博士論文

Essays on Empirical Analysis of Family Policy

（家族政策に関する実証研究）

鳥谷部貴大

Acknowledgements

I am deeply grateful to my supervisor, Professor Daiji Kawaguchi, for his generous support and advices as well as his enthusiasm for teaching. His insightful lectures on labor economics attracted me to empirical research, which eventually led me to this dissertation. I am also grateful to Daisuke Adachi, Carlos Carrillo-Tudela, Yuji Genda, Maria Guadalupe, Hiromi Hara, Hidehiko Ichimura, Keisuke Kawata, Katsuhiro Komatsu, Mizuki Komura, Ayako Kondo, Rasmus Lenz, Philip MacLellan, Shiko Maruyama, Masato Oikawa, Fumio Otake, Mai Seki, Michael Waldman, Masayuki Yagasaki and conference participants at Australian National University, the Asian and Australasian Society of Labor Economics and University of Technology, Sydney, Chuo University, Hitotsubashi University, the Japan Economics Association Meeting, the Kansai labor workshop, Musashi University, National Taiwan University, Osaka University, the Research Institute f Economy, Trade and Industry, Seoul National University, Sophia University, Tohoku University, and the University of Tokyo for their valuable comments. I would like to sincerely thank Taiyo Fukai for his support and encouragement as well as his brilliant comments. Of course, none of them is responsible for any remaining errors. All errors are my own.

I thank the Keio University Panel Data Research Center for permission to use the Keio Household Panel Survey, the Ministry of Internal Affairs and Communications for permission to use the microdata of Japanese Census, the Ministry of Health, Labour and Welfare for permission to use the Vital Statistics and for providing information on parental-leave takeup, and GESIS–Leibniz Institute for the Social Sciences for permission to use German scientific-use file of Programme for the International Assessment of Adult Competencies. This work is supported by the Research Institute of Economy, Trade and Industry and Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research Grant Numbers 15H05692, 16H03630, 16H06322, 16J08836 and 18J21486.

Finally, I would like to express the deepest appreciation to my family, for they created the best environment to concentrate on study. This dissertation would have been impossible, without their invaluable encouragement and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Developed countries have experienced low fertility rate and population aging for last decades. In this situation, female labor force is one of the most essential factors to sustain social welfare system and meet labor demand of firms. Today, women have similar or even higher college enrollment rate than men in many developed countries, but still they tend to get out of labor market at the time of marriage or childbirth, presumably due to family responsibility. From this view point, family policies such as parental leave and public childcare provision could help women balance work and family responsibility and expand their career opportunities after marriage or childbirth.

This dissertation investigates the causal relationship between family policies and household decision making, with particular focus on female labor supply, which consists of three studies: Chapter 2 serves to identify the mode of family decision making to understand the role of bargaining power. This is important to consider the impact of family policies, because policies can affect the bargaining power among family members, which potentially leads to unintended side effects. Chapters 3 and 4 analyze the impact of parental leave policy on female labor market outcomes, as parental leave policy is a popular policy instrument to encourage maternal career advancement.

Chapter 2 investigates a household's commitment to a resource allocation by exploiting a 2007 Japanese pension reform allowing divorced women to claim a portion of their husband's pension benefits while keeping the household's total benefits unchanged, without imposing functional form assumptions on preferences or technology. Although the reform should not affect a couple's decision-making under full commitment, we find that it increased wives' leisure and decreased their market and domestic work. This suggests that wives were able to increase their welfare by exploiting an improved outside option, and thus commitment to resource allocation is less than complete.

In Chapter 3, drawing on the micro data of Japanese population census, we evaluated the long-run impact of taking parental leave on maternal employment. To this end, we used the parental leave reforms in 1992 and 1995 as natural experiments, which introduces job-protected leave and cash benefits during the leave, respectively. We found that the job protection and cash benefits both increased full-time employment while decreasing part-time employment in long-run. Rescaling these policy impacts by the take-up rate of the parental leave, we found that the magnitude of the effect of taking the leave was more than 30 percentage points. Since the increase in the full-time employment was offset by the decrease in the part-time employment, the employment rate was unaffected. Therefore, the parental leave policy strengthened labor market attachment of mothers, allowing those who would otherwise engage in part-time jobs after childbirth to continue full-time jobs.

In Chapter 4, we analyze the effect of parental-leave policies on the gender gap in career advancement. We measure career advancement by using the newly developed index based on skill use questions in the Programme for the International Assessment of Adult Competencies and demonstrate that this index captures career advancement more accurately and succinctly than traditional labor-market outcomes. A cross-country comparison shows that a generous parental-leave policy narrows the gender gap in career advancement among lower-skilled but widens the gap among moderately higher-skilled workers. This finding is robust after controlling for international differences in other family policies, gender norms, and labor-market institutions. We assess several hypotheses to explain the heterogeneous policy impacts on career advancement in light of the obtained findings.[1]

[1]Chapter 3 is a joint work with Daiji Kawaguchi at the University of Tokyo, and Chapter 4 is a joint work with Taiyo Fukai at the Economic and Social Research Institute. I deeply appreciate their permission to use these works as a part of my dissertation.

Chapter 2

Empowerment effects and intertemporal commitment of married couples: Evidence from Japanese pension reform

2.1 Introduction

Economists have long been interested in the extensive economic gains from marriage, including the joint use of public goods, specialization and self-insurance of household members. The self-insurance is especially important for the personal security of household members, as the effects of business cycles can be severe when job displacement risks are not completely insured (Krebs, 2007). While these risks can be mitigated by conventional unemployment insurance, they can also be hedged by income pooling through marriage, although inter- and intra-family insurance would not completely absorb income shocks (Altonji et al., 1992; Hayashi et al., 1996). In addition, marriage also allows an individual to insure other risks such as longevity risks through the heritage of his/her partner (Kotlikoff and Spivak, 1981).

In order to realize these economic gains, however, marriage partners must be able to commit to their initial resource allocation plan, and the traditional unitary model of household decision making implicitly requires this by assuming the household is a single decision maker. However, when marriage partners have distinct preferences, this introduces a potential conflict in their incentives that can lead to a "hold-up" problem whereby one partner might deviate from the initial allocation plan by exploiting a changed situation to improve his or her individual welfare. Should this occur, a couple cannot fully enjoy the fruits of marriage, as Dufwenberg (2002) theoretically finds that a couple's failure to commit to a resource plan leads to a failure to specialize and an under-accumulation of human capital. Further, Voena (2015) finds that divorce and property division laws can affect both asset allocation and female labor force participation. From these studies, we see that a couple's lack of commitment can make it difficult for a household to achieve the first-best resource allocation, or *ex ante* efficiency. Economic gains from family formation thus depend on the degree of commitment.

Since the seminal work of Chiappori (1988), research in household decision making has shifted from the unitary model to the collective one in which household members have individual preferences and the resource allocation of the household is obtained through bargaining.[1] However, studies of bargaining in family decision-making have typically adopted a static collective framework that is silent about dynamic issues such as divorce or changes in the viability of an outside option for each household member. In adapting the collective model to a dynamic setting, Mazzocco (2007) argues that differing preferences among household members could result in a lack of commitment because the partners have an incentive to deviate from the initial plan as uncertainty is resolved. Since bargaining position is fixed under full commitment but is allowed to fluctuate under limited commitment so that household members may attempt to improve their situation when outside options change, the degree of commitment is testable by examining within-household variation in bargaining position.

These empirical tests of the degree of commitment are of interest not only in relation to the economic benefits of marriage, but also for their practical implications in the specification of life-cycle models. While the limited commitment model is more general than the full commitment model, its complexity leads to a computational burden in obtaining a solution (Chiappori and Mazzocco, 2017). Consequently, studies typically attempt to limit this burden through simplifying assumptions such as full commitment (Casanova, 2010), a single agent (Adda et al., 2017), or functional form impositions on the Pareto weight (van der Klaauw and Wolpin, 2008).[2] While full commitment may be a good approximation of reality if any resource allocation distortions associated with hold-up problems are relatively minor, researchers need to consider the extent to which the full commitment assumption drives their results. If distortions exist, the model should allow for the possibility of incomplete commitment.

Another reason why limited commitment should be considered is that full commitment has not been supported by some recent empirical studies. For example, in their studies of dynamic models nesting the full-commitment case, Mazzocco (2007) and Lise and Yamada (2018) both find the evidence inconsistent with full commitment, with Lise and Yamada (2018) finding that household members do update their bargaining position, but only when the participation condition for marriage is binding. In another study, Blau and Goodstein (2016) test degree of commitment by examining whether an unexpected inheritance affects the relative bargaining position and household behavior, controlling for household budget and any expected inheritance.

[1] Apps and Rees (1997), Browning and Chiappori (1998), Chiappori (1992), Chiappori (1997), and Blundell et al. (2005) developed the collective model, while others such as Angrist (2002), Attanasio and Lechene (2014), Aura (2005), Cherchye et al. (2012), Duflo (2003), Francis (2011), Ponczek (2011) have tested and rejected the unitary model. Aronsson et al. (2001) is one of the few studies that has not rejected the unitary model.

[2] Studies that do allow for incomplete commitment include Voena (2015) and Low et al. (2018)

Since a couple under full commitment fixes its bargaining position at marriage, it should not be affected by any unexpected shock when preferences and budget constraint are held constant. However, they find mixed results, with full commitment rejected in some specifications but not others.[3] Notwithstanding the mixed results, an advantage of the Blau and Goodstein (2016) approach, compared to the above two studies, is the identification strategy not relying on functional form assumption on preferences or household technology. For example, Mazzocco (2007) assumes the bargaining position to be independent of the level of assets, and Lise and Yamada (2018) assume an interior solution for time-allocation to estimate their structural model. In contrast, if unexpected shocks only affect bargaining position, it is not necessary to specify the forms of the household objective function or home production technology.

This study contributes to the literature on household decision-making by testing the degree of commitment without relying on any *a priori* assumptions about functional form and by using a Japanese dataset that allows us to exploit a major pension reform as a natural experiment using a difference-in-differences (DD) estimation methodology. The Japanese pension reform of 2007 allowed a couple to divide their pension benefits upon divorce. Prior to the reform, since the bulk of public pension benefits in Japan are proportional to labor earnings, a dependent spouse specializing in home production would have found it difficult to live only on her own pension benefits after divorce. The reform addressed this issue by allowing the spouse with fewer pension records to claim a portion of the partner's pension records tracked during the marital period. For our purposes, several features of the reform are beneficial for testing the degree of commitment of a household. Firstly, the reform allows for sharing of pension benefits while keeping the total amount of the benefits unaffected. The unexpected shock thus does not change the household budget, which facilitates our identification of the degree of commitment. Secondly, the pension division applies only to the public pension insurance that covers permanent employees, which leaves a dependent wife of a self-employed spouse unaffected. Thirdly, the maximum share of pension benefits after division is 50 percent, which means that if both spouses are permanently employed and of a similar age, there is very little room for pension balancing. This allows us to use these households as a control group for the counterfactual inference required for our difference-in-differences (DD) estimation strategy.

Following the framework of Mazzocco (2007), we first constructed a dynamic collective model to investigate the effect of the pension reform under both full and limited commitment, finding that only couples without full commitment were affected. We tested several of the model's predictions. First, as the reform is expected to mostly affect couples in which the husband's pension benefits are large, and since most young

[3]These mixed results are attributed to a lack of statistical power, as inheritance is a rare event. Because inheritance amount has been shown to suffer from serious measurement error (Laitner and Sonnega, 2010), the receipt of an inheritance is used instead, which lowers the statistical power. Additionally, as the study examines several full commitment null hypotheses one-by-one under various specifications, the statistical inference seems difficult to interpret.

couples have not yet accumulated substantial benefits, the reform is most likely to affect elderly couples.[4] Second, low net worth households are more likely to respond to the reform because pension benefits comprise only a small share of the total assets of high net worth households that would be divided upon divorce. Third, as noted by Chiappori and Mazzocco (2017), the dynamic collective model implies that household members re-bargain their resource allocation only when one member's participation condition is binding. We tested this prediction using young couples, whose current period participation condition is unlikely to bind because they cannot receive pension benefits for several decades.

We then conducted empirical analysis by using DD estimation with the *Keio Household Panel Survey* (KHPS), a household panel survey in Japan. The treatment group consists of households in which the husband was a permanent employee and the wife was not, the control group consists of all other households. The results of our DD estimation do not support the full-commitment model. Consistent with the model's prediction, we did not find any significant impact on young couples, but elderly wives aged 50–59 increased their leisure time by 5 hours per week (or 5 percent) by decreasing equally their market and domestic work. The elasticities of those outcomes to the life-time pension benefits received upon divorce are 0.05, 0.20 and 0.05, respectively. A subsample analysis of home ownership as a proxy for individual assets showed that those wives who were most affected by the reform were those who did not own a home, which is also consistent with the model's prediction. Finally, the model passed several tests of the common trend assumption of our identification strategy including a placebo test, a specification with a group-specific linear trend, and a triple-differences estimation using young households as an additional comparison group.

A limitation of the study is that the test of full commitment relies on the assumption that couples do not eventually divorce. This limitation is not unique to this study, however, and is typically imposed throughout the literature either explicitly or implicitly (Blau and Goodstein, 2016; Lise and Yamada, 2018; Mazzocco, 2007). When divorce is possible and seems on the horizon, spouses have an incentive to prepare for this future divorce (Mazzocco et al., 2006), which means that the household allocation plan can be contingent on the post-divorce economic situation even under full commitment. This concern would appear to be negligible for elderly couples in Japan, however, due to the low annual and life-time divorce rates of 0.3 and 3.0 percent, respectively, for Japanese wives aged 55. Therefore, we believe that this assumption cannot be the main driver of our findings.

Additionally, when we take the rejection of the full commitment model as a given, this study highlights the difficulty in making a commitment, as couples fail to achieve full commitment even within a stable marital relationship. This limited commitment implies that *ex ante* efficient allocation is not necessarily *ex post* efficient and, as a result of this inconsistency, a hold-up problem may occur. The resulting inefficiency

[4] In a study of the reform using relatively young couples of 48 years and younger, Sakamoto (2008) finds no policy impact.

could be more serious among young couples because their relatively higher divorce rates indicate that the participation condition is likely to bind. If this is the case, then household behavior including human capital accumulation and investment in children may be distorted by a fluctuation in the bargaining position, and this also makes risk-sharing among household members difficult.

2.2 Institutional background

2.2.1 Japanese pension system and the Reform of 2007

In order to better understand the context of the study, this section describes the Japanese public pension system and the reform of 2007. The Japanese public pension system consists of three insurance policies, the Employee pension (*kosei nenkin*), the Mutual Aid pension (*kyosai nenkin*) and the National pension (*kokumin nenkin*). The first two policies cover permanent employees (i.e., full-time workers not hired under a time-limited contract) in the private and public sectors, respectively,[5] but are otherwise identical, so we hereafter refer to them collectively as the Employee pension insurance. Within the Employee pension insurance, participants aged under 70 pay pension premiums as long as they earn labor income, with the amount of the premium proportional to their labor income. The age of eligibility for benefits is around 60, depending on sex and birth cohort, with the age of eligibility for men higher than women and for a recent cohort higher than an earlier cohort (for details, see Table 2.7). The pension benefits consist of two parts: a basic part and a proportional part, with the basic part depending only on the number of years for which the participant paid premiums, and the proportional part depending on earnings and duration of premium payments prior to retirement. The National pension covers those who are not covered by the Employee pension, which includes mainly the self-employed, part-time employees and dependent wives, all of whom pay premiums until they turn 60 and become eligible for benefits at age 65. The National pension is similar to the basic component of the Employee pension, except for the age of eligibility and the participants in each respective plan.

Before the pension reform of 2007, the Japanese pension system was thought to be inequitable because spouses were not allowed to claim any fraction of the pension benefits of their partners should they divorce. While a homemaker wife played an important role in enabling her husband to specialize in market work, she previously had no access to his pension benefits. As the average monthly National pension

[5] In addition to permanent employees, the pension plans cover part-time employees who work more than three-quarters of the hours per day and days per month worked by a full-time employee. For example, if a full-time employee works eight hours per day and twenty days per month, then a part-time employee working more than six hours per day and fifteen days per month participates in either the Employee or Mutual Aid pension.

benefit in 2007 was only about 540 U.S. dollars,[6] a homemaker would have difficulty living only on her own pension benefits after divorce.

In order to address this issue, a reform was approved in 2004 and enacted on April 1st, 2007 to permit divorcing spouses to divide the proportional part of the household's Employee pension records tracked during the marital period.[7] Although the proportion of the household pension record claimed by either party is determined by agreement between the spouses, the division must range between 0-50 percent, as the spouse with the smaller pension cannot claim more than half of the total and there is no room for division if the records of both spouses are equal. If the spouses fail to agree on division, a rate is provided by the courts. In 99 percent of cases in 2007, this rate was 50 percent[8], consistent with the asset division rule which divides assets accumulated during the marital period equally.[9] Since a divorced homemaker is now assured of at least some income beyond her basic pension, the post-divorce situation is said to have improved.

Several features of the 2007 pension reform help to facilitate our analysis. Firstly, as the reform does not change the total amount of pension benefits received by a household, it does not affect the household budget under marriage, which allows us to isolate changes in allocations due to bargaining. Secondly, as the reform provides a better outside option to dependent wives, we can investigate whether this change affects commitment. Thirdly, as the pension reform leaves some households unaffected, the policy design allows us to create treatment and control households for the counterfactual inferences required for our DD estimation. Since the pension reform applies only to Employee pension records, it does not affect the self-employed, who are covered by the National pension. Additionally, when both spouses are permanent employees, they each have their own Employee pension record and so the pension splitting opportunities are marginal. Hereafter, we refer to this latter type of household as a "dual-permanent" household.[10] As discussed below, households either "dual-permanent" or with one self-employed spouse were used to control for the economic trend in absence of the reform.

[6] The Employee pension paid out about 1,610 U.S. dollars per month on average. Source: Japanese Ministry of Health, Labour and Welfare: `http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000106808_1.html`

[7] The basic part of the pension is not based on income and so is not divisible.

[8] Source: Supreme Court of Japan; `http://www.courts.go.jp/vcms_lf/20513001.pdf`.

[9] A second pension reform was implemented in 2008 that allows a dependent wife (or husband) to claim half of the total household Employee pension records tracked after May 1st, 2008 should they divorce. This reform, which applies to households in which the wife's annual income is less than about 13,000 U.S. dollars, has a fixed division ratio of 50 percent. However, considering that the first pension reform applies to the entire Employee pension records tracked during the marital period, the function of this second reform is, at most, supplementary, as it does not apply to pension records before May 1st, 2008. As its impact on household behavior is likely to be negligible, our analysis focused only on the impact of the first reform.

[10] Explicitly excluded are those households in which the husband is a permanent employee and the wife works part-time, as part-time employees are typically not covered by the employee pension. Even if a wife working part-time is covered by the employee pension, however, the amount of her pension records is likely to be much less than that of a husband who is a permanent employee.

2.2.2 Divorce law in Japan

Although Japanese divorce law in principle requires mutual consent before a couple may divorce, enforcement is not strict and so in practice it may operate similar to a unilateral divorce. Under Article 770 of the Japanese Civil Code, a judicial divorce is permitted if a spouse: (i) has committed an act of unchastity; (ii) has been abandoned in bad faith; (iii) is of unknown whereabouts and for at least three years it has not been clear whether s/he is dead or alive; (iv) is suffering from severe mental illness and there is no prospect of recovery; or (v) has any other grave concern making it difficult to continue the marriage. While the fifth point is ambiguous, the Ministry of Justice has issued a guideline that allows a couple to divorce after five years of separation. As a result, a spouse who wishes to divorce may simply end cohabitation with his/her partner and file a suit for divorce after the required time has elapsed.

As for the distribution of assets upon divorce, the property division rule in Japan assures divorcees equal division of those assets for which they both "contributed" in obtaining, but it is not necessary for this contribution to be monetary, whereas other assets are divided on basis of the title. For example, if a wife specializes in home production and has no earnings, her domestic work is regarded as a contribution to purchasing housing. On the other hand, the asset division does not apply to assets accumulated before marriage or obtained via inheritance. Since asset division is implemented based on holdings at the time of divorce or at the end of cohabitation if a couple separates prior to divorce, separation can be an effective strategy to divorce without mutual consent, as a spouse can file for divorce several years after the end of cohabitation but asset division is implemented as if the couple had divorced at the beginning of the separation.

2.3 Model

In this section, we provide a model to describe how a couple responds to the pension reform under full and limited commitment. Consider the following 3-period problem for a 2-member household ($j = 1, 2$) in which the couple marries in the 1st period and may or may not choose to divorce in the 2nd or 3rd period. The household members supply market and domestic labor in the 1st and 2nd periods and retire in the 3rd period. A fraction τ of labor earnings is collected as a pension premium, and the couple receives the pension benefits b_{j3} in the 3rd period. Each spouse derives his or her welfare from private consumption c_{jt} and leisure l_{jt}, where the consumption good is produced by domestic labor h_{jt} and there is a market good g_t. We assume unilateral divorce, so divorce is possible without mutual consent. We denote member j's asset and pension benefits upon divorce as a^D_{jt} and b^D_{jt}, respectively.

We now define full and limited commitment. A couple achieves "full commitment" if the *participation condition* that the value of the marriage for each member is greater than or equal to the value of divorce is required only at the initial period, whereas the degree of commitment is said to be "limited" if the participation condition must be satisfied at each period. Thus, under limited commitment, both members must be satisfied with the marriage at each period; otherwise, an unsatisified member may threaten to divorce to increase his/her consumption or leisure, even if that member does not intend to carry through with the divorce. Here, we assume that the couple does not ever divorce. The case of actual divorce is discussed in Section 2.7.

The household problem under full commitment is

$$\max_{(c_{jt},l_{jt},b_{jt},g_t,a_{t+1})} \mu E_1\left[\sum_{t=1}^{3}\beta^{t-1}u_1(c_{1t},l_{1t},\theta_{1t})\right] + (1-\mu)E_1\left[\sum_{t=1}^{3}\beta^{t-1}u_2(c_{2t},l_{2t},\theta_{2t})\right], \tag{2.1}$$

$$\text{s.t. } c_{1t}+c_{2t} = F(h_{1t},h_{2t},g_t), \tag{2.2}$$

$$a_{t+1} = (1+r_t)a_t + \sum_{j=1}^{2}(1-\tau_{1t})(1-l_{jt}-h_{jt})w_{jt} - g_t \quad (t=1,2), \tag{2.3}$$

$$b_{j,\,t+1} = (1+r_t)b_{jt} + \tau_{1t}(1-l_{jt}-h_{jt})w_{jt}, \tag{2.4}$$

$$0 \le l_{jt}+h_{jt} \le 1, \quad l_{j3}+h_{j3} = 1, \tag{2.5}$$

where θ_{jt} is match-specific utility, w_{jt} is j's market wage, r_t is the risk-free interest rate,[11] and the total time is normalized to one. Since the couple marries in the 1st period, the initial-period participation condition is summarized in the Pareto weight μ. Solving this problem backwardly, we observe that the household allocation plan is contingent on total assets $a_t + b_{1t} + b_{2t}$ but not on the composition of those assets (a_t, b_{1t}, b_{2t}). Hence, the solution to the problem is

$$x_{jt}(\Omega_t) = \tilde{x}_{jt}(a_t+b_{1t}+b_{2t},\theta_{1t},\theta_{2t},w_{1t},w_{2t}) \quad (x \in \{c,l,h\}; j \in \{1,2\}), \tag{2.6}$$

where $\Omega_t = (a_t, a^D_{1t}, a^D_{2t}, b_{1t}, b_{2t}, b^D_{1t}, b^D_{2t}, \theta_{1t}, \theta_{2t}, w_{1t}, w_{2t})$ is the set of state variables.

The household problem under limited commitment is similar to the full commitment household problem (2.1) through (2.5) above, but additionally requires the participation conditions in periods 2 and 3:

$$u_j(c_{j2},l_{j2},\theta_{j2}) + \beta u_j(c_{j3},l_{j3},\theta_{j3}) \ge V^D_{j2}\left(a^D_{j2}+b^D_{j2},w_{j2}\right), \tag{2.7}$$

$$u_j(c_{j3},l_{j3},\theta_{j3}) \ge V^D_{j3}\left(a^D_{j3}+b^D_{j3}\right) \quad (j \in \{1,2\}), \tag{2.8}$$

[11] Although, for simplicity, the amount of pension benefits is assumed to be accumulated in the same way as savings, the model identification result does not rely on this assumption.

where $b^D_{jt} = b_{jt}$ under the pre-reform regime and $b^D_{jt} = \frac{b_{1t}+b_{2t}}{2}$ under the post-reform regime. The 3rd-period problem is

$$\begin{aligned} V_3(\Omega_3) = \max\ & \mu u_1(c_{13}, l_{13}, \theta_{13}) + (1-\mu) u_2(c_{23}, l_{23}, \theta_{23}), \\ \text{s.t. } & c_{13} + c_{23} = F(h_{13}, h_{23}, g_3), \\ & g_3 = a_3 + b_{13} + b_{23}; \quad l_{j3} + h_{j3} = 1 \quad (j \in \{1,2\}), \\ & u_j(c_{j3}, l_{j3}, \theta_{j3}) \geq V^D_{j3}\left(a^D_{j3} + b^D_{j3}\right) \quad (j \in \{1,2\}). \end{aligned}$$

Denoting the Lagrange multiplier on the participation conditions by λ_j (≥ 0), this problem is rewritten as

$$\begin{aligned} & \max\ (\mu + \lambda_1) u_1(c_{13}, l_{13}, \theta_{13}) + (1 - \mu + \lambda_2) u_2(c_{23}, l_{23}, \theta_{23}), \\ & \text{s.t. } c_{13} + c_{23} = F(h_{13}, h_{23}, g_3), \\ & \quad g_3 = a_3 + b_{13} + b_{23}; \quad l_{j3} + h_{j3} = 1 \quad (j \in \{1,2\}), \end{aligned}$$

where the $\lambda_j = 0$ if member j's participation condition is not binding. Thus, the couple re-bargains only when the participation condition of one member binds.

The main difference from the full commitment case is that the pension division rule upon divorce, b^D_{j3}, matters and this in turn implies that so does asset composition (b_{1t}, b_{2t}). Indeed, we observe in general that

$$\lambda_j = \lambda_j(a^D_{13}, a^D_{23}, b_{13}, b_{23}, \theta_{13}, \theta_{23}) \neq \lambda_j(a^D_{13}, a^D_{23}, b_{13} + b_{23}, \theta_{13}, \theta_{23}).$$

Since a similar argument applies to the 1st- and 2nd-period problems, the household allocation plan cannot be written as a function of total assets:

$$x_{jt}(\Omega_t) \neq \tilde{x}_{j,t}(a_t + b_{1t} + b_{2t}, \theta_{1t}, \theta_{2t}, w_{1t}, w_{2t}) \quad (x \in \{c, l, h\}; j \in \{1,2\}). \tag{2.9}$$

In order to consider the impact of the pension reform on the married couple, suppose that the reform is unexpectedly implemented at the beginning of the 2nd period. Under limited commitment, the initial allocation plan satisfies

$$\begin{aligned} & u_j(c_{j2}, l_{j2}, \theta_{j2}) + \beta u_j(c_{j3}, l_{j3}, \theta_{j3}) \geq V^D_{j2}\left(a^D_{j2} + b_{j2}, w_{j2}\right), \\ & u_j(c_{j3}, l_{j3}, \theta_{j3}) \geq V^D_{j3}\left(a^D_{j3} + b_{j3}\right) \quad (j \in \{1,2\}), \end{aligned}$$

but due to the change in pension division rule, the new participation conditions are

$$\begin{aligned} & u_j(c_{j2}, l_{j2}, \theta_{j2}) + \beta u_j(c_{j3}, l_{j3}, \theta_{j3}) \geq V^D_{j2}\left(a^D_{j2} + \frac{b_{12} + b_{22}}{2}, w_{j2}\right), \\ & u_j(c_{j3}, l_{j3}, \theta_{j,3}) \geq V^D_{j3}\left(a^D_{j3} + \frac{b_{13} + b_{23}}{2}\right) \quad (j \in \{1,2\}), \end{aligned}$$

and the 2nd-period participation condition is now more restrictive for member 2 if $b_{12} > b_{22}$.

Taking this into consideration, other things being equal, the greater is $(b_{12} - b_{22})$, the more likely the participation condition for member 2 will bind. Since the pension division upon divorce applies only to the pension records during the marital period, a young couple does not have many divisible pension benefits (i.e. b_{12} is small), suggesting that elderly households are more likely to be affected by the reform. On the other hand, suppose that the initial allocation plan satisfies the 2nd period participation condition but does not satisfy the 3rd period one. In this case, the main impact is on the 3rd period allocation because the 3rd period participation constraint cannot hold without adjusting the allocation at that period. The 2nd period allocation is then affected due to consumption smoothing.

The above discussion implies that the main impact of the pension reform should be observed after the reform is implemented, and young households are unlikely to show substantial change in allocation in the reform year even if they re-bargain their future resource allocation. Furthermore, given the concavity of the value function, the value of divorce is sensitive to the amount of pension benefits when the amount of other assets available upon divorce is small. This suggests that a couple with few assets other than pension benefits will be most affected by the reform.

2.4 Data and identification strategy

2.4.1 Data

This study utilized data from the *Keio Household Panel Survey* (KHPS) provided by the Keio University Panel Data Research Center. The KHPS is an annual household-level panel survey of households beginning in 2004 and consisting of 4,000 households (3,000 married and 1,000 single). Each year, the survey is conducted at the end of January, and respondents are asked about their usual time allocation as well as background information such as age, sex, family composition and employment status. As Japan's fiscal year begins in April, each implementation of the survey inquires about the previous fiscal year, with KHPS 2004, for example, asking about respondent behavior in 2003. Accordingly, KHPS 2004 includes data on the socio-economic status of respondents before the 2004 pension reform approval, and KHPS 2008 and succeeding waves represent household behavior after the enactment of the reform.

Although ten waves (KHPS 2004 through 2013) were available, we restricted our main sample to KHPS 2005 through 2008. As KHPS 2004 does not include information about time spent on childcare, the domestic labor supply in this wave is inconsistent with other waves, so we used this first wave only to obtain background information. We also excluded KHPS 2009 through 2013 from our sample because of two external events: the global financial crisis of 2008, the impact of which seems difficult to distinguish from that of the 2007 pension reform, and the Great East Japan Earthquake of 2011, which very likely caused heterogeneous effects across households. The set of households chosen for the analysis sample was selected according to the following criteria: (1) spouses who married before 2004[12], who live together, and who were both aged 30–59 in 2007; (2) a time allocation that meets the constraint of 168 hours; and (3) the lack of any missing key variables for our analysis. In empirical studies, we used family size, the number of children and a indicator variable for children aged less than or equal to six to control for household heterogeneity. The variables of interest are time allocation, leisure, market labor supply and domestic labor supply[13], which were measured as the average hours per week.

DD estimation, which separates the impact of a specific policy from the counterfactual time trend that would have been faced by the treatment group had there not been any treatment, requires that the analysis sample of households be divided into a treatment group and a control group. We defined the treatment group as households in which the husband was permanently employed but the wife was not because those households would have been most affected by the pension reform. The control group consisted of the remaining households; that is, "dual-permanent" households and households in which the husband was self-employed.[14] The "dual-permanent" households consist of households in which both members are permanent employees, so households in which the wife works for part-time jobs are excluded. The treatment status of each household was fixed throughout the sample period by using the employment status in 2003, one year before the approval of the reform.

2.4.2 Sample and Treatment Group Validity Checks

From the sample descriptive statistics shown in Table 2.1, we can see that there are some differences between the treatment and control households, with the average treatment household slightly younger and having more children than the control household. Further, while leisure time is similar, wives in treatment households work less in the market and longer in domestic production than those in control households, while husbands in treatment households work longer in the market and less in domestic production.[15] Although these differences may signal potential heterogeneity between the two groups, our identification strategy is robust to this heterogeneity, for DD estimation with household fixed effects controls for heterogeneity in preferences as long as the differences demonstrate the same trend over time. In addition, we also implemented a triple differences (DDD) estimation (discussed in Section 2.6.4) that is robust to a violation of this common trend assumption.

[12]We did not include newly married couples in our sample because in 2004, these couples may have known about the pension reform at the time of their marriage and so within-household variation in bargaining position is not identifiable for those households.

[13]Domestic labor supply consisted mainly of meal preparation, laundry, grocery shopping, cleaning and childcare.

[14]Households in which the wife was a permanent employee and the husband was not would also have been affected by the pension reform, but as the impact would be in the opposite direction of that of our treatment group, and since the size of this subset was very small, we included this type of household in the control group.

[15]See Table 2.8 for descriptive statistics of time-allocation before and after the reform.

Before implementing our difference-in-differences (DD) estimation strategy, we confirmed graphically the common trend assumption required to identify the average treatment effect on treated (ATT). Figure 2.1 shows trends in the time allocation of each spouse and highlights several points relevant to our analysis. First, as is also seen in the descriptive statistics, the amount of time allocated to various activities differs between the treatment and control groups. In particular, there are relatively large differences between the two groups in the market and domestic labor supplies of wives, with the typical wife in the treatment group working more hours in the household and fewer in the market than a typical wife in the control group. With control variables and fixed effects, DD estimation can account for these differences in the absolute levels of time allocation as long as the common trend assumption holds. Second, and more importantly, the treatment and control groups show a similar trend before the pension reform, which supports the common trend assumption required for our DD estimation.

While it is difficult to see any large impact of the pension reform from Figure 2.1, the policy impact becomes clear when heterogeneity in the age of the wife is considered. In particular, the impact of the reform on a wife's allocation to leisure is striking (Figure 2.2). In the most elderly group (wives aged 50-59), the time allocated to leisure increases by about 5 hours in the treatment group, while showing a pre-reform trend similar to the control group (Figure 2.2a). In contrast, in the other age groups, the treatment and control groups both show a similar time trend before and after the pension reform (Figures 2.2b and 2.2c). These observations support our contention that the pension reform appears to have had a substantial impact on elderly households but a limited impact on younger households.

2.4.3 Regression Framework for Difference-in-Differences Analysis

We applied a difference-in-differences (DD) estimation strategy to investigate the causal impact of the 2007 Japanese pension reform. The treatment group consisted of households in which the husband was a permanent employee in 2003 and the wife was not, whereas the control group consisted of all other households. The estimation equation was specified as

$$y_{it}^j = \delta_1^j After_t \cdot Treatment_i + x_{it}'\beta^j + \sum_{t=2005}^{2007} \gamma_t^j d_t + c_i^j + u_{it}^j, \tag{2.10}$$

where i, j and t denote each household, each household member and the year, respectively. Dummy variables include $After_t$, which takes one if $t = 2007$, $Treatment_i$, which takes one if household i is in the treatment group, and d_t, which is a year dummy variable. The control variables x_{it} are a constant and household characteristics, which include the squared ages of both spouses, family size, the number of children and a dummy variable indicating whether the household has children aged

6 or younger.[16] c_i^j is household fixed effects. Note that $After_t$ is equivalent to d_{2007}, and household fixed effects absorb the treatment dummy variable, $Treatment_i$, which was fixed according to employment status in 2003. The dependent variable, y_{it}^j, is the allocation to leisure, market labor supply, and domestic labor supply of spouse j of household i in year t.

The coefficient of interest is δ_1^j, which represents the household time allocation response to the pension reform and is key to test the degree of commitment. Under full commitment, a couple's consumption plan is contingent on its total assets and, given that amount, asset composition is irrelevant. As the couple does not respond to the reform, δ_1^j is expected to be zero. Under limited commitment, however, the couple's consumption plan is contingent on each member's asset share upon divorce, so δ_1^j is not equal to zero, assuming that the participation conditions are violated in some households due to the reform. Hence, if δ_1^j is different from zero, we reject the full commitment model. Since the theoretical model suggests that the pension reform will have heterogeneous effects according to the wife's age, we therefore estimated equation (2.10) by dividing the sample into three groups: an elderly group of households with wives aged 50–59 in 2007, a middle-aged group with wives aged 40–49, and a young group with wives aged 30–39.[17]

By estimating by the age of the wife, it allows us to investigate how spouses might update their bargaining positions. For example, if we find no policy impact on the time allocation of younger households, this suggests that those households display some degree of commitment, which is consistent with the model in which the re-bargaining occurs only when the participation condition is binding. We must remember, however, that this specific test may lack statistical power, for even if the younger household does re-bargain its resource allocation plan, the effect may be too small to detect due to the small policy impact on the outside option of a young wife.

2.5 Results

2.5.1 Baseline results

Table 2.2 reports the empirical results of the DD estimation showing the household response to the reform by age group. In the most elderly group, consisting of households with wives aged 50–59, the wife's allocation to leisure increased by 5.0 hours per week, or 4.8 percent. This increase in leisure was associated with a roughly equal

[16] Although we did not include wage rates in the empirical model since those of non-labor participants are unobserved, our estimates do not suffer from any bias as long as the (potential) wage rates are uncorrelated with the treatment status, conditional on household fixed effects, year dummy variables and other control variables. Since the treatment status was fixed over the period, we believe that this conditional independence assumption is not restrictive.

[17] Another approach would have been to categorize households by marriage duration instead of age, as the amount of divisible pension records depends on marriage duration, but this would have been problematic because the marital period is likely to be correlated with match quality. Indeed, in equations (2.7) and (2.8), we see that the higher the match-specific utility $\theta_{j,t}$, the less likely the participation conditions are to be binding.

decrease in market labor supply and domestic labor supply. Although the effect on labor supply was not statistically significant here, we found that this effect was statistically significant and the size of the estimates were similar in the triple-differences (DDD) estimation discussed in Section 2.6.4. Furthermore, the estimates in this age group were jointly different from zero at 10 percent significance level, so our statistical inference is not a consequence of testing multiple hypotheses one-by-one.

These estimation results imply that the elderly spouses fail to completely commit to their initial allocation plan due to conflicting incentives, which leads us to reject the full commitment model of household behavior. While one might consider complete commitment as a good approximation if re-bargaining effects were negligible, we found substantial re-bargaining in the form of a 5-percent change in the wife's leisure. We thus conclude that the model with limited commitment is a better approximation of actual household behavior. Additionally, we found that home production played a non-negligible role in re-bargaining. Specifically, if we had instead defined leisure as non-market hours, we would have missed about half of the change in the wife's time allocation. As shown in Blundell et al. (2005), the level of home production depends on the marginal willingness of each spouse to pay, and thus, contribution to home production is an important bargaining domain.[18]

In order to gauge the magnitude of the policy impact, we consider the amount of pension benefits that a wife can access from her husband upon divorce. Suppose that the annual value of the divisible pension benefits for a wife aged 55 is 4.8 thousand U.S. dollars, which was the average amount among those who were eligible for benefits in 2014 (i.e., wives who were older than 53 when the pension reform was enacted)[19], the wife's age of eligibility for pension benefits is 60, and she marries at 28 and dies at 86, which reflects the average lifespan of women in Japan.[20] Setting the interest rate $r = 0.01$, the present value of benefits for a wife aged 55 is 90.5 thousand U.S. dollars. Given that the annual amount of her National pension benefits is 6.5 thousand US dollars beginning at age 65, this pension reform changes the present value of her total pension benefits from 92.0 thousand to 182.6 thousand US dollars, which is almost a 100 percent increase. This back-of-the-envelope calculation indicates that the wife's leisure, market work and domestic work elasticities to the life-time pension benefits upon divorce are 0.05, 0.20 and 0.05, respectively. Since the reform does not change total benefits, these effects are not due to any wealth effect but can be attributed solely to the re-bargaining effect. Furthermore, since the wife may have other assets at her disposal, the elasticity to the assets at her disposal upon divorce could be even larger.

[18]Blundell et al. (2005) discuss the public good provision under a cooperative framework. Although we specified the home production good as a private good, we can easily modify our model so that both private and public goods are produced at home without changing the main predictions.

[19]Source: Japanese Ministry of Health, Labour and Welfare; `http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000106808_1.html`.

[20]Data Source: `http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai07/` and `http://www.mhlw.go.jp/toukei/saikin/hw/life/life07/index.html`.

Unlike with elderly couples, we did not find any statistically significant changes in time-allocation among the younger households, and the estimates for the two younger groups were also jointly insignificant (Table 2.2). It is worth noting that these two younger groups correspond to the households in the Sakamoto (2008) study, which does not find any policy impact of the reform. This is consistent with the model's prediction that the current-period participation condition is unlikely to bind for younger households. Firstly, the pension division applies only to the records tracked during marriage, which is short for young households. Secondly, as the pension benefits do not count as collateral, a divorcing wife is not able to immediately receive the present value of any future pension benefits but must wait until retirement.[21] Thirdly, the discounted future value of the divorce receipts for a young wife would be low, which again makes the current-period participation condition unlikely to bind. For all these reasons, the possible division of pension benefits is less relevant to a younger wife, resulting in almost no impact of the reform on her household resource allocation. Even if the participation conditions at future periods bind under the current allocation, the impact on the current-period resource allocation is only through consumption-smoothing, not through the direct impact of the re-bargaining. As a result, the resource allocation would be gradually adjusted to minimize the deviation from the *ex ante* efficient allocation.

This discussion is in line with Lise and Yamada (2018), who argue that small shocks do not trigger re-bargaining. Since any possible improvement in welfare is too small for a young wife to initiate divorce, her bargaining position is not updated. Consequently, the reform has virtually no impact on young households, for it affects neither the budget set nor the utility weight on each member. We must recall, however, that this view of the way to update bargaining power should be treated with caution, since it can possibly be due to a lack of statistical power in testing minute changes in time allocation.

2.5.2 Home ownership and net housing value

Another prediction from our model is that the impact of the reform depends on the amount of household assets other than pension benefits because the assets obtained through pension division would be negligible for high net worth households. In Japan, the property division rule assures divorcing spouses equal division of assets for which they both contributed in obtaining, and one of the most important non-financial household assets is the family home. The interpretation of the law is that a home purchased during the marital period would be divided among the spouses if they divorced, but if it was acquired either before marriage or through an inheritance or gift, it would belong to a single spouse and would not be divided.

[21] Exceptions to this include borrowing against future benefits through the Welfare and Medical Service Agency and the Japan Finance Corporation. For more detail, see `http://hp.wam.go.jp/home/tabid/36/Default.aspx` and `https://www.jfc.go.jp`.

The impact of the pension reform on a dependent wife's option outside of marriage is therefore likely to also depend on the value of any property that she would obtain upon divorce. If the value of property is high, the pension benefits divided from her husband would be only of marginal importance to her. As a result, the pension reform would be expected to have a negligible impact on a wife with other real assets and a substantial impact on a wife with no property. Another possible interpretation is that home ownership could also work as a commitment device that makes divorce undesirable relative to staying married. In such cases, a household with home ownership would likely have a higher degree of commitment than one without.

To test this prediction, we divided the most elderly age group into two subgroups according to the household's net housing value and home-ownership.[22] To this end, we first calculated the net value of all houses and plots of land using the self-evaluated value of these properties less the remaining debt from acquiring them.[23] We then assigned a property value to each spouse according to the property rights from housing and land, and then divided the most elderly group into two subgroups: wives with positive property values and those without. Given the above discussion, we would expect the reform to have a substantial impact only on wives without a positive property value.

Table 2.3 highlights the heterogeneous effects of the pension reform on wives according to the value of the property they own. The leisure of wives whose net housing value was non-positive substantially increased by 9 hours per week while their market and domestic labor supply decreased by 4 and 5 hours, and these estimates were jointly significantly different from zero. In contrast, the hours allocated to both leisure and production by wives with positive property value were not affected in any statistically significant way and, furthermore, the point estimates were almost zero. Although we could not reject the null hypothesis that the response to the reform was the same across the two groups in terms of time allocation at the 10 percent level (with p-value 0.16), we believe that this is due to the relatively small sample size. These estimation results thus seem to suggest that those who were mainly affected by the reform were wives with poor options outside of marriage, which is well explained by the collective model with limited commitment. To sum up, this subsample analysis further supports the rejection of the full commitment model.

[22] Although it would have been possible to conduct this subsample analysis on the basis of household assets or savings, we chose home ownership because the title to properties other than housing was not available in our dataset. Consequently, we could not distinguish the fraction of savings divided upon divorce from the fraction of savings accumulated before marriage or accumulated through inheritance.

[23] We omitted the top and bottom two percent of the property values to alleviate any influence of outliers.

2.6 Validity check of the identification assumption

2.6.1 Placebo test

While the common trend assumption required for DD estimation to identify the ATT is not directly testable, it is worth considering what might occur if this assumption was violated. For example, if the results in Table 2.2 were driven entirely by a business cycle unique to the treatment group, we would observe its impact throughout all age groups. However, we found significant changes in the time allocation in the most elderly age group while not in the other groups, so our results are unlikely to be explained by an economic shock that would have influenced only the treatment group.

In order to further confirm the validity of the common trend assumption, we estimated placebo effects of the policy by counterfactually assuming that the reform was enacted in 2005 and 2006 as well as in 2007. Specifically, we estimated the following equation for each age group:

$$y_{it}^j = \sum_{t=2005}^{2007} \delta_{1t}^j d_t \cdot Treatment_i + x_{it}'\beta^j + \sum_{t=2005}^{2006} \gamma_t^j d_t + c_i^j + u_{it}^j. \tag{2.11}$$

Since $\delta_{1,2005}^j$ and $\delta_{1,2006}^j$ represent placebo policy effects, they should be zero when time trends are common across the groups. If they are different from zero, the common trend assumption may be violated. In the following analysis, we focus on the time allocation of the wife since we found close to null effects for the husband in our DD analysis (Columns 4–6 in Table 2.2).[24]

Table 2.4 shows the estimates of δ_{1t}^j in equation (2.11) for the wives. Consistent with our baseline results, we found a significant impact on the most elderly age group but a smaller impact on the other age groups.[25] In terms of the coefficients on the placebo years, the estimates of these coefficients were not statistically different from zero in the most elderly and youngest households, supporting the common trend assumption. However, in the middle-aged group, the leisure of the wife significantly decreased in 2006, which could potentially indicate a violation of our identification assumption. Since leisure is defined as the residual hours after production activities, however, an increase in leisure and a decrease in domestic labor supply are merely systematic, as her market hours did not change. Further, as we tested a total of 12 hypotheses that each placebo coefficient is zero, it is not unlikely that one of them

[24] As the husband in the treatment group was typically a full-time employee, it appears to have been difficult for him to have changed his hours worked. Furthermore, there were many households in which the husband did not engage in any household production. As a result, changes in bargaining power caused by the pension reform may not have been substantial enough for these husbands to deviate from the corner solution. However, our estimation results do not immediately suggest that the pension reform had no impact on husbands, as their levels of consumption may have declined, both in terms of private and public goods.

[25] The estimates were jointly significant in the most elderly group. Moreover, we rejected the null hypothesis that the 9 coefficients on $Treatment \times d_{2007}$ are all zero, where the three restrictions are redundant since the time allocation sums up to 168.

might be rejected at a 5 or 10 percent significance level even if all of them are true. In fact, we could not reject the null hypothesis that the placebo coefficients are all zero, with p-value 0.36.

However, we still cannot completely negate the possibility of the violation of the common trend assumption, so to further address this issue, we explicitly allowed for a linear time trend specific to the treatment group. The first column in Table 2.5 shows the estimation results. We see that the estimate for leisure is quantitatively similar to the baseline result and although market and domestic labor supply estimates differ from the baseline, the standard errors tend to be large and the signs are the same as the baseline. Considering that we have only one treatment year out of the four sample period, it seems natural that the estimates would become imprecise, and so the robust result for leisure thus supports our contention that our estimation result is not driven by a violation of the identification assumption.

2.6.2 Sensitivity to specification and sample restriction

One potential cause of differing time trends between the treatment and control groups would be a change in family structure. In particular, the typical treatment household tends to have a larger family and younger children than the control household (Table 2.1). Although in the baseline specification we controlled for family size, the number of children, whether households had a child under 6, and household fixed effects, it is possible that these were insufficient to completely remove the impact of changes in family structure. Thus, we controlled for family structure by using fourth-order polynomials of family size and number of children, dummy variables indicating whether or not the household had a child aged 0–6, 7–12 and 13–18, and interaction terms between these polynomials and the dummy variables. Despite this flexible specification, we obtained estimation results that were qualitatively and quantitatively identical to the baseline results (Column 2 in Table 2.5). Therefore, any potential changes in family structure seem not to be a concern for our identification strategy.

A second potential violation of the common trend assumption could occur because in the baseline estimation the control group included wives who were permanent employees in 2003 whereas the treatment group did not. Since it is conceivable that the time allocation trend of wives with permanent employment differs from those with part-time or no employment, as an additional confirmation of our baseline results, we excluded wives with permanent employment from our sample and re-estimated the baseline DD equation (2.10). The third column of Table 2.5 shows that the leisure time of the wife in the most elderly age group increased while her market hours and domestic work hours decreased in response to the pension reform. Furthermore, the magnitude of the estimates is comparable with those from the baseline result. As before, none of the estimates for the younger two groups are statistically significant. These findings suggest that differences in time-use were not the main driver of our baseline results.

As discussed in Section 2.3, the key to identifying the degree of commitment of spouses to an initial resource allocation plan is intra-household variation in bargaining position caused by the pension reform. Though approved in 2004 and enacted into law in 2007, discussion of the reform began several years earlier, with published newspaper articles about the reform appearing as early as 1998. It is thus possible that some people aware of the potential reform before 2000 (Figure 2.4) may have married in the early 2000's already anticipating that the division of pension records upon divorce might be allowed in the future. If this were the case, then the pension reform of 2007 would be unsuitable as an identification strategy, as it would be difficult to separate the changes in bargaining position after marriage from the bargaining position at marriage. Consequently, as couples that potentially anticipated the pension division reform would need to be excluded from the analysis, we implemented a subsample analysis that limited the estimation sample to wives in the most elderly group who married in 1997 or earlier. As the results (Column 6 in Table 2.5) are identical to the baseline, the issue of couples anticipating the pension reform at the time of their marriage appears to be inconsequential to our analysis.

2.6.3 Another policy change: The mandatory retirement age

Another consideration in interpreting our estimation results is the enactment in 2006 of the Elderly Employment Stabilization Law (EESL) affecting the mandatory retirement system in Japan. From 2006, the Japanese government through the enactment of the EESL has required firms to comply with a scheme to raise the mandatory retirement age to 65 in order to fill the gap between the existing mandatory retirement age of 60 and the pension eligibility age of 65. Specifically, the EESL allows those born in 1946 to remain employed until age 63, those born in 1947 or 1948 until age 64, and those born in 1949 or later until age 65. The new law did not force firms to raise the mandatory retirement age but instead provided them with three options to continue to employ workers who would otherwise need to retire: (1) raise the mandatory retirement age; (2) extend or renew employment contracts; or (3) abolish mandatory retirement. According to the Ministry of Health, Labour and Welfare, more than 80 percent of firms chose option 2 rather than either raising or abolishing the mandatory retirement age.[26]

It is expected that the EESL would increase the labor force participation rate of elderly people hired as full-time employees. Kondo and Shigeoka (2017) show that it has indeed increased the ratio of salaried workers aged over 60, but the impact is rather small at only 3 percentage points and is seen only in large firms ($\geq$ 500 employees), perhaps because small and medium firms had already abolished or raised the retirement age. Additionally, they find that contract renewals tend to be associated with a substantial decline in wages, which discourages employees from continuing to participate in the labor force. Despite the apparent limited impact of the EESL, it

[26] Data Source: `http://www.mhlw.go.jp/toukei/list/11-23c.html`.

could nevertheless bias our estimates due to the similarity of the target populations. Those born in 1946 became 60 years old in 2006, and a main target of the EESL was full-time employees, so the pension reform enacted in 2007 presumably affected the same employees. Given that the effect of the EESL was to increase the future earnings of the husband, this wealth effect could potentially decrease the current market labor supply of the wife and increase her leisure. Since those affected by the EESL are employees at large firms, we excluded them from the sample and re-estimated the baseline empirical model (2.10). As this subsample analysis (Column 4 in Table 2.5) replicated the baseline results, the EESL also does not seem to be a main driver of our findings.

One further possibility is that the EESL may potentially have an indirect affect on younger employees. Since the EESL requires firms to continue to employ workers who would otherwise retire at age 60, it is conceivable that firms might accommodate this new requirement by decreasing the number of young employees, either by not hiring new graduates or by letting go of part-time employees. Conversely, if elderly workers and young workers are viewed as complementary, then firms could increase the number of young employees as a result of the EESL. Ohta (2012) and Kondo (2016) report negative correlations between the proportions of employees aged over 60 and of female part-time workers, potentially suggesting that the EESL has created a crowding-out effect. However, since our control group includes part-time employees, such an indirect impact of the EESL is controlled by our DD estimation to a certain degree. Furthermore, as long as the EESL affects young and old female part-time employees in a similar manner, DDD estimation (described next) partials out any potential indirect impact of the EESL so that potential biases caused by the EESL are likely to be small.

2.6.4 Triple differences estimation

As the characteristics of the 2007 pension reform suggest that it had little effect on anyone other than elderly households, this is an opportunity to implement triple differences (DDD) estimation, which is more robust than DD estimation since the third difference can partial out any time trend specific to the treatment group. We must be careful in interpreting DDD estimates, however, as DDD estimation removes any policy impact that is common across the young and elderly households in the treatment group. In particular, as some estimates for the oldest and middle-aged group were similar (Table 2.2), DDD estimation could underestimate the policy impact. Thus, the biases in DDD estimation would make the null hypothesis of full commitment difficult to reject, and so, the resulting hypothesis test would be rather conservative.

For DDD estimation, we introduced two younger age groups as additional controls and introduced a dummy variable, Old_i, which takes one if the wife was over 50 in

2007. Then, we estimated

$$y_{it}^j = \delta_1^j DDD_{it} + \delta_2^j After_t \cdot Treatment_i + \delta_3^j After_t \cdot Old_i + x_{it}'\beta^j + \sum_{t=2005}^{2007} \gamma_t^j d_t + c_i^j + u_{it}^j, \quad (2.12)$$

where $DDD_{it} = After_t \cdot Treatment_i \cdot Old_i$, and δ_1^j represents the impact of the pension reform. In order to accommodate any arbitrariness in the choice of the control group, we estimated equation (2.12) by using two subsamples as well as the entire sample, with one subsample consisting of the most elderly group and the second youngest group and the other consisting of the most elderly group and the youngest group.

In estimating equation (2.12), we found that the leisure of the wife increased more than 5 hours per week, which was statistically significant irrespective of the choice of control group (Table 2.6). While the estimate using wives aged 30–39 as the control group is relatively larger (Column 7), the size of other two estimates are comparable to the results of the DD estimation (Columns 1 and 4). Hence, we are confident that the results from both DDD and DD estimation provide sufficient evidence to reject the full commitment model. Meanwhile, the market labor supply of the typical wife decreased by 4–4.5 hours, a result that was robust to change in control groups. Moreover, the sign of the DDD estimate is in line with those of the baseline DD estimates, though its magnitude is somewhat greater than the baseline and similar to the DD estimate which allowed for a linear time trend specific to the treatment group. All in all, considering the robustness of our findings, it appears that the wives' market hours decreased due to the changes in the value of divorce to her.

In contrast to market work, we observed relatively large variation in the DDD estimates of the impact on domestic work, and the estimates were not statistically significant. However, the sign of the coefficient was still negative in all three cases, as it was with the DD estimate. Furthermore, the estimate became smaller when we used the second most elderly age group as the control group, and this is consistent with our earlier discussion that DDD estimation possibly underestimates the policy impact by filtering out effects common among elderly and younger households.

A final considerations is that while we maintained the identification assumption that there was no time trend specific to the treatment group in the most elderly age group, the DDD estimates would fail to recover the ATT if this assumption did not hold. Thus, to check the sensitivity of our estimates, we implemented a further DDD estimation while allowing for a linear time trend specific to the most elderly age group and also to the treatment group. Under the identification assumption, the resulting estimates would likely be similar to the DDD estimates obtained from the empirical model without those group-specific time trends and, indeed, we found an almost identical result with this specification (Column 7 in Table 2.5). To sum up the above discussion, a comprehensive series of the robustness checks further supports our rejection of the full commitment model.

2.7 Divorce

In Section 2.3, we assumed that a couple does not eventually divorce, and the identification result indeed relies on that assumption. In this section, we thus examine the case when divorce does occur. With a positive probability of divorce, the household objective function becomes

$$\mu E\left[\sum_{t=1}^{3}\beta^{t-1}\left\{(1-D_t)u_1(c_{1t},l_{1t};\theta_{1t})+D_tV_{1t}^D\left(a_{1t}^D+b_{1t}^D\right)\right\}\right]$$
$$+(1-\mu)E\left[\sum_{t=1}^{3}\beta^{t-1}\left\{(1-D_t)u_2(c_{2t},l_{2t};\theta_{2t})+D_tV_{2t}^D\left(a_{2t}^D+b_{2t}^D\right)\right\}\right],$$

where D_t is a decision to divorce, which is an absorbing state, and $D_1 = 0$ as we focus on a married couple. In this case, each spouse cares about their welfare after potential divorce even under full commitment, and consequently, this concern about divorce makes the household behavior contingent on the share of assets upon divorce. Note that this caveat is not unique to this study but is typical in the literature.[27]

One possible interpretation of our results is that the pension division system might have affected the wife's actions prior to divorce and thus our estimates reflect that rather than the effects of re-bargaining. For example, it seems plausible that a housewife planning to divorce may begin working in order to prepare for her life after divorce, and Mazzocco et al. (2006) suggest that wives in the United States tend to start working about two years before divorce so as to accumulate human capital. Since the Japanese pension reform provided an additional income source after divorce, this might eliminate the need for a wife to work before (and possibly after) the divorce. In this scenario, her market hours would decline after the reform and her leisure would increase, which is compatible with our estimation results but unrelated to changes in her bargaining position.

Reflecting on the above situation, however, it seems difficult to explain our findings entirely by such "divorce-concern" behavior due to the low probability of divorce in Japan, particularly among elderly households. Figure 2.3a illustrates the annual divorce rate of elderly households before and after the reform and shows, for example, that 3 in 1000 couples with wife aged 55 divorced the next year. This probability declines rapidly as the age of the couples rises, becoming less than 0.1 percent at age 70. Although we are referring here to cross-sectional data, from this we calculated a life-time divorce rate at each age of the wife (Figure 2.3b). While the life-time probability of divorce is relatively high for young wives because both the annual divorce rate and life expectancy are high, this is not true of elderly wives, whose life-time divorce rate at age 55 is only about 3 percent. Since the value of divorce under full commitment

[27] Our limitation is the same as that of Blau and Goodstein (2016), while Mazzocco (2007) does not allow assets to be used for bargaining, and Lise and Yamada (2018) do not incorporate endogenous divorce behavior or human capital accumulation in their structural model so the "divorce-concern" behavior discussed by Mazzocco et al. (2006) is not addressed.

affects household behavior only through the decision to divorce, these low probabilities of divorce suggest that our baseline model without divorce approximates reality well and our estimation results do not seem to driven entirely by "divorce-concern" behavior.

We also calculated the weight ω_2^D put on the divorce value in the wife's value function. The weight of the wife aged t would be approximated by

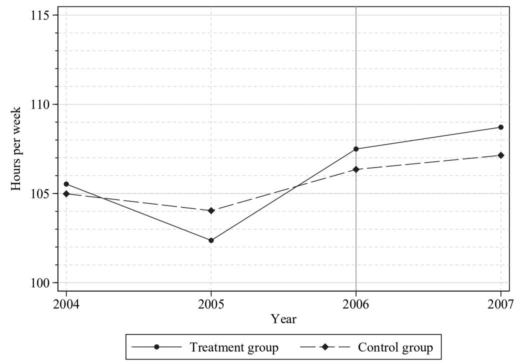
$$\begin{aligned}\omega_2^D(t) &= \frac{\sum_{s=t}^{T} \beta^{s-1} \Pr(D_s = 1 | D_{s-1} = 0)}{\sum_{s=t}^{T} \beta^{s-1} \Pr(D_s = 1 | D_{s-1} = 0) + \sum_{s=t}^{T} \beta^{s-1} \Pr(D_s = 0 | D_{s-1} = 0)} \\ &= \frac{\sum_{s=t}^{T} \beta^{s-1} \Pr(D_s = 1 | D_{s-1} = 0)}{\sum_{s=t}^{T} \beta^{s-1}},\end{aligned}$$

with the imprecision due to the decision to divorce depending on a distribution of uncertainties. Figure 2.3c plots ω_2^D at each age, showing that this weight is minute at all age points. It is still possible, however, that the weight on the divorce state could be large particularly among couples on the verge of divorce. To partially check if those couples are the driving force of our estimation results, we re-estimated our baseline regression excluding couples that divorced between 2005 and 2012, but the results were unchanged (Column 5 in Table 2.5). Hence, our results seem difficult to be explained by divorce concern behavior.

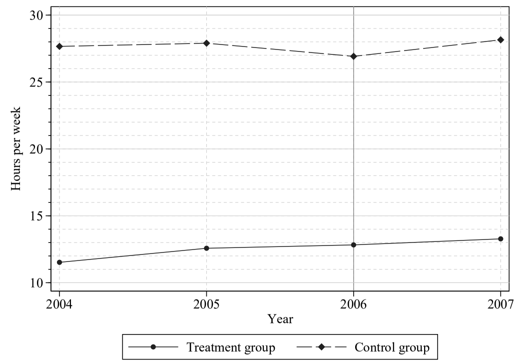
If we can accept that the full commitment model is rejected, our findings underline how difficult it is to make a commitment. Although the marital relationship of elderly couples in Japan tends to be stable and divorce is rare, even they fail to achieve full commitment. While a lack of substantial exogenous variation did not allow us to test the degree of commitment of young couples, we suspect that their comparatively less stable marital relationships would make commitment even more difficult. In fact, a high divorce rate may suggest that the participation conditions are likely to bind, so a young couple may have more opportunities to threaten divorce or to bring about a hold-up problem. We thus believe that a lack of commitment is a key feature of family decision making, and its economic consequences are worth considering in future empirical studies.

2.8 Conclusion

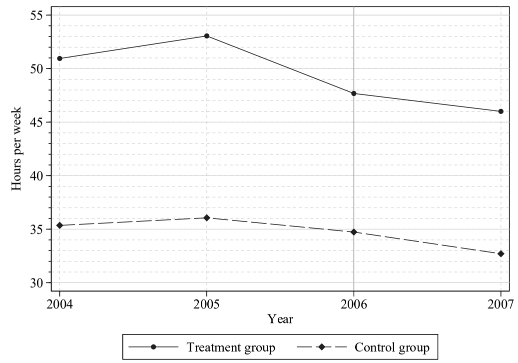
It appears that couples have difficulty committing to an initial resource allocation plan. By using a pension reform in Japan to create a natural experiment, we filled a gap in the literature by testing a full commitment model of household decision-making without imposing any specific functional form assumptions on preferences or home production technology. The results led us to reject the full commitment model, as elderly wives exploited their improved outside option to enjoy more leisure by reducing market and domestic labor. Consistent with the model's prediction, this impact was most striking among low net worth couples. Incomplete commitment, however, does not mean no commitment, and we found suggestive evidence that a couple does not respond to small shocks unless the participation condition binds. A lack of commitment thus seems to distort resource allocation and typically makes the first-best allocation difficult, if not impossible, to achieve. Therefore, future work should address the size of distortion in long-term decision making such as human capital accumulation and investment in children as well as risk sharing. Finally, given the non-negligible impact of re-bargaining, we believe that model-based studies need to incorporate this feature of family decision making.



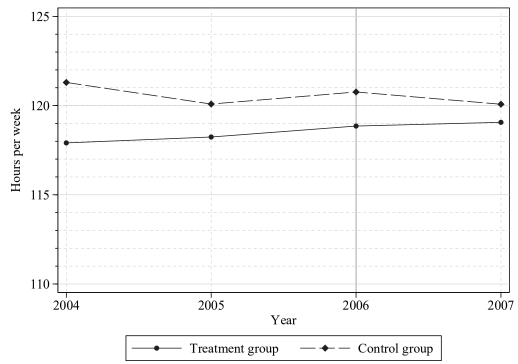
(A) Leisure: Wife



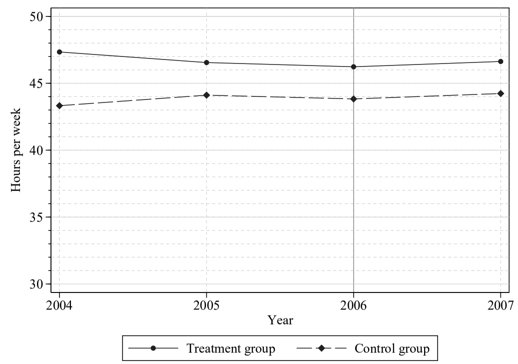
(C) Market Labor Supply: Wife



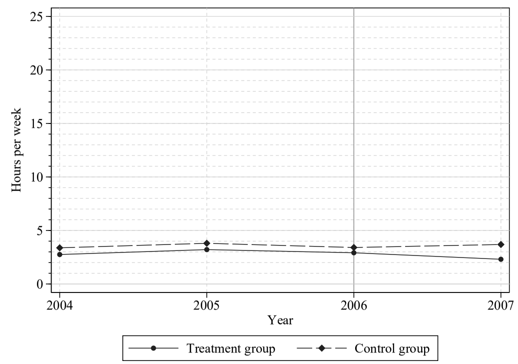
(E) Domestic Labor Supply: Wife



(B) Leisure: Husband



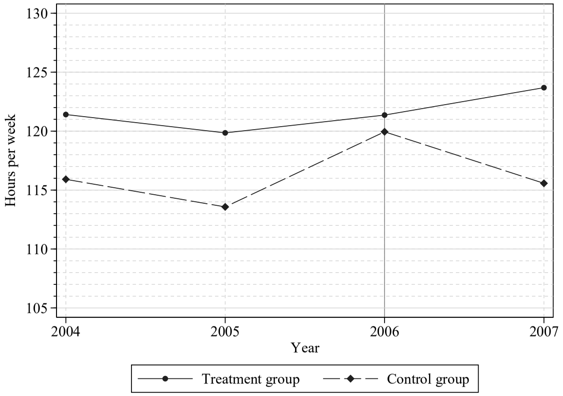
(D) Market Labor Supply: Husband



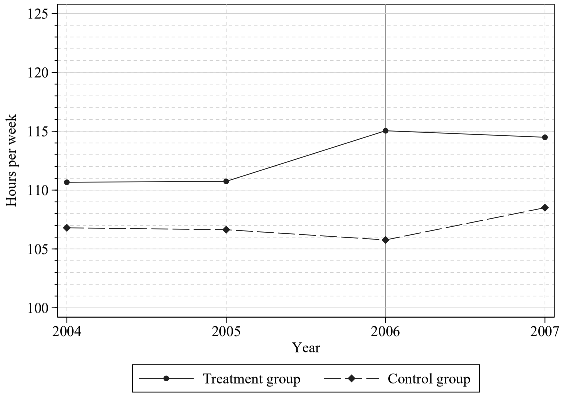
(F) Domestic Labor Supply: Husband

FIGURE 2.1: Time trend in market labor supply, domestic labor supply and leisure

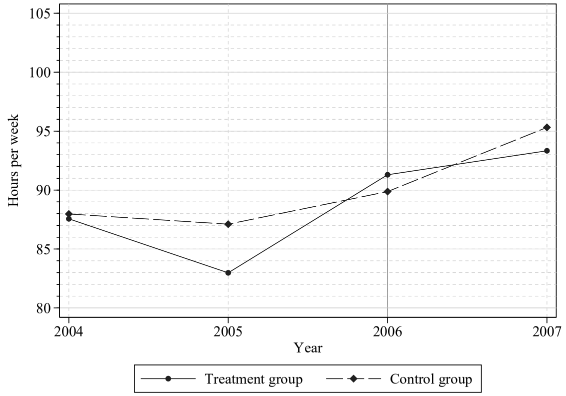
Data Source: The *Keio Household Panel Survey*
Note: This figure shows the yearly trend in the time allocation of wives and husbands. The solid lines represent the time allocation of individuals in the treatment group while the dotted lines represent individuals in the control group. The vertical lines highlight 2006, one year before the pension reform was enacted.



(A) Wife Aged 50-59



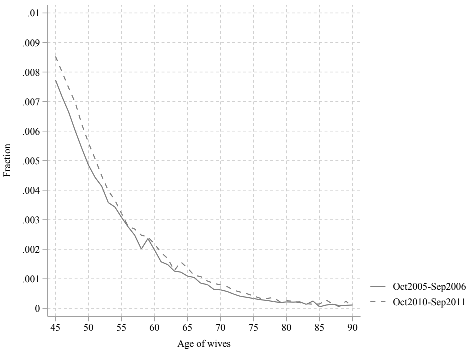
(B) Wife Aged 40-49



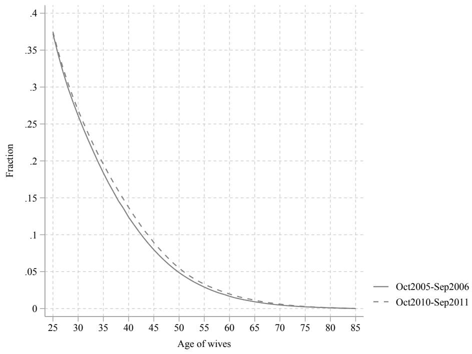
(C) Wife Aged 30-39

FIGURE 2.2: Time trend in time allocated to leisure by age of wife

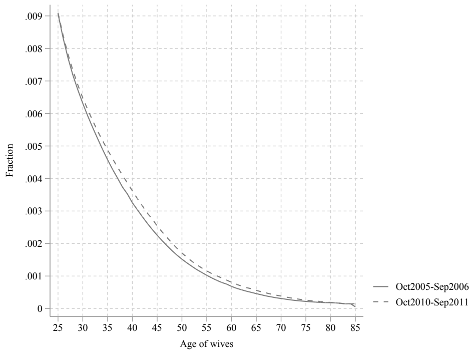
Data Source: The *Keio Household Panel Survey*
Note: Yearly trend in the time allocated to leisure by wives are presented for each age group. Solid lines represent the treatment group and dotted lines represent the control group. The vertical lines indicate 2006, one year before the pension reform was enacted.



(A) Yearly divorce probability



(B) Life-time divorce probability



(C) Utility weight on the divorce state

FIGURE 2.3: Divorce probability and utility weight put on the divorce value

Source: *Vital Statistics* and *Census*.

Table 2.1: Descriptive statistics of KHPS sample

	Wife		Husband		Household	
	Treatment (1)	Control (2)	Treatment (3)	Control (4)	Treatment (5)	Control (6)
Leisure	105.86	105.51	118.46	120.59		
	(32.63)	(29.02)	(19.14)	(21.96)		
Market labor supply	12.49	27.65	46.72	43.85		
	(15.30)	(21.71)	18.41	(20.79)		
Domestic labor supply	49.65	34.84	2.82	3.56		
	(35.23)	(29.22)	(5.82)	(7.16)		
Age	42.87	44.72	45.06	47.24		
	(7.64)	(7.67)	(8.01)	(7.83)		
Family size					4.04	4.07
					(1.21)	(1.39)
Number of children					1.72	1.52
					(0.96)	(1.00)
Children under 6 years old					0.24	0.17
					(0.43)	(0.38)
Marital period					15.77[i]	17.53 [ii]
					(8.25)	(8.66)
Observations	2,868	1,812	2,868	1,812	2,868	1,812

Note: The table shows means and standard deviations, with the latter in parentheses. Time allocation was measured as the average hours per week. i) Sample size was 2836. ii) Sample size was 1776.

Table 2.2: The effect of pension reform: Difference-in-differences estimation

		Wife		
Age Group	Leisure (1)	Market Labor Supply (2)	Domestic Labor Supply (3)	Observations [Households]
50–59	4.976**	-2.363	-2.613*	1,333
	(2.509)	(2.136)	(1.435)	[442]
40–49	-0.322	2.166	-1.844	1,804
	(1.878)	(1.491)	(1.578)	[575]
30–39	-3.265	1.475	1.790	1,543
	(3.445)	(1.637)	(3.153)	[490]
30–59	0.797	0.695	-1.492	4,680
	(1.530)	(0.997)	(1.288)	[1,507]

		Husband		
Age group	Leisure (4)	Market labor supply (5)	Domestic labor supply (6)	Observations [Households]
50–59	1.014	-0.171	-0.842	1,333
	(2.180)	(2.109)	(0.567)	[442
40–49	-0.935	1.259	-0.324	1,804
	(2.056)	(2.004)	(0.504)	[575]
30–39	1.615	-0.811	-0.805	1,543
	(2.361)	(2.293)	(0.874)	[490]
30–59	0.376	0.312	-0.688*	4,680
	(1.281)	(1.243)	(0.371)	[1,507]

Note: The table shows the estimation results of equation (2.10) by the age group of the wife. Only the estimated values of δ_1^j are reported, with standard errors clustered by each household in parentheses.
$***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.

TABLE 2.3: Subsample analysis by housing value and property rights (Wives aged 50–59)

	Housing ≤ 0	Housing > 0
	(1)	(2)
Dependent variable		
Leisure	9.116**	-0.593
	(4.342)	(4.487)
Market work	-4.083	-0.314
	(3.319)	(4.200)
Domestic work	-5.033*	0.906
	(2.753)	(2.006)
p-value of joint test	0.086	0.900
p-value of H_0: same impact in (1) and (2)	0.156	
Observations	495	500
Households	165	164

Note: The table shows the estimation results of equation (2.10) for the most elderly age group of the wife, by subsamples defined by net housing value, where the top and bottom two percents of the property values of wives are excluded to eliminate outliers. Only the estimated value of δ_1^j is reported. Standard errors clustered by each household are in parentheses. The p-value for the joint test against the null hypothesis that the response to the reform is, in terms of time allocation, the same across two groups is less than 0.1.
$*** p < 0.01$, $** p < 0.05$, $* p < 0.1$.

TABLE 2.4: Placebo test for the wives

Age group: 50–59	Leisure (1)	Market labor supply (2)	Domestic labor supply (3)
$Treatment \times d_{2007}$	5.733*	-1.659	-4.074**
	(2.998)	(2.527)	(1.882)
$Treatment \times d_{2006}$	-0.992	2.957	-1.965
	(2.587)	(1.944)	(2.009)
$Treatment \times d_{2005}$	3.199	-0.832	-2.367
	(2.656)	(2.158)	(1.890)
Observations	1333	1333	1333
Households	442	442	442
Age group: 40–49	Leisure (4)	Market labor supply (5)	Domestic labor supply (6)
$Treatment \times d_{2007}$	1.138	3.103	-4.240*
	(2.641)	(1.894)	(2.193)
$Treatment \times d_{2006}$	4.705*	0.378	-5.084**
	(2.751)	(1.729)	(2.376)
$Treatment \times d_{2005}$	-0.431	2.412	-1.981
	(2.546)	(1.718)	(2.109)
Observations	1804	1804	1804
Households	575	575	575
Age group: 30–39	Leisure (7)	Market labor supply (8)	Domestic labor supply (9)
$Treatment \times d_{2007}$	-3.809	2.032	1.777
	(4.623)	(2.201)	(4.135)
$Treatment \times d_{2006}$	0.902	0.745	-1.647
	(4.342)	(2.109)	(4.022)
$Treatment \times d_{2005}$	-2.635	0.882	1.753
	(3.887)	(1.895)	(3.679)
Observations	1543	1543	1543
Households	490	490	490

Note: The table shows the estimation results of equation (2.11), checking the pre-time trend for the wives. Standard errors clustered by each household are in parentheses. Only the estimated values of the interaction terms between the treatment dummy variable and year dummy variables are reported.
$*** \; p < 0.01$, $** \; p < 0.05$, $* \; p < 0.1$.

TABLE 2.5: Robustness checks using households with wives aged 50–59

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable							
Leisure	5.837*	5.198**	5.833*	4.930*	4.976**	5.890**	5.654*
	(3.531)	(2.530)	(3.052)	(2.574)	(2.509)	(2.509)	(3.162)
Market work	-5.229*	-2.608	-2.524	-2.164	-2.363	-2.935	-4.361*
	(2.863)	(2.201)	(2.697)	(2.225)	(2.136)	(2.180)	(2.400)
Domestic work	-0.608	-2.590*	-3.309**	-2.767*	-2.613*	-2.955**	-1.293
	(2.223)	(1.378)	(1.681)	(1.464)	(1.435)	(1.411)	(2.280)
p-value of joint test	0.173	0.069	0.076	0.088	0.090	0.035	0.146
Method	DD	DD	DD	DD	DD	DD	DDD
Group-specific linear trend	X						X
Flexible family structure		X					
Wives with permanent employment	X	X		X	X	X	X
Wives with husbands in large firms	X	X	X		X	X	X
Couples divorced btw. 2005 and 2012	X	X	X	X		X	X
Couples married before 1998	X	X	X	X	X		X
Observations	1333	1333	1104	1219	1328	1287	4680
Households	442	442	364	413	439	426	1507

Note: The table shows the estimation results of equation (2.10) and (2.12) with various specifications and sample restrictions, for the most elderly age group. Only the estimated value of δ_1^j is reported. Standard errors clustered by household are in parentheses. In the DDD analysis, the two young groups were used as the control.
$*** \ p < 0.01$, $** \ p < 0.05$, $* \ p < 0.1$.

TABLE 2.6: The effect of the pension reform on the wives: Triple differences estimation

Control group: Wife aged 30–49			
Dep. Var	Leisure (1)	Market labor supply (2)	Domestic labor supply (3)
DDD	5.658* (3.147)	-4.278* (2.399)	-1.380 (2.265)
Observations	4,680	4,680	4,680
Households	1,507	1,507	1,507
Control group: Wife aged 40–49			
Dep. Var	Leisure (4)	Market labor supply (5)	Domestic labor supply (6)
DDD	5.233* (3.113)	-4.565* (2.597)	-0.669 (2.117)
Observations	3,137	3,137	3,137
Households	1,017	1,017	1,017
Control group: Wife aged 30–39			
Dep. Var	Leisure (7)	Market labor supply (8)	Domestic labor supply (9)
DDD	7.636* (4.380)	-3.911 (2.701)	-3.725 (3.665)
Observations	2,876	2,876	2,876
Households	932	932	932

Note: The table shows the estimated values of δ_1^j in equation (2.12) for wives, with standard errors clustered by each household in parentheses.
$***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.

2.A Supplementary figures and tables

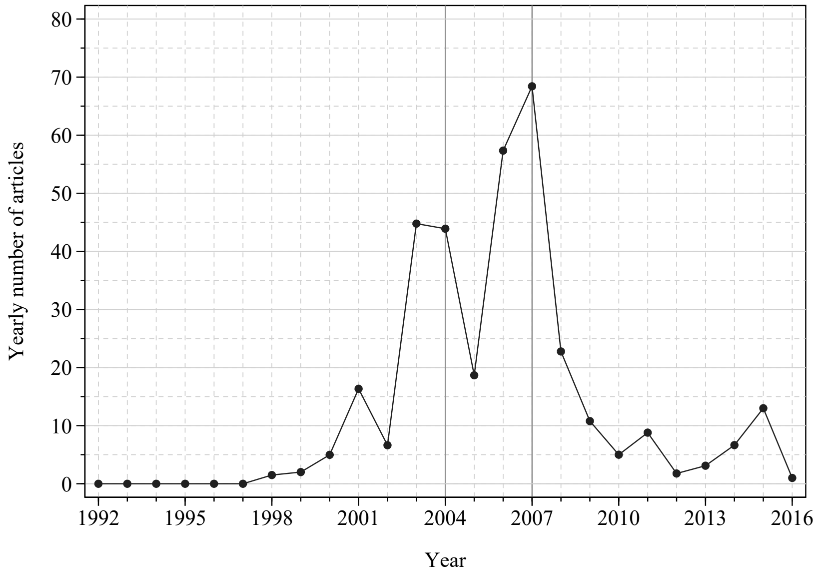


FIGURE 2.4: The yearly number of newspaper articles about pension division

Data Source: *Kikuzo II Visual*, *Maisaku*, *Semi-Annual Newspaper Issuer Report*, *Nikkei Telecom* and *Yomidasu Rekishikan*
Note: This figure shows the yearly number of newspaper articles about pension division on divorce published in the *Asahi*, *Mainichi*, *Nikkei* and *Yomiuri* newspapers, which are the prime national newspapers in Japan. The vertical axis shows the weighted sum of the number of the articles, where the annual yearly circulation of each newspaper was used as the weight and the weight on the number of the articles of *Yomiuri* newspaper was normalized to one. Due to data availability, the total number of articles in 2015 and 2016 was not weighted. The pension reform was voted in 2004 and enacted in 2007.

TABLE 2.7: Age of eligibility for Employee Pension Insurance

Birth cohort	Basic part		Proportional part	
	Men	Women	Men	Women
1940	60	60	60	60
1941	61	60	60	60
1942	61	60	60	60
1943	62	60	60	60
1944	62	60	60	60
1945	63	60	60	60
1946	63	61	60	60
1947	64	61	60	60
1948	64	62	60	60
1949	65	62	60	60
1950	65	63	60	60
1951	65	63	60	60
1952	65	64	60	60
1953	65	64	61	60
1954	65	65	61	60
1955	65	65	62	60
1956	65	65	62	60
1957	65	65	63	60
1958	65	65	63	61
1959	65	65	64	61
1960	65	65	64	62
1961	65	65	65	62
1962	65	65	65	63
1963	65	65	65	63
1964	65	65	65	64
1965	65	65	65	64
1966	65	65	65	65

Table 2.8: Changes in the time allocation of the treatment group and the control group

	Wife				Husband			
Leisure	Before (1)	After (2)	D (3)	DD (4)	Before (1)	After (2)	D (3)	DD (4)
Treatment group	105.09	108.71	3.63**	1.57	118.3	119.06	0.76	1.41
	[33.48]	[29.10]	(1.49)	(2.30)	[19.41]	[18.09]	(0.87)	(1.49)
Control group	105.09	107.14	2.06		120.73	120.08	-0.65	
	[29.60]	[26.70]	(1.68)		[22.10]	[21.45]	(1.27)	
Market labor supply	Before (5)	After (6)	D (7)	DD (8)	Before (5)	After (6)	D (7)	DD (8)
Treatment group	12.27	13.28	1.01	0.37	46.74	46.63	-0.11	-0.60
	[15.34]	[15.15]	(0.70)	(1.33)	[18.55]	[17.89]	(0.84)	(1.43)
Control group	27.51	28.15	0.64		43.74	44.23	0.49	
	[21.57]	[22.22]	(1.25)		[20.84]	[20.61]	(1.20)	
Domestic labor supply	Before (9)	After (10)	D (11)	DD (12)	Before (9)	After (10)	D (11)	DD (12)
Treatment group	50.64	46.01	-4.64***	-1.94	2.96	2.31	-0.65**	-0.81*
	[36.26]	[30.92]	(1.60)	(2.43)	[6.16]	[4.26]	(0.27)	(0.47)
Control Group	35.4	32.71	-2.70		3.53	3.69	0.16	
	[30.06]	[25.74]	(1.69)		[7.19]	[7.05]	(0.41)	

Note: The table shows the means of leisure before and after the pension reform, their difference within each group, and a DD estimate, with standard deviations in brackets and standard errors in parentheses.
$*** \; p < 0.01$, $** \; p < 0.05$, $* \; p < 0.1$.

Chapter 3

Introduction of parental leave policies and maternal employment in long-run

3.1 Introduction

The expansion of family policies coincides with shrinking gender gaps in market outcomes in the last decades. However, its causal relationship is not obvious, and family policies sometimes reinforce the traditional gender roles (Albrecht et al., 2003; Blau and Kahn, 2013; Fernández-Kranz and Rodríguez-Planas, 2013; Canaan, 2019). At the same time, the gender gaps are still non-negligible, and one remarkable example is "child penalty," which is a sharp and persistent drop in earnings after the childbirth, despite the prenatal earnings trajectory similar to that of women without a child (Kleven et al., 2019). This dip is associated with changes in occupation, shift to "family-friendly" work environment, reduced work time and efforts of mothers, and/or shift to "mommy track" (Fernández-Kranz et al., 2013; Kleven et al., 2018). In line with these observations, Goldin (2014) points out that the gender wage gap is closely related with compensated wage differential in terms of work flexibility. These studies suggest that the childbirth changes women's work attitude and environment to reconcile work and family responsibilities. What is an effective family policy to promote women's career after childbearing?

One of the most popular policy instruments for this purpose is parental leave. In fact, according to the OECD family database, most developed countries have expanded the parental leave program for decades. A parental leave policy typically guarantees a mother's right to be reinstated in her original or its equivalent post after the leave, and provides cash benefits during the leave to smooth income fluctuations. Although a generous parental leave policy is expected to facilitate career continuation of mothers, it is costly for employers as it requires additional human resource management, potentially inducing them to shift (non-) pecuniary costs of parental leave provision onto the wage of female workers (Gruber, 1994), or statistically discriminate them (Thomas, 2018). In addition, long duration of the leave may result in human capital depreciation. As a result of these potential negative aspects, the policy impact depends on its generosity in theory; too generous one is harmful to maternal market opportunities whereas too limited one is useless for mothers.

Following the expansion of parental leave programs, a bunch of empirical studies have evaluated the policy effect on maternal labor market outcomes. Cross-country studies find mixing results that the parental leave policy increases the female employment rate but widens the gender wage gap, and these policy impacts are heterogeneous across the educational levels. (Ruhm, 1998; Thévenon and Solaz, 2013; Olivetti and Petrongolo, 2017). On the other hand, a majority of recent studies use a discontinuous feature of policy reforms to ensure clean identification, and find the policy impact large in short-run but negligible in long-run. In short-run, the extension of parental leave duration causes delay in return-to-work behavior of mothers and increase maternal time with their children (Baker and Milligan, 2008; Lalive and Zweimüller, 2009; Dustmann and Schönberg, 2012; Schönberg and Ludsteck, 2014; Dahl et al., 2016).[1][2] Despite the delay in return-to-work behavior, its impact is attenuated over time, and their market outcomes catch up in long-run (Lalive and Zweimüller, 2009; Schönberg and Ludsteck, 2014; Carneiro et al., 2015; Dahl et al., 2016).

However, these negligible long-run impacts do not necessarily imply that the parental leave is useless for the career continuation of mothers, because most previous studies evaluate the effect of extending parental leave. In other words, the effect through parental-leave take up is not well considered in the literature, and this extensive margin seems crucial particularly when the parental leave program is introduced or when the limited parental leave program is expanded. Although some studies evaluate the extension of short parental leave, the effect through the extensive margin is not investigated due to the high take-up even under the pre-reform regimes (Carneiro et al., 2015), or sample restriction focusing on those who took maternity leave (Schönberg and Ludsteck, 2014). In fact, Kluve and Schmitz (2018) estimate the effect of a German policy reform as a mixture of the extensive and intensive margins and find a positive long-run impact. To summarize, the policy introduction seems to have a distinct impact from the policy extension, and this extensive margin is understudied in the literature.

To investigate the effect of taking parental leave, this study evaluated the long-run impacts of the introduction of job protection and cash benefits on long-run maternal labor market outcomes. To identify the causal impact, we exploited the parental leave reforms in Japan, namely, the job protection policy in 1992 and the paid leave policy in 1995. The 1992 reform introduced unpaid leave with job protection until the first birthday of children. The 1995 reform introduced cash benefits and the exemption of social insurance premiums during the leave, which amount to about 37 percent of the

[1]The policy impact on child development tends to negligible (Gregg et al., 2005; Baker and Milligan, 2010; Rasmussen, 2010; Dustmann and Schönberg, 2012; Dahl et al., 2016), while some studies in the U.S. find positive impacts (Berger et al., 2005; Rossin, 2011).

[2]Baum (2003) and Baker and Milligan (2008) find no short-run impact of short job protection policy. They attribute their null results to private arrangement that mothers could make even in absence of the legitimated job protection.

pre-leave earnings. Additionally, the eligibility of the parental leave was expanded in 1995. The first reform is particularly useful to separate the effect of job protection from the effect of cash benefits. On the other hand, the second reform serves to estimate the impact of taking parental leave as cash benefits and eligibility expansion increase the take-up rate of the parental leave.

Our empirical analysis drew on the micro data of the population Census in Japan between 1990 and 2010, and each Census data has more than a million of newborns. The large data set is indispensable in the Japanese context, for a fraction of mothers taking parental leave was limited to around 5 percent in 1990's. In fact, Asai (2015) evaluates the short-run impact of the 1995 reform in Japan using the Employment Status Survey, whose sampling rate is about 1 percent, but does not find any policy impact, presumably due to lack of statistical power. Thus, our population data helps to detect the policy impact. To evaluate short- and long-run policy impacts, the maternal market outcomes are measured by the employment status 0–15 years after the child birth.

In order to estimate the policy impact, we exploited two different identification strategies. First, we employed the DID estimation, in which those who gave a birth several years after the policy enactment were compared with those who gave a birth in 1991, one year before the job policy reform, using the 2005 Census. Since the child age in 2005 depends on the birth year, the policy effect is not separated from the child age effect in this simple comparison. We thus controlled for the child age effect by using the mothers with the same age of children in the 2000 Census as they were not eligible for the parental leave. The estimation result demonstrated that the job protection policy did not affect the maternal market outcome in the reform year, but had positively positive impacts 1–2 years after the enactment and furthermore, the second reform seems to amplify the policy impact. In particular, the parental leave policies increased the full-time employment in long-run while decreasing the part-time employment. Since these two impacts offset each other, the employment rate was unaffected in long-run. Therefore, our empirical evidence suggests that the provision of unpaid leave helps a mother continue her full-time job, and a main beneficially for the parental leave policy is those who would, in absence of the policy, quit their jobs at the child birth but return to labor market later. Given the rigid labor market in Japan where the port of entry to a career job is concentrated around school graduation (Imai and Itami, 1984; Genda et al., 2010), the mother once quitting her job have difficulty getting a job equivalent to her original post and is likely to work for a part-time job. Thus, the parental leave policy, protecting her job position with income compensation, enabled her to continue a full-time job.

Since this DID method is invalidated if the mothers giving births several years after the reform are selective in terms of market outcomes, we also conducted the RDD-DID analysis, which is often used in the literature (Lalive and Zweimüller (2009) for example). Given that the accessibility to the paid and unpaid leave depends on the birth timing, the policy impact is identified by the regression discontinuity design (RDD) using the birth timing as running variable, combined with difference-in-differences (DID) method to control for potential seasonality in birth timing as well as the relative age of children at the survey timing. While this RDD-DID method provides sharper identification than the first estimation method, it tends to result in imprecise estimates as it focuses on local part of the data. In this sense, these two methods complement each other. As a result of the RDD-DID estimation, we found suggestive evidence that the second reform increased the take-up rate of the parental leave in short-run. In long run, it increased the full-time employment in long-run while decreasing the part-time employment with no impact on employment. Furthermore, the two-sample Wald estimate constructed from these short-run and the long-run impacts indicates that the take-up of the parental leave increased (decreased) the probability to be a full-time (part-time) worker in long-run by more than 30 percentage points. Therefore, this study makes a sharp contrast with the previous studies finding the negligible effect of extending the leave duration.

Our study is closely related to Stearns (2018), who also evaluates the effect of parental leave in Britain, focusing on the extensive margin. She finds that the introduction of job-protected leave with some cash benefits increases maternal employment but does not affect work hours in long-run. She also finds some negative impact on promotion particularly among the high-skilled group. One advantage of the Japanese reform is that the first reform provides the job protection but not cash benefits. While Stearns (2018) argues that the cash benefit is not the main channel of her results as its amount is small, our reform helps understand the effect of unpaid job-protected parental leave. Another difference between the Japanese policy and British policy seems the eligibility. While most working mothers are eligible in Britain, Japanese system requires mothers to have non-fixed term contract and tenure greater than one year. Therefore, the type of mothers affected by the policy reforms seems different between these two country, and indeed, the Japanese parental leave reform had different impacts on maternal employment.

Asai (2015) and Yamaguchi (2019) are also related to this study, which evaluate the effect of parental leave policy using Japanese data. As mentioned above, Asai (2015) exploits the same policy reform as ours as a natural experiment but with much smaller data than ours, and as a result, she does not find a short-run policy impact on maternal employment. Due to advantage in the population census data, we detected the policy impact both in short-run and long-run. On the other hand, Yamaguchi (2019) estimates the dynamic discrete choice model, and his simulation result shows that the legislation of job-protected leave for a year is effective to maternal labor supply but additional duration of job protection on top of one year is ineffective. Since our study took design-based approach using the policy reforms as a natural experiment while Yamaguchi (2019) takes model-based approach to reveal the mechanism of the policy impact, we believe that these two studies complement each other.

3.2 Background information

3.2.1 Parental leave system in Japan

This section describes the maternity leave policy and parental leave policy in Japan as background information. Japanese leave system has clear distinction between maternity leave and parental leave. The maternity leave policy dates back to early 20th century, and was first legislated by the amended Factory Acts in 1926. This maternity protection was also legitimated by its successor, the Labor Standard Acts in 1947. It prohibits a woman from working 6 weeks before and 8 weeks after childbirth (Article 65). Furthermore, an employer is not allowed to dismiss a woman during maternity leave or within 30 days thereafter (Article 19). During the maternity leave, a woman receives cash benefits that amount to two-thirds of the average earnings in the past 12 months if she is covered by the Health Insurance (Health Insurance Act, Articles 99 and 102).[3]

The parental leave system is relatively new. In 1972, the Working Women Welfare Act required employers to use best efforts to adjust maternal work conditions but not mandatory. In 1975, job protection was provided for women working in public educational and medical sectors until the first birthday of children. The first universal job protection was legitimated in 1991 and enacted in April 1992 (Act on Childcare Leave, Caregiver Leave, and Other Measures for the Welfare of Workers Caring for Children or Other Family Members).[4] This legislation allows a mother to take leave until the first birthday of her child if her tenure is one year or longer.[5] As an exception, this legislation was not applied to the establishment with 30 employees or less, and this establishment size restriction is abolished in 1st April 1995, at the same timing of the second reform. Importantly, those who gave a birth before the enactment of the law are eligible for the job protection as long as they satisfy the above eligibility conditions, though they need to wait until 1st April 1992 to take the job-protected leave. The duration of the job protection was extended from 10 to 16 months in April 2005, and to 22 months in October 2017 only for the mothers whose children cannot enroll childcare centers due to capacity constraint. This study focuses on the 1992 reform, as opposed to these two policy reforms.

While the parental leave legitimated by the 1992 reform was unpaid, cash benefits were introduced in April 1995. A mother can receive the benefits during the leave period after 1st April 1995 and until the first birthday of her child if she was employed as a full-time worker for 12 months within the past two years.[6] The amount of the

[3] The Health Insurance typically covers full-time employees whose labor contract is expected to last for more than a year while not covering most part-time employees (Health Insurance Act, Article 3).

[4] This applied to establishments with 30 or more employees at this time, and extended to all establishments in April 1995.

[5] This legislation is for private employees, but quite similar laws were implemented for public employees.

[6] More precisely, the eligibility is determined by the total period covered by the Employment Insurance. A worker is covered by this insurance if his/her weekly hours worked are more than or equal to 20 hours and his/her employment is expected to continue for one month. If a mother was

benefits is a quarter of the average monthly payment during the past 6 months. In addition to income replacement, a mother is exempted from the premiums of the pubic health and pension insurance during the leave, and the parental leave benefits are not taxed (Health Insurance Act, Article 76 and Welfare Pension Insurance Act, Articles 12, 82-2, 139 (5) and (6)). According to the Annual Health and Welfare Report 1995 of the Ministry of Health, Labour and Welfare (MHLW), the premium of these insurances, on average, amount to 12 percent of monthly earnings. Therefore, the effective replacement rate is around 0.37.[7]

Similarly to the job protection policy, those who gave a birth before the enactment of the law are also eligible as long as they satisfy the eligibility conditions. As a result, the effective (or average) replacement rate depends on the birth timing and the duration of the leave period (Figure 3.1), because the leave period after April 1995 is associated with cash benefits while the leave period before the enactment is still unpaid. Thus, this policy reform does not provide sharp discontinuity but rather provides gradual transition of the policy regime, which contrasts to the parental leave system in some other countries where the eligibility is not given to those who gave a birth before the policy enactment. Since the policy regime is still discontinuous across birth months and the difference in the effective replacement rate leads to different take-up rate across child birth months, this parental leave reform serves to identify the causal impact.

In addition to the legitimated leave, some firms had their own maternity or parental leave system, but its availability was limited before the introduction of parental leave (Ministry of Labour, 1991, 1993). Ministry of Labour (1996) reports that most establishments did not provide additional leave on top of the legitimated leave; about 90 percent of establishments simply followed the legislation, whereas 5 percent of establishments provided additional unpaid leave and another 5 percent provided additional paid leave. Therefore, in contrast to the U.S., where the private parental leave system is prevalent (Baum, 2003), the parental leave policies were in effect for most firms.

3.2.2 Data

We used the Japanese Census 1990 through 2010, which takes place in October every five years. The Census data cover all individuals in Japan and collect basic demographic characteristics such as age, gender, marital status and household structure as well as employment status, occupation and industry. Additionally, a decennial census

covered by the Employment Insurance for 12 months within the past 2 years, she is eligible for the cash benefits (Employment Insurance Act, Articles 4, 6 and 61-4).

[7]One-fifth of the cash benefits are paid 6 months after the leave conditional on return to work (Employment Insurance Act, Article 61-5), but this aspect does not seem to affect the interpretation of the policy impact, as the MHLW reports that about 85 percent of mothers taking the leave returned to work even before the introduction of the cash benefits and this fraction did not change after the cash benefit policy. This high fraction is presumably because a mother is required to commit to return to work when applying to the parental leave. This delayed payment system was abolished in 2010.

collects years of education. Employment status is about the last week of the survey (i.e., from 24th to 30th in September), and either of (1) mostly worked; (2) worked besides doing housework; (3) worked besides attending school; (4) absent from work; (5) unemployed; or (6) out of labor force. The survey instruction categorizes full-time workers as "mostly worked" while part-time workers as "worked besides doing housework" or "worked besides attending school." If a mother is on parental leave, her employment status is supposed to be "absent from work."[8] Hereinafter, we called the first category as "full-time," the second and third categories as "part-time," and the third category as "on leave." We defined "employed" as the union of these three categories. The analysis sample is restricted to mothers in non-institutional households who were not schooling at the time of the survey and gave a birth at age between 16 and 45.[9]

3.3 Effect of parental leave policy

3.3.1 Identification strategy

Our first identification strategy relies on the assumption that timing of the birth is random. In particular, we compare a mother who gave a birth in 1992 or after with a mother who gave a birth in 1991, a year before the introduction of the job protection, and our focus is the employment status in the Census 2005. Since the employment status is measured in 2005, those mothers face the same labor market condition, irrespective to their treatment status, and thereby, macro economic shocks are unlikely to confound this comparison.

A major concern in this comparison is, however, the child age in 2005, which differs depending on the birth year. Since children of the treated mothers are younger than children of the untreated mothers, and since mothers with small children may shorten their work hours, the effect on the full-time employment could be downwardly biased, whereas the effect on the part-time employment upwardly biased. The one-to-one relationship between the treatment status and the child age prevents us from identifying the policy impact separately from the effect of the child age.

We address this issue by using the cohorts 5 years before, which facilitates to partial out the child age effect. Letting $Y_{t,\,a}$ be the employment status in year t of the mother with an a-year-old child, then the policy effect on a mother with a-year-old child is

$$\text{Policy effect} = \underbrace{\left(E\left[Y_{2005,\,a}|X\right] - E\left[Y_{2005,\,14}|X\right]\right)}_{\text{Policy effect + child age effect}} - \underbrace{\left(E\left[Y_{2000,\,a}|X\right] - E\left[Y_{2000,\,14}|X\right]\right)}_{\text{child age effect}}$$

[8] This category also includes employees and self-employed workers whose absence from work did not exceed 30 days up to the census date or who received or expected to receive wage or salary during the week before the census date.

[9] For the detailed definition of non-institutional household, see the following web page: `http://www.stat.go.jp/english/data/kokusei/2010/pdf/ex.pdf`

where X is some control variables, and $a \leq 13$, as the age of a child born in 1991 is 14. The essential assumption here is that the child age effect is the same across the two cohorts, and this assumption is analogous to the parallel trend assumption in the standard difference-in-differences analysis.

Given the increasing trend of the parental leave take-up (Table 3.1), the policy impact is likely to evolve overtime. Therefore, we estimate the following dynamic DID equation:

$$y_{it} = \sum_{\substack{9 \leq a \leq 18 \\ a \neq 14}} \delta_a D_i(a) \cdot T_t + \sum_{9 \leq a \leq 18} \tau_a D_i(a) + x_i' \gamma + \eta_t + u_{it}, \tag{3.1}$$

where T_t is a dummy variable that takes 1 if a mother gives a birth between 1987 and 1999, η_t is survey-year fixed effects, which absorb macroeconomic shocks in 2000 and 2005, and x_i is the collection of demographic characteristics, including the age of mothers, the number of household members, the number of children (household members aged 15 or less), an indicator for a nuclear household, an indicator for a single mother, and immigrant status. $D_i(a)$ is an indicator that takes 1 if the child is a years old. The parameter of interest is δ_a, which represents the policy impact on a mother with an a-year-old child for $a \leq 13$ and the placebo impact for $a \geq 15$ as the parental leave is not available in 1990 or before. (The mothers of 14-year-old children are the baseline.) The treatment status, T_t and $D_i(a)$, was assigned on the basis of the birthday of the first child in this analysis, to avoid using the same mother both in the treatment and control groups.

3.3.2 Estimation result

We first show the descriptive evidence to confirm the assumption that the child age effect is constant across cohorts. Figure 3.2 plots the full-time employment rate relative to the baseline cohort, i.e., $\hat{E}[Y_{2005,\,a}] - \hat{E}[Y_{2005,\,14}]$ (solid line) and $\hat{E}[Y_{2000,\,a}] - \hat{E}[Y_{2000,\,14}]$ (dashed line). This figure has two notable features. First, the solid and dashed lines have downward slopes, which is natural because a mother with an immature child is less likely to work for full time than a mother with a mature child. Therefore, the child age effect is so substantial that a simple comparison, $\hat{E}[Y_{2005,\,a}] - \hat{E}[Y_{2005,\,14}]$, cannot be regarded as the causal impact of the parental leave policy. Second, the solid line is almost identical with the dashed line before the policy introduction, but deviates from it after the policy introduction. It suggests that the child age effect is likely to be constant across cohorts; otherwise, the solid line would deviates from the dashed line even before 1992. Furthermore, consistent with the increasing trend in take-up rate, the difference between the two lines become larger overtime.

To complete the descriptive analysis, we repeat the same exercise in terms of the part-time employment and employment. In contrast to the full-time employment, the fraction of the part-time employment is decreased after the introduction of the parental leave (Figure 3.3), whereas the employment rate at each age relative to the baseline cohort is the same across the two cohorts (Figure 3.4). This finding suggests that the positive impact on the full-time employment is completely offset by the negative impact on the part-time employment, resulting in the null impact on the employment rate.

Although the graphical analysis is suggestive for the policy impact, it does not control for any demographic characteristics or provide statistical inference, and hence, we turn to the regression analysis. We estimate equation (3.1) and plot the estimated coefficients, $\hat{\delta}_a$, with their 95 percent confidence intervals (Figure 3.5). In 1992, when the policy was first introduced, the job protection policy did not affect the maternal employment status (13 years after child birth). The policy impact, growing overtime, was found to be positive on the full-time employment among mothers giving a birth in 1993 and 1994. Thus, it implies that the job-protection policy increases the full-time employment. While the size of impact appears small, 0.4–0.8 percentage points, this is likely to be attributed to be low fraction of mother taking-up the parental leave. Although we do not have the take-up information in 1993 or 1994, it would be less than the fraction in 1996, 0.048, and thus, we suspect that the job-protection has a non-negligible impact on those who taking-up the leave. The policy impact continued to grow after 1995, when the second reform was conducted. Similarly, the negative policy impact on the part-time employment grew over time. As a result of these two effects, the employment rate itself was unaffected. Therefore, the graphical findings are robustly found in the regression analysis.

Our empirical evidence implies that the parental leave policy works in favor of mothers who would otherwise once quit their jobs after childbearing and return to labor market as part-time workers, but not affecting mothers who would never return to labor market without the policy. While the size of the policy impact is at most 1.5 percentage points, it does not immediately mean that the parental leave is not quite useful for maternal career continuation. In fact, we need to consider the take-up rate to evaluate the magnitude of the treatment effect of taking parental leave, as the policy take-up is far from 100 percent. Since the formal parental leave system did not exist until 1992 and the firm-provided parental leave was not prevalent (Ministry of Labour, 1991, 1993), the policy impact on the parental leave take-up is well approximated by the take-up rate of the formal parental leave. Hence, the effect of parental-leave take-up is obtained as the Wald estimate, namely, the reduced-form estimates, $\hat{\delta}_a$, divided by the take-up rate. Calculating it using the 1996 values, it amounts to the substantial impact of parental leave; taking up parental leave increases the full-time employment by about 30 percentage points while decreasing the part-time employment by the same amount.

3.3.3 Threat to identification

Violation of the common trend assumption

Unfortunately, we found some statistically significant estimates in terms of the placebo years, though the magnitude was small. Since it potentially indicates the violation of the identification assumption, we confirmed the robustness of the results by controlling for the group-specific linear age effect. In particular, we estimated

$$y_{it} = \sum_{9 \le a \le 13} \delta_a D_i(a) \cdot T_t + \sum_{9 \le a \le 18} \tau_a D_i(a) + \beta ChildAge_i \cdot T_t + x_i' \gamma + \eta_t + u_{it}, \quad (3.2)$$

where β addresses the child age effect specific to the treatment cohorts. As β is identified by the mothers with 15–18-year old children in the treatment cohort, the placebo terms $(\delta_{15}, \ldots, \delta_{18})$ are excluded from the model.

The dotted lines in Figure 3.6 demonstrates the estimation results with and without the linear age effect. The estimated policy impact on the full-time employment became smaller while the impact on the part-time employment became larger (in absolute value) than the baseline estimates. These results are natural, because the placebo estimates in the baseline model have upward slopes both in terms of full-time and part-time employment. Since the policy impacts remain significant after controlling for the child age by a linear term, our result is not simply explained by the violation of the assumption of the constant child age effect.

Composition change in mothers

Another essential assumption for the identification of the treatment effect is that the childbirth decision of the mother is unaffected by the policy, at least during the analysis period. This assumption seems restrictive, because the parental leave policy may induce career-oriented women to have a child. In such a case, the work propensity of those who got pregnant after the reforms may be higher than that of those who got pregnant before the reforms, and this selection effect makes the counterfactual market outcomes between the treated and control non-comparable, leading to a spurious positive impact on maternal labor market outcomes. In this case, we will observe increase in the 1st births (relative to 2nd or 3rd births) particularly in region with high parental-leave take-up. We thus divided the prefectures into two groups using the take-up rate in 2000 calculated from the Census, and plotted the fraction of the 1st births by those groups (Figure 3.7). While the fraction has an increasing trend due to the decreasing fertility rate, this trend is not specific high-take-up region but common across the groups. We also conducted the DID analysis using these two groups to confirm that the selection effect is statistically and economically negligible (Figure 3.7).[10]

[10] In this DID analysis, we controlled for the birth cohort of mothers and their marital status, and the standard errors were clustered by municipality $\times$ mother's birth cohort to address serious correlation of birth outcomes.

We are aware that the DID analysis performed in this section is vulnerable to the selection problem and that it is difficult to completely negate the potential biases in this research design. As discussed in the literature,[11] we can alleviate the difference in maternal (un)observed characteristics between the treatment and control groups by focusing on a small window around the policy implementation, due to difficulty in controlling birth timing. We thus confirm the robustness to the selection issue by using 3-month and 6-month windows in the next section. Finally, we note that, aside from the selection issue, the DID analysis presented in this section is preferable in our context, because the maternal response to the policy is not immediate, unlike some European countries. The dynamic DID framework can capture change in policy impact overtime, but the regression discontinuity design misses it by limiting its focus to a small window.

3.4 Effect of parental leave policy: More robust approach

3.4.1 Reform in 1992

The preceding analysis shows the substantial long-run impact of the parental leave policy, but the identification assumption seems restrictive particularly because it rules out the possibility that the policy effect on fertility is heterogeneous depending on labor market attachment. As pointed out in the literature (Lalive and Zweimüller, 2009), this selection issue is alleviated by focusing on mothers who got pregnant without knowing the policy change but one is facing to new policy regime whereas another is facing to old policy regime. While this argument suggests regression discontinuity design as preferable identification strategy, the literature also concerns about the seasonality in births, for the RDD window is relatively wide.[12] As a result, a popular approach combines the RDD with the DID to partial out such potential seasonality. In the subsequent analysis, we apply this RDD-DID method to confirm the robustness of our empirical findings.

We first use the policy cutoff of the job protection, and then investigate that of the paid leave in the next subsection. Since it seems essential for mothers to access to the job protection immediately after the maternity leave, the mothers with a child born in February 1992 or after were the main population treated. In addition, the mothers with a child born in January were also continuously accessible to the job protection, because the Labor Standards Act prohibits employers to dismiss their employees during 30 days after maternity leave. On the other hand, a mother with a child born in December 1991 or earlier is eligible for the job protected leave from the next April, but her job is not protected until April. Therefore, the earlier the child was born, the more difficult utilizing the legitimated leave system. In this sense, the reform does not give a sharp threshold but fuzzy one, and we exploit, as

[11] For example, Lalive and Zweimüller (2009).

[12] For example, the RDD window is 3 months in the baseline specification of Schönberg and Ludsteck (2014).

a source of identification, the accessibility to the job protection depending on child birth months. In particular, we set the policy cutoff in January 1991, which is the cutoff to continuously utilize the job protection.

The identification assumption is that counterfactual maternal market outcomes (in absence of the reform) are identical between two mothers giving a birth just before and after the policy implementation (RDD assumption), or if not identical, their differences are the same as the corresponding differences of the cohort five years before (DID assumption). Given those assumptions, the policy impact is estimated from the following empirical model:

$$y_{ijt} = \beta_0 + \beta_1 D_i + \beta_2 D_i \cdot T_t + x'_{ijt}\gamma + \xi_j + \tau_t + u_{ijt}, \tag{3.3}$$

where i, j and t indicate individuals, prefectures and years, respectively, and ξ_j and τ_t are prefecture- and year-fixed effects, respectively. The vector of control variables, x_{ijt} is the same as before. An indicator variable T_t takes one if mother is reform-year cohort, and D_i takes one if the child birth month is after the threshold month, i.e., January. The sample is restricted to 3 months before and after the threshold month in baseline analysis. We also conducted the robustness check using 6-month window, which gave similar magnitude of the estimates with smaller standard errors. The outcome variable, y_{ijt}, is employment status 3–13 years after child birth. For the job protection policy, we do not analyze the impact of the job protection policy on the take-up of parental leave, for the Census was not conducted in the reform year, 1992.

To validate the RDD assumption, we compared demographic characteristics of the mothers across the child birth quarters, because the quarter of delivery could be correlated with market outcomes and such self-selection complicates the causal inference. Tables 3.2 and 3.3 describe demographic characteristics of mothers in the sample for the analysis of the unpaid and paid leave, respectively. The average age of the mothers at childbirth is around 29, and the 2nd quarter group was the oldest while the group of the 1st quarter group was the youngest by construction of these groups. To see this, suppose, for example, that all mothers gave a birth when they become 29 years old, and then, the 3rd quarter group would be younger than the 2nd quarter group by 0.25 years (or 3 months). The remaining rows report the household structure, and its difference is minor. Although the fraction of single mothers is different across birth-quarter groups, it seems natural because cumulative divorce probability is increasing in time passed from marriage or child birth. To sum up, this descriptive statistics suggests that demographic characteristics are balanced across the child birth quarters.

Table 3.4 presents the estimation result on the maternal employment 3–13 years after childbirth, and we found no impact on maternal employment, irrespective to the timing of evaluation. Although the estimate for part-time employment 3 years after childbirth with 6-month window is significant at 5 percent level, it is natural since we have 24 estimates in the table. Considering that the estimation sample consists of those who gave a birth between the 3rd quarter of 1991 and the 2nd quarter of 1992, the null (or tiny) policy impact seems consistent with the earlier DID result. In fact, the estimates for 1992 were not statistically significant in Figure 3.5 and the estimate gradually becomes larger from 1993.

There are two possibilities for the null policy effect immediately after the reform. First, it is possible that the provision of job-protected leave has no impact on the maternal market outcomes, for example because the public parental leave system crowds out private parental leave system (Baum, 2003). However, this explanation is unlikely to apply to our case because the firm-specific parental leave programs were rare when the parental leave policy was introduced. Furthermore, this explanation is not compatible with the growing policy impact in Figure 3.5. Second, the policy has no impact at the time of introduction due to low policy take-up. Although we do not have any direct evidence, we speculate that social pressure or uncertainty about the consequence of parental leave take-up lead to low take-up rate, which is observed in parental leave for fathers in European countries (Dahl et al., 2014). Given that the mothers in the estimation sample are the first cohort that has access to the formal parental leave system, non-pecuniary cost for taking up parental leave could be substantial without any “peers” taking up parental leave before.

3.4.2 Reform in 1995

We next apply the RDD-DID method to the reform in 1995. Since the paid leave reform is not “sharply” designed as well, those who gave a birth before the policy enactment can receive the cash benefits from April 1995 to the first birthday of the children (as long as work experience condition is satisfied). As a result, the effective replacement rate depends on the child birth month. For example, the mother with children born in the 4th quarter of 1994 and the 1st quarter of 1995 faced, on average, the replacement rates of 0.26 and 0.36, respectively if the parental leave is fully taken, and the difference in the effective replacement rates is widened when the parental leave is not fully taken (Figure 3.1). We set the policy cutoff in January 1995 to make the analysis comparable with the analysis job protection policy. In this comparison, the parental leave is fully associated with cash benefits in the treatment group (except for January), whereas initial months of parental leave is unpaid in the control group. The amount of unpaid benefits is not negligible: for example, if a mother gives a birth in December 1994 and starts the leave from February 1995, her first two months of the leave are unpaid, which amounts to about 3 quarters of her monthly earnings. In addition, those who are working in small establishments can continuously access to the job protection provided maternity and parental leave if they give a birth in the 1st quarter of 1995 or after, while their jobs are not protected during several months after the maternity leave if they give a birth in the 4th quarter of 1994 of before. The estimation equation is the same as before.

Table 3.5 presents the RDD-DID estimates. We evaluated the policy impact 5–15 years after child births. The effect on the full-time employment tends to be stable regardless of the age of the child. This implies that the mother taking the parental leave continues to work for full-time after childbirth, and she would found some difficulty finding or returning to full-time job in absence of the parental leave. While the estimates from 3-month window range between 0.0022–0.0049, this seems to be due to lack of precision, and indeed, 6-month window results range within this range with less variability. On the other hand, the impact on part-time is small when the child is 5 years old. Since this child is still pre-school age, those who quit her job around the childbirth may not return to labor market but take care of her child. In such a case, the policy impact on part-time employment is negligible because the affected margin is working for full-time or not working. In contrast, when the child get mature, we found negative impact on the part-time employment, possibly because those mothers start to work for part-time job after their children start schooling. Since these two effects offset each other at least in long-run, the employment rate is not affected.

While the magnitude in this analysis is smaller than the earlier DID result in Figure 3.5, it is due to difference in the baseline group. The baseline in the earlier analysis is mothers who gave a birth in 1991, a year before the introduction of the formal parental leave. On the other hand, the baseline in the current analysis is those who gave a birth between the 3rd or 4th quarter of 1994, and they have access to the formal parental leave with lower effective replacement rate than the treatment group. Thus, the current analysis focuses on the subtle margin given by the effective replacement rate, resulting in smaller estimates than the previous analysis. Since this small impact masks a substantial impact of “taking-up” parental leave due to a small difference in take-up rate between the treatment and control group, we construct the two-stage least square estimator in Section 3.4.3.

3.4.3 Short-run impact and effect of taking parental leave

Our reduced-form evidence gives the statistically significant impacts of the paid leave policy, but the estimated coefficients appear small and difficult to interpret. As discussed earlier, those small estimated coefficients do not necessarily indicate a negligible impact of taking up parental leave, and we need the first-stage estimate to infer the effect of parental leave take-up. Nonetheless, its estimation is not straightforward due to a lack of direct information on the parental leave take-up in 1994 and 1995. Instead, available from the Census 1995 is whether or not a mother is taking leave in October 1995. Though not ideal, this information still helps us infer the first-stage impact.

One way to estimate the first stage is to regard the leave status in the Census as parental leave take-up, but this approach is unfavorable because it results in an artificial difference in “take-up rate” across the birth month. As an illustration, compare a mother who gave a birth in June 1995 with a mother who gave a birth in October 1994, and consider the minimum duration of the parental leave required to be observed as leaving in the Census 1995. The former needs 2 months of parental leave (after 2 months of maternity leave), whereas the latter does 10 months. Put it differently, the latter is not counted as the take up in the Census unless the parental leave is fully taken. Furthermore, the leave status for those who gave a birth in the 3rd quarter of 1994 cannot be observed in the Census, irrespective to their parental leave duration, for their children are older than 1 year old in October 1995. As a result, the first-stage impact is over-estimated, and the degree of the bias is serious in the case of 6-month window, compared to 3-month window.

Given the first-stage estimate, we calculate the 2nd-stage estimate by dividing the reduced-form estimate by the first-stage estimate, which is referred to as the two-sample Wald estimator (Angrist, 1990; Dee and Evans, 2003). Since our first-stage estimate is upwardly biased, the resulting 2nd-stage estimate is the lower bound (in absolute value) of the 2nd-stage impact, namely, the effect of taking the parental leave. The standard errors are calculated by the delta method.[13]

Table 3.6 shows the policy impact on the leave status in Census. As expected, we obtained much larger estimate with 6-month window than with 3-month window. Given the relatively similar reduced-form estimates regardless of the window choice, the 6-month result for the first-stage seems to be severely biased. Aside from the leave status, we found suggestive evidence for positive impact on employment in short-run, which is marginally significant when using 6 month window. This positive short-run impact on employment is consistent with our view that the parental leave prevent mothers from quitting the job at childbirth. Since "employment" includes leave status as well as full-time and part-time employment, the mis-measurement of parental leave take-up does not bias the employment result as long as those who take-up the parental leave does not quit the job after the leave period, which is actually rare because the mother is required to commit to return to work when applying to the parental leave. (See footnote 7.)

Using the first-stage and reduced-form estimates, we calculated the two-sample Wald estimate (Table 3.7), which is interpreted as the effect of taking the parental leave on maternal employment. The size of the impact is non-negligible. For the full-time employment, the positive impact is between 33–75 percentage points and for the part-time employment, the negative impact is around 50 percentage points, with 3-month window. The variability of the estimates is due to the lack of precision in reduced-form estimates. We obtained more stable results by using the the 6-month window. Although the 6-month window results could be severely downwardly biased (in magnitude), we still found the sizable impact on both the full-time (about 18–20 percentage points) and part-time employment (12–23 percentage points) in long-run.

[13] In calculation of the standard error, we relied on the assumption of the null correlation between the first-stage and second-stage estimates as in Dee and Evans (2003). However, our estimates may be correlated since we used the whole population in Japan for both estimates, though we could not match individuals across years. In such cases, the standard error may be over- or under-evaluated, depending on the correlation between u_1 and u_r as well as the sign of the coefficients. If one concerns that the OLS estimate obtained by regressing y on D would be over-estimated (given β_1 and β_r are positive), then the standard error would be over-evaluated, and the statistical inference based on our assumption becomes rather conservative.

Furthermore, the size of these estimates is compatible with the estimate from the DID estimation in Figure 3.5. In 1996, the size of the estimates for full-time and part-time employment is 0.015 and -0.014, respectively, and rescaling these estimates by the take-up rate, 0.048, we obtain 0.318 for full-time and -0.303 for part-time, which are comparable to the Wald estimate evaluated at age 10 in Table 3.7.

3.4.4 Placebo test

To check the validity of our RDD-DID framework, we estimated equation (3.3) by using less policy-relevant birth quarters, the 2nd and 3rd quarter of 1994 and 1989, where the 3rd quarter is regarded as the treatment group (Table 3.10). In contrast to the positive impact on the 1st-quarter group, we did not find any policy impacts in this exercise. The estimates were not statistically significant but also did not show the systematic pattern found in the 1st-quarter group. Indeed, the estimates for the full-time employment and part-time employment are sometimes positive and sometimes negative. This finding is consistent with the policy design in which the effective replacement rate depends on the birth month of children and the duration of parental-leave take-up. For example, if a mother gives a birth on September 1994, her first 5 months of parental leave is unpaid (after 2 months of post-birth maternity leave), which makes the parental leave less affordable. Furthermore, if she is working in small-size establishment, her job is not protected between December 1994 and March 1995. Consequently, we did not find any impact on this less policy-relevant group.

3.4.5 Anticipation of the policy reform

Although the RDD-DID analysis is more robust than our initial DID analysis in terms of the identification assumption, it is potentially invalidated by the anticipation of the policy changes. If the reform were predictable and mothers could completely control the timing of the childbearing, those who gave a birth just after the cutoff date would not be comparable with those who did so just before the cutoff. In particular, such comparison leads to over-estimation of the policy impact, as the treated are likely to have higher propensity to work than the untreated in that case. Thus, similarly to the initial analysis, the identification assumption may be violated due to the self-selection.

In practice, the introduction of cash benefits seems difficult to anticipate at the time of conception if the child birth month was around the date of the policy reforms. The bill on the job protection was first deliberated in the Diet on 12th April 1991 and promulgated 15th May 1991, while the bill on the cash benefits was first deliberated on 31st May 1994 and promulgated 29th June 1994. In order to give a birth on February 1995 and start taking the leave from the next April, she needs to conceive around May 1994, namely, before the bill was deliberated in the Diet. Although it was still possible that she had anticipated the policy change on 11th March 1994, when it was submitted to the Diet, she would have had only several months to get pregnant.

However, the eligibility expansion for those who are working in small establishments was announced at the time of the reform in 1992, so this is completely predictable. Although we cannot completely negate the possibility of self-selection, we believe the RDD-DID analysis with 3-month window is relatively safe from potential biases. As discussed in Lalive and Zweimüller (2009), even if the policy change is predictable, manipulation of birth timing is difficult. If a mother try to avoid giving a birth in the 4th quarter of 1994, she has only 3 months to be conceived to give a birth in the 1st quarter of 1995. Furthermore, such a “selective” mother is likely to give a birth after February to fully receive cash benefits. In this case, the probability that she gives a birth in the 1st quarter of 1995 becomes further low. In fact, we confirmed this argument by investigating the number of births around the cutoff. Figure 3.9 shows the average number of births per day in each ten days for each fiscal year, and we have no evidence of increasing the number of births after the cut-off. Therefore, we believe that the self-selection, if any, would have a minor impact on our analysis.

3.5 Other results

In the analysis heretofore, we focused on the average effects of the paid leave policy across the maternal population, but the treatment effects could vary across household situation depending on the availability of childcare and financial support. For example, mothers co-residing with their parents may receive assistance for child care or financial support even without the paid leave. Thus, the paid leave policy can be more effective for nuclear households than for other households. To investigate this potential heterogeneity, we re-ran the analysis by each household type, restricting the sample to mothers with children born between the 3rd quarter of 1994 and the 2nd quarter of 1995. In the subsample analysis hereafter, we prefer 6-month specification as the 3-month specification tends to provide relatively imprecise estimates, which makes difficult to compare the policy impact across groups. The estimation results with 3-month window are reported in Appendix 3.A.

The subsample analysis demonstrates that propensity to take up the leave tends to be higher among the nuclear households than others, while not showing substantial heterogeneity in the long-run effects (Table 3.8). Although the estimated short-run impact was identical across the household type in terms of percentage points, it does not mean that the nuclear households have the same demand for the parental leave as other types of households, because a fraction of households eligible for the parental leave is different between the two groups. In fact, 1990 Census shows that a fraction of married women without children working for full time is 0.37 among households consists of a wife and husband while the corresponding fraction is 0.45 among households consists of a wife and husband and some other members. Thus, the eligibility rate is likely to be lower among the nuclear households than other households, and the identical estimates rather indicate higher take-up rate by the nuclear households among those eligible. In terms of long-run policy impacts, the estimates was similar across the household types, and thus, the effect of taking the parental leave is similar across household types. This finding possibly suggests that those who do not have access to informal care tend to take the parental leave, irrespective of the households types; the informal care is not crowded out by the parental leave policy.

We also investigated the heterogeneous policy impact across mothers' educational attainment for the two reasons. First, the opportunity costs of quitting jobs (or taking non-paid leave) after childbirth depends on the level of human capital. Second, mothers with high educational attainment may put higher value on time spent with their children (Guryan et al., 2008), and this preference can result in higher take-up rate among them. Given these possibilities, we divided the sample into two groups depending on whether mothers graduated four-years/junior college or not, to reveal the heterogeneity in terms of maternal education. Since the Census asks the years of education every ten years, we investigated the policy impact on the maternal outcome of 5 and 15 years after childbirth.[14]

The subsample analysis by the education group provides suggestive evidence that a main beneficiary of the paid leave policy is college graduates (Table 3.9). In fact, magnitude of the policy impacts are larger for the college graduates than for the high-school graduates in terms of both full-time and part-time employment and 5 and 15 years after childbirth. In addition, we also found this pattern with the specification with 3-month window (Table 3.12), though the difference in the estimates is not statistically significant with 6- or 3-month window. Since the wage of the college graduates is likely to be higher than the high-school graduate, the costs to quit the job seems higher as well, leading them to higher take up rates. Additionally, the college graduates may have higher take-up rate due to strong preferences for the time with children. The long-term effect of taking the parental leave also depends on labor market attachment, because those who continue their jobs after childbirth can possibly get out of labor market after 10 or 15 years. If such tendency is more common among the high-school graduates than among the college graduates, the long-run impact on the high-school graduates becomes less striking than the impact on the college graduates. Although we found a larger impact on educated mothers, this is not necessarily interpreted as evidence that the parental leave policy reduces the gender gap in market opportunities, and indeed, Stearns (2018) finds positive impact of the parental leave policy on the employment rate but finds the negative impact on promotion. Thus, further research is necessary to conclude the policy impact on maternal career.

[14] Also, due to this survey design, we used the mother with children born in 1984/1985 as control group in this analysis to eliminating seasonality in maternal employment by birth quarters. As shown in the next subsection, use of this cohort did not change our baseline results.

3.6 Conclusion

This study evaluated the parental leave policies introduced in 1990s in Japan. The estimation results from the RDD-DID estimation indicated the null impact of the introduction of the job protection, while the introduction of the cash benefits associated with eligibility expansion positively affected the maternal employment. In particular, this reform leads mothers to take-up the leave in short-run, and increases the full-time employment while decreasing the part-time employment in long-run. As a result, the overall employment rate is unaffected, suggesting that the main beneficiary of the policy is those who would quit their jobs at child birth and return to the market when their children become mature in absence of the policy. According to the two-sample Wald estimates, the long-run effect of taking the paid leave on full-time employment is more than 30 percent. Therefore, the policy impact at this extensive margin is substantial, in contrast to the intensive margin (or the extension of the duration) that is often studied in the literature. Although the RDD-DID did not find the effect of the job protection, the DID analysis using the birth year of children indicated that the job-protected leave was not immediately taken up by the mothers but gradually prevailed, and the positive policy impact of the job protection was indeed found 1–2 years after the enactment of the policy. Therefore, the unpaid leave also encourages a mother to continue her full-time job. While this study shows that the introduction of the parental leave program is useful to strengthen labor market attachment of mothers, the policy impacts on gender gaps in labor market opportunities are still understudied and future study will address this issue.

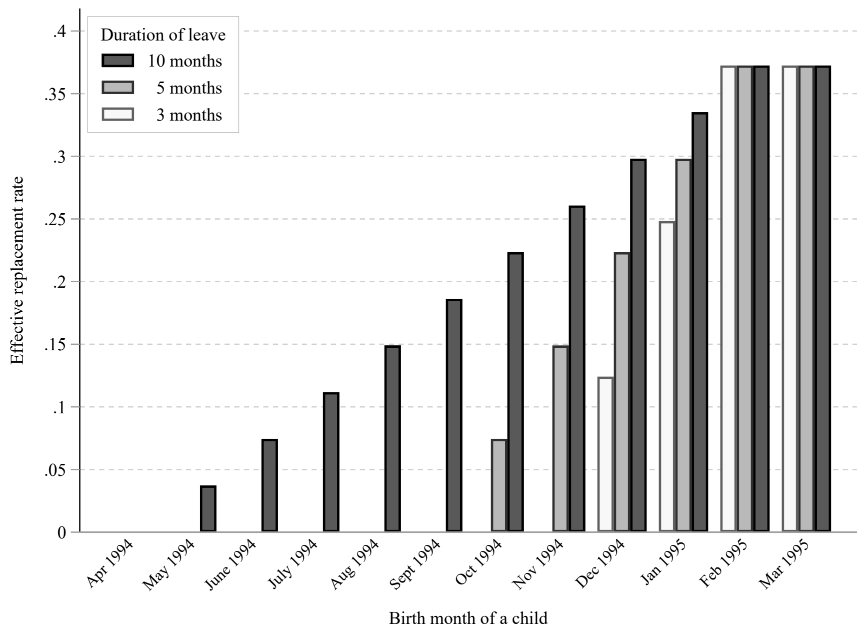


FIGURE 3.1: Effective replacement rate across the birth month of a child

Note: This figure shows the effective replacement rates when a mother takes parental leave for 10, 5 and 3 months.

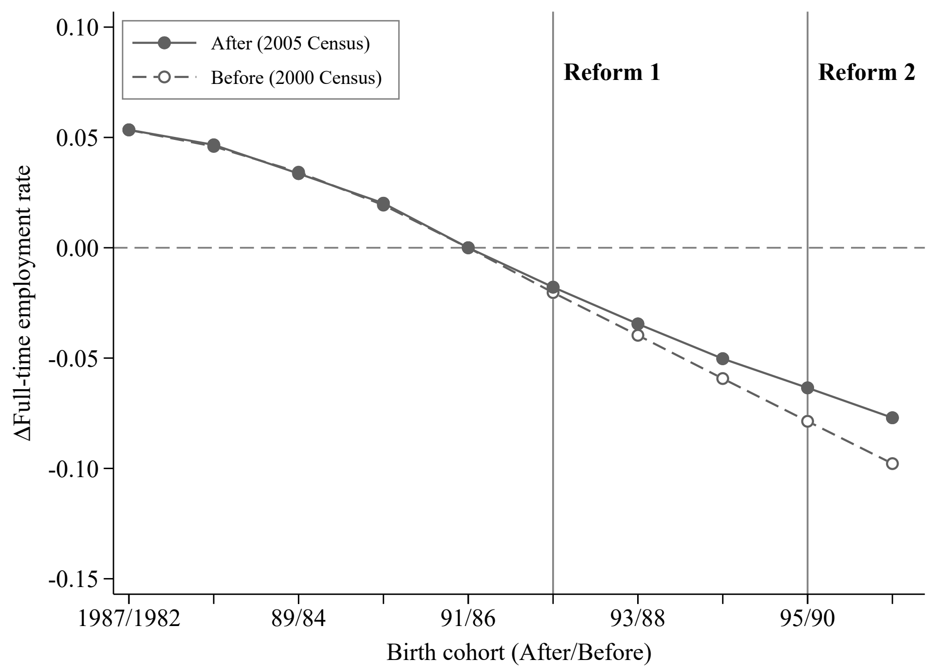


FIGURE 3.2: Trend in full-time employment rate relative to the baseline cohort

Note: This figure plots $\hat{E}[Y_{2005,a}] - \hat{E}[Y_{2005,14}]$ and $\hat{E}[Y_{2000,a}] - \hat{E}[Y_{2000,14}]$, where Y is the full-time employment and child age a is between 6 and 18.

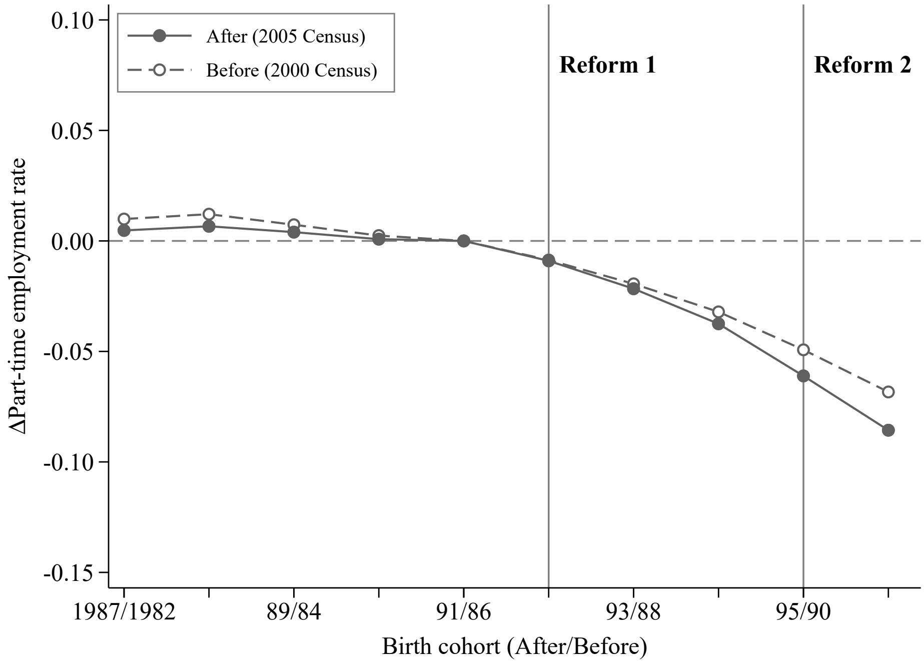


FIGURE 3.3: Trend in part-time employment rate relative to the baseline cohort

Note: This figure plots $\hat{E}[Y_{2005,\,a}] - \hat{E}[Y_{2005,\,14}]$ and $\hat{E}[Y_{2000,\,a}] - \hat{E}[Y_{2000,\,14}]$, where Y is the part-time employment and child age a is between 6 and 18.

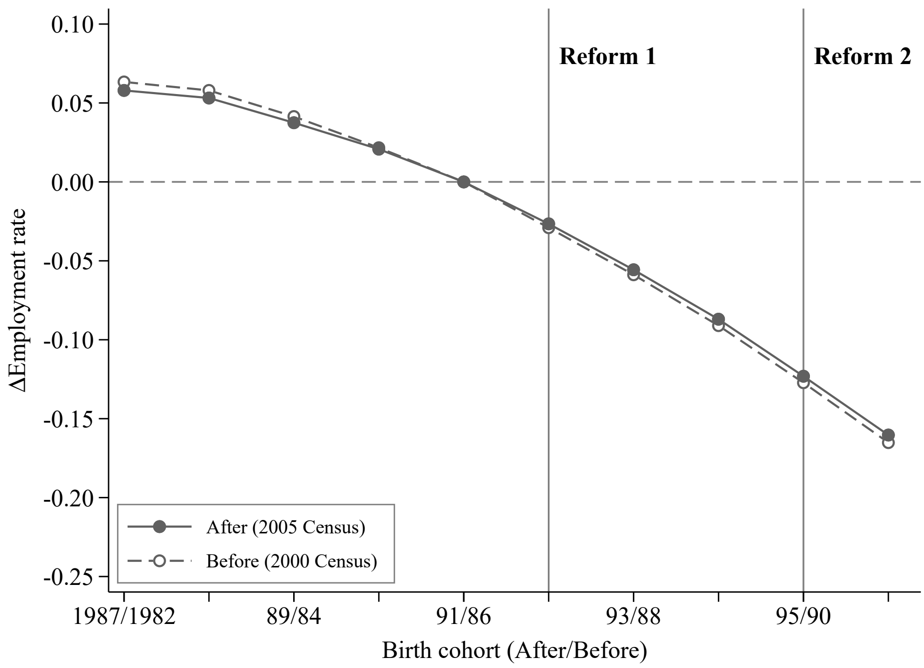


FIGURE 3.4: Trend in employment rate relative to the baseline cohort

Note: This figure plots $\hat{E}[Y_{2005,\,a}] - \hat{E}[Y_{2005,\,14}]$ and $\hat{E}[Y_{2000,\,a}] - \hat{E}[Y_{2000,\,14}]$, where Y is the employment and child age a is between 6 and 18.

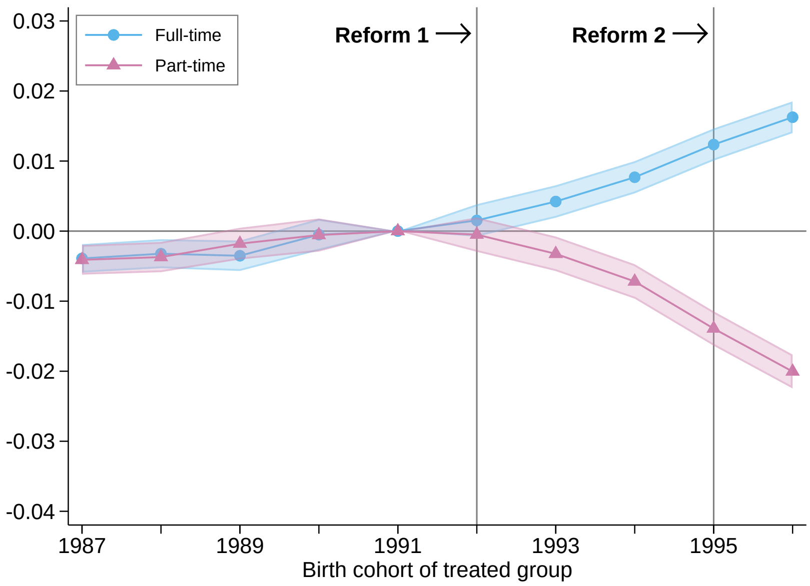


FIGURE 3.5: The effects of the parental leave policies by child birth years

Note: This figure shows the estimation result of equation (3.1). The shaded areas represent the 95 confidence confidence interval of each estimate.

Table 3.1: Fraction of mothers taking parental leave between 1996 and 1999

Year	1996	1997	1998	1999	2000
	4.8%	5.4%	5.8%	6.3%	7.0%

Data source: Monthly Report on Employment Insurance Services (*Koyou hoken jigyou geppou*), the Ministry of Health, Labour and Welfare.

Table 3.2: Maternal characteristics (Analysis sample for the reform in 1992)

Birth quarter of child	Q2 (1)	Q3 (2)	Q4 (3)	Q1 (4)
Mother's age	32.69	32.49	32.29	32.02
	[4.16]	[4.21]	[4.22]	[4.25]
N. of household member	4.68	4.67	4.63	4.59
	[1.23]	[1.23]	[1.23]	[1.23]
N. of children	2.23	2.21	2.18	2.14
	[0.73]	[0.74]	[0.75]	[0.76]
Nuclear household(%)	73.04	73.05	73.57	73.83
	[44.37]	[44.37]	[44.09]	[43.96]
Single mother(%)	3.32	3.22	3.05	3.00
	[17.90]	[17.66]	[17.21]	[17.07]
Twin(%)	0.71	0.70	0.71	0.73
	[8.38]	[8.33]	[8.40]	[8.51]
Immigrant(%)	1.13	1.18	1.18	1.25
	[10.59]	[10.79]	[10.82]	[11.12]
Observations	575473	603670	635262	613243

Note: This table demonstrates the demographic characteristics of the mothers whose children were born in 1986–1987 and 1991–1992. the census data on 1990 and 1995 were used. The standard deviations are in the brackets.

TABLE 3.3: Maternal characteristics (Analysis sample for the reform in 1995)

Birth quarter of child	Q2 (1)	Q3 (2)	Q4 (3)	Q1 (4)
Mother's age	29.92	29.69	29.49	29.29
	[4.27]	[4.31]	[4.31]	[4.34]
N. of household member	4.22	4.18	4.15	4.14
	[1.25]	[1.25]	[1.23]	[1.24]
N. of children	1.80	1.77	1.75	1.74
	[0.81]	[0.82]	[0.82]	[0.82]
Nuclear household(%)	76.56	76.85	77.63	77.74
	[42.36]	[42.18]	[41.67]	[41.60]
Single mother(%)	1.82	1.65	1.53	1.40
	[13.35]	[12.73]	[12.28]	[11.76]
Twin(%)	0.74	0.74	0.78	0.73
	[8.57]	[8.60]	[8.81]	[8.53]
Immigrant(%)	1.50	1.49	1.60	1.64
	[12.14]	[12.13]	[12.54]	[12.69]
Observations	549613	574255	600773	582612

Note: This table demonstrates the demographic characteristics of the mothers whose children were born in 1989–1990 and 1994–1995. the census data on 1990 and 1995 were used. The standard deviations are in the brackets.

TABLE 3.4: The effects of the reform in 1992 on maternal employment

Yrs since birth	3 years		8 years		13 years	
Window	3m	6m	3m	6m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Employed	-0.0011	-0.0014	0.0002	-0.0008	-0.0005	-0.0005
	(0.0017)	(0.0012)	(0.0017)	(0.0013)	(0.0017)	(0.0012)
Full-time	0.0004	0.0007	0.0005	-0.0005	0.0005	0.0006
	(0.0013)	(0.0009)	(0.0015)	(0.0011)	(0.0017)	(0.0012)
Part-time	-0.0017	-0.0025	-0.0006	-0.0005	-0.0012	-0.0013
	(0.0014)	(0.0010)	(0.0016)	(0.0012)	(0.0017)	(0.0013)
On leave	0.0002	0.0004	0.0003	0.0001	0.0002	0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
Observations	1188715	2349336	1196756	2283789	1183119	2148622

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1992. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between October 1986 and March 1987 or between October 1991 and March 1992 for 3-month window, and between July 1986 and June 1987 or between July 1991 and June 1992 for 6-month window.

TABLE 3.5: The effects of the reform in 1995 on maternal employment

Yrs since birth	5 years		10 years		15 years	
Window	3m	6m	3m	6m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Employed	0.0023	0.0019	-0.0012	-0.0003	0.0017	0.0012
	(0.0017)	(0.0013)	(0.0017)	(0.0013)	(0.0017)	(0.0013)
Full-time	0.0024	0.0030	0.0022	0.0038	0.0049	0.0034
	(0.0014)	(0.0010)	(0.0016)	(0.0012)	(0.0018)	(0.0013)
Part-time	0.0000	-0.0012	-0.0036	-0.0043	-0.0031	-0.0022
	(0.0015)	(0.0011)	(0.0017)	(0.0013)	(0.0018)	(0.0014)
On leave	-0.0001	0.0001	0.0002	0.0003	-0.0002	-0.0000
	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
Observations	1143993	2195841	1142816	2090993	1103755	1887301

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between October 1989 and March 1990 or between October 1994 and March 1995 for 3-month window, and between July 1989 and June 1990 or between July 1994 and June 1995 for 6-month window.

TABLE 3.6: The effects of the reform in 1995 on maternal employment

Yrs since birth	0 years	
Window	3m	6m
	(1)	(2)
On leave	0.0066	0.0188
	(0.0006)	(0.0004)
Employed	0.0024	0.0019
	(0.0015)	(0.0011)
Full-time	-0.0014	-0.0106
	(0.0011)	(0.0008)
Part-time	-0.0028	-0.0063
	(0.0011)	(0.0008)
Observations	1132223	2257598

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between October 1989 and March 1990 or between October 1994 and March 1995 for 3-month window, and between July 1989 and June 1990 or between July 1994 and June 1995 for 6-month window.

Table 3.7: The two-sample Wald estimate of the effect of taking parental leave

Yrs since birth	5 years		10 years		15 years	
Window	3m	6m	3m	6m	3m	6m
	(1)	(2)	(3)	(4)	(5)	(6)
Employed	0.3511	0.0995	-0.1848	-0.0140	0.2511	0.0631
	(0.2666)	(0.0675)	(0.2651)	(0.0691)	(0.2564)	(0.0679)
Full-time	0.3622	0.1592	0.3319	0.2014	0.7500	0.1819
	(0.2199)	(0.0546)	(0.2473)	(0.0631)	(0.2757)	(0.0716)
Part-time	0.0066	-0.0631	-0.5518	-0.2308	-0.4691	-0.1186
	(0.2297)	(0.0583)	(0.2644)	(0.0683)	(0.2707)	(0.0736)

Note: This table shows the two-sample Wald estimate using the 1st-stage and reduced-form estimation results. The standard errors are calculated by the delta method. In this analysis, we used mothers who gave birth between October 1989 and March 1990 or between October 1994 and March 1995 for 3-month window, and between July 1989 and June 1990 or between July 1994 and June 1995 for 6-month window.

Table 3.8: The effects of the reform in 1995 by household type (6-month window)

		Nuclear households		Other households	
		Reduced form	Wald estimates	Reduced form	Wald estimates
Years after birth		(1)	(2)	(3)	(4)
0 year	Employed	0.0005		0.0034	
		(0.0011)		(0.0026)	
	Full-time	-0.0118		-0.0113	
		(0.0008)		(0.0022)	
	Part-time	-0.0067		-0.0039	
		(0.0008)		(0.0019)	
	On leave	0.0190		0.0186	
		(0.0005)		(0.0009)	
	N	1753606		503992	
5 years	Employed	0.0010	0.0525	0.0034	0.1807
		(0.0015)	(0.0769)	(0.0026)	(0.1388)
	Full-time	0.0021	0.1099	0.0037	0.1993
		(0.0011)	(0.0581)	(0.0024)	(0.1311)
	Part-time	-0.0013	-0.0705	0.0001	0.0032
		(0.0012)	(0.0658)	(0.0023)	(0.1244)
	On leave	0.0002	0.0132	-0.0004	-0.0217
		(0.0003)	(0.0134)	(0.0005)	(0.0253)
	N	1651050		544791	
10 years	Employed	-0.0004	-0.0204	0.0004	0.0228
		(0.0015)	(0.0814)	(0.0024)	(0.1277)
	Full-time	0.0033	0.1743	0.0046	0.2482
		(0.0013)	(0.0698)	(0.0026)	(0.1403)
	Part-time	-0.0040	-0.2100	-0.0045	-0.2416
		(0.0015)	(0.0785)	(0.0025)	(0.1377)
	On leave	0.0003	0.0153	0.0003	0.0163
		(0.0002)	(0.0121)	(0.0004)	(0.0204)
	N	1560553		530440	
15 years	Employed	0.0017	0.0907	-0.0001	-0.0068
		(0.0015)	(0.0802)	(0.0023)	(0.1244)
	Full-time	0.0035	0.1855	0.0027	0.1477
		(0.0015)	(0.0807)	(0.0028)	(0.1524)
	Part-time	-0.0017	-0.0882	-0.0032	-0.1741
		(0.0016)	(0.0849)	(0.0027)	(0.1467)
	On leave	-0.0001	-0.0066	0.0004	0.0196
		(0.0002)	(0.0122)	(0.0004)	(0.0198)
	N	1417578		469723	

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between July 1989 and June 1990 or between July 1994 and June 1995.

Table 3.9: The effects of the reform in 1995 by mother's education (6-month window)

Years after birth	5 years		15 years	
	HS	College	HS	College
Employed	0.0016	0.0014	0.0022	0.0012
	(0.0016)	(0.0020)	(0.0016)	(0.0020)
Full-time	0.0021	0.0038	0.0021	0.0032
	(0.0013)	(0.0017)	(0.0017)	(0.0021)
Part-time	-0.0008	-0.0028	0.0000	-0.0022
	(0.0014)	(0.0016)	(0.0017)	(0.0021)
Observations	1500025	8871396	1266597	770451

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between July 1989 and June 1990 or between July 1994 and June 1995.

3.A Supplementary figures and tables

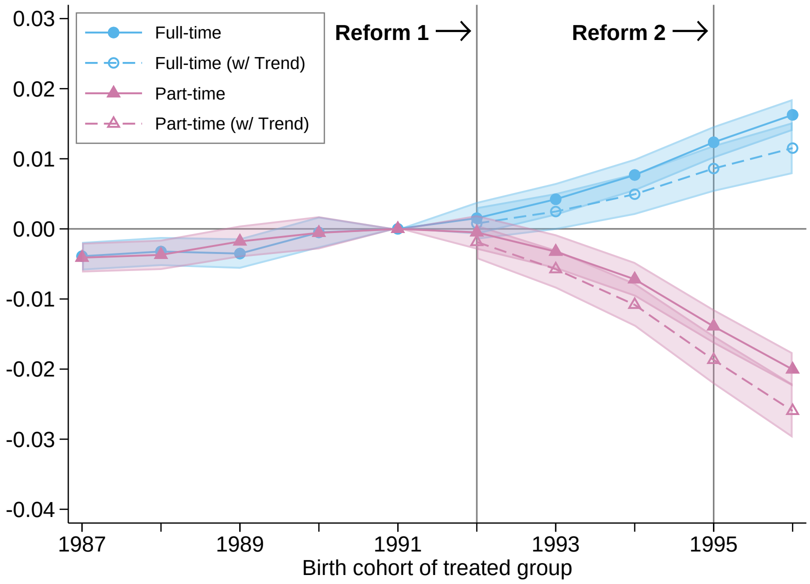


FIGURE 3.6: The effects of the parental leave policies by child birth years with trend

Note: This figure shows the estimation result of equation (3.1) with and without the cohort-specific linear age effect. The shaded areas represent the 95 confidence confidence interval of each estimate.

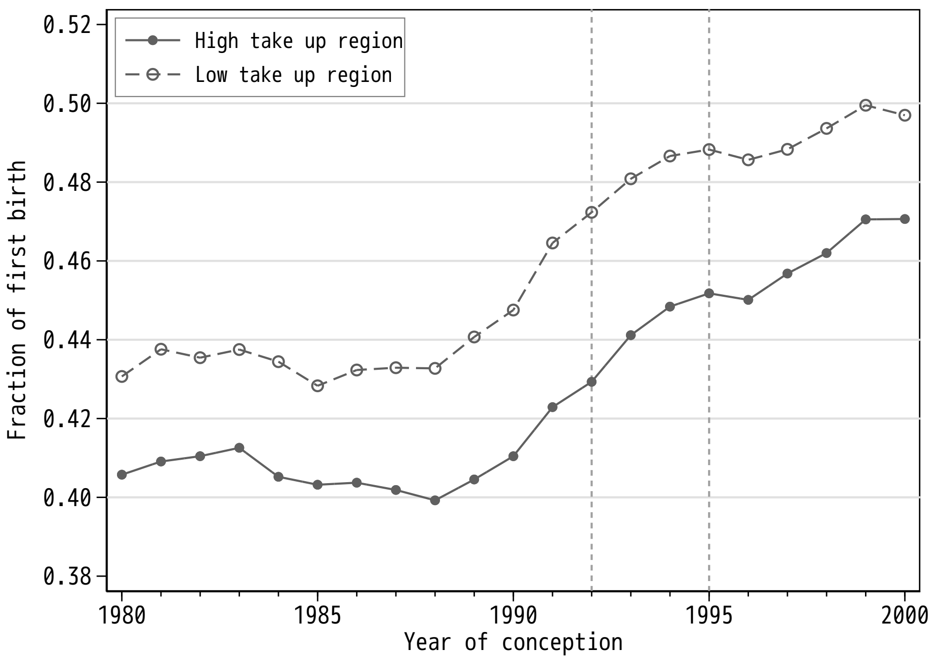


FIGURE 3.7: Fraction of the 1st births by regional take-up rate

Data source: *Vital Statistics* (Ministry of Health, Labour and Welfare)
Note: This figure shows the fraction of the 1st births. The regional group is defined by the median of prefecture-level take-up rate in 2000 calculated from the 2000 Census data.

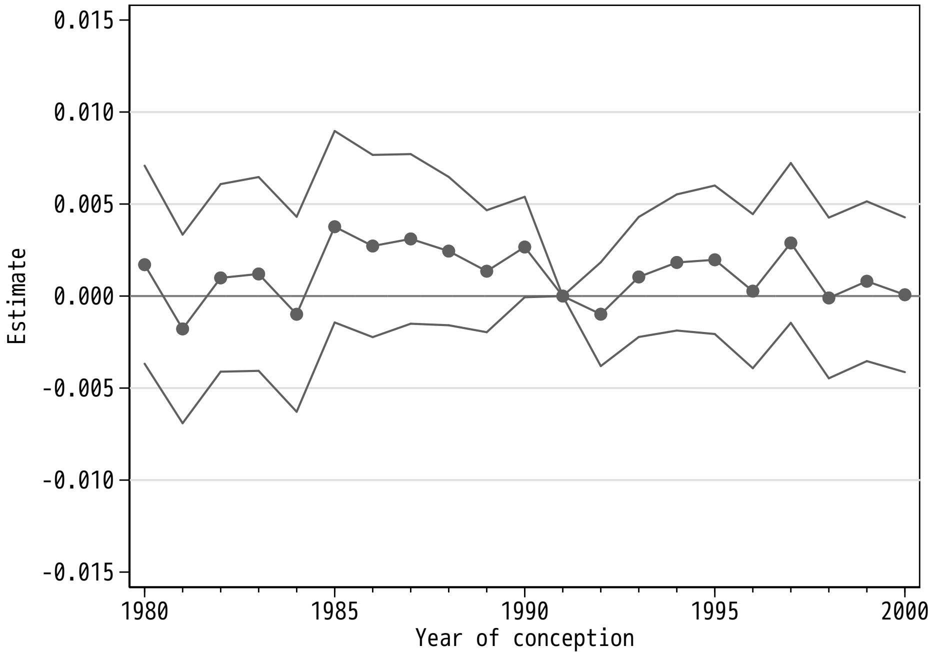


Figure 3.8: DID estimation on fraction of the 1st births

Data source: *Vital Statistics* (Ministry of Health, Labour and Welfare)
Note: This figure shows the DID estimate on the fraction of the 1st births, in which the treatment group is the high take-up prefectures and the control group is the low take-up prefectures. The unit of analysis is year-prefecture, and the base year is 1991. The control variables used in this analysis are cohort dummy variables of the mothers, dummy variables of the conception year, prefecture dummy variables, and indicator for the single mother.

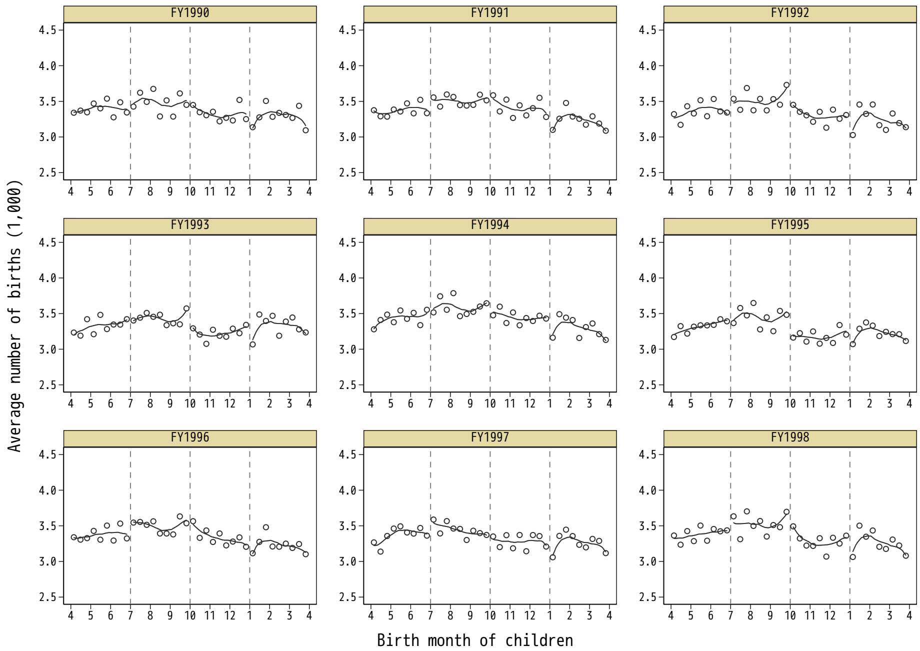


Figure 3.9: Number of childbirths around policy cut-offs

Data source: *Vital Statistics* (Ministry of Health, Labour and Welfare)
Note: This figure shows the average number of births per day for each 10 days between 1990 and 1999. Since the parental leave is taken after the 56 days of maternity leave, the number of births is calculated on the birth date plus 56 days.

TABLE 3.10: The effects of the reform in 1995 on maternal employment

	1994Q3 vs 1994Q2		
Years after birth	6 years (1)	11 years (2)	16 years (3)
Employed	-0.0012	-0.0018	-0.0014
	(0.0017)	(0.0017)	(0.0016)
Full-time	-0.0008	-0.0024	0.0009
	(0.0014)	(0.0016)	(0.0017)
Part-time	-0.0009	0.0004	-0.0020
	(0.0015)	(0.0017)	(0.0017)
On leave	0.0005	0.0003	-0.0003
	(0.0003)	(0.0003)	(0.0003)
Observations	1181893	1180752	1142872

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between April 1989 and September 1989 or between April 1994 and September 1994.

TABLE 3.11: The effects of the reform in 1995 by household type (3-month window)

		Nuclear households		Other households	
		Reduced form	Wald estimates	Reduced form	Wald estimates
Years after birth		(1)	(2)	(4)	(5)
0 years	Employed	0.0012		0.0054	
		(0.0016)		(0.0037)	
	Full-time	-0.0028		0.0014	
		(0.0011)		(0.0031)	
	Part-time	-0.0026		-0.0025	
		(0.0012)		(0.0026)	
	On leave	0.0066		0.0066	
		(0.0007)		(0.0013)	
	Observations	879541		252671	
5 years	Employed	0.0015	0.2265	0.0044	0.6735
		(0.0020)	(0.3021)	(0.0036)	(0.5582)
	Full-time	0.0016	0.2422	0.0037	0.5646
		(0.0015)	(0.2347)	(0.0034)	(0.5234)
	Part-time	-0.0001	-0.0208	0.0012	0.1750
		(0.0017)	(0.2583)	(0.0032)	(0.4919)
	On leave	0.0000	0.0052	-0.0004	-0.0661
		(0.0004)	(0.0535)	(0.0007)	(0.1002)
	Observations	867976		276007	
10 years	Employed	-0.0011	-0.1652	-0.0012	-0.1841
		(0.0020)	(0.3092)	(0.0033)	(0.5010)
	Full-time	0.0021	0.3190	0.0023	0.3447
		(0.0018)	(0.2742)	(0.0036)	(0.5473)
	Part-time	-0.0034	-0.5085	-0.0039	-0.5899
		(0.0020)	(0.3001)	(0.0035)	(0.5472)
	On leave	0.0002	0.0243	0.0004	0.0611
		(0.0003)	(0.0491)	(0.0005)	(0.0810)
	Observations	869497		273306	
15 years	Employed	0.0023	0.3449	0.0001	0.0107
		(0.0020)	(0.2984)	(0.0032)	(0.4887)
	Full-time	0.0052	0.7928	0.0034	0.5104
		(0.0020)	(0.3095)	(0.0039)	(0.5924)
	Part-time	-0.0028	-0.4210	-0.0030	-0.4556
		(0.0020)	(0.3063)	(0.0037)	(0.5662)
	On leave	-0.0002	-0.0269	-0.0003	-0.0441
		(0.0003)	(0.0483)	(0.0005)	(0.0793)
	Observations	855488		248256	

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between October 1989 and March 1990 or between October 1994 and March 1995.

TABLE 3.12: The effects of the reform in 1995 by mother's education (3-month window)

Years after birth	5 years		15 years	
	HS	College	HS	College
Employed	0.0005	0.0037	0.0018	0.0037
	(0.0021)	(0.0028)	(0.0021)	(0.0027)
Full-time	0.0004	0.0041	0.0034	0.0067
	(0.0018)	(0.0024)	(0.0023)	(0.0028)
Part-time	0.0002	-0.0006	-0.0015	-0.0030
	(0.0019)	(0.0023)	(0.0022)	(0.0028)
Observations	773251	444393	712068	427674

Note: This table shows the estimation results of equation (3.3), focusing on the effects of the reform in 1995. We only reported the estimates of interest, that is, the estimated coefficients on $D_i \times After_i$. The heteroskedasticity robust standard errors are reported in the parentheses. In this analysis, we used mothers who gave birth between April 1984 and March 1985 or between April 1994 and March 1995. The reported mean value is among the treatment group in the post-reform year.

Chapter 4

Parental leave and the gender gap in career advancement

4.1 Introduction

To promote women's labor-force participation without sacrificing family formation, most developed countries adopt parental-leave policies that either legally protect jobs or pay cash benefits during the leave. As of 2011, women are eligible for a one-year or longer paid parental leave in Austria, Denmark, Germany, Japan, Korea, and Sweden.[1] The goal of the policy is to ease mothers' career continuation and advancement after child bearing. Indeed, many studies show that introducing short-term parental leave encourages the maternal labor supply.[2] In contrast, several studies warn that generous parental leave, such as that in Nordic countries, could threaten the career advancement of higher-skilled women by placing them in less demanding jobs (Albrecht et al., 2003, 2015).

A few studies indeed report adverse effects of parental-leave policies on women's career advancements by exploiting international differences in the generosity of the policy. Blau and Kahn (2013) report that generous parental leave policy increases female labor-force participation but decreases full-time employment among women, as well as the fraction of managers and professionals, based on cross-country data from 22 Organisation for Economic Co-operation and Development (OECD) countries. Olivetti and Petrongolo (2017) report that parental-leave policies increase the employment of low-educated women but decrease the wages of high-educated women, based on cross-country data from 30 OECD countries. Although these studies suggest a potential trade-off of the policy consequences, namely, favorable effects on less career-oriented women at the cost of unintended negative effects on career-oriented women, their analyses are limited to country-level aggregated data, which prevent a detailed analysis on the heterogeneous policy impacts on career advancements across skill levels and its underlying mechanism.

[1] In addition to these countries, former communist countries, including Czech Republic, Estonia, and Slovakia, mandate a paid parental leave of one and a half years or longer.

[2] Studies include Ruhm (1998); Baum (2003); Berger and Waldfogel (2004); Baker and Milligan (2008); Lalive and Zweimüller (2009); Dustmann and Schönberg (2012); Thévenon and Solaz (2013); Schönberg and Ludsteck (2014); Carneiro et al. (2015); Dahl et al. (2016); Olivetti and Petrongolo (2017). These studies are reviewed in Kunze (2016) and Rossin-Slater (2018).

In the specific context of Sweden, Albrecht et al. (2003, 2015) attribute the large gender wage gap at the higher end of the wage distribution to its generous parental-leave policy, hypothesizing that the generous policy unintentionally hinders the career advancement of women through their human capital depreciation during the parental-leave period or triggering statistical discrimination against them. Without significant variation in the parental-leave policies within a country, however, one cannot derive a definitive conclusion about how parental-leave policies affect women's career advancement by skill level. Thus, despite the growing attention, we have scant systematic evidence on any potential drawbacks of parental-leave policies.

For a credible assessment of parental-leave policies on the career advancement of women across skill levels, we need significant variation in the policies and good measurements of the skill levels and career advancement of individual workers. To fulfill this aim, we rely on micro data from the Programme for the International Assessment of Adult Competencies (PIAAC) compiled by the OECD, covering more than 30 countries (of which we use 24 countries) that differ substantially in the length of their respective parental-leave periods. The PIAAC is the best-suited micro data set for our purpose, because it includes measurements on both skills and skill uses: The PIAAC measures the literacy and numeracy of adults based on an on-site test, as well as the frequencies of implementing certain tasks requiring a specific skill, such as reading manuals/reference sources or calculating prices/costs. The skill-use measurements enable us to construct objective measures of skill use on the job for each individual that we demonstrate succinctly and precisely measures career advancement.

This skill-use index has several strengths for measuring the career advancement of women over wages, which are the conventional proxy variable for workers' productivity. First, factors other than skill under-utilization can explain why women's wages are lower than men's wages. For example, Becker-type taste discrimination explains the lower wages of women relative to men's. Furthermore, the increase of women's labor supply induced by generous parental-leave policies depresses women's wage relative to men's if women and men are not perfect substitutes (Ruhm, 1998), or if the cost of providing parental leave could be shifted onto women's wage (Gruber, 1994). Second, our skill-use index constructed in a uniform method across countries serves as an internationally comparable measure of job assignments. Third, unlike employment or hours worked, our skill-use score measures both quantity and quality inputs. The quantity is measured by the frequency of engaging in a certain task, and the quality is measured by the type of that task. Finally, our skill-use measure corresponds exactly to the skill measure; this exact correspondence helps us avoid any seeming skill-underutilization due to the gap between the measured skill and its use.

Scrutiny of the PIAAC reveals substantial gender gaps in literacy-skill use in some countries, regardless of small gender gaps in literacy skill in most countries, suggesting that women's skill is under-utilized in such countries. In contrast, the gender gap in numeracy is not negligible in the first place, as found in previous studies (Guiso et al., 2008; Fryer and Levitt, 2010; Nollenberger et al., 2016), but the gender gaps in numeracy use are even more substantial in some countries. Thus, we find underutilizations of both literacy and numeracy in some countries, but the tendencies are similar across the two skills. Therefore, to avoid repetition, we mainly report the results based on literacy skill. We make this choice because literacy is arguably a more general skill, forming a foundation for many tasks, and literacy use is indeed found to have a stronger correlation with wages than numeracy use.

We next link each country's gender gap in skill utilization to the generosity of its parental-leave policy. To this end, we construct the measure of paid parental-leave policies from the OECD family database and the International Labour Organization (ILO) legal database. The length of paid parental leave varies substantially across countries, ranging from 0 in the US to 1.3 years in Austria, as of 2011 (and even longer for former socialist countries). We then regress skill use on the length of the parental-leave policy by skill quartiles. The analysis shows that generous parental-leave policies narrow the gender gap in skill use among the least-skilled workers (1st quartile in the skill distribution) but widens the gap among moderately skilled workers (3rd quartile). The positive policy impact on the least-skilled workers is mainly through promoting their employment, implying that parental leave helps unskilled workers who would otherwise drop out of the labor force to stay employed through job protection. In contrast, the negative policy impact on moderately skilled workers is found at both the extensive and intensive margins. This implies that parental leave discourages the skill use of moderately skilled women through driving them out of the labor force and hindering their career development. The heterogeneous impacts of parental leave on skill use are quite robust, even after controlling for other family policies, gender norms, and other labor-market institutions. According to the most-preferred specification, a one-year-longer job protection period narrows the literacy-use gender gap by 0.054 standard deviation among the lowest-skilled workers, while a one-year-longer paid leave period widens the gender gap by 0.301 standard deviation among moderately higher-skilled workers.

To demonstrate the necessity of drawing on the skill-use index as the measurement of career development, we report the analysis results using conventional measures of labor-market outcomes, such as employment status, hours worked, and hourly wages, as the outcome variables. We obtain some evidence that generous parental leave policies affect women's career development; however, most of the impacts are only imprecisely estimated, and we cannot derive definitive conclusions from these estimates. Capturing the policy impact through these conventional labor-market outcomes is difficult, most probably because these outcomes, particularly wages, are determined by many factors other than career development.

Our research design, which exploits the cross-country variation in parental-leave policies, complements the literature that draws on a natural experiment in a single country. Recent influential studies exploiting the discontinuous extension of the parental-leave period by the birth day of the child in countries credibly identifies the local effect of the extension on mothers' labor-market outcomes, but they do not identify the indirect policy impacts through the market equilibrium. For example, Adda et al. (2017) emphasize the importance of the market equilibrium in assessing family policies, because family policies could change forward-looking decision making, such as human capital investment in youth. This equilibrium effect is captured by a cross-country comparison if we assume that the observed outcome is the long-run equilibrium outcome.

As an additional benefit, cross-country comparison allows us to estimate the impact of parental-leave policies through policy take up. Previous studies that estimate the impacts of extending the parental leave length among those who already had taken up the leave tend to find minor impacts on employment and wage in the mid-run (Lalive and Zweimüller, 2009; Schönberg and Ludsteck, 2014; Carneiro et al., 2015; Dahl et al., 2016). In contrast, Lalive et al. (2014) find that job protection without cash benefits is not effective due to low take up, and Kluve and Schmitz (2018) find a positive mid-run impact on maternal employment by including the effect through take up. The latter study, in particular, contrasts with Schönberg and Ludsteck (2014), as both studies stand on German policy reforms. All these studies suggest the importance of looking at the extensive margin through take up. Thus, relying on cross-country comparisons of countries with various take-up rates is important to learn the expected policy impacts in the countries with less generous parental-leave policies.

While cross-country studies are often criticized for their vulnerability to omitted variable bias, all our specifications allow for gender-neutral country × skill-level-specific unobserved determinants of skill use. Thus, our identification does not simply depend on the cross-country variation of the generosity of parental leave; rather, it exploits the heterogeneous impacts of parental leave on skill use, conditional on skill level across genders. Furthermore, a simple omitted variable bias cannot explain our non-monotonic findings that generous parental leave promotes skill-use among lower-skilled women but demotes it among moderately skilled women.

Most previous studies pointing to a negative consequence of parental leave on women's career advancement list two probable mechanisms: human capital depreciation during the leave and statistical discrimination against women (Albrecht et al., 2003, 2015). Between the two explanations, human capital depreciation is not consistent with our findings that generous parental-leave policies decrease the skill use of women conditional on the current skill, because skill depreciation should be captured by a decrease in the current skill score. Furthermore, it does not explain the finding that generous parental leave hinders the career development of moderately skilled women but not the most-skilled women. To further test if longer parental leave entails the human capital depreciation of relatively higher skilled women, we control for the actual years of leave from the labor market, which can be calculated in our data set. We find that the actual years of leave is negatively associated with literacy use on the job but that the association does not differ between genders. Furthermore, the negative impact of generous paid-leave policies on skill use of modestly skilled women is the same regardless whether the actual years of leave from the labor market are controlled. These findings suggest that human capital depreciation is not a main mechanism for the unintended consequence of parental leave.

Our findings at least do not contradict the hypothesis that parental-leave policies strengthen statistical discrimination against moderately skilled women. Thomas (2018) sets up a theoretical model in which the firm invests in the firm-specific skill accumulation of its workers when only workers know their future labor-market attachment; she shows that the firm trains only those workers who send a costly signal for future commitment by working long hours. In the model, the introduction of mandatory parental leave encourages family-oriented workers to behave as if they are career-oriented and makes firms reluctant to train and promote female workers. An analogous argument implies that parental leave policies lower the expected value of worker's skill level by encouraging the labor-force participation of lower-skilled workers, which in turn reduces the expected returns to training, from the employer's viewpoint. As a result, the employer provides fewer training opportunities for moderately skilled workers. Thomas (2018) confirmed her theoretical prediction using hours worked at early career stages as a measure of a signal for labor-market attachment and promotion as the outcome variable. Our findings that the expansion of parental-leave policies hinders the skill utilization of moderately higher skilled, not the highest-skilled, women are compatible with the hypothesis.

4.2 Data

4.2.1 Main data set

We draw on the Programme for the International Assessment of Adult Competencies (PIAAC) by the OECD, which aims to measure adults' cognitive and workplace skills. Twenty-four countries participated in the PIAAC Round 1 (2008–2013), and 9 countries participated in Round 2 (2012–2016); participating countries in each round are tabulated in Table 4.1. Our analysis sample consists of all participating countries in Rounds 1 and 2, except for Australia and Indonesia, whose data sets are not provided for public use, and Russia, whose data set does not include Moscow residents. Accordingly, our analysis sample includes individuals from 30 countries, but 6 countries are excluded because they lack some social-institution indices (See Section 4.2.4 for those indices).[3] Hence, our main analysis sample consists of the remaining 24 countries. The survey targets individuals ages 16–65 and collects basic background information, such as age, sex, and educational attainment.

A distinguishing feature of the PIAAC is that it tests literacy, numeracy, and problem-solving skills in technology-rich environments. None of the respondents completed all three test sections; rather, they completed two at most, where the sections

[3] We obtained the German scientific-use file from GESIS.

are randomly assigned; possible combinations are "literacy and numeracy," "literacy and problem solving 1," "literacy and problem solving 2," "numeracy and problem solving 1," "numeracy and problem solving 2," and "problem solving 1 and problem solving 2." The fraction of respondents taking the problem-solving section tends to be small, because its assigning probability is lower than that of the other two test sections and because some countries opted out of it (including France, Italy, and Spain). We thus decided not to use the problem-solving section.

The PIAAC data set contains plausible values (PV), which are computed based on the test results and background information, such as sex and educational attainment (OECD, 2013). Since sex, which is the variable of interest in our analysis, is used to impute the PVs, we do not rely on those PVs, and instead calculate test scores based on Item Response Theory (IRT) by ourselves, as described in detail in the next section.

We restrict the sample to prime-age adults, those between 25 and 59 at the time of the survey, while all individuals taking the computer-based assessment are used to estimate the skill and skill-use indices.[4] We restrict the age range so that the sample construction is relatively free from school enrollment and retirement decisions. We did not restrict our analysis sample to the age range of those for whom parental-leave policies are directly relevant (e.g., 25–40), because parental-leave policies could affect the gender gaps in life-time career tracks. We exclude full-time students and the permanently disabled from the sample. Also, we exclude observations with missing values in the variables necessary for our analysis.

4.2.2 Calculation of skill and skill-use indices

To obtain test scores that purely capture performance on the examination, we construct our own test scores that depend only on the responses to questions on the examination, drawing on the IRT instead of using the built-in PVs. The way the latent score is calculated in the IRT is different from our daily grading routine, in which the allotment of marks to each question is pre-determined. In contrast, the IRT characterizes each question by its "difficulty" and "discrimination," which are estimated from test takers' response patterns.

The two-parameter logistic model of the IRT specifies the probability of making the correct response as

$$\Pr(y_{ij} = 1 \mid a_j, b_j, \theta_i) \equiv \frac{\exp\left(a_j(\theta_i - b_j)\right)}{1 + \exp\left(a_j(\theta_i - b_j)\right)}, \tag{4.1}$$

where y_{ij} takes one if respondent i correctly answers test item j, and zero otherwise, and θ_i is the latent trait of respondent i. Each test item j is characterized

[4]If a respondent does not have the basic ability to use a computer or if he/she refuses to use a computer, he/she takes a paper-based assessment (PBA). OECD (2013) suggests that the computer-based tests and the paper-based tests are comparable. Our results are robust to the use of the PBA sample.

by two parameters: a_j, the "discrimination" parameter of item j, which represents the sensitivity of being correct to the ability; and b_j, which represents the "difficulty" that shifts the probability of being correct irrespective of the ability. This specification assumes that test items measure the uni-dimensional latent trait summarized by θ_i, and that observed item responses are independent, conditional on the latent trait, θ_i. In fact, test items in the PIAAC are designed to apply this model, such that each question is independent of each other. Letting $y_i = (y_{i1}, \cdots, y_{iJ})$ and $B = (a_1, \cdots, a_J, b_1, \cdots, b_J)$, the conditional distribution for respondent i is denoted as

$$f(y_i \mid B, \theta_i) = \prod_{j=1}^{J} \left[\Pr(Y_{ij} = 1 \mid a_j, b_j, \theta_i)\right]^{y_{ij}} \left[1 - \Pr(Y_{ij} = 1 \mid a_j, b_j, \theta_i)\right]^{1-y_{ij}}. \tag{4.2}$$

Given the prior distribution of the latent trait θ_i, which is assumed to follow the standard normal distribution, $\hat{B}$ is chosen to maximize the log-likelihood,

$$\ln L(B) = \sum_{i=1}^{N} \ln \left(\int f(y_i \mid B, \theta) d\Phi(\theta) \right), \tag{4.3}$$

where Φ is the standard normal distribution function.

Finally, the latent trait parameter θ_i, is estimated using Bayes' theorem; its immediate application gives the posterior distribution of the latent trait, θ_i, conditional on the estimated parameters and response patterns. Then, the empirical Bayes mean (or posterior mean) of θ_i is

$$\tilde{\theta}_i = \int_{-\infty}^{\infty} \theta \phi(\theta \mid y_i, \hat{B}) d\theta = \int_{-\infty}^{\infty} \theta \frac{f(y_i \mid \hat{B},\ \theta)\phi(\theta)}{\int f(y_i \mid \hat{B},\ \theta)\phi(\theta) d\theta}\, d\theta. \tag{4.4}$$

We estimate the latent parameters for each country, allowing discrimination and difficulty parameters to differ across countries. To facilitate the interpretation, we normalize the estimated skill indices so that they each have exactly zero mean and one standard deviation. A set of 49 test items is used to estimate the literacy skill score, and another set of 49 test items is used to estimate the numeracy skill score.

In addition to skill possession, respondents in the PIAAC report their skill use at work with well-defined responses, which enable us to compute the latent traits for skill use. For example, they are asked, "In your job, how often do you usually read directions or introductions?" for use of literacy skill, and "In your job, how often do you usually calculate prices, costs or budgets?" for use of numeracy skill. Respondents answer these questions using a five-point frequency scale: (1) Never, (2) Less than once a month, (3) Less than once a week but at least once a month, (4) At least once a week but not every day, or (5) Every day. There are 8 items for literacy use and 6 items for numeracy use. (See Appendix 4.A for details.) These responses are more objective than responses such as "often" and "rare," because the measurement units are well defined.

Using this information, we apply the general partial credit model (GPCM; Muraki, 1992) which is an extension of the two-parameter logistic model to the polytomous items (ordered responses) to each set of skill-use items. Then, we obtain two skill-use indices for each respondent as the empirical Bayes means of the posterior distribution of latent skill-use intensity; i.e., skill use of literacy and skill use of numeracy. The skill-use indices are normalized to have a zero mean and one standard deviation.

Figure 4.1 summarizes the gender differences in skill and skill use, where each point is the gender gap of skill or skill use and the bars indicate the 95% confidence intervals. Literacy scores are roughly the same across genders, though women's scores tend to be slightly lower than men's in some countries. In contrast, the gender gaps in literacy use scores are substantially different across countries: Women use literacy more in Poland and Slovenia and use it less in Japan, Korea, Netherlands, and Norway. In terms of numeracy, women tend to score lower and use it less at work than men. From casual observation, gender gaps in skill use tend to be small or reversed in ex-communist countries, such as Poland and Slovakia.[5] de Haan (2012) documents that these countries encourage women to participate in the labor market by providing opportunities for education and training, in order to meet the demands of labor-intensive industries under socialist regimes.

Although the international variation in gender skill-use gaps is notable, the gender gaps in skill use in this figure should be interpreted with caution due to self-selection into the labor force, as only market participants are asked about their skill use at work. In the main analysis, the skill-use scores for non-participants are considered to be less than the minimum score value among market participants in each country, as their skill is actually not used in the labor market. This assignment of non-participants' skill use is justified by regarding our skill-use measure as actual skill use rather than potential skill use that would be attained if one participated in the labor market. As a result, our skill-use measure is a mixture of extensive and intensive margins, that is, participation in the market and skill-use levels in the market, respectively. We check to see the extent to which our analysis is driven by this imputation (or the extensive margin) by restricting the sample to labor-market participants.

For brevity of exposition, the following analysis focuses on literacy skill and its utilization, instead of those of numeracy. The choice of literacy over numeracy is partially based on the concern that numeracy skill is acquired by taking labor-market prospects into consideration. The usage of numeracy seems limited to market production, in comparison with the usage of literacy, which applies to both market and household production. As a result, women with high numeracy skill might differ from other women in unobserved ways, such as attitudes toward work (Guiso et al., 2008; Fryer and Levitt, 2010; Nollenberger et al., 2016). Furthermore, items to measure numeracy skill use do not seem to be as general as items to measure literacy skill use (e.g., use of algebra). In fact, Table 4.2 shows that literacy skill use is more closely correlated with wage rates than numeracy skill use. We report the results from the

[5] We define ex-communist countries as including Czech, Estonia, Poland, and Slovakia.

analyses on numeracy skill in Appendix 4.C; despite the aforementioned concerns, the relationships between numeracy skill use and parental leave are qualitatively similar to the relationships between literacy skill use and parental leave, though the relationships of the former tend to be less precisely estimated. Note that we did not choose to use both literacy and numeracy tests because only a portion of respondents take both literacy and numeracy tests and thus the effective sample size decreases significantly.

4.2.3 Parental-leave policies

We collected parental-leave policies in 2011 from the relevant laws in each country, as well as the Working Conditions Laws Database of the International Labour Organization (ILO) and the OECD family database. See Appendix 4.B for a full description of the data sources. We define the duration of parental leave as the sum of maternity and parental leave duration, in years, in a particular country. To be sure, these two policies are distinct, in the sense that maternity leave is given only to women, while parental leave is gender neutral; in reality, however, the parental leave is most likely to be taken by women in many countries.

Since parental-leave policies have two functions, job protection and income compensation, we measure these aspects by the duration of paid leave and the duration of job protection. Figure 4.6 summarizes the duration of parental leave of each country in 2011 in terms of the paid parental-leave period and the job-protection period. We confirm that the paid-leave policy has sufficient variation across countries, and many countries support substantially long job-protection periods that extend more than three years, while some of them, such as Finland, France, and Spain, provide cash benefits for less than one year.

4.2.4 Social institutions other than parental-leave policies

Since we implement cross-country comparisons that associate the length of parental leave with women's skill utilization, the correlation may be driven by gender norms or other market institutions that affect both the policy and the outcome. To control for those institutions, we construct a quantitative measure of the strength of traditional gender norms using internationally comparable social surveys: the World Values Survey Wave 6 and the European Values Study 2008.[6] We further collect other quantitative indicators for social institutions, such as tax policy[7], child care policy, the strength of employment protection, and the unionization rate from the OECD database. In addition, following Blau and Kahn (2013), we construct the indicators for right to part-time work and equal treatment of part-time workers from

[6] Both surveys asked "When jobs are scarce, should men have more right to a job than women?" with possible responses "Agree" (= 1), "Neither" (= 0) and "Disagree" (= −1). We defined the index as the average of individual responses within each country.

[7] Since characteristics of the tax system depend on the levels of earnings, the OECD evaluates it at 133% and 200% of the mean earnings of a single household. Although we collect the index evaluated at 200%, the differences associated with this choice are minor and the qualitative argument was unaffected.

OECD (2010). Since the industrial structure could affect both the policy and the outcome, we control for the fractions of public-sector employment and service sector employment, respectively, which are calculated using the PIAAC. Appendix 4.C gives summary statistics for these indices.

4.3 Validation of skill and skill-use indices

Before conducting a detailed analysis using these skill and skill-use indices, we check their validity by examining whether they are correlated with conventional proxy variables for each worker's productivity or career advancement. We restrict the analysis sample to men to abstract gender issues away and to mitigate possible selection biases. Figures 4.2 and 4.3 illustrate the relationship between the occupation-average hourly wages and literacy skill and skill use in each country, where the size of the circles indicates the number of observations in each occupation. The figures demonstrate the positive correlation in all countries, suggesting that occupations with skilled workers or intensive skill use are associated with decent wages. This positive correlation between wages and skill and skill use ensures that skill and skill-use measures carry substantive information correlated with wages, the conventional proxy for productivity.

To further confirm the correlations of skill, skill use, and log hourly wages, we estimate the following equation:

$$\ln(wage)_{ij} = \beta^s Skill_{ij} + \beta^{su} SkillUse_{ij} + X_{ij}\beta^x_j + \lambda_{s(i),j} + u_{ij}, \tag{4.5}$$

where i and j indicate each individual and country; X_{ij} include age indicators, years of education, and dummy variables, indicating that the test language is the same as the respondent's native language and that parents are immigrants; and $\lambda_{s(i),j}$ is country-occupation fixed effects, with $s(i)$ indicating individual i's occupation in country j. We estimate the model with and without $\lambda_{s(i),j}$. The estimates demonstrate, for example, that one-standard-deviation increases in literacy skill and skill use are associated with 6.0% and 9.8% increases in hourly wages, respectively, and this positive correlation is still observed after controlling for occupation (Table 4.2). These partial associations imply that skill use conditional on skill possession is worth considering to better understand the underlying mechanism of the observed gender gaps in market outcomes.

We further show that literacy use is closely related to productivity and career advancement. First, the positive relationship between wage and literacy use is observed not only at the mean but also across the distribution. In fact, their joint distribution has a high density in the diagonal region, and in particular, it is dense at the bottom-left and top-right corners (Figure 4.4). Furthermore, the literacy-use distribution is substantially distinct across occupations (Figure 4.5). Professionals stochastically dominate other occupations, followed by managers, armed forces, and technicians, whereas the elementary occupations use literacy use least frequently. Hence, our literacy-use score well reflects across-occupational differences. Finally, the positive and sizable correlation between wage and literacy use after controlling for the 2-digit occupation code suggests that the literacy-use score reflects within-occupation differences as well (Table 4.2). Since the skill-use score is constructed from tasks implemented on the job, it presumably better captures productivity or career advancement than conventional labor market outcomes, such as wages. We will demonstrate this in Section 4.6.

4.4 Parental leave and women's skill utilization

4.4.1 Literacy and literacy use in each country

Our main goal is to unpack the relationship between the under-utilization of women's skills and parental-leave policies, conditional on the current skill level, and to distinguish it from other social institutions and social norms. As a step toward documenting gender differences in skill utilization at the extensive margin, we examine the gender difference in the employment rate over skill distribution. We define those who engage in paid work or unpaid work for their own business in the week prior to the interview as those who are in employment; we also define those who are away from their job but will return, including those who are on parental leave, as those who are in employment.

Figure 4.7 illustrates the employment rate over the skill distribution in the sample countries, which are ordered by the duration of the paid-leave duration. There are generally positive correlations between literacy skill and the employment rate across countries. As a whole, men's employment rate is higher than women's at any given literacy score, but the magnitude of the gender gaps varies across countries. The gaps are smaller in Scandinavian countries, such as Denmark, Finland, Norway, and Sweden, while the gaps are significant in Southern European and Asian countries, such as Greece, Italy, Japan, and Korea. In addition, the slopes of female labor-force participation profiles vary across countries. In some Northern European countries, such as Belgium, Denmark, and Sweden, skilled women are more likely to participate in the labor force than non-skilled women, while we do not find this tendency in Asian countries. Particularly in Japan, higher-skilled women are *less* likely to work.

We now repeat the same exercise focusing on the intensive margin; we examine the gender difference in literacy use by skill level among those who work. Figure 4.8 draws the relationship between skill and its use among labor-force participants. If workers and jobs are matched assortatively, based on only literacy skill and its requirement on the job, we should observe 45-degree lines for all countries. In reality, the literacy scores and utilization scores are positively associated, but the slope is less than unity.[8] Women's skill-use is less intensive than men's at each skill level in most countries, with ex-communist countries, such as Poland, as exceptions. The size of the gender gaps in skill use varies significantly across countries; for instance, the gaps are large in Austria, Chile, Japan, and Norway. The slopes are upward, indicating that those who have high skill levels tend to use their skill more frequently, but the literacy-use/literacy gradients differ across genders in some countries. Hence, we study how parental-leave policies explain gender differences in skill use by skill levels.

4.4.2 Parental leave and literacy use

Figure 4.8 suggests that there are important differences in the gender gaps in literacy use across countries. Moreover, the sizes of the gender gap in literacy use differ substantially across the various extents of literacy skill; the gender gaps are uniform, irrespective of the literacy levels in some countries, but in contrast, the gender gaps are larger at the high literacy skill level in the other countries. We now attempt to connect the gender gap in literacy use with the length of paid parental leave, paying attention to the heterogeneous gender gaps by literacy skill. To do so, we first estimate each country's gender gap in literacy use by literacy quartiles. Second, we examine the relationship between the country- and quartile-specific gender gaps in literacy use and the length of paid parental leave.

Figure 4.9 displays the relationship between the literacy-use gender gap and the paid-leave policy by each literacy quartile group, where these quartile groups are defined by each country. The gender gap is measured by the difference between the average skill-use levels of women and men, and thus, the negative value indicates that women tend to use less skill than men. In addition, we exclude ex-communist countries, because their social institutions are different from those of other countries (de Haan, 2012). We include these countries in the main analysis, which controls for various aspects of social institutions.

Figure 4.9 shows that the relationships between the length of paid parental leave and the gender gaps in literacy use differ across the literacy quartiles. Among workers in the lowest-skilled group, the length of paid parental leave is positively correlated with the country-specific gender gap in literacy use. In contrast, among workers with higher levels of literacy (literacy levels: Q3 and Q4), the longer the length of paid parental leave, the larger are the gender gaps in literacy use. In sum, the figure shows the association of the length of paid parental leave and skill under-utilization among higher-skilled women. A caveat here is that this analysis includes only those who work; we will address this limitation in Section 4.4.3.

[8] The argument here implicitly assumes that the skill and skill-use scores have "similar" distributions. For example, if one distribution is uni-modal while another is bimodal, we do not necessarily observe a 45-degree line under perfectly assortative matching. If we transform them into percentiles, however, the argument is validated as long as both distributions are continuous. Indeed, the less-than-45-degree line is found in terms of percentiles as well.

4.4.3 Pooled country analysis

We next investigate whether the relationships observed in Figure 4.9 still hold after partialing out the effects of other institutions. We also incorporate the non-working population in the analysis. Since those who are not working use no skills for market production, their skill-use scores are lower than the lowest values observed among those in the labor force. Note that we do not attempt to measure the *potential* literacy use of non-participants that would be realized if they worked in the market. We instead measure *actual* skill use in the labor market. Hence, the skill-use indices are considered to be left-censored, where the threshold, the minimum value of literacy use among labor-force participants, varies across countries. Since the censored Tobit model takes into account non-utilized skill due to non-participation as well as skill use within the market, it captures both the extensive and intensive margins. Furthermore, we check that our estimation results are not fully explained by the extensive margin by restricting the analysis sample to those who work.

Using the Tobit model, we estimate the effect of the literacy score on literacy use by regressing the literacy-use score on the dummy variables, indicating the literacy-score quartile. We examine the difference of the relationship between the literacy score and literacy use by gender and the length of parental leave by interacting the female dummy variable and the length of leave with the dummy variables for the literacy-score quartiles. Specifically, we estimate the following model, pooling all individuals from the sample countries:

$$\begin{aligned} y_{ijs}^{*} = \sum_{q=1}^{4} 1\{q = s\} \cdot (\beta_{0q} + \beta_{1q} Female_i + \beta_{2q} Female_i \times PL_j \\ + \beta_{3q} Female_i \times Inst_j + x_i' \beta_{4q} + c_{js}) + u_{ijs}, \quad (4.6) \end{aligned}$$

and the latent skill-use level is observed if it exceeds a certain threshold;

$$y_i = \begin{cases} y_{ijs}^{*} & \text{if} \quad y_{ijs}^{*} > y_j^L, \\ y_{ijs}^{L} & \text{if} \quad y_{ijs}^{*} \leq y_j^L, \end{cases} \qquad u_{ijs} \mid Female_i, s, x_i, c_{js} \sim N(0, \sigma_j^2), \qquad (4.7)$$

where i, j, and $s \in \{1,2,3,4\}$ indicate individuals, countries, and skill quartile groups, respectively. The threshold y_j^L is the minimum of skill-use score among those who are employed in country j. The indicator function, $1\{q = s\}$, takes one if individual i's literacy skill belongs to the literacy quartile q; PL_j is the duration of the parental leave of country j measured in three ways: paid-leave length, job-protection period, and the full-replacement-equivalent length; and $Inst_j$ is the vector of institutional variables of country j, including an ex-communist dummy variable, the childcare center utilization rate, an index of the tax system, public sector size, service sector size, an index of employment protection policy, and union density. The vector x_i is a collection of individual characteristics, which include age indicators, years of education, and immigrant status. Country × skill quartile group fixed effects, c_{js}, captures the country-specific relationship between the literacy-skill quartile and literacy use.

While our sample is randomly collected from each country, the error terms of the above model could be correlated within each country due to country-specific unobserved factors. In particular, mis-specification regarding country-level variables would produce an error term common across individuals within a country. To allow for this correlation, we report the clustering of robust standard errors by the cell, defined by country times skill-quartile group.

Table 4.3 shows the estimation results of the Tobit model consisting of equations (4.6) and (4.7), using the duration of paid leave as the measurement of parental-leave policies. The basic specification in Column 1 that does not control for the gender-specific institutional term (i.e., $Female_i \times Inst_j$) except for parental leave and the ex-communist dummy variable, shows that one-year-longer parental leave narrows the gender gap in skill use by 0.15 standard deviation at the lowest skill quartile. In contrast, one-year-longer parental leave widens the gender gap in literacy usage by 0.12 standard deviation at the third quartile of the literacy distribution, though this result is not statistically significant (with p-value 0.101). Column 2 controls for the size of the public and service sectors, and it changes the estimated coefficients in a negative direction. Since the gender gap in literacy use is smaller in the service and public sectors than in other sectors, and the countries with longer parental leave tend to have larger service/public sectors, including the shares of these two sectors affects the size of the estimates. In contrast, the impact of other institutional variables tends to be relatively minor (Columns 3–5). For example, the estimated coefficients for the 2nd and 3rd skill groups range between −0.10 and −0.16, and −0.30 and −0.25, respectively. As for the first quartile, the positive impact becomes small and insignificant after controlling for a suite of social institution indicators.

The estimation results reported in Columns 1–5 are the Tobit results using the sample that includes those who do not work. Thus, the estimated coefficients capture the mixture of the extensive and intensive margins. To focus on the impact at the intensive margin, we estimate the same model by OLS, using those who work as the analysis sample. The estimation results are reported in Column 6 of Table 4.3. The estimated coefficient for the first quartile is negative and not statistically significant. At the second and third quartiles, the estimated coefficients become attenuated, but the estimated coefficients remain statistically significant with about two-thirds and one-third of the full sample estimate in Column 5, respectively. This implies that longer parental leave suppresses the skill use of higher-skilled women not only at the extensive margin but also at the intensive margin. According to the most preferred specification reported in Column 5, a one-year-longer paid leave decreases literacy-skill use by 0.301 SD.

While the paid-leave duration represents some features of the parental-leave system, it may miss other features, such as the job protection. Indeed, the job-protection duration has some variation independent of the paid-leave duration (Figure 4.6). Hence, we rerun the same analysis but using the job-protection duration instead of the paid-leave duration as the measurement of parental-leave policy, as reported in Table 4.4. While we find a negative impact on women in the second and third skill groups, in terms of both the intensive and extensive margins, the magnitude is smaller than the paid-leave impact (Columns 5 and 6). Given that the legitimated duration of job protection tends to be longer than that of paid leave, it might be explained by the actual take-up behavior, in which some mothers may fully take up the paid leave but not the unpaid leave. Another result specific to the job-protection duration is the positive impact on women in the least-skilled group. According to Column 5 of Table 4.4, a one-year-longer job protection period increases literacy-skill use by 0.054SD. This result is reasonable, because job protection enables women with weak market attachment to continue their jobs. Consistent with this explanation, the positive impact disappears when the sample is restricted to those employed, as reported in Column 6 of Table 4.4.

Overall, the impact of parental leave on the gender gap in literacy use differs substantially across skill levels; in particular, it suppresses the skill use of moderately skilled women. Since the analysis based on the job-protection policy tends to qualitatively replicate the analysis based on the paid-leave policy, we mainly use the paid-leave policy in the subsequent analysis. The analysis results based on the job-protection policy are reported in the Appendix.

4.4.4 Addressing policy endogeneity

Because we exploit the cross-country variation in parental-leave policies, our estimates could be confounded by unobserved heterogeneity across countries. We have shown that the estimates are not sensitive to controlling for market and social institutions (Table 4.3). One may still suspect that parental-leave policies are closely correlated with gender norms, but the correlation between the duration of the paid leave and our gender norm index is moderate, 0.22. Indeed, the estimated coefficients of our interest did not change after controlling for the gender-norm index (Columns 2 and 3 in Table 4.3). We note that this is not due to an "imprecise" measure of gender norms, because it does explain international differences in gender gaps in skill use and wages.[9] Thus, the contemporaneous omitted variable bias would be of less concern.

Reverse causality is a potential threat to our identification strategy, however, because parental leave is sometimes regarded as a countermeasure against gender gaps in market outcomes. One possible scenario is that policy makers enhance parental-leave policies to meet the demand for child care induced by the increase in labor-force participation among less-skilled women; in this scenario, we would observe a positive correlation between the duration of the parental leave and skill use among lower-skilled women. Alternatively, policy makers might enhance parental-leave policies to alleviate large gender gaps between skilled men and women, which would generate a

[9] As expected, the traditional gender norm is associated with a large gender gap in skill use.

negative correlation between the duration of parental leave and the skill use of skilled women. Thus, our findings could possibly be explained by such alternative scenarios.

To address this concern, we rerun the analysis using past paid-leave policies in place of the current one. In particular, we use the duration of paid leave between 1971 and 2011 that is collected by the OECD. If reverse causality is a dominant source of the contemporaneous correlation of the policy and women's skill use, past policies should not be correlated with women's current skill use. In contrast, if the parental policy affects women's career development, the past policy variables should affect the current skill use, because the policies implemented decades ago would be relevant to some of the population who were ages 25–59 in 2011.

Table 4.5 demonstrates the estimation results. First, to confirm that generous parental-leave policies suppress the skill use of skilled women, we re-estimate our model by using the current policy variable collected by OECD (Column 2 in Panel A). Although the size of the estimated coefficient for the 3rd-quartile group is smaller than the baseline one reported in Column 1, this is driven by two countries, Finland and Norway, where our measure of parental-leave duration substantially deviates from that of the OECD.[10] Once these countries are excluded from the estimation sample, our baseline result is quantitatively replicated by the OECD measure (Columns 1 and 2 in Panel B).

Columns 3–6 examine the extent to which our estimation result is robust to the use of past parental-leave duration. We find similar estimation results using the paid-leave duration in 2001 and 1991. In contrast, the estimates with the policy variable in 1981 and 1971 are not in line with our baseline findings, but those estimates are imprecise compared to the others. Since the current population is less relevant to the policy implemented 30 or 40 years ago, the imprecise estimates seem natural. To sum up, our findings are robust to the use of past policy variables to some extent, and it seems difficult to explain those findings only by reverse causality. Of course, using lagged independent variables does not completely rule out the potential reserve causality. To overcome this limitation, we would need to rely on the exogenous change of parental-leave policy, but such an exogenous policy variation is hard to come by with an internationally comparable data set like the PIAAC, and thus implementing such an analysis is left for future research.

4.4.5 Impact of an outlier

Since our parental-leave measure has only country-level variation and is limited to 24 countries, it is possible that our estimation result is driven by a few countries. Thus, we confirm the robustness of the result against an "outlier" by conducting the baseline analysis with the sample excluding each country, one at a time. Figure 4.10 summarizes the estimation result. At the top of each panel, the estimate from the

[10] The OECD counts the parental-leave duration that is available conditional on not using public childcare centers, while we do not count it, because we do not know if other countries have a similar system.

full sample and its 95 percent confidence interval is reported, and the subsequent points and bars represent the estimates from the sample excluding a country that is labeled on y-axis. The negative impact on the second and third quartile groups is robustly found and the magnitude is stable for most cases. Thus, this figure suggests that no single country has a substantial impact on the estimation result, and thus our findings are unlikely to be driven by a small set of countries.

4.5 Effects on other skill use

This section extends our findings on literacy use to other workplace tasks and skill use. Career advancement seems to involve non-routine tasks, for example, adapting to new environments or collaborating with others. Related to these tasks, the PIAAC asks about the frequency with which the respondent learns new work-related things from co-workers or through learning-by-doing, as well as the frequency of influencing others via instruction, presentations, advice, or negotiation. In addition, the survey also asks the frequency with which the respondent engages in writing tasks. We thus, using these items, construct skill-use scores for learning, influence, and writing by the Generalized Partial Credit Model (See Appendix 4.A for the full set of items used to construct each skill-use score).

Estimation of equations (4.6) and (4.7) demonstrates that generous paid-leave policies reduce the learning opportunities of moderately skilled women, as well as the frequency with which they influence others (Table 4.6). We also found that they complete writing and numeracy tasks less often. A variety of task-related skill-use measures help generalize our finding on literacy-skill use to the broader dimension of cognitive and non-cognitive skill use. Thus, these findings reinforce our claim that generous parental-leave policies unintentionally result in the under-utilization of women's skill among moderately skilled women.

4.6 Effects on employment, hours, and wages

To highlight the benefit of using the skill-use index as the measurement of career development, we estimate the policy impact using conventional market outcomes, the employment, hours worked, and hourly wage. Table 4.7 tabulates the estimation results of the same model using employment status, hours worked, and hourly wage as the dependent variables. Column 1 shows that the paid-leave policy does not have a significant impact on employment in either the economic or statistical sense. Here it is worth noting that using the length of job protection instead of the length of paid leave shows that a longer protection period promotes the employment of the lowest skilled women, as shown in Appendix Table 4.4. This result suggests that job protection plays a relatively more important role for the employment continuation of less skilled women, who presumably weakly attach to the labor market, as discussed before.

Column 2, in contrast, shows that the policy prolongs work hours by about 2 hours among working women with the 1st, 2nd, and 4th quartile literacy skill. This result is consistent with the notion that a generous parental-leave policy enables women to stay on full-time jobs. The estimated coefficients for the wage equation reported in Column 3 of Table 4.7 are negative for all quartiles, but the effects are only imprecisely estimated.[11]

The nuanced results from these conventional outcomes suggest that they are not as informative as the skill-use score in measuring the degree of career development. For instance, "employment," as a binary variable, does not have any information about the tasks in which he/she engages, and similarly, the hours worked measures the quantity of labor input but not its quality. Although the wage rate could be seen as productivity in the perfectly competitive market, this one-to-one relationship does not hold in reality for various reasons, including discrimination, monopsony, search friction, internal labor-market consideration, collective bargaining, and labor market interventions, such as the minimum wages. Our skill-use score, in contrast, summarizes both quantity and quality inputs; the quantity is measured by the frequency to do a certain task, and the quality is measured by the content of that task. Furthermore, since the items used to construct the score are directly related to the production process, it seems less sensitive to market structure, as long as we focus on engagement in the production process.

4.7 Discussion

4.7.1 Selection into jobs as a mediator

We now shed light on the mechanism between parental leave and skill utilization by examining the impact of parental leave on the selection into jobs. Blau and Kahn (2013) suspect that generous family policies push women out of management occupations and professional occupations, except for teachers and nurses. More generally, generous family policies may lead women to routine task-intensive jobs, because employers' costs of parental leave vary across occupations or jobs; a routine worker can be easier to substitute with other workers while she is on leave.

To examine the extent to which occupation choice explains the gender skill-use gap, we rerun our baseline analysis with 2- or 4-digit occupation codes added as control variables and with the sample restricted to the employed population. Since the detailed occupation code is unavailable in 6 countries, they are excluded from the estimation sample.[12] As a result of dropping one quarter of the sample countries, the estimated coefficients seem somewhat different from the original ones, but they still show a negative impact on moderately skill women (Column 6 in Table 4.3

[11] In this analysis, we used the Heckman sample selection correction method without any variables excluded from the wage equation, and hence, non-random sample selection issues if any, are unlikely to be mitigated. In fact, the resulting estimate was almost identical to the OLS estimate in Panel B.

[12] Austria, Estonia, Finland, Ireland, Sweden and United States.

and Column 1 in Table 4.8). Columns 2 and 3 demonstrate the estimation results with 2- and 4-digit occupation codes, respectively, and we examine the extent to which the occupational choice or segregation explains the observed impacts.[13] If the negative policy impact on moderately skilled women is driven by the occupation choice, it is expected to disappear after controlling for occupation. The occupational choice explains most of the negative impact on the 1st quartile group, and about one half of the impact on the 2nd quartile group. The estimate for the 3rd group also shrinks after controlling for occupational code, though it is not significantly different from the baseline.[14] At the same time, however, the negative impact on moderately skilled women is not completely explained by the occupational choice or segregation, and thus, a non-negligible fraction of the negative impact is attributed to the within-occupation gender gap, which is consistent with findings of the gender wage gap study by Goldin (2014).

4.7.2 Potential mechanisms

We discuss the potential mechanisms behind the robust finding that generous parental-leave policies affect the skill-use of women heterogeneously across skill levels. Parental-leave policies could affect women's skill utilization through at least three channels; 1) job protection, 2) human capital depreciation, and 3) statistical discrimination. First, the job protection provided by parental-leave policies allows women who would otherwise drop out of the labor force to continue working. Considering that lower-skilled workers generally have weaker labor-force attachment than higher-skilled workers, the job protection would be effective for lower-skilled workers. Indeed, job-protection policies narrow the gender gap of literacy use at the bottom skill quartile, but this effect disappears when we focus on the employed population (Column 6 in Table 4.4). We observe a similar tendency in terms of other market outcomes. Thus, the positive effect of parental-leave policies on the least-skilled group is arguably driven by job protection, while other mechanisms could also work among those in the market.

The second possible channel is human capital depreciation during parental-leave periods. This effect is likely to be relevant among higher-skilled women and may seem consistent with our empirical evidence. We should note, however, that we examined skill utilization conditional on the *current* skill level measured by the literacy-skill score, which is a novel feature of our approach compared with the approach that relies on a proxy for the skill level at a fixed point in time, such as educational attainment. In this setting, skill depreciation caused by parental-leave policies should be reflected in the conditioning variable and should not affect skill use conditional on the skill level. Regardless of this feature, one might argue that unobserved skill, not measured by skill score, may depreciate at different rates across the skill distribution and that those depreciated skills lead to a gender gap in skill use.

[13] Relatively uncommon occupations (those with less than 50 observations in the sample) are grouped into a single category.

[14] These results are basically replicated by using the job protection policy (Table 4.5).

To address this reasonable concern, we attempt to directly capture skill depreciation by constructing a new variable that captures the years absent from the labor market after school graduation, drawing on the actual years of labor-market experience included in the PIAAC. We estimate the basic model controlling for the years absent from the labor market to see how the estimated coefficients of interest change. If human capital depreciation is a prime mechanism, the coefficient on "Female×PL×Literacy skill" should attenuate substantially. The estimation results reported in Table 4.9 show that the estimated coefficients do not change much after controlling for years absent from labor market, at least as far as the 3rd quartile group is concerned. In terms of human capital depreciation terms, an additional year absent from the labor market reduces literacy use by 0.05 standard deviation, while that "penalty" tends to be higher among skilled women than the corresponding men.[15]

Finally, generous parental-leave policies can potentially encourage employers to statistically discriminate against a certain type of women. To illustrate the basic mechanism, we briefly summarize the model of Thomas (2018). Suppose that an employer chooses a worker to offer training opportunities without observing her future labor-market attachment (particularly after childbearing), but she can signal the degree of her market attachment with some costs. Since the returns to the training depend on the market attachment, the employer offers the training when observing a signal that is higher than a certain threshold. Without a parental-leave policy, most workers quit their jobs after childbearing due to the substantial costs of supplying labor while taking care of their children, and thus, they have no incentive to send costly signals, whereas strongly attached workers can afford the signals and collect the returns after childbearing. As a result, the signals are informative to distinguish the strongly attached workers from weakly attached ones. Maternity-leave policies lower the cost of the labor supply after childbearing, thereby preventing workers with weaker market attachments from getting out of the market (but they do not work full time). Furthermore, their career consideration leads them to behave similarly to the strongly attached workers; i.e., they send the signal to get training opportunities. Consequently, maternity-leave policies make it more difficult for employers to distinguish strongly attached workers from weakly attached workers, which heightens the threshold to offer training and lowers the probability to get training opportunities, given each level of the signal. Therefore, maternity leave takes away training opportunities from some workers who otherwise would have such opportunities, particularly workers on the verge of the training threshold.

The essence of Thomas (2018)'s model is the asymmetric information setting, where the employer makes the training decision without observing the exact type of a worker, and the maternity-leave policy makes the worker's signal less informative. In

[15] This larger penalty may be because of a difference in the skill sets possessed by men and women or other reasons. Sarsons (2017), for example, suggests that the same signal brings different information, depending on whether it comes from a man or a woman; bad information from a man tends to be interpreted in a more optimistic way than bad information from a woman.

our context, the source of asymmetric information is the worker's skill level, which is presumably well related to the labor-market attachment.[16] The analogous argument to Thomas (2018) suggests that parental-leave policies allow lower-skilled workers to participate in the market after childbearing, but make skilled workers less distinguishable from others, because the expected skill level conditional on each value of the signal becomes lower due to the parental-leave policy. Although the skilled worker has an incentive to send a higher signal under the regime with the parental-leave policy, her optimal signal would not be high enough to get the same training opportunities as in the regime without the parental-leave policy, because sending the higher signal is costly. Thus, this statistical discrimination process pushes some skilled women out of career tracks, leading to the under-utilization of their skill. This explanation at least does not contradict our empirical findings on skill use, indicating that parental-leave policies have a negative impact, particularly on moderately skilled women.

To further demonstrate that parental leave suppresses women's career development through statistical discrimination, we analyze the effect of parental leave on promotion. The challenge of implementing a promotion analysis is to come up with an internationally comparable measure of promotion. As a proxy for promotion, we use the number of subordinates. As documented in Appendix 4.C, the distributions of the number of subordinates among men differ across countries but have a certain similarity (Figure 4.4). Tables 4.16–4.19 consistently show that parental-leave policies, both the period length of paid leave and employment protection, increase the probability of less-skilled women, measured in either literacy or numeracy, to have no subordinate. These findings are consistent with the notion that parental-leave policies increase the employment of less-skilled women who would have dropped out of the labor force if there had been no parental-leave policy. In contrast, the longer paid-leave period decreases the probability of higher-skilled women, measured in either literacy or numeracy, to have 25 or more subordinates. This finding confirms that generous paid leave supresses the promotion of higher-skilled women.

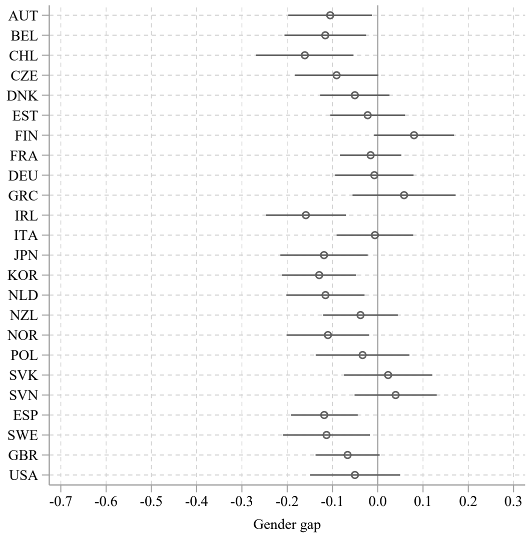
While our empirical results suggest that statistical discrimination serves as a mechanism to produce the negative impact of parental-leave policies on the career development of moderately skilled women, the full-blown test of the hypothesis would require detailed personnel panel data that records the clearly defined job rank within a company and a natural experiment, such that the change in parental policy affects only a part of workers in the company. Such a full-blown test with rich data and credible research design is left for future research.

[16] In this case, a worker is not characterized by the marginal utility of leisure, as in Thomas (2018), but by the productivity and/or returns to the training. If the productivity (or returns to the training) of some workers is in the range where they participate in the market after childbearing only under the regime with the parental leave policy, then the same prediction as Thomas (2018) is obtained.

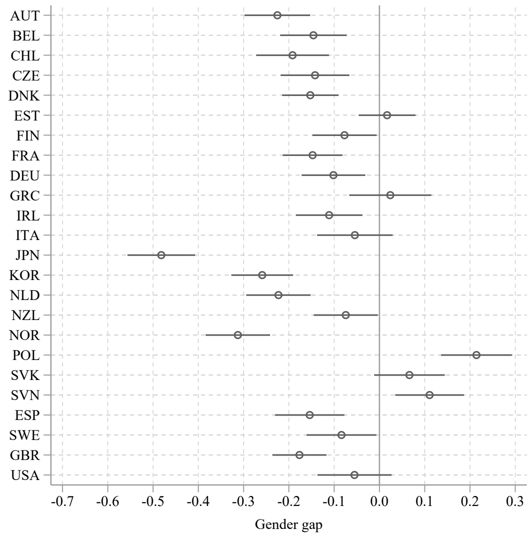
4.8 Conclusion

This study investigated the impact of parental-leave policy on the employment and career development of women relative to men through the lens of skill utilization. Drawing on the PIAAC covering 24 OECD countries, we constructed objective measures of literacy and its use on the job. Exploiting the cross-country variation in the length of paid leave and job protection, we found substantial heterogeneity in the impacts of parental-leave policy on the gender gap in skill utilization; for the lowest-skilled group, longer job protection increases the skill utilization of women through encouraging employment. In contrast, for the moderately skilled group, a longer paid leave period suppresses their skill utilization, through depressing both employment and skill use, conditional on employment. These findings are robust against controlling for other cross-country differences, such as other family policies, gender norms, and labor-market institutions. These findings suggest that expanding parental-leave policies entails a trade-off; on one hand, parental leave promotes the employment of the least-skilled women, who would otherwise drop from the labor market, while on the other hand, the policy hinders the career advancement of moderately skilled women.

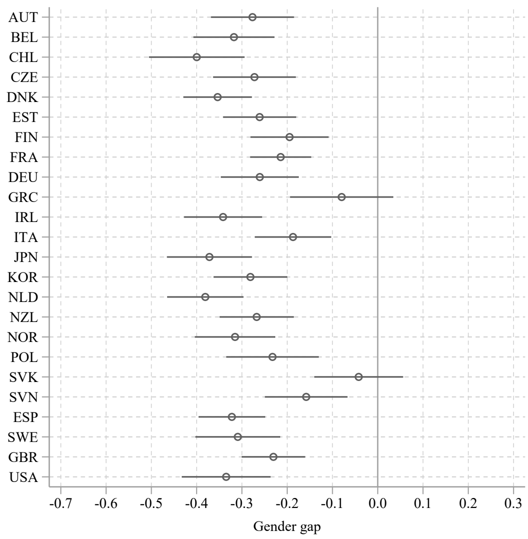
We investigated why moderately skilled women suffer from the expansion of parental leave, focusing on two plausible hypotheses: human capital depreciation during the leave and statistical discrimination by employers. Between the two alternatives, human capital depreciation contradicts our examination of skill utilization conditional on skill score, because skill depreciation should be captured by a lower skill score. Further examination based on the actual years of leave from the labor market also negates skill depreciation as an explanation. In contrast, our robust finding that women with moderately high skill, not the ones with the highest skill, suffer from generous parental-leave policies corroborates with the theoretical consequence of statistical discrimination. The full-blown test of the hypothesis, however, requires a personnel data set with a panel structure and exogeneous variation of policy treatment within a company; it is thus left for future research.



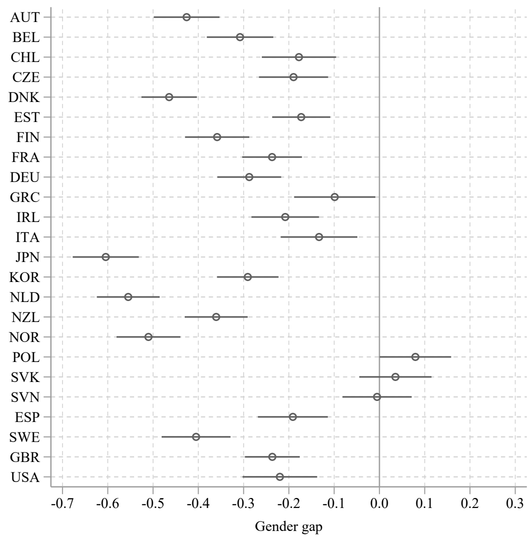
(A) Literacy skill



(B) Literacy skill use



(C) Numeracy skill



(D) Numeracy skill use

FIGURE 4.1: Gender gaps in skill and skill use

Note: This figure shows unconditional gender gaps in skill and skill use. Each point represents the gender gap, and the bars indicate its 95 percent confidence interval.

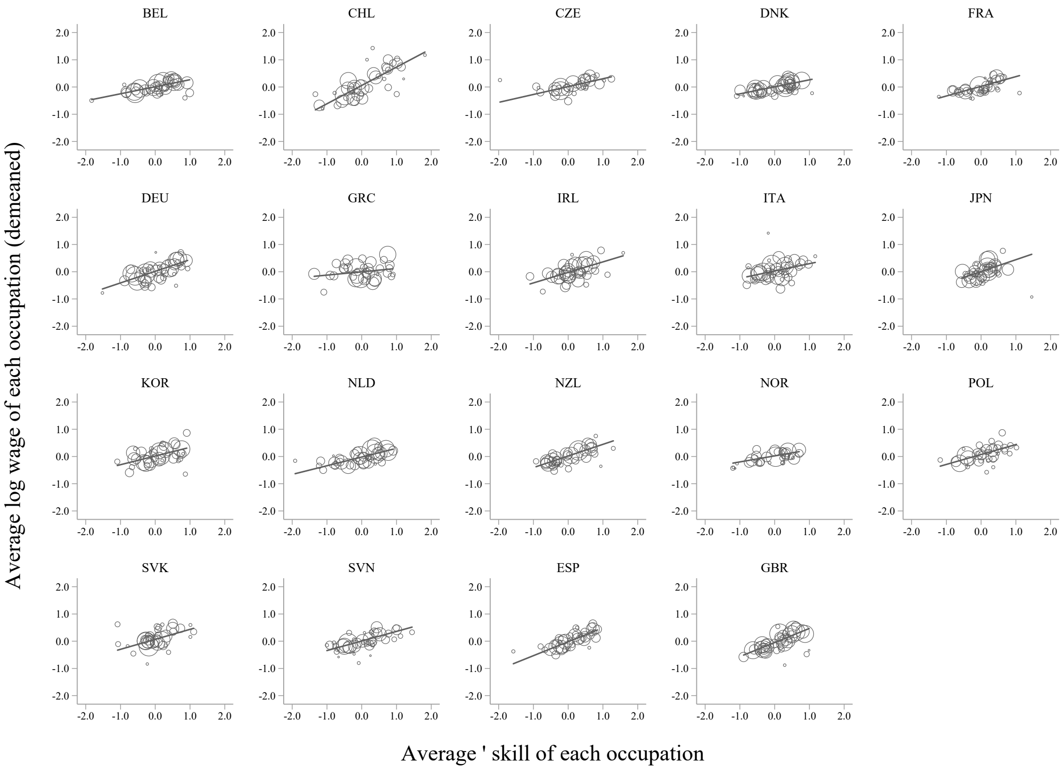


FIGURE 4.2: Occupation-average wage rates and literacy skill

Note: This figure shows the correlation between occupation-average wage rates and average literacy skill. The size of each circle indicates the number of observations engaging in each occupation. The line is the fitted value by the weighted least squares, where the number of observations in each occupation is used as a weight.

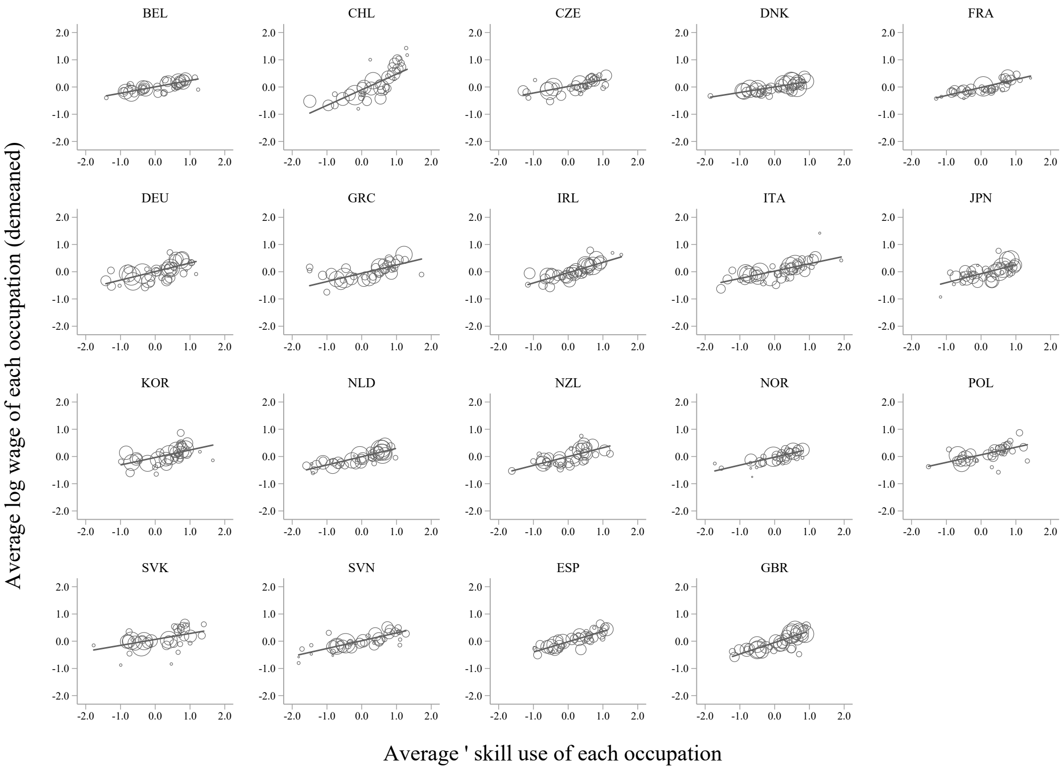


FIGURE 4.3: Occupation-average wage rates and literacy skill use

Note: This figure shows the correlation between occupation-average wage rates and average literacy skill use. The size of each circle indicates the number of observations engaging in each occupation. The line is the fitted value by the weighted least squares, where the number of observations in each occupation is used as a weight.

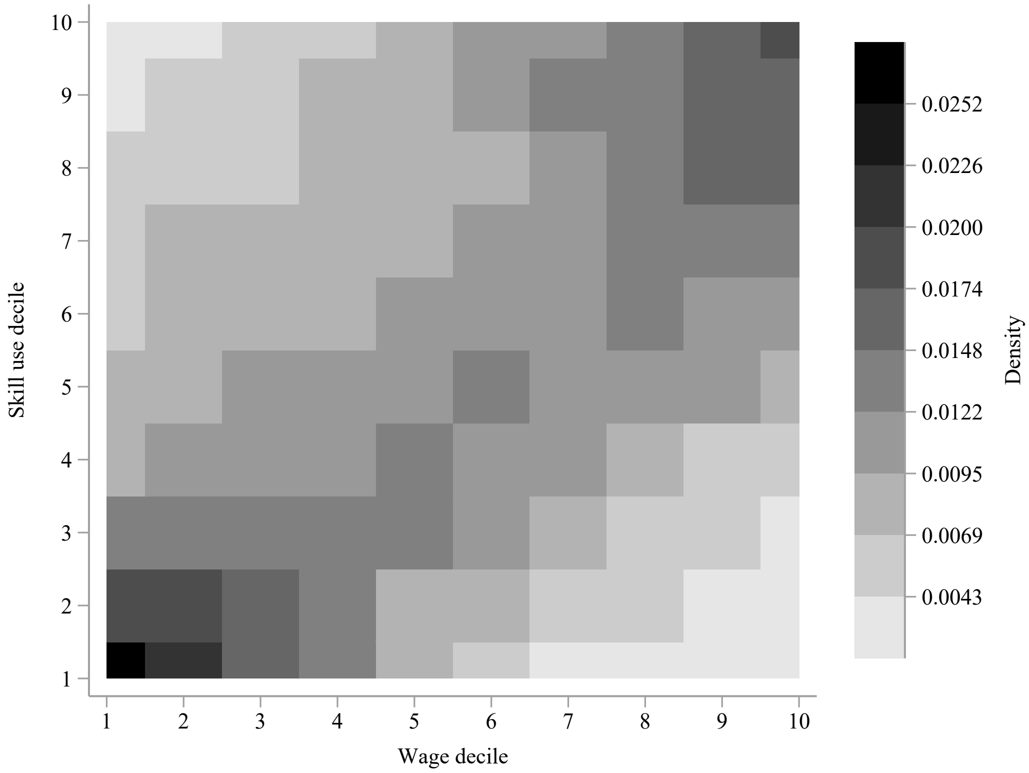


FIGURE 4.4: Joint distribution of literacy use at work and wage

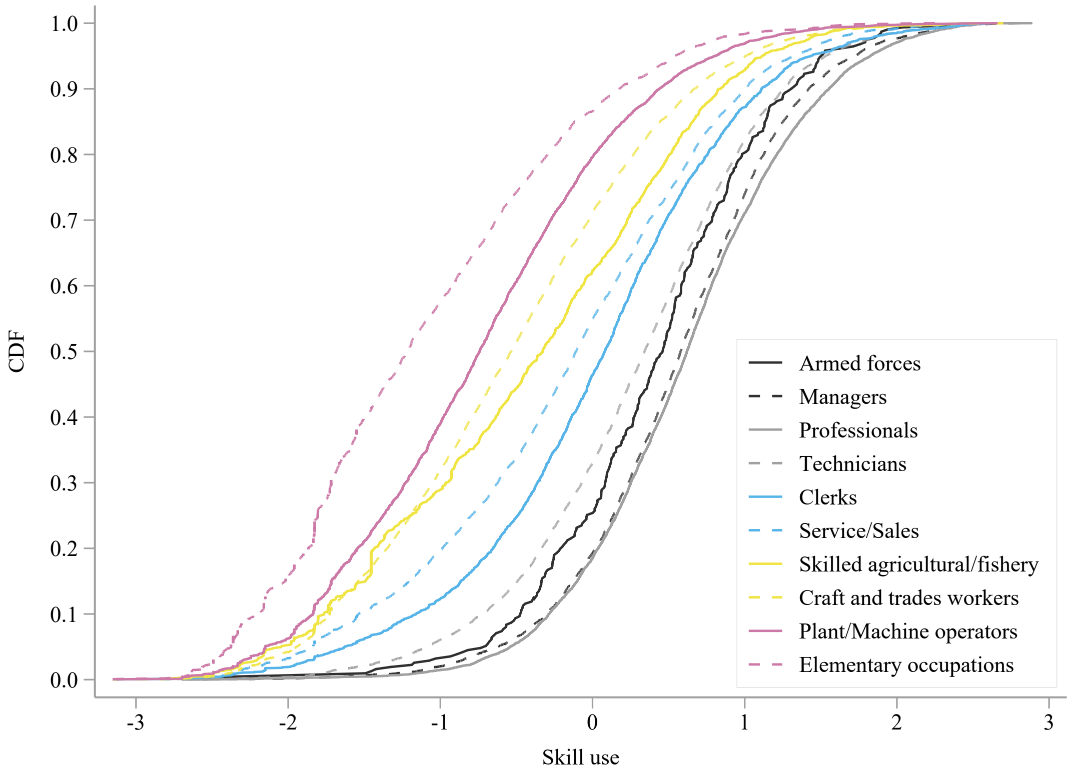
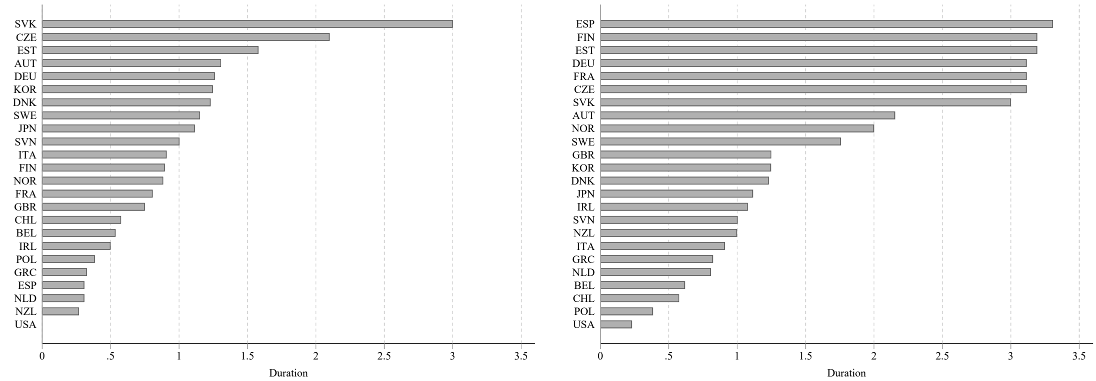


FIGURE 4.5: CDF of literacy use at work by occupation



(A) Paid parental leave in 2011 (B) Job protection in 2011

FIGURE 4.6: Summary of parental-leave policies

Data source: The Working Conditions Laws Database of the ILO and the OECD family database. See Appendix 4.B for a full description of the data sources.

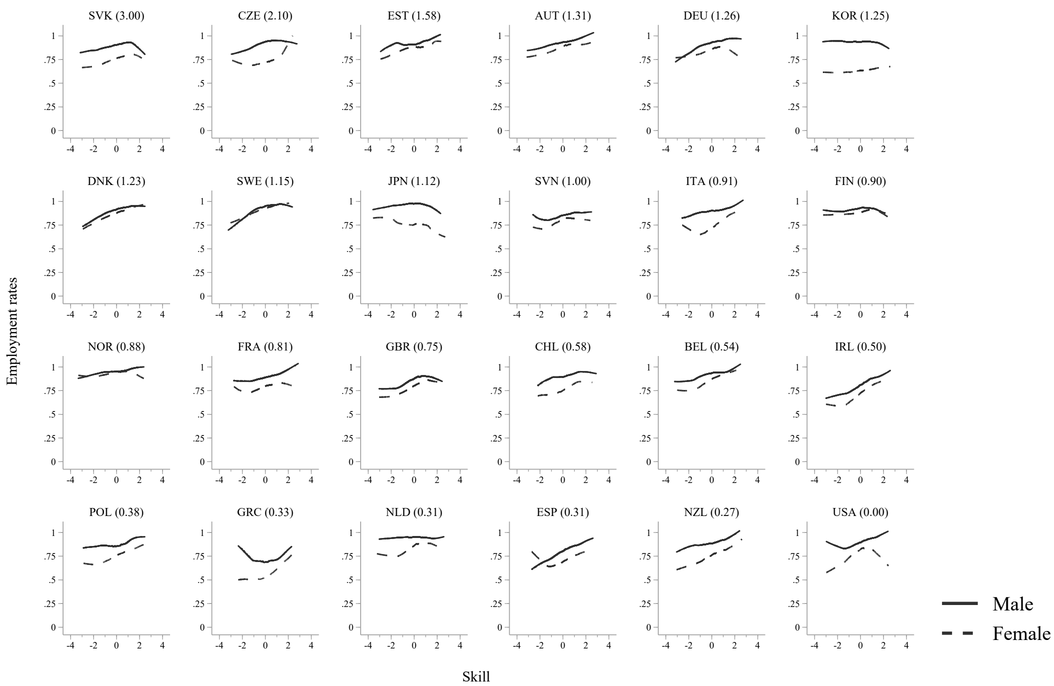


FIGURE 4.7: Literacy and employment rate

Note: Length of paid parental leave in year is in parentheses next to the country name.

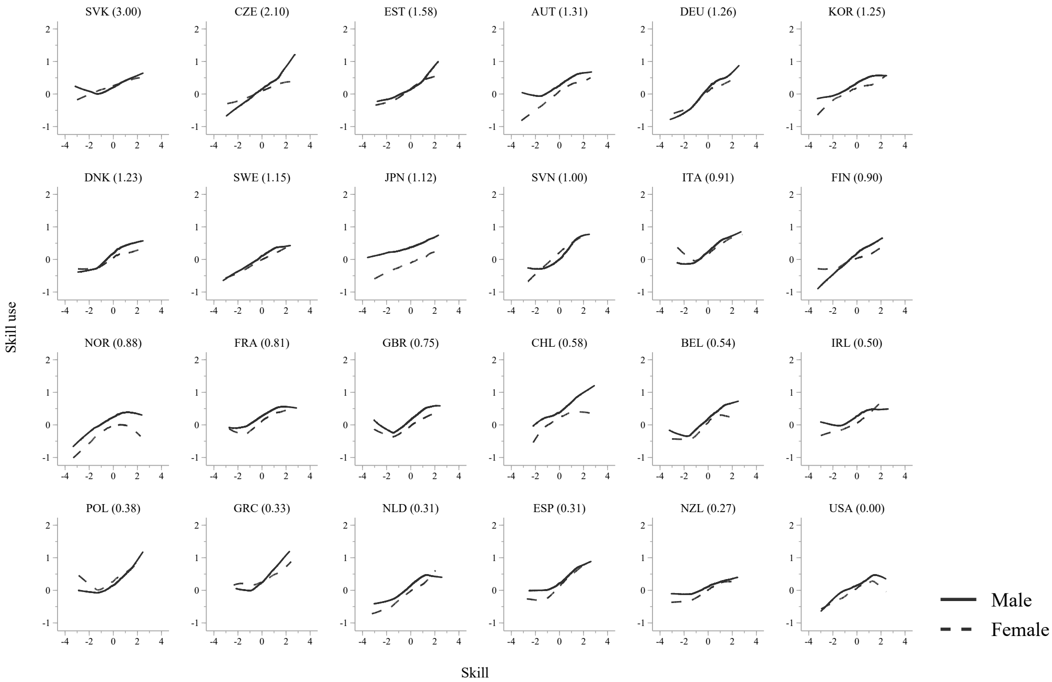


Figure 4.8: Literacy and literacy-use among labor-force participants

Note: Length of paid parental leave in year is in parentheses next to the country name.

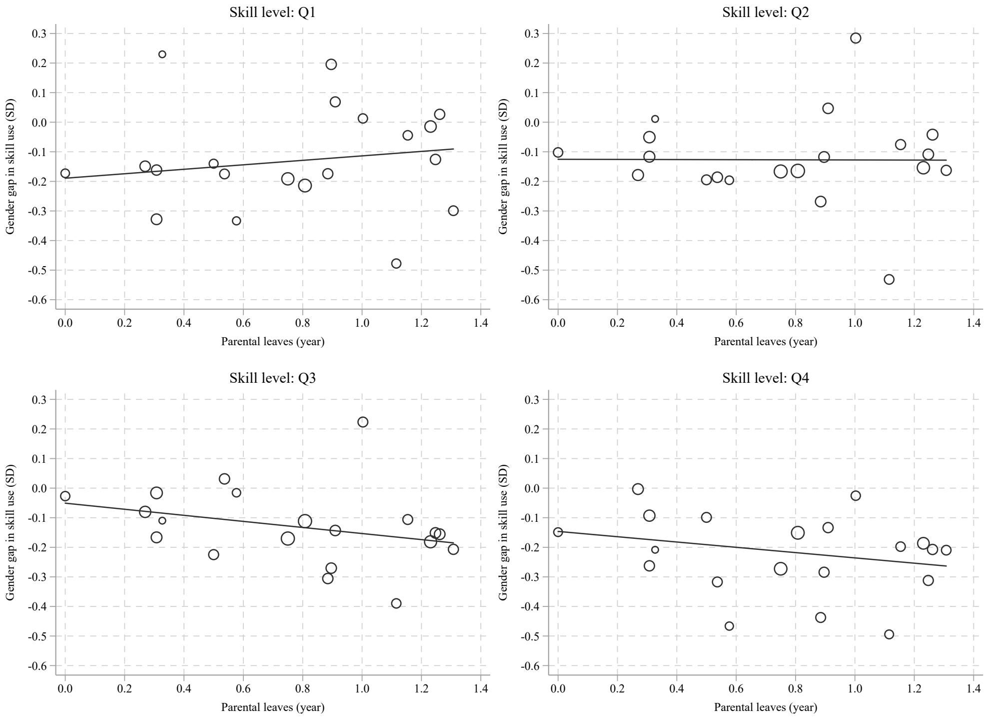


FIGURE 4.9: Unconditional gender gap in literacy skill use and the paid-leave policy

Note: This figure shows relationship between the gender gap in literacy skill use and the paid-leave policy. The gender gap in each country is calculated as a raw difference in average skill-use levels between employed women and men. The line is the fitted value by the weighted least squares, where the number of observations in each country is used as a weight. In this figure, ex-communist countries are excluded, because their social institutions tend to differ from those of other countries (de Haan, 2012).

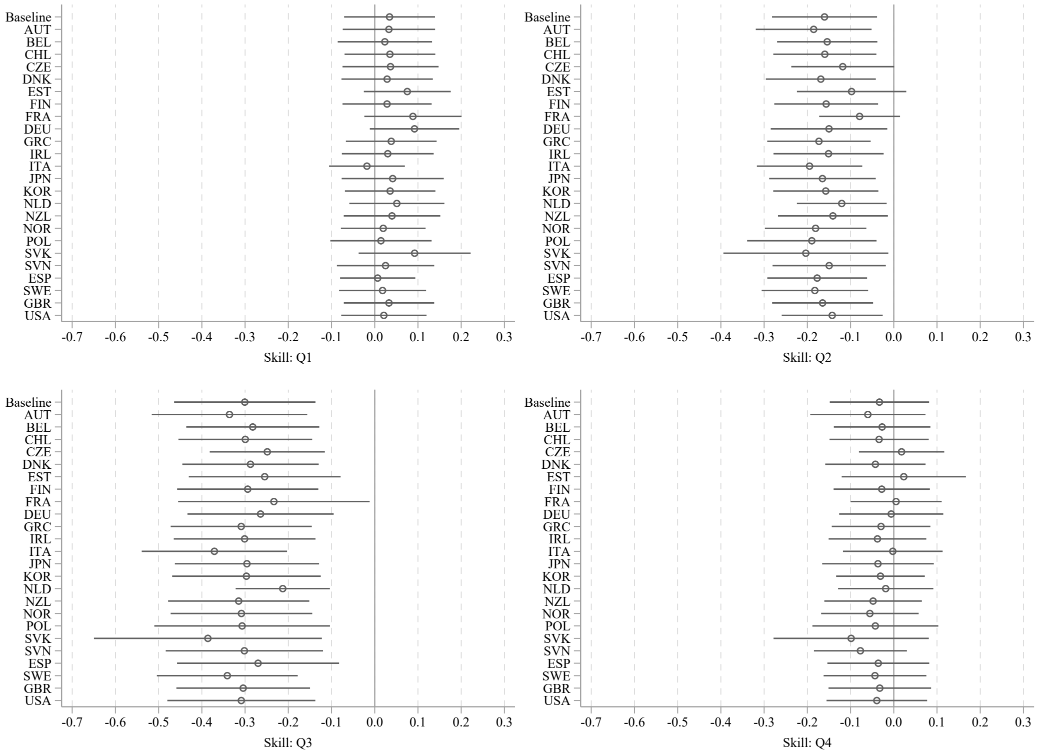


FIGURE 4.10: Impact of excluding one country from the sample on the estimates

Table 4.1: Participating countries in PIAAC

Round 1 (2008–2013)	Australia, Austria, Belgium, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Russian Federation, Slovak Republic, Spain, Sweden, United Kingdom, United States
Round 2 (2012–2016)	Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, Turkey
Round 3 (2016–2019)	Ecuador, Hungary, Kazakhstan, Mexico, Peru, United States

Table 4.2: Regression estimates of hourly wages on skill and skill use

Dep.Var. $ln(wage)$	Skill: Literacy		Skill: Numeracy	
	(1)	(2)	(3)	(4)
Skill	0.060***	0.048***	0.057***	0.042***
	(0.005)	(0.005)	(0.005)	(0.004)
Skill-use	0.098***	0.066***	0.071***	0.040***
	(0.005)	(0.006)	(0.004)	(0.004)
Occupation	No	Yes	No	Yes
Observations	12955	12790	12800	12642
Countries	21	21	21	21

Note: This table shows the estimation results of equation (4.5). We did not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language was the same as the native language of the respondent, or that parents were immigrants. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p< 0.1$, ** $p< 0.05$, *** $p < 0.01$.

TABLE 4.3: The paid leave policy and utilization of literacy skill at work

Dep.var. literacy skill use	Full sample					Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Literacy skill: Q1	0.147**	0.072	0.008	0.056	0.034	-0.043
	(0.071)	(0.060)	(0.034)	(0.045)	(0.054)	(0.049)
Female×PL×Literacy skill: Q2	-0.056	-0.173***	-0.196***	-0.171**	-0.160***	-0.095*
	(0.060)	(0.064)	(0.066)	(0.080)	(0.062)	(0.053)
Female×PL×Literacy skill: Q3	-0.121	-0.238***	-0.292***	-0.250***	-0.301***	-0.099*
	(0.074)	(0.064)	(0.064)	(0.070)	(0.083)	(0.051)
Female×PL×Literacy skill: Q4	-0.021	-0.109*	-0.095	-0.037	-0.033	0.003
	(0.063)	(0.058)	(0.059)	(0.064)	(0.059)	(0.027)
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure		X	X	X	X	X
Female×Skill×Family policies			X	X	X	X
Female×Skill×Gender norm				X	X	X
Female×Skill×Market institutions					X	X
Countries	24	24	24	24	24	24
Observations	48970	48970	48970	48970	48970	41223

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for literacy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.4: The job protection policy and utilization of literacy skill at work

Dep.var. literacy skill use	Full sample					Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Literacy skill: Q1	0.113***	0.094***	0.058***	0.054***	0.054***	-0.009
	(0.033)	(0.027)	(0.016)	(0.013)	(0.013)	(0.016)
Female×PL×Literacy skill: Q2	0.023	0.006	-0.023	-0.038	-0.041**	-0.032**
	(0.027)	(0.031)	(0.034)	(0.031)	(0.020)	(0.015)
Female×PL×Literacy skill: Q3	-0.020	-0.036	-0.076***	-0.101***	-0.106***	-0.061***
	(0.036)	(0.038)	(0.026)	(0.023)	(0.023)	(0.016)
Female×PL×Literacy skill: Q4	0.020	0.003	-0.004	-0.018	-0.020	0.005
	(0.022)	(0.019)	(0.023)	(0.020)	(0.018)	(0.017)
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure		X	X	X	X	X
Female×Skill×Family policies			X	X	X	X
Female×Skill×Gender norm				X	X	X
Female×Skill×Market institutions					X	X
Countries	24	24	24	24	24	24
Observations	48970	48970	48970	48970	48970	41223

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for literacy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates of the coefficients of the interaction terms associated with the literacy skill index and the indicators for social institutions and social norms. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.5: Robustness checks against the reverse causality using the past paid leave policies

Dep.var. literacy skill use	Panel A: All available countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Literacy skill: Q1	0.067	0.025	0.026	0.067***	0.106	0.040
	(0.057)	(0.021)	(0.027)	(0.024)	(0.075)	(0.075)
Female×PL×Literacy skill: Q2	-0.084	-0.088***	-0.087***	-0.059	0.103	0.081
	(0.077)	(0.029)	(0.032)	(0.038)	(0.118)	(0.086)
Female×PL×Literacy skill: Q3	-0.272***	-0.135***	-0.150***	-0.128**	0.072	0.021
	(0.095)	(0.048)	(0.045)	(0.052)	(0.145)	(0.143)
Female×PL×Literacy skill: Q4	-0.027	-0.098***	-0.093***	-0.068**	0.095	0.078
	(0.055)	(0.024)	(0.026)	(0.031)	(0.077)	(0.080)
Parental leave policy year	2011	2011	2001	1991	1981	1971
Source of parental leave policy	Original	OECD	OECD	OECD	OECD	OECD
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X	X
Female×Skill×Family policies	X	X	X	X	X	X
Female×Skill×Market institutions	X	X	X	X	X	X
Countries	21	21	21	21	21	21
Observations	43387	43387	43387	43387	43387	43387
Dep.var. literacy skill use	Panel B: Exclude Finland and Norway					
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Literacy skill: Q1	0.042	0.036	0.053	0.084**	0.096	0.081
	(0.053)	(0.049)	(0.040)	(0.035)	(0.064)	(0.077)
Female×PL×Literacy skill: Q2	-0.108	-0.106	-0.066	-0.042	0.065	0.021
	(0.067)	(0.066)	(0.054)	(0.050)	(0.106)	(0.114)
Female×PL×Literacy skill: Q3	-0.267***	-0.194*	-0.197***	-0.110	0.046	-0.118
	(0.086)	(0.100)	(0.067)	(0.081)	(0.130)	(0.156)
Female×PL×Literacy skill: Q4	-0.045	-0.014	-0.021	-0.006	0.044	-0.040
	(0.041)	(0.044)	(0.029)	(0.030)	(0.065)	(0.064)
Parental leave policy year	2011	2011	2001	1991	1981	1971
Source of parental leave policy	Original	OECD	OECD	OECD	OECD	OECD
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X	X
Female×Skill×Family policies	X	X	X	X	X	X
Female×Skill×Gender norm	X	X	X	X	X	X
Female×Skill×Market institutions	X	X	X	X	X	X
Countries	19	19	19	19	19	19
Observations	39599	39599	39599	39599	39599	39599

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for literacy score. As policy parental leave variables, we use the duration of paid leave between 1971 and 2011 which are collected by the OECD as well as the duration in 2011 in our database. In column 1, to ease comparison, we restrict the sample to countries where the OECD database is available. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4.6: The paid leave policy and utilization of literacy skill at work

Dep.var.	Learning (1)	Influence (2)	Writing (3)	Numeracy (4)
Female×PL×Literacy skill: Q1	0.191**	0.016	-0.078	-0.043
	(0.079)	(0.062)	(0.072)	(0.059)
Female×PL×Literacy skill: Q2	-0.154**	-0.095*	-0.195***	-0.091*
	(0.061)	(0.057)	(0.070)	(0.049)
Female×PL×Literacy skill: Q3	-0.364***	-0.234***	-0.230**	-0.201**
	(0.090)	(0.075)	(0.101)	(0.094)
Female×PL×Literacy skill: Q4	-0.110*	-0.014	-0.027	-0.028
	(0.062)	(0.066)	(0.055)	(0.079)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	48966	48966	48966	48966

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for skill use scores other than literacy use. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The variables used to construct the dependent variables relying on the partial credit model is listed in Appendix 4.A.

TABLE 4.7: The paid leave policy and market outcomes

Dep.var.	Employment (1)	Work hours (2)	ln(*wage*) (3)
Female×PL×Literacy skill: Q1	0.002	2.971	-0.014
	(0.009)	(1.844)	(0.065)
Female×PL×Literacy skill: Q2	0.003	2.233*	-0.059
	(0.005)	(1.278)	(0.055)
Female×PL×Literacy skill: Q3	-0.009	0.765	-0.030
	(0.007)	(0.758)	(0.057)
Female×PL×Literacy skill: Q4	0.001	1.794**	-0.040
	(0.007)	(0.872)	(0.036)
Mean value among men	0.99	42.19	3.81
Method	OLS	Tobit	Heckit
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	24	23	21
Observations	35410	33919	31515

Note: This table shows estimation results regarding market outcomes. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.8: The paid leave policy and utilization of literacy skill at work (Control for occupation)

Dep.var. literacy skill use	Baseline (1)	2-digit code (2)	4-digit code (3)
Female×PL×Literacy skill: Q1	-0.069*	0.019	-0.017
	(0.039)	(0.042)	(0.030)
Female×PL×Literacy skill: Q2	-0.149***	-0.073**	-0.081**
	(0.035)	(0.036)	(0.040)
Female×PL×Literacy skill: Q3	-0.063	-0.041	-0.044
	(0.048)	(0.055)	(0.039)
Female×PL×Literacy skill: Q4	0.023	0.003	-0.014
	(0.035)	(0.027)	(0.029)
Diffrence from baseline: Q1		0.088***	0.052**
Diffrence from baseline: Q2		0.077***	0.069***
Diffrence from baseline: Q3		0.022	0.019
Diffrence from baseline: Q4		-0.020*	-0.037**
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	18	18	18
Observations	31308	31308	31308

Note: This table shows estimation results regarding the literacy use, where the estimation sample was restricted to the employed population. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.9: The paid leave policy, years leaving labor market and utilization of literacy skill at work

Dep.var. literacy skill use	Full sample		Employed	
	(1)	(2)	(3)	(4)
Female×PL×Literacy skill: Q1	-0.008	0.031	-0.020	-0.007
	(0.036)	(0.047)	(0.044)	(0.043)
Female×PL×Literacy skill: Q2	-0.175***	-0.115	-0.062	-0.044
	(0.067)	(0.074)	(0.053)	(0.053)
Female×PL×Literacy skill: Q3	-0.360***	-0.299***	-0.080	-0.065
	(0.089)	(0.076)	(0.048)	(0.044)
Female×PL×Literacy skill: Q4	-0.082	-0.021	0.003	0.017
	(0.064)	(0.065)	(0.028)	(0.028)
AL×Literacy skill: Q1		-0.054***		-0.015***
		(0.006)		(0.003)
AL×Literacy skill: Q2		-0.049***		-0.009**
		(0.006)		(0.004)
AL×Literacy skill: Q3		-0.047***		-0.013***
		(0.006)		(0.004)
AL×Literacy skill: Q4		-0.049***		-0.013***
		(0.005)		(0.004)
Female×AL×Literacy skill: Q1		0.000		-0.001
		(0.004)		(0.004)
Female×AL×Literacy skill: Q2		-0.009		-0.006
		(0.007)		(0.005)
Female×AL×Literacy skill: Q3		-0.018**		-0.003
		(0.008)		(0.004)
Female×AL×Literacy skill: Q4		-0.018**		-0.005
		(0.009)		(0.005)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	21	21	21	21
Observations	42483	42483	36273	36273

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for literacy score. AL is the actual years leaving labor market. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.A Skill use items

4.A.1 Literacy skill use

1. Read directions or instructions
2. Read letters, memos or e-mails
3. Read articles in newspapers, magazines, or newsletters
4. Read articles in professional journals or scholarly publications
5. Read books
6. Read manuals or reference materials
7. Read bills, invoices, bank statements or other financial statements
8. Read diagrams, maps or schematics

4.A.2 Numeracy skill use

1. Calculate prices, costs, or budgets
2. Use or calculate fractions, decimals, or percentages
3. Use a calculator – either hand-held or computer-based
4. Use simple algebra or formulas
5. Use more advanced math or statistics, such as calculus, complex algebra, trigonometry, or regression techniques
6. Prepare charts, graphs, or tables

4.A.3 Learning opportunities

1. In your own job, how often do you learn new work-related things from co-workers or supervisors?
2. How often does your job involve learning-by-doing from the tasks you perform?
3. How often does your job involve keeping up to date with new products or services?

4.A.4 Influencing others

1. How often does your job usually involve instructing, training, or teaching people, individually or in groups?
2. How often does your job usually involve making speeches or giving presentations in front of five or more people?

3. How often does your job usually involve advising people?

4. How often does your job usually involve planning the activities of others?

5. How often does your job usually involve persuading or influencing people?

6. How often does your job usually involve negotiating with people either inside or outside your firm or organization?

4.A.5 Writing skill use

1. Writing skills at work: In your job, how often do you usually write letters, memos, or e-mails?

2. Writing skills at work: In your job, how often do you usually write articles for newspapers, magazines, or newsletters?

3. Writing skills at work: In your job, how often do you usually write reports?

4. Writing skills at work: In your job, how often do you usually fill in forms?

4.B Data source of parental-leave policies

TABLE 4.B1: Data source of parental leave policies

Country	Paid leave	Job protection	Full-rate equivalence	Source	Note
Austria	1.308	2.154	1.108	Maternity Protection Act: 3, 5, 14.4, 15.1. General Social Insurance Act: 162. Child Care Benefit Act 14.1.	For parental leave benefits, there are some alternatives: 14.53 EUR/day for 30 months, 20.80 EUR/day for 20 months, 26.60 EUR/day for 15 month, 33 EUR/day for 12 months, or for 12 months with replacement rate, 0.8. We employed the last one, which is also the one employed in the OECD family database.
Belgium	0.537	0.619	0.297	Labour Act Art: 39, Royal Decree Regarding the Establishment of a Parental Leave in the Framework of Interruption of Professional Career Art: 2, 10, Royal Decree to Execute the Act Respecting Compulsory and Indemnity Insurance Scheme Art, 114, 115, 216, 217	The duration of job protection is 4 months. Parental leave benefits are flat-rate (679.59 EUR/month) for 3 months. We calculated the replacement rate using the median value of female monthly earnings (= 679.59/2187.625).
Chile	0.577	0.577	0.577	Labour Code: 195, 197	
Czech	2.100	3.115	1.340	Labour Code: 195, 196	The duration of parental leave is up to 3 years of a child, after the end of the maternity leave. Parental leave benefits are fixed amount with four alternatives: 11,400 CZK until the child is 24 months old, 7,600 CZK until 36 months old, 7,600 CZK until 9 months old and after it 3,800 CZK until 48 months old, and lower rate with 3,000 CZK for some periods. We employed the first one, and calculated the replacement rate using median female wage in PIAAC.
Denmark	1.231	1.231	1.231	Consolidation Act no. 1084 of 13 November 2009 on Entitlement to Leave and Benefits in the Event of Childbirth: 6, 7, 9, 10, 21, 33, 35, 36, 37	The weekly maternity leave benefits are capped by DKK 3,332. Since parents are allowed to prolong the parental leave (from 32 weeks) up to 46 weeks, we employed 46 weeks as the duration of parental leave.
Estonia	1.580	3.192	1.580	Holidays Act: 27, 30, 31, Health Insurance Act: 58, 84, Parental Benefits Act: 2, 3, 4	Although there is no paid parental leave, parents have right to receive parental benefits for 435 days. If the mother takes unpaid parental leave, the amount of the benefits is calculated on the basis of her wage. If she does not take parental leave, the amount may be calculated on the basis of his spouse's wage. The amount is capped by three time the average income.

Country	Paid leave	Job protection	Full-rate equivalent	Source	Note
Finland	0.896	3.192	0.619	Health Insurance Act: 9, 10, 11	The replacement rate of maternity leave benefits is progressive: 70 percent up to 26,720 EUR, 40 percent up to 41,100 EUR, and then 25 percent. The replacement rate of maternity leave benefits was evaluated at median of female earnings using PIAAC. Finland may have home-care leave until the child becomes age 3, if that child does not enroll public childcare. We excluded this type of leave from our definition of paid parental leave. Note that home-care leave benefit is 327.46 EUR/month, which is relatively small amount (female median monthly earning in PIAAC is 2,457 EUR, so if we take the replacement rate into account, equivalent weeks will be similar, regardless of inclusion of home care benefits).
France	0.808	3.115	0.380	Labour Code: 1225-17, 1225-18, 1225-19, 1225-2020, 1225-47, 1225-48, 1225-54, Social Security Code: 331-3, 331-4, 331-5, 331-6, 323-4, OECD family database	According to the OECD family database, parental leave benefits may be available, but not referred to in the ILO database. Thus, we followed the OECD family database. While the duration of paid leave was of 2011, the replacement rate was of 2016, due to data availability.
Germany	1.262	3.115	0.934	Maternity Protection Act: 3.2, 6.1, 13.1, 14.1. Parental Allowance and Parental Leave Act: 1, 2.1, 2.2, 2.5, 15.2, 15.3, 16.1, 16.3.	The parental leave benefits are capped by EUR 1,800 (monthly).
Greece	0.327	0.823	0.163	Social Security Programs Throughout the World: Europe, 2010	OECD family database suggests that Greece has 26 weeks special parental leave with flat amount, which is not shown in ILO database, and we did not take it into account. (Note that this value from the OECD family database was about policy in 2016.)
Ireland	0.500	1.077	0.400	Maternity Protection Act: 8, 10-11, 14, 16. Parental Leave Act: 6. Social Welfare Consolidation Act: 6, 47, 49.	The maternity leave benefits are capped by EUR 280 per week.
Italy	0.910	0.910	0.480	Legislative Decree No. 151 of 2001: 16, 20, 22, 26, 32, 34	
Japan	1.115	1.115	0.603	Labor Standards Act 1947: 65, Employment Insurance Act: 61. National Health Insurance Law: 8. Childcare and Elderly care Act: 5, 9	Information about parental leave policy in the ILO database is incorrect.
Korea	1.247	1.247	0.647	Labor Standards Act: 20, 74, 76. Act on Equal Employment and Support for Work-Family Reconciliation: 19. Enforcement Decree of the Employment Insurance Act: 95, 100, 101.	The parental leave benefits should be between 500,000 won/month and 1 million won/month.
Netherlands	0.308	0.808	0.308	Work and Care Act: 3.1.2, 3.1.3, 3.8, 6.1.1, 6.2	

Country	Paid leave	Job protection	Full-rate equivalent	Source	Note
New Zealand	0.269	1.000	0.269	Parental Leave and Employment Protection Act 1987: 9.1, 26, 28, 29, 30, 71J.	The rate of maternity leave payments is the lesser of USD 325 or 100 percent of weekly payment.
Norway	0.885	2.000	0.885	Working Environment Act: 12. National Insurance Act 14.7.	Norway has two alternatives 46 weeks with 100 percent replacement rate and 56 weeks with 80 percent replacement rates. Although we employed the first one, these two has little difference in terms of full-rate equivalent. Norway has home care leave benefits, which are available for 23 months (3,303 NOK/month for 23 months from 13 months old) if the child does not use public early childhood education and care services, but we did not include this in the definition of paid parental leave.
Poland	0.385	0.385	0.385	Social Security Program throughout the World, Europe 2010.	Although Poland has care leave for 60 days per year if the child is younger than 8 years old, we did not include this because this seems to be a temporal leave, say, when a child gets sick.
Slovak Republic	3.000	3.000	1.084	Labour Code: 166.1, Act on Social Insurance 48.2, 48.3, 53, 55. Social Security Programs throughout the World, Europe 2010, 2012.	In terms of the duration of maternity leave, the ILO database seems incorrect, which may be the duration in 2009 but not 2011. The amount of parental leaves benefit is fixed, 164.22 EUR/month in 2010. The replacement rate was evaluated at the median female wage in PIAAC.
Slovenia	1.003	1.003	1.003	Parental Protection and Family Benefits Act: 17, 26, 29, 31, 41, 43, 44	
Spain	0.308	3.308	0.308	Decree No.1/1995 enacting the Worker's Charter Art: 46, 48. Royal Decree No295/2009 on Cash benefits of the Social Security System concerning Maternity, Paternity, Risk during Pregnancy and Risk during Breastfeeding Art: 3, 8	
Sweden	1.154	1.758	0.874	Parental Leave Act: 4, 5. Public Insurance Act: 4.3, 4.5, 4.6.	Sixty days out of 480 days are given to each parent as exclusive right and the remaining 300 days can be divided between them however they choose. Thus, we defined 420 days as paid leave period for the mother. Since the parental leave benefits exceeding 390 days are the flat amount, 180 SEK/day. The replacement rate of this flat part was evaluated at female mean monthly earnings in 2011 (SEK 26,200) from "Sweden and gender equality."
United Kingdom	0.750	1.250	0.354	Employment Rights Act 1996: 71. The Maternity and Parental Leave Regulations 1999: 7, 14, 15. Social Security Contributions and Benefits Act 2002: 165, 166. Statutory Maternity Pay (General) Regulations 1986: 2	The replacement of the maternity leave benefits is 90 percent for the first 6 weeks (without ceiling), and the lower of either 128.73 pounds or 90 percent of average weekly earnings for the remaining 33 weeks. The replacement rate was evaluated at the median female wage in PIAAC.
United States	0.000	0.231	0.000	Family and Medical Leave Act: 102	While there is no Federal-level paid leaves, some states have paid leave systems.

Country	Paid leave	Job protection	Full-rate equivalent	Source	Note

Note: One year was counted by 52 weeks and one month was counted by 4.3 weeks.

4.C Supplemental figures and tables

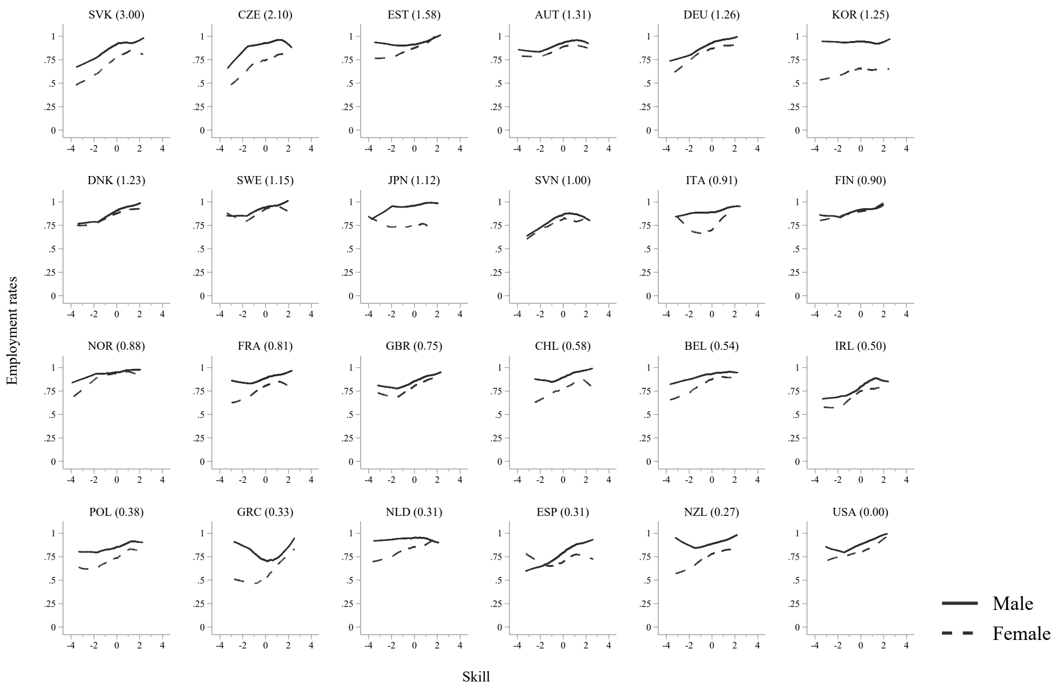


FIGURE 4.1: Employment rates at each numeracy skill level

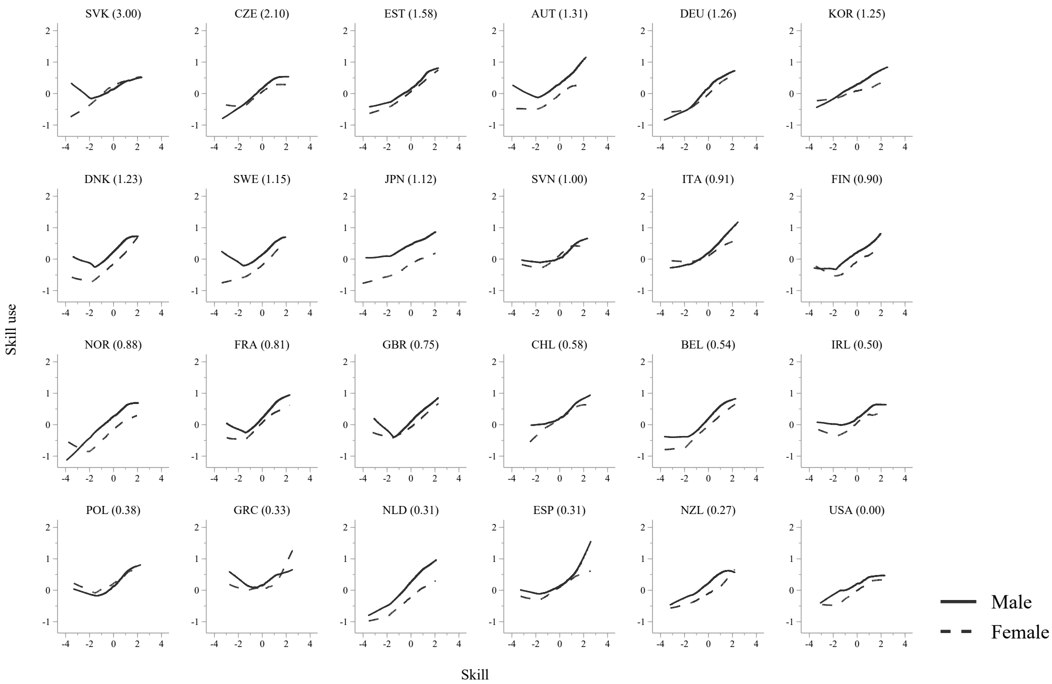


Figure 4.2: Skill use and skill within labor-force participants (Numeracy)

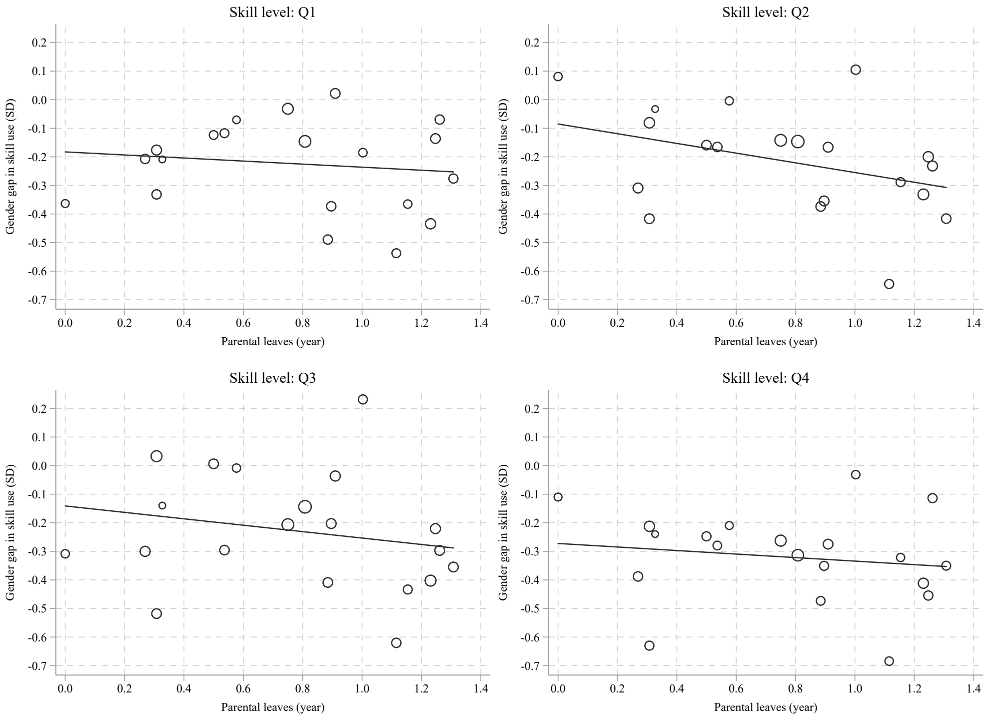


FIGURE 4.3: Unconditional gender gap in numeracy skill use and the paid-leave policy

Note: This figure shows relationship between the gender gap in numeracy skill use and the paid-leave policy. The gender gap in each country is calculated as a raw difference in average skill-use levels between employed women and men. The line is the fitted value by the weighted least squares, where the number of observations in each country is used as a weight. In this figure, ex-communist countries are excluded, because their social institutions tend to differ from those of other countries (de Haan, 2012).

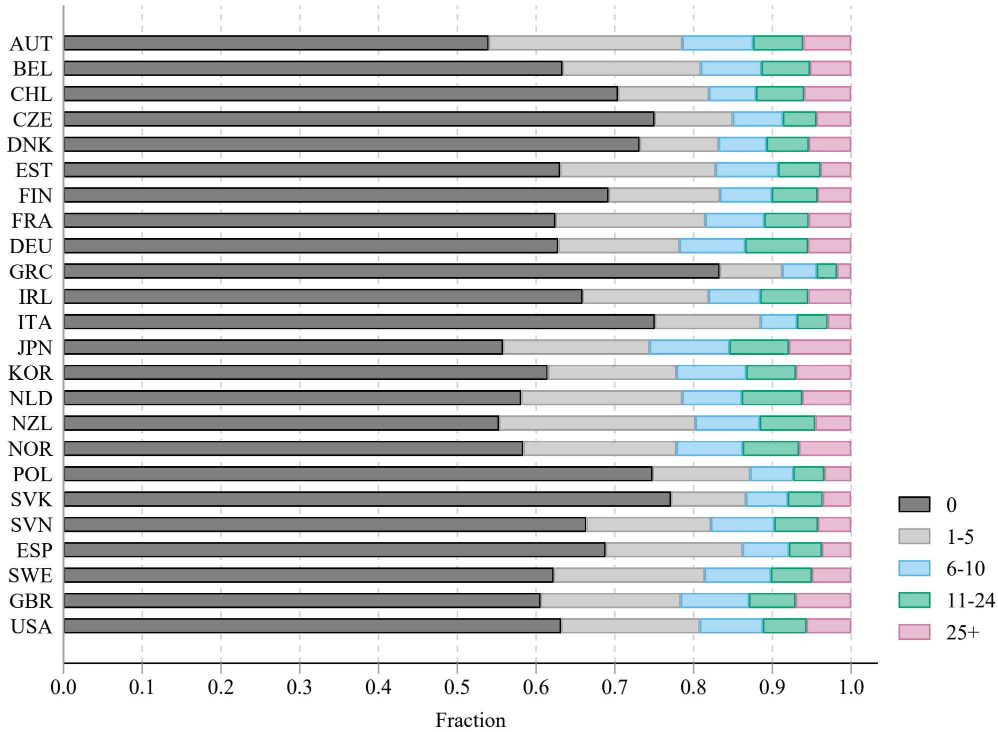


FIGURE 4.4: Distribution of the number of subordinates among men

Note: This figures shows the distribution of the number of subordinates in each country. The sample is restricted to employed men.

Table 4.2: Summary statistics of institutional indices

	Dual earner penalty	Childcare enrollment	Equal right part-time	Right for part-time	Gender norms	Public sector	Service sector	Emp. protect.	Union density
AUT	-0.128	0.172	1.000	0.000	-0.571	0.228	0.706	2.440	0.284
BEL	-0.089	0.413	1.000	1.000	-0.611	0.257	0.730	3.131	0.551
CHL	-0.071	0.176	1.000	0.000	-0.348	0.140	0.710	1.800	0.153
CZE	0.011	0.062	0.000	0.000	-0.305	0.218	0.652	2.751	0.158
DNK	-0.145	0.675	0.000	0.000	-0.934	0.359	0.768	2.320	0.664
EST	0.000	0.217	1.000	0.000	-0.335	0.257	0.656	2.066	0.070
FIN	-0.271	0.285	1.000	0.000	-0.818	0.316	0.733	2.167	0.696
FRA	0.029	0.510	1.000	1.000	-0.712	0.235	0.737	2.823	0.077
DEU	0.172	0.252	1.000	0.000	-0.478	0.206	0.695	2.842	0.185
GRC	-0.254	0.229	1.000	1.000	-0.263	0.217	0.738	2.440	0.228
IRL	-0.446	0.229	1.000	0.000	-0.549	0.267	0.777	1.978	0.326
ITA	-0.244	0.272	1.000	0.000	-0.467	0.215	0.689	3.032	0.363
JPN	-0.175	0.266	1.000	1.000	0.359	0.120	0.697	2.085	0.190
KOR	-0.216	0.290	1.000	0.000	0.060	0.122	0.682	2.168	0.099
NLD	-0.259	0.596	0.000	1.000	-0.527	0.257	0.794	2.884	0.184
NZL	-0.300	0.381	0.000	0.000	-0.606	0.212	0.743	1.010	0.209
NOR	-0.209	0.551	1.000	0.000	-0.913	0.358	0.804	2.310	0.535
POL	-0.012	0.080	1.000	1.000	-0.236	0.193	0.610	2.391	0.136
SVK	0.020	0.046	1.000	0.000	-0.239	0.241	0.614	2.635	0.141
SVN	-0.095	0.410	1.000	1.000	-0.658	0.296	0.613	2.670	0.220
ESP	-0.164	0.397	1.000	1.000	-0.581	0.205	0.747	2.558	0.169
SWE	-0.399	0.479	1.000	0.000	-0.895	0.351	0.778	2.517	0.675
GBR	-0.243	0.391	1.000	0.000	-0.654	0.303	0.818	1.759	0.258
USA	0.000	0.280	0.000	0.000	-0.595	0.204	0.802	1.171	0.113

Table 4.3: The job protection policy and utilization of literacy skill at work

Dep.var.	Learning (1)	Influence (2)	Writing (3)	Numeracy (4)
Female×PL×Literacy skill: Q1	0.105***	0.064***	0.036*	-0.019
	(0.020)	(0.019)	(0.020)	(0.027)
Female×PL×Literacy skill: Q2	-0.010	-0.033*	-0.024	-0.040**
	(0.015)	(0.018)	(0.017)	(0.019)
Female×PL×Literacy skill: Q3	-0.078**	-0.058***	-0.078***	-0.092***
	(0.033)	(0.021)	(0.027)	(0.022)
Female×PL×Literacy skill: Q4	-0.034*	-0.017	-0.023**	-0.025
	(0.018)	(0.018)	(0.011)	(0.019)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	48966	48966	48966	48966

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for skill use scores other than literacy use. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.4: The job protection policy and market outcomes

Dep.var.	Employment (1)	Work hours (2)	ln(*wage*) (3)
Female×PL×Literacy skill: Q1	0.005*	1.230**	-0.053***
	(0.003)	(0.566)	(0.010)
Female×PL×Literacy skill: Q2	0.004	0.412	-0.062***
	(0.003)	(0.403)	(0.004)
Female×PL×Literacy skill: Q3	-0.003	-0.099	-0.049***
	(0.003)	(0.279)	(0.008)
Female×PL×Literacy skill: Q4	-0.002	0.104	-0.014
	(0.002)	(0.336)	(0.012)
Mean value among men	0.99	42.19	3.81
Method	OLS	Tobit	Heckit
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	24	23	21
Observations	35410	33919	31515

Note: This table shows estimation results regarding market outcomes. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.5: The job protection policy and utilization of literacy skill at work (Control for occupation)

Dep.var. literacy skill use	Baseline (1)	2-digit code (2)	4-digit code (3)
Female×PL×Literacy skill: Q1	-0.035*	-0.029*	-0.011
	(0.018)	(0.017)	(0.017)
Female×PL×Literacy skill: Q2	-0.044*	-0.030**	-0.024
	(0.023)	(0.015)	(0.016)
Female×PL×Literacy skill: Q3	-0.064***	-0.076***	-0.055***
	(0.017)	(0.015)	(0.013)
Female×PL×Literacy skill: Q4	0.028	0.012	0.017
	(0.025)	(0.021)	(0.026)
Diffrence from baseline: Q1		0.006	0.024*
Diffrence from baseline: Q2		0.014	0.020*
Diffrence from baseline: Q3		-0.012	0.009
Diffrence from baseline: Q4		-0.016***	-0.011**
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	18	18	18
Observations	31308	31308	31308

Note: This table shows estimation results regarding the literacy use, where the estimation sample was restricted to the employed population. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.6: The job protection policy, years leaving labor market and utilization of literacy skill at work

Dep.var. literacy skill use	Full sample		Employed	
	(1)	(2)	(3)	(4)
Female×PL×Literacy skill: Q1	0.032**	0.044***	-0.009	-0.005
	(0.013)	(0.013)	(0.017)	(0.017)
Female×PL×Literacy skill: Q2	-0.046**	-0.030	-0.027**	-0.023*
	(0.018)	(0.021)	(0.013)	(0.013)
Female×PL×Literacy skill: Q3	-0.121***	-0.100***	-0.057***	-0.053***
	(0.025)	(0.025)	(0.016)	(0.016)
Female×PL×Literacy skill: Q4	-0.029*	0.004	0.009	0.018
	(0.015)	(0.016)	(0.014)	(0.014)
AL×Literacy skill: Q1		-0.054***		-0.015***
		(0.006)		(0.003)
AL×Literacy skill: Q2		-0.049***		-0.009**
		(0.006)		(0.004)
AL×Literacy skill: Q3		-0.048***		-0.013***
		(0.006)		(0.004)
AL×Literacy skill: Q4		-0.048***		-0.013***
		(0.005)		(0.004)
Female×AL×Literacy skill: Q1		0.000		-0.001
		(0.004)		(0.004)
Female×AL×Literacy skill: Q2		-0.009		-0.006
		(0.007)		(0.005)
Female×AL×Literacy skill: Q3		-0.017**		-0.003
		(0.008)		(0.004)
Female×AL×Literacy skill: Q4		-0.019**		-0.005
		(0.009)		(0.005)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	21	21	21	21
Observations	42483	42483	36273	36273

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for literacy score. *AL* is the actual years leaving labor market. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates of the coefficients of the interaction terms associated with the literacy skill index and the indicators for social institutions and social norms. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.7: The paid leave policy and utilization of numeracy skill at work

Dep.var. numeracy skill use	Full sample					Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Numeracy skill: Q1	-0.044	-0.103	-0.179***	-0.073	-0.136**	-0.114
	(0.070)	(0.075)	(0.069)	(0.055)	(0.064)	(0.076)
Female×PL×Numeracy skill: Q2	-0.064	-0.135	-0.212**	-0.124	-0.073	-0.082
	(0.068)	(0.084)	(0.106)	(0.089)	(0.084)	(0.093)
Female×PL×Numeracy skill: Q3	-0.046	-0.176**	-0.210***	-0.132*	-0.119*	0.006
	(0.071)	(0.072)	(0.079)	(0.079)	(0.062)	(0.079)
Female×PL×Numeracy skill: Q4	-0.011	-0.095	-0.183***	-0.126*	-0.151**	-0.039
	(0.075)	(0.079)	(0.069)	(0.069)	(0.077)	(0.065)
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure		X	X	X	X	X
Female×Skill×Family policies			X	X	X	X
Female×Skill×Gender norm				X	X	X
Female×Skill×Market institutions					X	X
Countries	24	24	24	24	24	24
Observations	49039	49039	49039	49039	49039	41311

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for numeracy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.8: The job protection policy and utilization of numeracy skill at work

Dep.var. numeracy skill use	Full sample					Employed
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Numeracy skill: Q1	0.061*	0.054	0.010	-0.017	-0.025	-0.061***
	(0.034)	(0.037)	(0.037)	(0.025)	(0.026)	(0.022)
Female×PL×Numeracy skill: Q2	0.024	0.013	-0.035	-0.065**	-0.065**	-0.073***
	(0.029)	(0.035)	(0.037)	(0.032)	(0.028)	(0.023)
Female×PL×Numeracy skill: Q3	0.041	0.024	-0.007	-0.032	-0.034	-0.016
	(0.029)	(0.030)	(0.032)	(0.027)	(0.022)	(0.024)
Female×PL×Numeracy skill: Q4	0.034	0.020	-0.031	-0.053**	-0.058***	-0.040*
	(0.031)	(0.033)	(0.027)	(0.021)	(0.017)	(0.022)
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure		X	X	X	X	X
Female×Skill×Family policies			X	X	X	X
Female×Skill×Gender norm				X	X	X
Female×Skill×Market institutions					X	X
Countries	24	24	24	24	24	24
Observations	49039	49039	49039	49039	49039	41311

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for numeracy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.9: Robustness checks against the reverse causality using the past paid leave policies (Numeracy skill)

Dep.var. numeracy skill use	Panel A: All available countries					
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Numeracy skill: Q1	-0.143*	-0.091*	-0.096**	-0.050	-0.075	-0.106
	(0.078)	(0.047)	(0.047)	(0.042)	(0.131)	(0.110)
Female×PL×Numeracy skill: Q2	-0.026	-0.067**	-0.106***	-0.073**	-0.319***	-0.131*
	(0.070)	(0.030)	(0.030)	(0.033)	(0.091)	(0.079)
Female×PL×Numeracy skill: Q3	-0.085	-0.070**	-0.072**	-0.086**	-0.239**	-0.187**
	(0.062)	(0.033)	(0.030)	(0.036)	(0.109)	(0.088)
Female×PL×Numeracy skill: Q4	-0.133*	-0.087***	-0.103***	-0.081***	0.018	-0.173*
	(0.075)	(0.027)	(0.024)	(0.027)	(0.115)	(0.093)
Parental leave policy year	2011	2011	2001	1991	1981	1971
Source of parental leave policy	Original	OECD	OECD	OECD	OECD	OECD
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X	X
Female×Skill×Family policies	X	X	X	X	X	X
Female×Skill×Market institutions	X	X	X	X	X	X
Countries	21	21	21	21	21	21
Observations	43469	43469	43469	43469	43469	43469
Dep.var. numeracy skill use	Panel B: Exclude Finland and Norway					
	(1)	(2)	(3)	(4)	(5)	(6)
Female×PL×Numeracy skill: Q1	-0.203***	-0.230***	-0.104*	-0.093*	-0.133	-0.156
	(0.065)	(0.067)	(0.063)	(0.055)	(0.126)	(0.134)
Female×PL×Numeracy skill: Q2	-0.062	-0.022	-0.101**	-0.078	-0.376***	-0.238***
	(0.071)	(0.071)	(0.051)	(0.049)	(0.046)	(0.082)
Female×PL×Numeracy skill: Q3	-0.109*	-0.144*	-0.082	-0.132**	-0.270**	-0.266**
	(0.064)	(0.084)	(0.052)	(0.054)	(0.115)	(0.105)
Female×PL×Numeracy skill: Q4	-0.163**	-0.099	-0.098**	-0.085*	-0.022	-0.289***
	(0.069)	(0.073)	(0.045)	(0.048)	(0.108)	(0.084)
Parental leave policy year	2011	2011	2001	1991	1981	1971
Source of parental leave policy	Original	OECD	OECD	OECD	OECD	OECD
Country×Skill quartile FE	X	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X	X
Female×Skill×Family policies	X	X	X	X	X	X
Female×Skill×Gender norm	X	X	X	X	X	X
Female×Skill×Market institutions	X	X	X	X	X	X
Countries	19	19	19	19	19	19
Observations	39585	39585	39585	39585	39585	39585

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for numeracy score. As policy parental leave variables, we use the duration of paid leave between 1971 and 2011 which are collected by the OECD as well as the duration in 2011 in our database. In column 1, to ease comparison, we restrict the sample to countries where the OECD database is available. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.10: The paid leave policy and market outcomes (Numeracy skill)

Dep.var.	Employment (1)	Work hours (2)	ln(*wage*) (3)
Female×PL×Numeracy skill: Q1	0.010	2.010	-0.011
	(0.007)	(1.239)	(0.087)
Female×PL×Numeracy skill: Q2	0.007	2.377*	-0.073
	(0.006)	(1.424)	(0.063)
Female×PL×Numeracy skill: Q3	-0.005	0.574	-0.002
	(0.007)	(1.230)	(0.047)
Female×PL×Numeracy skill: Q4	-0.002	1.408	-0.055
	(0.006)	(0.906)	(0.033)
Mean value among men	0.99	42.18	3.79
Method	OLS	Tobit	Heckit
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	24	23	21
Observations	35427	33939	31506

Note: This table shows estimation results regarding market outcomes. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.11: The job protection policy and market outcomes (Numeracy skill)

Dep.var.	Employment (1)	Work hours (2)	ln(*wage*) (3)
Female×PL×Numeracy skill: Q1	0.001	0.308	-0.061***
	(0.004)	(0.467)	(0.013)
Female×PL×Numeracy skill: Q2	-0.001	0.324	-0.065***
	(0.003)	(0.484)	(0.010)
Female×PL×Numeracy skill: Q3	-0.000	0.056	-0.048***
	(0.003)	(0.382)	(0.007)
Female×PL×Numeracy skill: Q4	-0.001	0.150	-0.026**
	(0.003)	(0.387)	(0.011)
Mean value among men	0.99	42.18	3.79
Method	OLS	Tobit	Heckit
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	24	23	21
Observations	35427	33939	31506

Note: This table shows estimation results regarding market outcomes. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.12: The paid leave policy and utilization of numeracy skill at work (Control for occupation)

Dep.var. numeracy skill use	Baseline (1)	2-digit code (2)	4-digit code (3)
Female×PL×Numeracy skill: Q1	-0.126***	-0.064	-0.093**
	(0.046)	(0.040)	(0.042)
Female×PL×Numeracy skill: Q2	-0.013	-0.031	-0.027
	(0.042)	(0.045)	(0.039)
Female×PL×Numeracy skill: Q3	0.161***	0.167***	0.162***
	(0.029)	(0.046)	(0.036)
Female×PL×Numeracy skill: Q4	-0.002	0.063**	0.026
	(0.037)	(0.028)	(0.034)
Diffrence from baseline: Q1		0.062*	0.033
Diffrence from baseline: Q2		-0.018	-0.013
Diffrence from baseline: Q3		0.007	0.001
Diffrence from baseline: Q4		0.065**	0.028
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	18	18	18
Observations	31303	31303	31303

Note: This table shows estimation results regarding the numeracy use, where the estimation sample was restricted to the employed population. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.13: The job protection policy and utilization of numeracy skill at work (Control for occupation)

Dep.var. numeracy skill use	Baseline (1)	2-digit code (2)	4-digit code (3)
Female×PL×Numeracy skill: Q1	-0.076**	-0.057**	-0.055**
	(0.030)	(0.024)	(0.027)
Female×PL×Numeracy skill: Q2	-0.025	-0.034	-0.033
	(0.028)	(0.022)	(0.020)
Female×PL×Numeracy skill: Q3	0.008	-0.014	-0.009
	(0.035)	(0.030)	(0.029)
Female×PL×Numeracy skill: Q4	0.010	0.001	-0.020
	(0.024)	(0.020)	(0.023)
Diffrence from baseline: Q1		0.019	0.021
Diffrence from baseline: Q2		-0.009	-0.008
Diffrence from baseline: Q3		-0.022	-0.017
Diffrence from baseline: Q4		-0.009	-0.029***
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	18	18	18
Observations	31303	31303	31303

Note: This table shows estimation results regarding the numeracy use, where the estimation sample was restricted to the employed population. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.14: The paid leave policy, years leaving labor market and utilization of numeracy skill at work

Dep.var. numeracy skill use	Full sample (1)	(2)	Employed (3)	(4)
Female×PL×Numeracy skill: Q1	-0.145**	-0.097	-0.103	-0.085
	(0.069)	(0.067)	(0.069)	(0.066)
Female×PL×Numeracy skill: Q2	-0.013	0.017	0.010	0.028
	(0.082)	(0.089)	(0.070)	(0.069)
Female×PL×Numeracy skill: Q3	-0.121	-0.088	0.090	0.099*
	(0.080)	(0.073)	(0.060)	(0.056)
Female×PL×Numeracy skill: Q4	-0.132**	-0.089	0.022	0.033
	(0.063)	(0.063)	(0.052)	(0.050)
AL×Numeracy skill: Q1		-0.046***		-0.015***
		(0.005)		(0.004)
AL×Numeracy skill: Q2		-0.043***		-0.012***
		(0.005)		(0.003)
AL×Numeracy skill: Q3		-0.042***		-0.016***
		(0.004)		(0.003)
AL×Numeracy skill: Q4		-0.041***		-0.013***
		(0.004)		(0.005)
Female×AL×Numeracy skill: Q1		0.001		-0.001
		(0.004)		(0.003)
Female×AL×Numeracy skill: Q2		-0.009*		-0.009**
		(0.005)		(0.004)
Female×AL×Numeracy skill: Q3		-0.015***		-0.001
		(0.005)		(0.004)
Female×AL×Numeracy skill: Q4		-0.013***		-0.004
		(0.005)		(0.006)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	21	21	21	21
Observations	42545	42545	36350	36350

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for numeracy score. AL is the actual years leaving labor market. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.15: The job protection policy, years leaving labor market and utilization of numeracy skill at work

Dep.var. numeracy skill use	Full sample		Employed	
	(1)	(2)	(3)	(4)
Female×PL×Numeracy skill: Q1	-0.047*	-0.031	-0.078***	-0.072***
	(0.029)	(0.026)	(0.023)	(0.022)
Female×PL×Numeracy skill: Q2	-0.066**	-0.050	-0.058**	-0.051**
	(0.033)	(0.034)	(0.023)	(0.022)
Female×PL×Numeracy skill: Q3	-0.050**	-0.034	-0.012	-0.008
	(0.021)	(0.022)	(0.021)	(0.021)
Female×PL×Numeracy skill: Q4	-0.054***	-0.029	-0.031	-0.024
	(0.016)	(0.018)	(0.026)	(0.026)
AL×Numeracy skill: Q1		-0.046***		-0.015***
		(0.005)		(0.004)
AL×Numeracy skill: Q2		-0.043***		-0.012***
		(0.005)		(0.003)
AL×Numeracy skill: Q3		-0.042***		-0.016***
		(0.004)		(0.003)
AL×Numeracy skill: Q4		-0.041***		-0.013***
		(0.004)		(0.005)
Female×AL×Numeracy skill: Q1		0.001		-0.001
		(0.004)		(0.003)
Female×AL×Numeracy skill: Q2		-0.009*		-0.009**
		(0.005)		(0.004)
Female×AL×Numeracy skill: Q3		-0.015***		-0.001
		(0.005)		(0.004)
Female×AL×Numeracy skill: Q4		-0.012***		-0.003
		(0.005)		(0.005)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	21	21	21	21
Observations	42545	42545	36350	36350

Note: This table shows estimation results of the censored Tobit model consisting of equations (4.6) and (4.7) for numeracy score. AL is the actual years leaving labor market. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis.
* p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.16: The paid leave policy and number of subordinates (Literacy skill)

Number of subordinates	0 (1)	≤ 5 (2)	≤ 10 (3)	≤ 24 (4)
Female×PL×Literacy skill: Q1	0.059***	0.060***	0.022*	0.001
	(0.022)	(0.014)	(0.012)	(0.010)
Female×PL×Literacy skill: Q2	0.057***	0.019	0.015	0.012
	(0.021)	(0.024)	(0.021)	(0.017)
Female×PL×Literacy skill: Q3	0.039	0.042*	0.010	0.038***
	(0.038)	(0.025)	(0.015)	(0.013)
Female×PL×Literacy skill: Q4	0.036	0.020	0.016	0.020*
	(0.029)	(0.015)	(0.020)	(0.011)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	35590	35590	35590	35590

Note: This table shows estimation results regarding the number of subordinates, in which the analysis sample was restricted to the employed. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 4.17: The job protection policy and number of subordinates (Literacy skill)

Number of subordinates	0 (1)	≤ 5 (2)	≤ 10 (3)	≤ 24 (4)
Female×PL×Literacy skill: Q1	0.019***	0.014**	0.005	-0.003
	(0.007)	(0.006)	(0.003)	(0.003)
Female×PL×Literacy skill: Q2	0.012**	0.001	-0.001	-0.003
	(0.005)	(0.009)	(0.008)	(0.005)
Female×PL×Literacy skill: Q3	0.028***	0.019***	0.006	0.000
	(0.011)	(0.005)	(0.005)	(0.004)
Female×PL×Literacy skill: Q4	0.005	0.005	0.005	0.004
	(0.009)	(0.005)	(0.006)	(0.004)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	35590	35590	35590	35590

Note: This table shows estimation results regarding the number of subordinates, in which the analysis sample was restricted to the employed. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 4.18: The paid leave policy and number of subordinates (Numeracy skill)

Number of subordinates	0 (1)	≤ 5 (2)	≤ 10 (3)	≤ 24 (4)
Female×PL×Numeracy skill: Q1	0.072***	0.052***	0.046***	0.013
	(0.022)	(0.018)	(0.013)	(0.011)
Female×PL×Numeracy skill: Q2	0.054**	0.029	0.004	0.011
	(0.021)	(0.018)	(0.023)	(0.012)
Female×PL×Numeracy skill: Q3	0.012	0.004	0.014	0.018
	(0.027)	(0.022)	(0.024)	(0.012)
Female×PL×Numeracy skill: Q4	0.035	0.053*	0.034***	0.021**
	(0.041)	(0.029)	(0.013)	(0.010)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	35654	35654	35654	35654

Note: This table shows estimation results regarding the number of subordinates, in which the analysis sample was restricted to the employed. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4.19: The job protection policy and number of subordinates (Numeracy skill)

Number of subordinates	0 (1)	≤ 5 (2)	≤ 10 (3)	≤ 24 (4)
Female×PL×Numeracy skill: Q1	0.021***	0.014*	0.011*	-0.000
	(0.008)	(0.007)	(0.006)	(0.004)
Female×PL×Numeracy skill: Q2	0.024***	0.005	0.000	0.001
	(0.006)	(0.004)	(0.005)	(0.003)
Female×PL×Numeracy skill: Q3	0.016	0.016*	0.017*	0.003
	(0.010)	(0.008)	(0.008)	(0.002)
Female×PL×Numeracy skill: Q4	0.006	0.002	-0.001	0.002
	(0.009)	(0.010)	(0.006)	(0.003)
Country×Skill quartile FE	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X
Female×Skill×Family policies	X	X	X	X
Female×Skill×Gender norm	X	X	X	X
Female×Skill×Market institutions	X	X	X	X
Countries	24	24	24	24
Observations	35654	35654	35654	35654

Note: This table shows estimation results regarding the number of subordinates, in which the analysis sample was restricted to the employed. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Bibliography

Adda, Jérôme, Christian Dustmann, and Katrien Stevens, "The Career Costs of Children," *Journal of Political Economy*, 2017, *125* (2), 293–337.

Albrecht, James, Anders Bjorklund, and Susan Vroman, "Is there a glass ceiling in Sweden?," *Journal of Labor Economics*, January 2003, *21* (1), 145–177.

__ **, Peter Skogman Thoursie, and Susan Vroman**, "Parental leave and the glass ceiling in Sweden," *Research in Labor Economics*, 2015, *41*, 89–114.

Altonji, Joseph G, Fumio Hayashi, and Laurence J Kotlikoff, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," *The American Economic Review*, 1992, *82* (5), 1177–1198.

Angrist, Josh, "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *The Quarterly Journal of Economics*, 2002, *117* (3), 997.

Angrist, Joshua D., "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *The American Economic Review*, 1990, *80* (3), 313–336.

Apps, Patricia F. and Ray Rees, "Collective Labor Supply and Household Production," *Journal of Political Economy*, 1997, *105* (1), 178–190.

Aronsson, Thomas, Sven-Olov Daunfeldt, and Magnus Wikström, "Estimating Intrahousehold Allocation in a Collective Model with Household Production," *Journal of Population Economics*, 2001, *14* (4), 569–584.

Asai, Yukiko, "Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from Japan," *Labour Economics*, 2015, *36*, 72 – 83.

Attanasio, Orazio P. and Valérie Lechene, "Efficient Responses to Targeted Cash Transfers," *Journal of Political Economy*, 2014, *122* (1), 178–222.

Aura, Saku, "Does the Balance of Power within a Family Matter? The Case of the Retirement Equity Act," *Journal of Public Economics*, 2005, *89* (9), 1699 – 1717.

Baker, Michael and Kevin Milligan, "How does job-protected maternity leave affect mothers' employment?," *Journal of Labor Economics*, 2008, *26* (4), 655–691.

__ **and** __ , "Evidence from maternity leave expansions of the impact of maternal care on early child development," *Journal of Human Resources*, 2010, *45* (1), 1–32.

Baum, Charles L., "The effect of state maternity leave legislation and the 1993 Family and Medical Leave Act on employment and wages," *Labour Economics*, 2003, *10* (5), 573 – 596.

Berger, Lawrence M. and Jane Waldfogel, "Maternity leave and the employment of new mothers in the United States," *Journal of Population Economics*, 2004, *17* (2), 331–349.

_ **, Jennifer Hill, and Jane Waldfogel**, "Maternity leave, early maternal employment and child health and development in the US," *The Economic Journal*, 2005, *115* (501), F29–F47.

Blau, David M. and Ryan M. Goodstein, "Commitment in the Household: Evidence from the Effect of Inheritances on the Labor Supply of Older Married Couples," *Labour Economics*, 2016, *42*, 123 – 137.

Blau, Francine D. and Lawrence M. Kahn, "Female labor supply: Why is the United States falling behind?," *The American Economic Review*, 2013, *103* (3), 251–256.

Blundell, Richard, Pierre-André Chiappori, and Costas Meghir, "Collective Labor Supply with Children," *Journal of Political Economy*, 2005, *113* (6), 1277–1306.

Browning, Martin and Pierre-André Chiappori, "Efficient Intra-household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 1998, *66* (6), 1241–1278.

Canaan, Serena, "Parental Leave, Household Specialization and Children's Well-Being," *IZA Discussion Papers*, 2019.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes, "A flying start? Maternity leave benefits and long-run outcomes of children," *Journal of Political Economy*, 2015, *123* (2), 365–412.

Casanova, Maria, "Happy Together: A Structural Model of Couples' Joint Retirement Choices," 2010.

Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen, "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information," *American Economic Review*, December 2012, *102* (7), 3377–3405.

Chiappori, Pierre-André, "Rational Household Labor Supply," *Econometrica*, 1988, *56* (1), 63–90.

_ , "Collective Labor Supply and Welfare," *Journal of Political Economy*, 1992, *100* (3), 437–467.

Chiappori, Pierre-André, "Introducing Household Production in Collective Models of Labor Supply," *Journal of Political Economy*, 1997, *105* (1), 191–209.

Chiappori, Pierre-Andre and Maurizio Mazzocco, "Static and Intertemporal Household Decisions," *Journal of Economic Literature*, September 2017, *55* (3), 985–1045.

Dahl, Gordon B., Katrine V. Loken, and Magne Mogstad, "Peer effects in program participation," *American Economic Review*, July 2014, *104* (7), 2049–74.

— **, Katrine V. Løken, Magne Mogstad, and Kari Vea Salvanes**, "What Is the case for paid maternity leave?," *The Review of Economics and Statistics*, 2016, *98* (4), 655–670.

de Haan, Francisca, "Women as the motor of modern life," *Women and Gender in Postwar Europe: From Cold War to European Union*, 2012, pp. 87–103.

Dee, Thomas S. and William N. Evans, "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates," *Journal of Labor Economics*, 2003, *21* (1), 178–209.

Duflo, Esther, "Grandmothers and Granddaughters: Old-age Pensions and Intrahousehold Allocation in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 1.

Dufwenberg, Martin, "Marital Investments, Time Consistency and Emotions," *Journal of Economic Behavior and Organization*, 2002, *48* (1), 57–69.

Dustmann, Christian and Uta Schönberg, "Expansions in marternity leave coverage and children's long-term outcomes," *American Economic Journal: Applied Economics*, July 2012, *4* (3), 190–224.

Fernández-Kranz, Daniel, Aitor Lacuesta, and Núria Rodríguez-Planas, "The Motherhood Earnings Dip: Evidence from Administrative Records," *Journal of Human Resources*, 2013, *48* (1), 169–197.

— **and Núria Rodríguez-Planas**, "Can parents' right to work part-time hurt childbearing-aged women? A natural experiment with administrative data," *IZA Discussion Paper*, 2013.

Francis, Andrew M., "Sex Ratios and the Red Dragon: Using the Chinese Communist Revolution to Explore the Effect of the Sex Ratio on Women and Children in Taiwan," *Journal of Population Economics*, 2011, *24* (3), 813–837.

Fryer, Roland G. Jr. and Steven D. Levitt, "An empirical analysis of the gender gap in mathematics," *American Economic Journal: Applied Economics*, April 2010, *2* (2), 210–40.

Genda, Yuji, Ayako Kondo, and Souichi Ohta, "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States.," *Journal of Human Resources*, 2010, *45* (1), 157 – 196.

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

Gregg, Paul, Elizabeth Washbrook, Carol Propper, and Simon Burgess, "The effects of a mother's return to work decision on child development in the UK," *The Economic Journal*, 2005, *115* (501), F48–F80.

Gruber, Jonathan, "The incidence of mandated maternity benefits," *American Economic Review*, June 1994, *84* (3), 622–641.

Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales, "Culture, gender, and math," *Science*, 2008, *320* (5880), 1164–1165.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney, “Parental Education and Parental Time with Children,” *Journal of Economic Perspectives*, September 2008, *22* (3), 23–46.

Hayashi, Fumio, Joseph Altonji, and Laurence Kotlikoff, “Risk-Sharing between and within Families,” *Econometrica: Journal of the Econometric Society*, 1996, *64* (2), 261–294.

Imai, Ken-ichi and Hiroyuki Itami, “Interpenetration of organization and market: Japan’s firm and market in comparison with the U.S.,” *International Journal of Industrial Organization*, 1984, *2* (4), 285 – 310.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, “Children and gender inequality: Evidence from Denmark,” *NBER Working Paper*, 2018.

_ , _ , Johanna Posch, Andreas Steinhauer, and Josef Zweimüller, “Child penalties across countries: Evidence and explanations,” *NBER Working Paper*, 2019.

Kluve, Jochen and Sebastian Schmitz, “Back to Work: Parental Benefits and Mothers’ Labor Market Outcomes in the Medium Run,” *ILR Review*, 2018, *71* (1), 143–173.

Kondo, Ayako, “Effects of Increased Elderly Employment on Other Workers’ Employment and Elderly’s Earnings in Japan,” *IZA Journal of Labor Policy*, 2016, *5* (1), 2.

_ and Hitoshi Shigeoka, “The Effectiveness of Demand-side Government Intervention to Promote Elderly Employment: Evidence from Japan,” *ILR Review*, 2017, *70* (4), 1008–1036.

Kotlikoff, Laurence J. and Avia Spivak, “The Family as an Incomplete Annuities Market,” *Journal of Political Economy*, 1981, *89* (2), 372–391.

Krebs, Tom, “Job Displacement Risk and the Cost of Business Cycles,” *American Economic Review*, June 2007, *97* (3), 664–686.

Kunze, Astrid, “Parental leave and maternal labor supply,” *IZA World of Labor*, 2016, *279.*

Laitner, John and Amanda Sonnega, “Intergenerational Transfers in the Health and Retirement Study Data,” 2010. Working Paper 2010-238, Michigan Retirement Reseach Center.

Lalive, Rafael, Analía Schlosser, Andreas Steinhauer, and Josef Zweimüller, “Parental leave and mothers’ careers: The relative importance of job protection and cash benefits,” *The Review of Economic Studies*, 2014, *81* (1), 219–265.

_ and Josef Zweimüller, “How does parental leave affect fertility and return to work? Evidence from two natural experiments,” *Quarterly Journal of Economics*, 2009, pp. 1363–1402.

Lise, Jeremy and Ken Yamada, “Household Sharing and Commitment: Evidence from Panel Data on Individual Expenditures and Time Use,” *The Review of Economic Studies*, 2018, p. rdy066.

Low, Hamish, Costas Meghir, Luigi Pistaferri, and Alessandra Voena, "Marriage, Labor Supply and the Dynamics of the Social Safety Net," Technical Report, National Bureau of Economic Research 2018.

Mazzocco, Maurizio, "Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment," *The Review of Economic Studies*, 2007, *74* (3), 857.

_, Claudia Ruiz, and Shintaro Yamaguchi, "Labor Supply, Wealth Dynamics, and Marriage Decisions," *California Center for Population Research*, 2006.

Ministry of Labour, "Report on Basic Survey of Employment Management of Women 1991," 1991. (*Heisei 3 Nenndo Joshi Koyou Kanri Kihon Chosa Kekka Houkokusho* in Japanese).

_, "Report on Basic Survey of Employment Management of Women 1993," 1993. (*Heisei 5 Nenndo Joshi Koyou Kanri Kihon Chosa Kekka Houkokusho* in Japanese).

_, "Report on Basic Survey of Employment Management of Women 1996," 1996. (*Heisei 8 Nenndo Joshi Koyou Kanri Kihon Chosa Kekka Houkokusho* in Japanese).

Muraki, Eiji, "A generalized partial credit model: Application of an EM algorithm," *ETS Research Report Series*, 1992, *1992* (1), i–30.

Nollenberger, Natalia, Núria Rodríguez-Planas, and Almudena Sevilla, "The math gender gap: The role of culture," *American Economic Review*, May 2016, *106* (5), 257–61.

OECD, "OECD Employment Outlook 2010," 2010.

OECD, "Technical report of the Survey of Adult Skills (PIAAC)," Technical Report, OECD 2013.

Ohta, Soichi, "Young and Elderly People in Labor Force: Reconsideration of the Substitutability," *The Japanese Journal of Labour Studies*, 2012, *54* (9), 60–74. (in Japanese).

Olivetti, Claudia and Barbara Petrongolo, "The economic consequences of family policies: Lessons from a century of legislation in high-income countries," *Journal of Economic Perspectives*, February 2017, *31* (1), 205–30.

Ponczek, Vladimir, "Income and Bargaining Effects on Education and Health in Brazil," *Journal of Development Economics*, 2011, *94* (2), 242 – 253.

Rasmussen, Astrid Wurtz, "Increasing the length of parents' birth-related leave: The effect on children's long-term educational outcomes," *Labour Economics*, 2010, *17* (1), 91 – 100.

Rossin, Maya, "The effects of maternity leave on children's birth and infant health outcomes in the United States," *Journal of Health Economics*, 2011, *30* (2), 221 – 239.

Rossin-Slater, Maya, "Maternity and family leave policy," in Laura M. Argys Susan L. Averett and Saul D. Hoffman, eds., *Oxford Handbook on the Economics of Women*, Oxford University Press, 2018.

Ruhm, Christopher J, "The economic consequences of parental leave mandates: Lessons from Europe," *Quarterly Journal of Economics*, 1998, *113* (1).

Sakamoto, Kazuyasu, "Does the System of the Employee's Pension Division Empower Wives?," *Japanese Journal of Research on Household Economics*, 2008, *80*, 17–30. (in Japanese).

Sarsons, Heather, "Interpreting signals in the labor market: Evidence from medical referrals," 2017. *Working paper.*

Schönberg, Uta and Johannes Ludsteck, "Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth," *Journal of Labor Economics*, 2014, *32* (3), 469–505.

Stearns, Jenna, "The long-run effects of wage replacement and job protection: Evidence from two maternity leave reforms in Great Britain," 2018. (Unpublished).

Thévenon, Olivier and Anne Solaz, "Labour market effects of parental leave policies in OECD countries," OECD Social, Employment and Migration Working Papers 141, OECD Publishing January 2013.

Thomas, Mallika, "The impact of mandated maternity benefits on the gender differential in promotions: Examining the role of adverse selection," 2018. Mimeo, Cornell University.

van der Klaauw, Wilbert and Kenneth I. Wolpin, "Social Security and the Retirement and Savings Behavior of Low-Income Households," *Journal of Econometrics*, 2008, *145* (1), 21 – 42. The use of econometrics in informing public policy makers.

Voena, Alessandra, "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?," *American Economic Review*, August 2015, *105* (8), 2295–2332.

Yamaguchi, Shintaro, "Effects of Parental Leave Policies on Female Career and Fertility Choices," *Quantitative Economics*, 2019, *10* (3), 1195–1232.