015). While Oswald et al. (2015) conducted laboratory experiments targeting elite university students, we conducted the experiments among actual workers in a large scale using a web survey. To our knowledge, this is the first study providing a causal evidence of the relationship between happiness and productivity among workers using an RCT.

Our RCT results show that the positive emotion induced by a comedy clip raises the productivity of those who live in Tokyo about 9 to 12 percent. Also, the natural experiment results show that real-life negative shocks lower the productivity of workers by about 11 to 17 %. The magnitudes of the impacts are consistent with the findings from Oswald et al. (2015) even though backgrounds of participants are different.

Our study implies that the investments on the happiness of employees can be beneficial for firms since it can improve productivity. Therefore, the results have the real-world contributions to provide firms broaden choices to improve their productivity. However, our RCT results also show the difficulty of improving the happiness of the employees as the intervention was not overall successful but successful for only participants from Tokyo. The accumulation of empirical studies on strategies to improve happiness and program evaluations of actual efforts of firms to improve happiness and productivities as well as theoretical studies exploring the possible mechanisms behind the link which can include the role of stress release (Fredrickson, 2011) and laughter (Bennet et al. 2003; Ikeda et al. 2020; Miller et al, 2009) would be important to achieve efficient field applications.

Reference

Angrist, J.& Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Bellet, C. S., Neve, J-E. D. & Ward, G. (2020). Does Employee Happiness Have an Impact on Productivity? *CEP Discussion Paper*, No 1655.

Bennet, M. P., Zeller, J. M., Rosenberg, L. & McCann, J. (2003). The Effect of Mirthful Laughter on Stress and Natural Killer Cell Activity. *Alternative Therapies in Health and Medicine*, 9(2), 38-45.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S. & Garcia-Retamero, R. (2012). Measuring Risk Literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, 7(1), 25–47

Fredrickson B. L. (2001). The Role of Positive Emotions in Positive Psychology. The Broaden-and-Build Theory of Positive Emotions. *The American psychologist*, 56(3), 218–226.

Ikawa, M. & Kusumi, T. (2018). The Effects of Numeracy on Change in Fear of Risk about Radioactive Substances in Food and Tokyo Inland Earthquake. *Japanese Journal of Risk Analysis*, 28(1), 23-30.

Ikeda, S., Ikeda, A., Yamagishi, K., Hori, M., Kubo, S., Sata, M., Okada, C., Umesawa, M., Sankai, T., Kitamura, A., Kiyama, M., Ohira, T., Tanigawa, T. & Iso, H. (2020). Longitudinal Trends in Blood Pressure Associated with the Frequency of Laughter: Longitudinal Study of Japanese General Population: the Circulatory Risk in Communities Study (CIRCS). *Journal of Epidemiology*, Accepted Article.

Kurokawa, H., Inukai, K. & Ohtake, F. (2014). Kanjyo no Henka ga Jikansenkou ni Oyobosu Eikyou. *Journal of Behavioral Economics and Finance*, 7, Proceedings, 45-49.

Miller, M. & Fry, W. F. (2009). The Effect of Mirthful Laughter on the Human Cardiovascular System. *Medical hypotheses*, 73(5), 636–639.

Oswald, A., Proto, E. & Sgroi, D. (2015). Happiness and Productivity, *Journal of Labor Economics*, 33(4) 789 - 822.

Senuma, F. (2015). Daigakusei no Warai wo Sketch Suru: Touzai no Anketo Chosa kara. *Japanese Journal of Laughter and Humor Research*, 22, 75-92.

Spahni, S., Morselli, D., Perrig-Chiello, P. & Bennett, K. M. (2015). Patterns of Psychological Adaptation to Spousal Bereavement in Old Age, *Gelotology*, 61(5), 456-468.

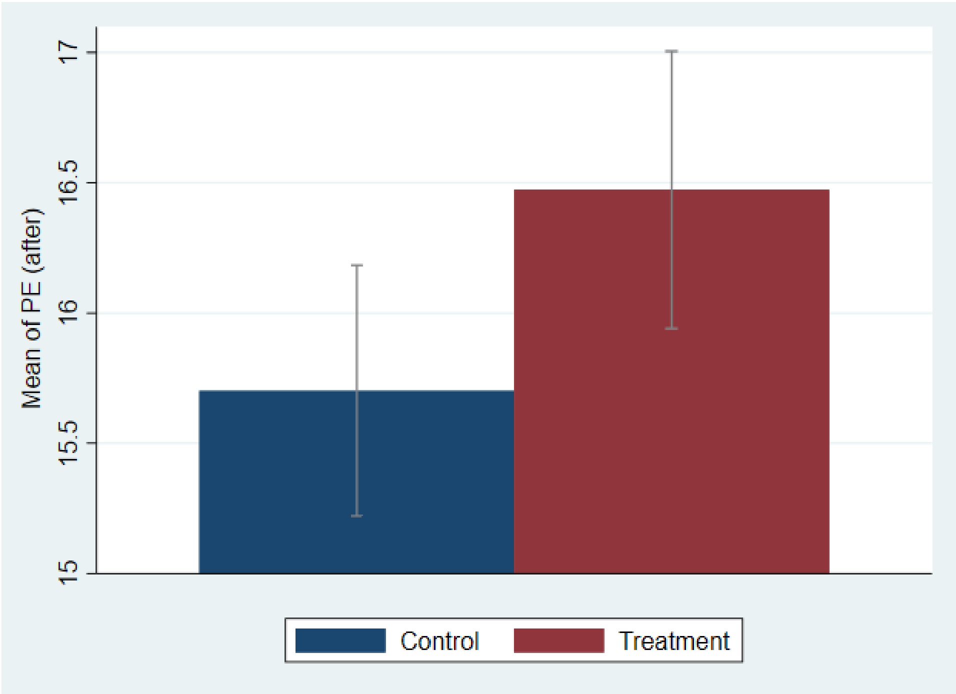
Tanaka. K. (2008). The Reliability & Validity of Brief, Momentary Mood Checklists?: Japanese Version (BMC-J). *Journal of Osaka University of Economics*, 58(7), 271-275.

Thomas, D. L. & Diener, E. (1990). Memory Accuracy in the Recall of Emotions. *Journal of Personality and Social Psychology*, 59(2), 291 - 297.

Ura, K. (2018). Naniwa no “Warai” wo Kangaeru. *Naniwa Osaka no “Warai” ni Kansuru Chosa to Kenkyu*, Research Center for Naniwa-Osaka Studies, 1-12.

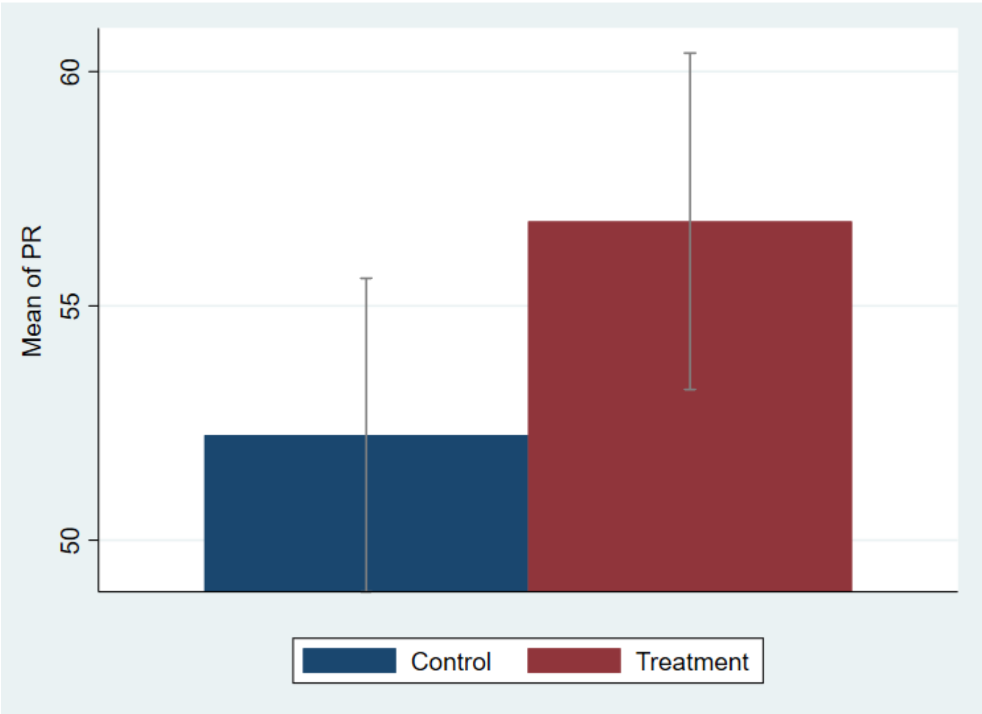
Figures

Figure 1. Treatment Effect on Positive Emotion (*PE*): Tokyo



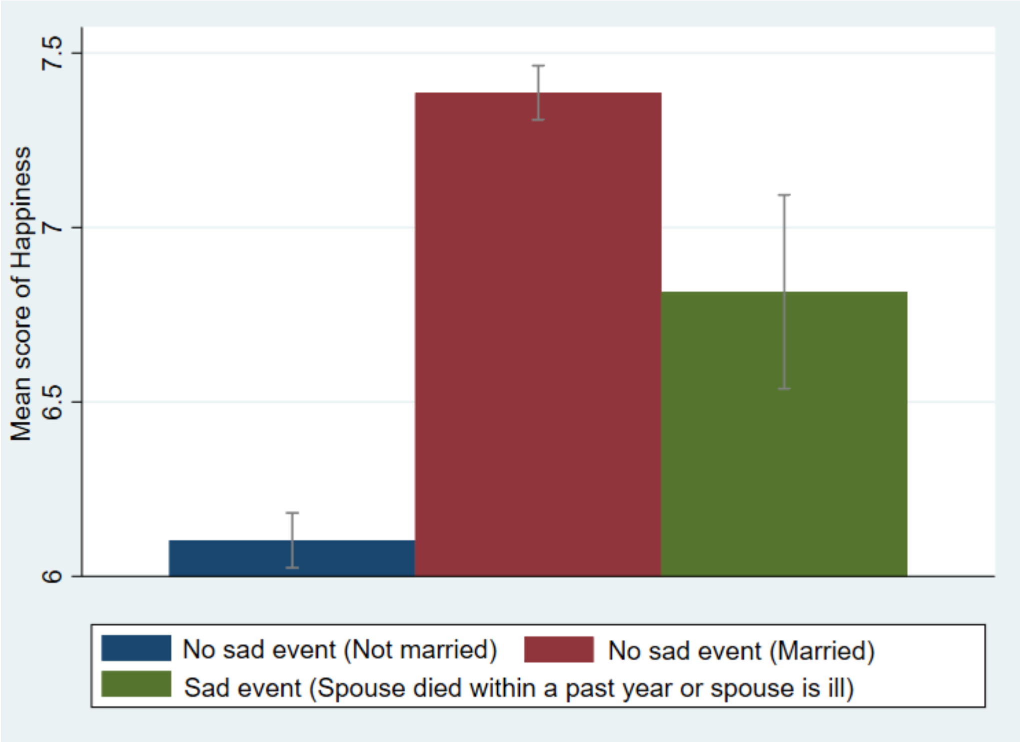
Notes: N=494. 90 % confidence intervals presented.

Figure2. Treatment Effect on Productivity (*PR*): Tokyo



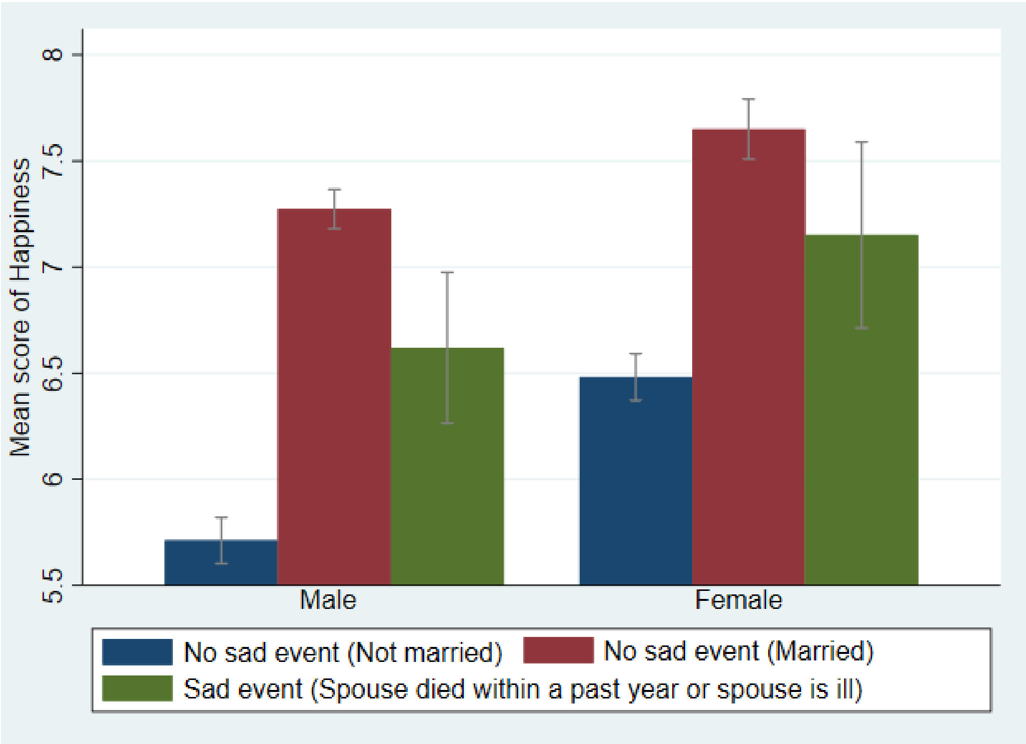
Notes: N=494. 90 % confidence intervals presented.

Figure 3. The Impact of *Sad event* on *Happiness*



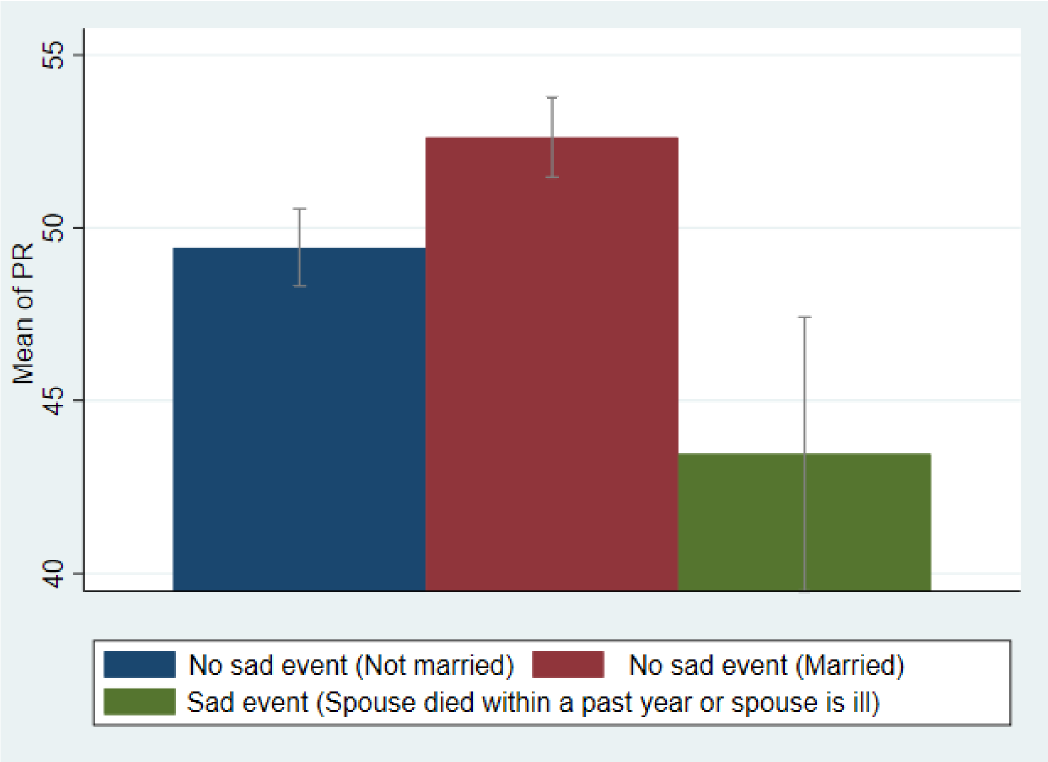
Notes: N=4,568. 90 % confidence intervals presented.

Figure 4. Gender Difference of the Impact of *Sad event* on *Happiness*



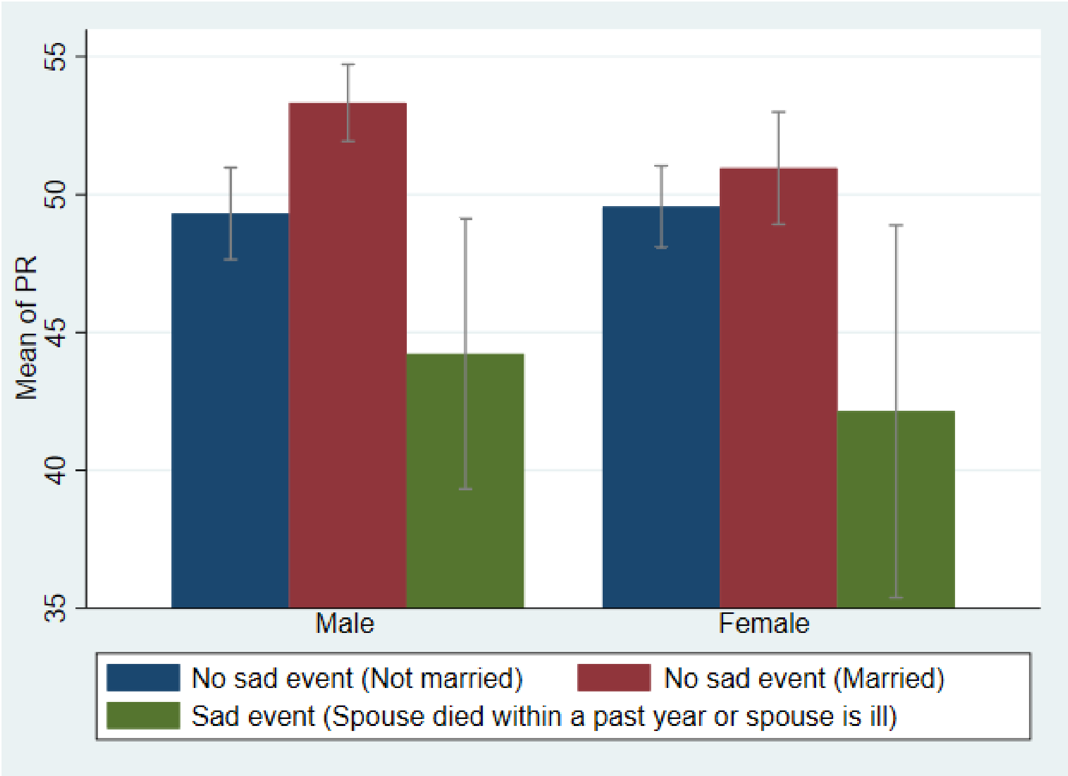
Notes: N=4,568. 90 % confidence intervals presented.

Figure 5. The Impact of *Sad event* on Productivity (*PR*)



Notes: N=4,568. 90 % confidence intervals presented.

Figure 6. Gender Difference of the Impact of *Sad event* on Productivity (*PR*)



Notes: N=4,568. 90 % confidence intervals presented.

Tables

Table 1. Summary Statistics

Variable	Obs	Mean	Std.D	Min	Max	Explanation of the variable
Treatment	4,568	0.503	0.500	0	1	Dummy for those who are in the treatment group (watched the comedy clip)
PE (after)	4,568	15.812	5.143	4	28	Positive emotion score after watching a clip
PE (before)	4,568	15.682	5.205	4	28	Positive emotion score before watching a clip
PR	4,568	50.599	32.343	1	159	Productivity measured by the number of the correct answers of additions
Skipped clip	4,568	0.139	0.346	0	1	Dummy for those who skipped the clips (incorrect answer to the confirmation quiz) *Skipped clip* = 1 is dropped from analysis of the Experiment 1.
Spouse death	4,568	0.001	0.033	0	1	Dummy for those whose spouse died within a past year
Spouse ill	4,568	0.033	0.179	0	1	Dummy for those whose spouse is ill
Sad event	4,568	0.039	0.194	0	1	Dummy for those "*Spouse death* = 1" or "*Spouse ill* =1"
Happiness	4,568	6.695	2.320	1	11	11 points scale Happiness score (11=Very Happy…1=Miserable)
Female	4,568	0.412	0.492	0	1	Female dummy
Age	4,568	42.571	11.319	18	64	Age
Numeracy	4,568	1.401	1.276	0	4	Number of correct answers of four mathematical questions from the Berlin Numeracy test
Income: Less than 3 mill JPY	4,568	0.273	0.446	0	1	Dummy for those whose yearly income is less than 3 mill JPY
Income: 3 - 7 mill JPY	4,568	0.447	0.497	0	1	Dummy for those whose yearly income is between 3 mill to 7 mill JPY
Income: 7 - 10 mill JPY	4,568	0.112	0.316	0	1	Dummy for those whose yearly income is between 7 mill to 10 mill JPY
Income: 10 - 15 mill JPY	4,568	0.039	0.195	0	1	Dummy for those whose yearly income is between 10 mill to 15 mill JPY
Income: 15 mill JPY or more	4,568	0.007	0.082	0	1	Dummy for those whose yearly income is 15 mill JPY or more
Income: 0 JPY	4,568	0.001	0.036	0	1	Dummy for those who don't have income
Income: Don't know	4,568	0.120	0.324	0	1	Dummy for those who don't know or don't want to answer his/her yearly income

Table 1. Summary Statistics (Continued)

Variable	Obs	Mean	Std.D	Min	Max	Explanation of the variable
Language grade: Low	4,568	0.112	0.315	0	1	Dummy for those whose junior high school language grade was low
Language grade: Relatively low	4,568	0.114	0.318	0	1	Dummy for those whose junior high school language grade was relatively low
Language grade: Middle	4,568	0.294	0.456	0	1	Dummy for those whose junior high school language grade was middle
Language grade: Relatively high	4,568	0.230	0.421	0	1	Dummy for those whose junior high school language grade was relatively high
Language grade: High	4,568	0.216	0.412	0	1	Dummy for those whose junior high school language grade was high
Language grade: Don't know	4,568	0.033	0.178	0	1	Dummy for those who doesn't know junior high school language grade
Math grade: Low	4,568	0.173	0.378	0	1	Dummy for those whose junior high school math grade was low
Math grade: Relatively low	4,568	0.138	0.345	0	1	Dummy for those whose junior high school math grade was relatively low
Math grade: Middle	4,568	0.247	0.431	0	1	Dummy for those whose junior high school math grade was middle
Math grade: Relatively high	4,568	0.180	0.385	0	1	Dummy for those whose junior high school math grade was relatively high
Math grade: High	4,568	0.230	0.421	0	1	Dummy for those whose junior high school math grade was high
Math grade: Don't know	4,568	0.032	0.175	0	1	Dummy for those who doesn't know junior high school math grade
Subjective health	4,568	2.225	0.811	1	4	Question: "Are you healthy? Answer: 1=Strongly agree … 4=Not at all
Married	4,568	0.478	0.500	0	1	Dummy for those who are married

Table 2. Impact of Positive Emotion (*PE*) on Productivity (*PR*)

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	*PE(after)*	*PE(after)*	*PR*	*PR*	*PR*
Model:	OLS	OLS	OLS	OLS	2SLS
Treatment	0.771*	0.573**	4.560+	6.043**	
	(0.435)	(0.240)	(2.977)	(2.615)	
PE (after)					10.54*
					(6.183)
PE (before)		0.822***		-0.961***	-9.624*
		(0.0376)		(0.276)	(5.287)
Numeracy		-0.118		9.588***	10.84***
		(0.120)		(1.073)	(1.911)
Female		-0.500**		-2.379	2.887
		(0.249)		(3.066)	(4.647)
Age		0.0198**		-0.215*	-0.424**
		(0.0100)		(0.115)	(0.187)
N	494	494	494	494	494
adj. R-sq	0.004	0.726	0.003	0.255	-0.408
Kleibergen-Paap rk Wald F statistic	-	-	-	-	5.704
(Maximal IV size)	-	-	-	-	(<25%)

Notes: Robust standard errors in parenthesis. The constant term is not presented. Other omitted variables from column (2), (4) and (5) are Subjective health, Income dummies, Language grade dummies and Math grade dummies. Those coefficients are not reported in the table but are available from the corresponding author upon request. Column (2) is the first stage estimation results of column (5).
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level
*** Significant at the 1% level

Table 3. Impact of *Happiness* on Productivity (*PR*)

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	*Happiness*	*Happiness*	*PR*	*PR*	*PR*
Model:	OLS	OLS	OLS	OLS	2SLS
Sad event	-0.549***	-0.564***	-9.123***	-6.262***	
	(0.174)	(0.161)	(2.497)	(2.383)	
Happiness					11.10**
					(5.061)
Married	1.284***	1.124***	3.208***	1.206	-11.26**
	(0.0670)	(0.0684)	(0.973)	(0.957)	(5.599)
Female		0.504***		1.215	-4.373+
		(0.0717)		(1.002)	(2.812)
Numeracy		0.0492*		7.974***	7.427***
		(0.0263)		(0.366)	(0.533)
Age		0.00335		-0.259***	-0.296***
		(0.00293)		(0.0395)	(0.0545)
N	4,568	4,568	4,568	4,568	4,568
adj. R-sq	0.073	0.210	0.004	0.198	-0.274
Kleibergen-Paap rk Wald F statistic	-	-	-	-	12.323
(Maximal IV size)	-	-	-	-	(<15%)

Notes: Robust standard errors in parenthesis. The constant term is not presented. Other omitted variables from column (2), (4) and (5) are Subjective health, Income dummies, Language grade dummies, and Math grade dummies. Those coefficients are not reported in the table but are available from the corresponding author upon request. Column (2) is the first stage estimation results of column (5).
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level
*** Significant at the 1% level

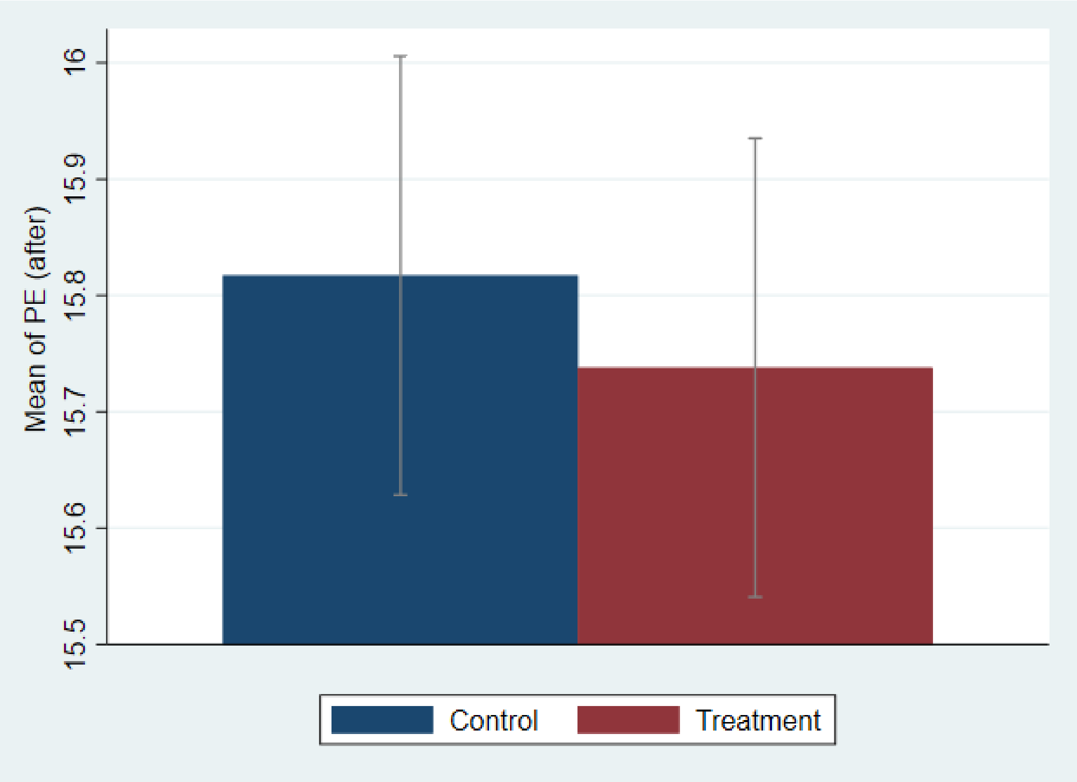
Appendix A: Overall Insignificance of the *Treatment*

Table A1. Balancing test using all sample

	(1)	(2)
PE (before)	-0.000524	0.000387
	(0.00151)	(0.00160)
Female dummy		-0.0282+
		(0.0183)
Age		-0.00168**
		(0.000732)
Numeracy		0.0123*
		(0.00686)
Income: Less than 3 mill JPY		Reference
Income: 3 mill - 7 mill JPY		-0.0139
		(0.0203)
Income: 7 mill - 10 mill JPY		-0.00639
		(0.0305)
Income: 10 mill - 15 mill JPY		-0.0325
		(0.0457)
Income: 15 mill JPY or more		-0.00142
		(0.104)
Income: 0 JPY		-0.0501
		(0.208)
Income: Don't know		-0.0130
		(0.0291)
Language grade: Low		Reference
Language grade: Relatively low		-0.00669
		(0.0371)
Language grade: Middle		0.0248
		(0.0339)
Language grade: Relatively high		0.0194
		(0.0347)
Language grade: High		0.0435
		(0.0359)
Language grade: Don't know		0.243+
		(0.148)
Math grade: Low		Reference
Math grade: Relatively low		0.0324
		(0.0320)
Math grade: Middle		-0.0217
		(0.0297)
Math grade: Relatively high		-0.0234
		(0.0315)
Math grade: High		-0.0510+
		(0.0318)
Math grade: Don't know		-0.205
		(0.149)
Subjective health		0.0137
		(0.0105)
Constant	0.474***	0.499***
	(0.0249)	(0.0544)
N	3,932	3,932
adj. R-sq	-0.000	0.000

Notes: Dependent variable is *Treatment*. Robust standard errors in parenthesis. +Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Figure A1. Treatment Effect on Positive Emotion (*PE*) Using All Sample



Notes: N=4,568. 90 % confidence intervals presented.

Table A2. Treatment Effect on *PE* Using All Sample

	(1)	(2)	(3)
Treatment	-0.0792	-0.0304	-0.0169
	(0.166)	(0.0920)	(0.0891)
PE (before)		0.835***	0.815***
		(0.0128)	(0.0115)
Female dummy			0.0310
			(0.103)
Age			0.00209
			(0.00408)
Numeracy			0.0235
			(0.0378)
N	3,932	3,932	3,932
adj. R-sq	-0.000	0.716	0.719

Notes: Dependent variable is *PE (after)*. Robust standard errors in parenthesis. The constant term is not presented. Other omitted variables from column (3) are Subjective health, Income dummies, Language grade dummies and Math grade dummies. Those coefficients are not reported in the table but are available from the corresponding author upon request.
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level
*** Significant at the 1% level

Appendix B: Searching for Areas Where the Treatment was Successful

Table A3. Treatment Effect Depending on Regions

	(1)	(2)
Region (Hokkaido) × *Treatment*	0.619+	0.602+
	(0.384)	(0.383)
Region (Tohoku) ×*Treatment*	-0.444	-0.471
	(0.366)	(0.367)
Region (Minami-Kanto) ×*Treatment*	0.120	0.162
	(0.159)	(0.158)
Region (Kita-Kanto) ×*Treatment*	-0.0445	-0.0790
	(0.340)	(0.339)
Region (Chubu) ×*Treatment*	-0.134	-0.125
	(0.221)	(0.217)
Region (Kinki) ×*Treatment*	0.0655	0.0573
	(0.222)	(0.225)
Region (Chugoku) ×*Treatment*	-0.0968	-0.0715
	(0.317)	(0.316)
Region (Shikoku) ×*Treatment*	-0.479	-0.319
	(0.522)	(0.533)
Region (Kyusyu) ×*Treatment*	-0.0939	-0.0476
	(0.325)	(0.320)
Region (Okinawa) ×*Treatment*	-1.427+	-1.621+
	(0.985)	(1.001)
PE (before)	0.835***	0.814***
	(0.0106)	(0.0116)
Age	0.00119	0.00245
	(0.00407)	(0.00411)
Numeracy	0.0503+	0.0263
	(0.0343)	(0.0379)
N	3,932	3,932
adj. R-sq	0.715	0.718

Notes: Dependent variable is *PE (after)* Robust standard errors in parenthesis. The constant term is not presented. Other omitted variables from both columns are dummies for each region. In addition, from column (2), Subjective health, Income dummies, Language grade dummies and Math grade dummies are omitted. Those coefficients are not reported in the table but are available from the corresponding author upon request.
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Table A4. Treatment Effect Depending on Prefectures

	(1)	(2)
Pref (Hokkaido) × *Treatment*	0.621+	0.600+
	(0.388)	(0.387)
Pref (Aomori) × *Treatment*	-1.973**	-1.980**
	(0.785)	(0.791)
Pref (Iwate) × *Treatment*	-0.216	-0.183
	(0.989)	(0.948)
Pref (Miyagi) × *Treatment*	0.197	0.204
	(0.584)	(0.593)
Pref (Akita) × *Treatment*	-0.959	-1.009
	(0.971)	(0.974)
Pref (Yamagata) × *Treatment*	-1.453	-1.432
	(1.159)	(1.149)
Pref (Fukushima) × *Treatment*	0.162	0.0567
	(0.995)	(0.983)
Pref (Ibaraki) × *Treatment*	-0.474	-0.575
	(0.691)	(0.685)
Pref (Tochigi) × *Treatment*	-0.860	-0.926
	(0.912)	(0.918)
Pref (Gunma) × *Treatment*	0.771	0.860
	(0.766)	(0.738)
Pref (Saitama) × *Treatment*	0.0183	0.139
	(0.416)	(0.418)
Pref (Chiba) × *Treatment*	0.193	0.226
	(0.377)	(0.374)
Pref (Tokyo) × *Treatment*	0.494**	0.520**
	(0.234)	(0.235)
Pref (Kanagawa) × *Treatment*	-0.475	-0.437
	(0.340)	(0.337)
Pref (Niigata) × *Treatment*	-0.0371	0.0159
	(0.538)	(0.518)
Pref (Toyama) × *Treatment*	-1.348+	-1.328+
	(0.851)	(0.836)
Pref (Ishikawa) × *Treatment*	-1.049	-1.181+
	(0.801)	(0.781)
Pref (Fukui) × *Treatment*	1.044	0.933
	(1.921)	(1.855)
Pref (Yamanashi) × *Treatment*	1.391+	1.295
	(0.966)	(0.984)
Pref (Nagano) × *Treatment*	-0.358	-0.349
	(0.549)	(0.554)
Pref (Gifu) × *Treatment*	0.291	0.286
	(0.948)	(0.951)
Pref (Shizuoka) × *Treatment*	-0.210	-0.284
	(0.463)	(0.459)
Pref (Aichi) × *Treatment*	0.162	0.204
	(0.334)	(0.327)
Pref (Mie) × *Treatment*	-0.828	-0.807
	(0.870)	(0.851)
Pref (Shiga) × *Treatment*	-0.469	-0.577
	(0.761)	(0.781)
Pref (Kyoto) × *Treatment*	-0.297	-0.324
	(0.694)	(0.696)

Table A4. Treatment Effect Depending on Prefectures (Continued)

	(1)	(2)
Pref (Osaka) × *Treatment*	-0.138	-0.179
	(0.307)	(0.313)
Pref (Hyogo) × *Treatment*	0.587	0.601
	(0.466)	(0.468)
Pref (Nara) × *Treatment*	0.463	0.578
	(0.937)	(0.922)
Pref (Wakayama) × *Treatment*	0.579	0.635
	(1.137)	(1.114)
Pref (Tottori) × *Treatment*	-0.377	-0.0515
	(0.857)	(0.854)
Pref (Shimane) × *Treatment*	0.0547	-0.269
	(0.607)	(0.695)
Pref (Okayama) × *Treatment*	1.162+	1.188*
	(0.721)	(0.704)
Pref (Hiroshima) × *Treatment*	-0.318	-0.358
	(0.480)	(0.483)
Pref (Yamaguchi) × *Treatment*	-1.147*	-1.019*
	(0.589)	(0.593)
Pref (Tokushima) × *Treatment*	-0.103	-0.0223
	(1.282)	(1.396)
Pref (Kagawa) × *Treatment*	-0.530	-0.380
	(0.864)	(0.886)
Pref (Ehime) × *Treatment*	-0.491	-0.316
	(0.814)	(0.825)
Pref (Kochi) × *Treatment*	-0.644	-0.424
	(1.601)	(1.532)
Pref (Fukuoka) × *Treatment*	-0.536	-0.497
	(0.490)	(0.484)
Pref (Saga) × *Treatment*	-0.538	-0.468
	(1.642)	(1.542)
Pref (Nagasaki) × *Treatment*	0.348	0.502
	(1.120)	(1.135)
Pref (Kumamoto) × *Treatment*	0.222	0.226
	(1.117)	(1.098)
Pref (Oita) × *Treatment*	-0.474	-0.585
	(0.980)	(0.932)
Pref (Miyazaki) × *Treatment*	0.511	0.721
	(1.272)	(1.125)
Pref (Kagoshima) × *Treatment*	0.886	1.121
	(0.803)	(0.801)
Pref (Okinawa) × *Treatment*	-1.430	-1.610+
	(0.994)	(1.011)
PE (before)	0.834***	0.814***
	(0.0106)	(0.0115)
Age	0.00193	0.00337
	(0.00407)	(0.00411)
Numeracy	0.0496	0.0269
	(0.0350)	(0.0386)
N	3,932	3,932
adj. R-sq	0.716	0.719

Notes: Dependent variable is *PE (after)* Robust standard errors in parenthesis. The constant term is not

presented. Other omitted variables from both columns are dummies for each prefecture. In addition, from column (2), Subjective health, Income dummies, Language grade dummies and Math grade dummies are omitted. Those coefficients are not reported in the table but are available from the corresponding author upon request.
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

Appendix C. Balancing Test for Participants from Tokyo

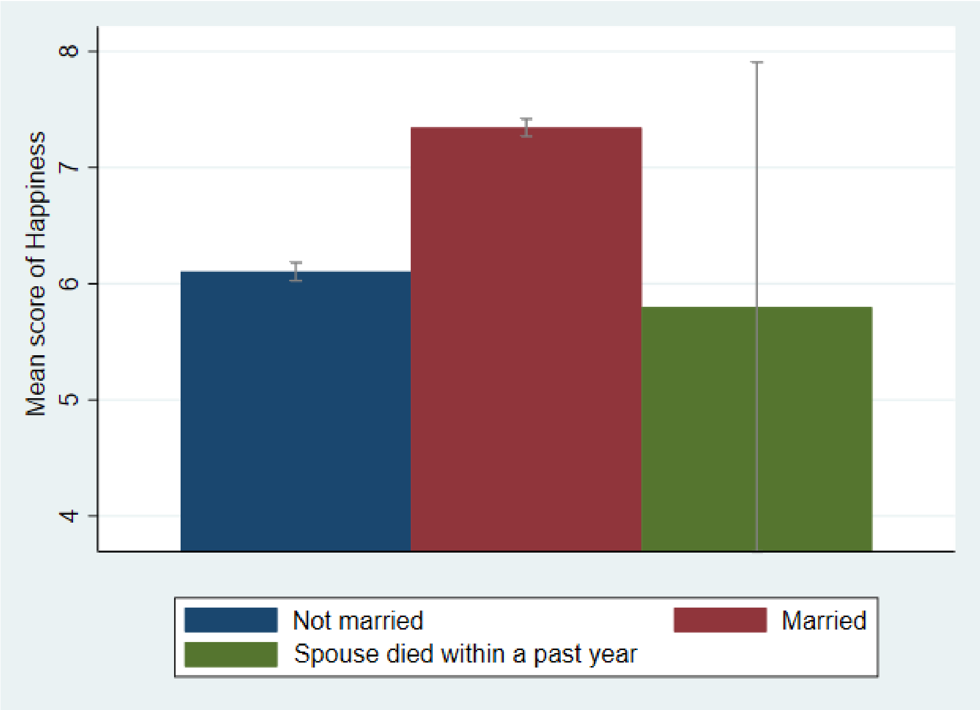
Table A5. Balancing Test: Tokyo

	(1)	(2)
PE (before)	0.00356	0.00397
	(0.00465)	(0.00500)
Numeracy		-0.00924
		(0.0195)
Female dummy		-0.0623
		(0.0519)
Age		-0.00400*
		(0.00215)
Income: Less than 3 mill JPY		Reference
Income: 3 mill - 7 mill JPY		0.0263
		(0.0629)
Income: 7 mill - 10 mill JPY		-0.00156
		(0.0810)
Income: 10 mill - 15 mill JPY		0.00375
		(0.107)
Income: 15 mill JPY or more		0.0449
		(0.366)
Income: Don't know		0.0637
		(0.0924)
Language grade: Low		Reference
Language grade: Relatively low		-0.188*
		(0.102)
Language grade: Middle		-0.0609
		(0.0973)
Language grade: Relatively high		0.0641
		(0.101)
Language grade: High		-0.0724
		(0.106)
Language grade: Don't know		0.475***
		(0.124)
Math grade: Low		Reference
Math grade: Relatively low		0.0239
		(0.0956)
Math grade: Middle		0.0248
		(0.0884)
Math grade: Relatively high		0.0360
		(0.0987)
Math grade: High		0.0990
		(0.0966)
Math grade: Don't know		-0.306*
		(0.168)
Subjective health		0.0403
		(0.0321)
Constant	0.427***	0.514***
	(0.0775)	(0.165)
N	494	494
adj. R-sq	-0.001	0.008

Notes: Dependent variable is *Treatment*. Robust standard errors in parenthesis.
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level

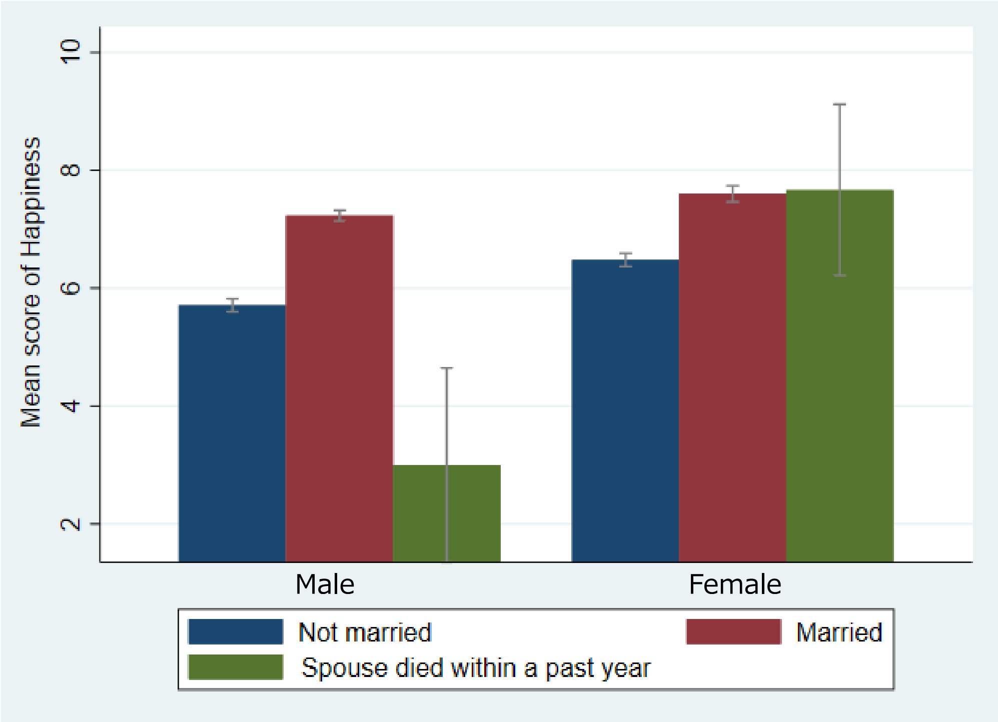
Appendix D. Estimations Treating the Death of the Spouse and the illness of the Spouse as Separate Variables

Figure A2. The Impact of the Death of the Spouse on Happiness



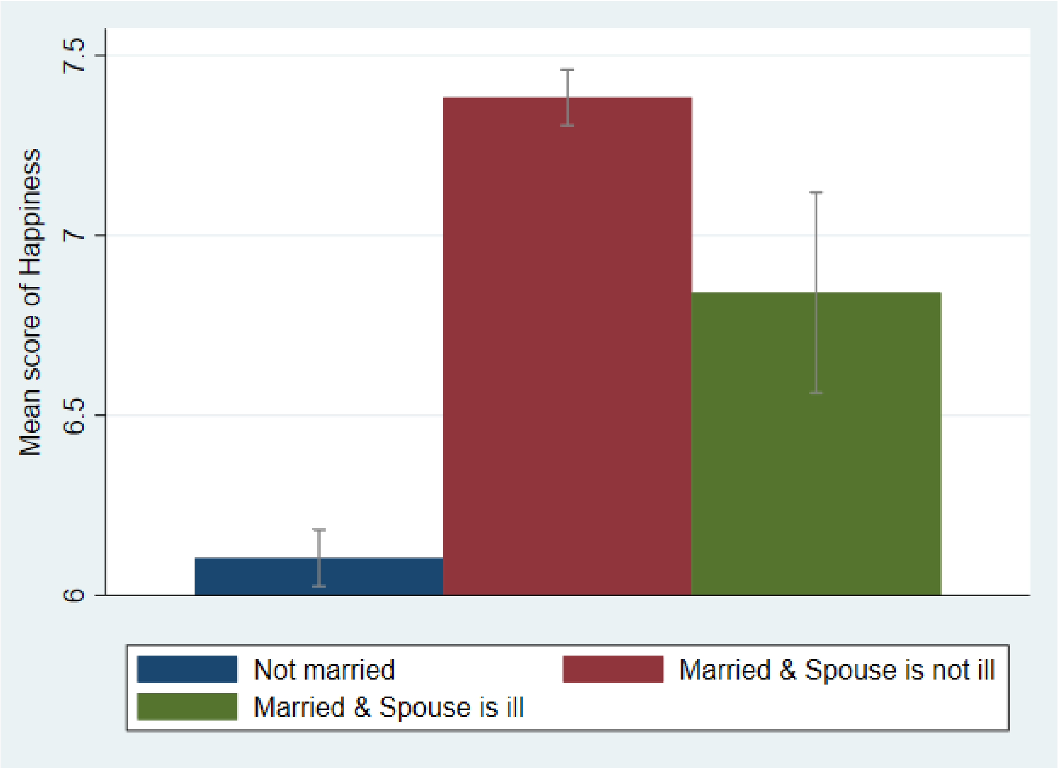
Notes: N=4,568. 90 % confidence intervals presented.

Figure A3. Gender Difference of the Impact of the Death of the Spouse on Happiness



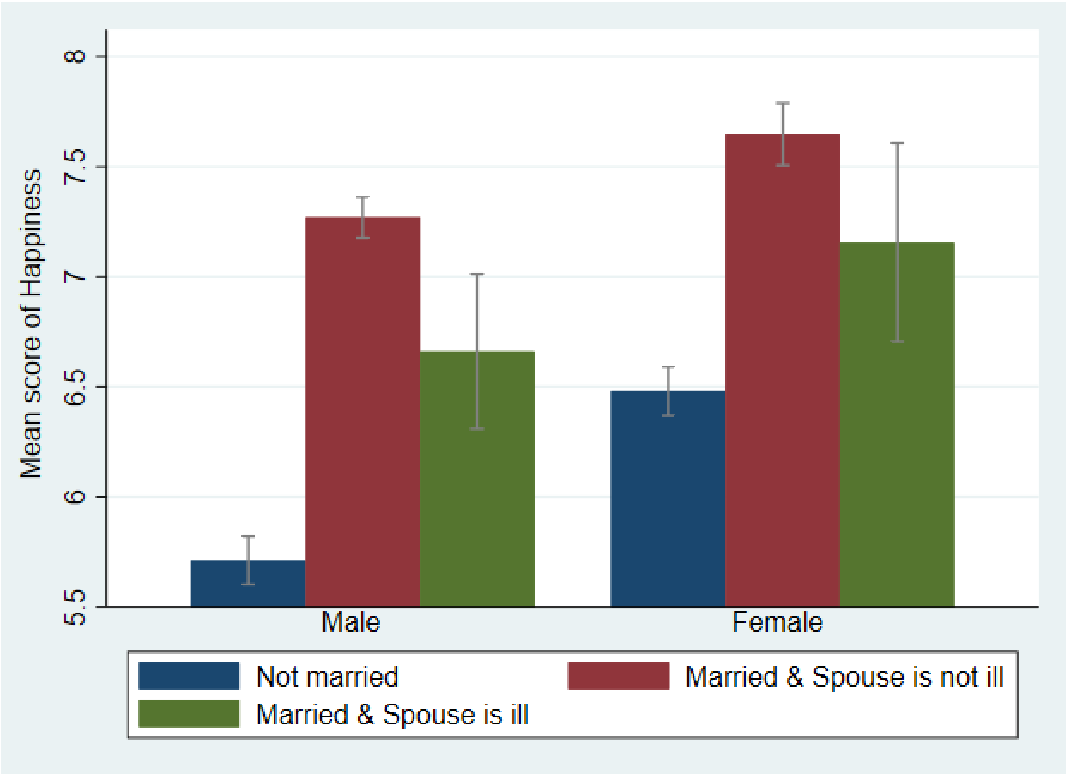
Notes: N=4,568. 90 % confidence intervals presented.

Figure A4. Impact of the Illness of the Spouse on Happiness



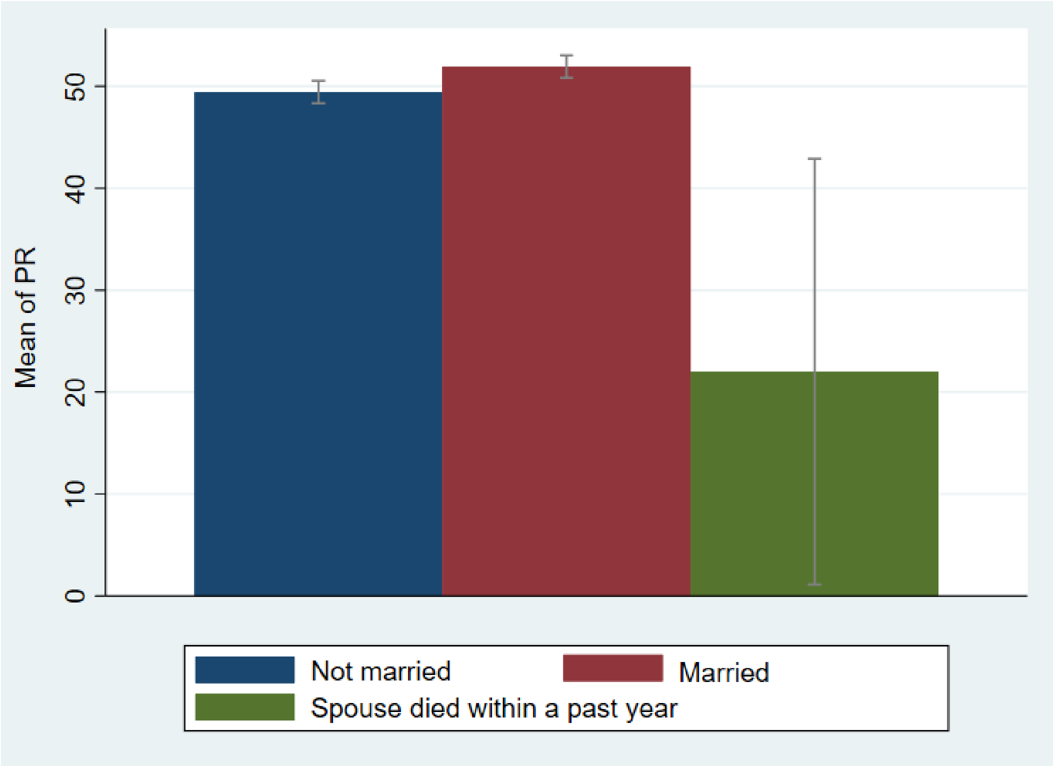
Notes: N=4,568. 90 % confidence intervals presented.

Figure A5. Gender Difference of the Impact of the Illness of the Spouse on Happiness



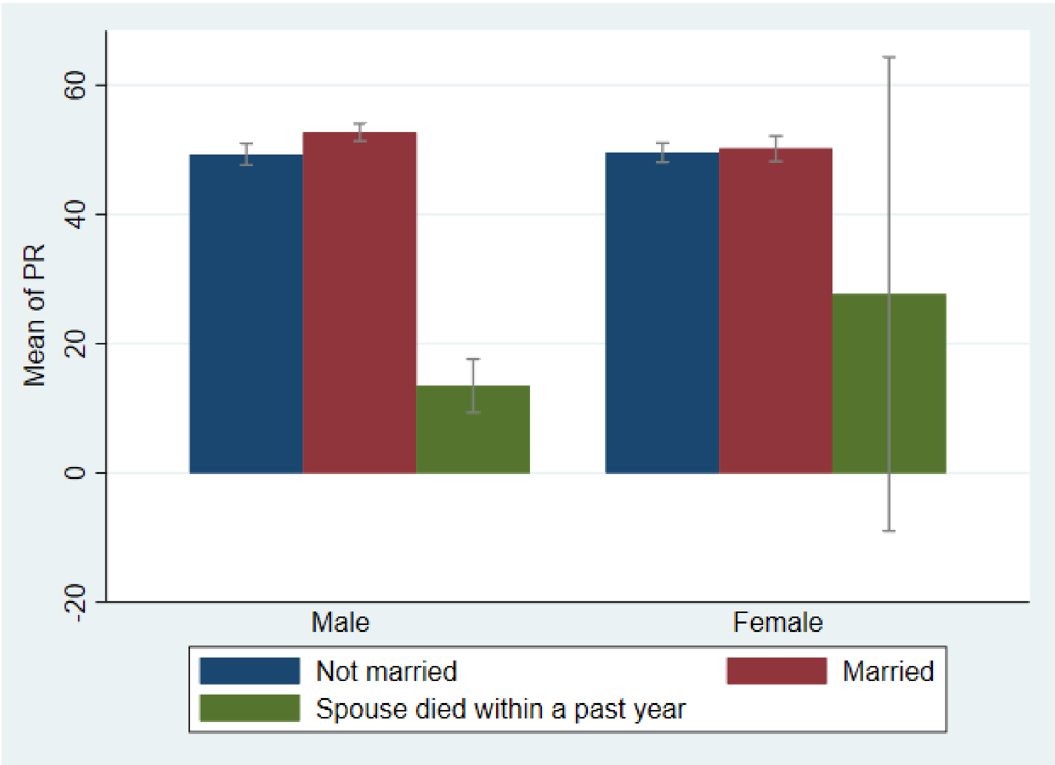
Notes: N=4,568. 90 % confidence intervals presented.

Figure A6. The Impact of the Death of the Spouse on Productivity (*PR*)



Notes: N=4,568. 90 % confidence intervals presented.

Figure A7. Gender Difference of the Impact of the Death of the Spouse on Productivity (*PR*)

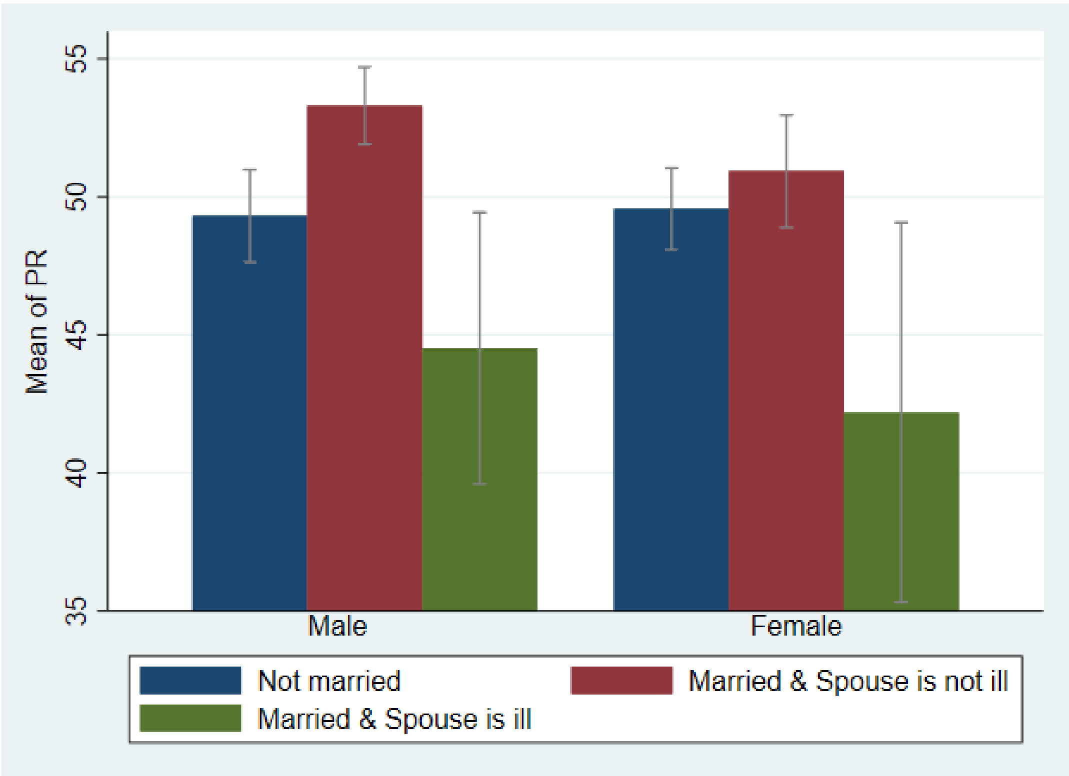


Notes: N=4,568. 90 % confidence intervals presented.

Figure A8. The Impact of the Illness of the Spouse on Productivity (*PR*)

Notes: N=4,568. 90 % confidence intervals presented.

Figure A9. Gender Difference of the Impact of the Illness of the Spouse on Productivity (*PR*)



Notes: N=4,568. 90 % confidence intervals presented.

Table A6. Impact of *Spouse death* & *Spouse ill* on Productivity (*PR*)

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	*Happiness*	*Happiness*	*PR*	*PR*	*PR*
Model	OLS	OLS	OLS	OLS	2SLS
Spouse death	-3.216***	-3.089***	-33.59***	-20.29***	
	(0.421)	(0.648)	(3.980)	(4.397)	
Spouse death ×*Female dummy*	4.206***	4.227***	14.11	6.738	
	(0.705)	(0.720)	(17.78)	(19.64)	
Spouse ill	-0.571***	-0.551***	-8.587***	-5.969**	
	(0.174)	(0.162)	(2.519)	(2.407)	
Happiness					8.625**
					(3.403)
Female dummy	0.601***	0.500***	-0.879	1.223	-3.149+
	(0.0684)	(0.0717)	(0.990)	(1.003)	(2.054)
Married	1.410***	1.123***	2.990***	1.179	-8.596**
	(0.0679)	(0.0684)	(1.003)	(0.958)	(3.849)
Numeracy		0.0492*		7.969***	7.552***
		(0.0263)		(0.366)	(0.464)
Age		0.00327		-0.258***	-0.288***
		(0.00293)		(0.0395)	(0.0485)
N	4,568	4,568	4,568	4,568	4,568
adj. R-sq	0.089	0.211	0.004	0.197	-0.082
Hansen J stat					1.362
(P-value)					(0.5062)
KPW F statistic					15.822
(Maximal IV size)					(<15%)

Notes: Robust standard errors in parenthesis. The constant term is not presented. Other omitted variables from column (2), (4) and (5) are Subjective health, Income dummies, Language grade dummies and Math grade dummies. Those coefficients are not reported in the table but are available from the corresponding author upon request. Column (2) is the first stage estimation results of column (5).
+Significant at the 15% level * Significant at the 10% level ** Significant at the 5% level
*** Significant at the 1% level