WORK-LIFE BALANCE AND GENDER EQUALITY IN JAPAN

A Dissertation

Submitted to the National Graduate Institute for Policy Studies (GRIPS)

in Partial Fulfillment of the Requirements for the Degree of

Ph.D. in Public Economics

by

Chie Aoyagi

September 2018

Abstract

The traditional Japanese work style appears to negatively affect the work-life balance of workers and to contribute to persistent gender inequality in the work place and at home. Regular workers in Japan are often expected to follow the ideal worker image, which often requires them to put work before personal life. As a result, men are more likely to struggle to meet the ideal worker image and women are more likely to be in non-regular employment so as to cut back on work and take care of home and family. The aim of this study is to determine if offering of limited-regular contracts, with lower pay and security but reduced requirements for overtime, relocation and transfers compared to regular contracts, could improve workers' welfare (in Chapter 2), and whether an offer of opportunity for males to engage in housework or childrearing would improve gender equality within couples (in Chapter 3). These are two main chapters in this dissertation. Chapter 2 presents an online Choice Experiment (CE) and Chapter 3 an economic experiment with married couples, investigating worker acceptance of and the effectiveness of those policy measures to address the issues of work-life balance and gender equality in Japan. This study, presented in Chapter 2, is the first use of a CE in a Japanese labor market study. More than 1000 subjects participated in the CE, in which they made a series of best-worst choices from a set of three jobs described by five attributes such as annual wage, overtime, employment security, transfer possibility, and relocation possibility. It was found that people are willing to forfeit a significant portion of their wage to avoid extreme overtime and job transfer. The results of the study also suggest that willingness to pay (WTP) varies across gender, family structure (i.e. presence of children), and emotional state (i.e. guilt level) of workers. The

study advances the policy argument about the use of limited-regular contracts, by quantifying the tradeoff between wage and non-wage job attributes for workers with various characteristics. Towards a complementary policy to improve gender equality, this study also explores the effectiveness of prior task allocation for improving equality in task division among couples, by examining experimentally testing the effect of additional experience of tasks on preference for doing those tasks, presented in Chapter 3. This is the first economic experiment conducted with Japanese married couples. In the experiment, a total of 51 Japanese couples performed two kinds of tasks, one paid task and one unpaid task. In each couple, one person was randomly assigned to get more experience of the paid/unpaid task. Then the husband and the wife separately indicated their preference regarding the division of work with their partner. Contrary to expectations, despite subjects' different prior task experience and gender, their indicated preferences of task division were similar: subjects gave higher preference to the Traditional task pair (i.e. male breadwinner option) over the Reverse task pair (i.e. female breadwinner option) or the Mixed task pair (i.e. half-time dual earner option with equal unpaid task sharing). However, when taking into account subject performance of the paid task, the results of regression analysis suggest that while more productive males favor traditional task division more than less productive males, more productive females prefer the traditional option less than less productive females. This could be a reminder that the success of labor market reform may depend at least in part on whether the reform package addresses gender stereotypes.

Dedication

This dissertation is dedicated to my husband who encouraged me to pursue my dreams.

Table of contents

List of Tables

List of Figures

Acknowledgement

Firstly, I would like to express my foremost gratitude to my main advisor Prof. Alistair Munro for the continuous support of my Ph.D. study, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study. I would also like to thank the rest of my thesis committee: Prof. Masako Kurosawa, Prof. Makoto Tanaka, Prof. Hitoshi Ieda, and Prof. Sachiko Kuroda, for their kind and generous support in helping me to improve my research despite their busy schedules.

I would also like to express my sincere gratitude to Mr. Lawrence Hunter who patiently guided me throughout my dissertation writing. My appreciation goes to all participants at the PA workshop whose suggestions and criticisms helped me to improve the dissertation.

I thank my colleagues at GRIPS, particularly Joseph Upile Matola, Sharif Mosharraf Hossain, Hailegabriel Abebe Fenta and Ayibor Raphael Edem for the fruitful discussions and kind support during my study in GRIPS. This work was supported by JSPS KAKENHI Grant Number JP17H02498, ’Towards a behavioral theory of the household.'

Last but not least, I would like to express my gratitude to my family for their love, support, and encouragement. I am grateful to my husband, Yosuke Aoyagi. Without his precious support, it would not be possible to conduct my Ph.D. study and pursue my dreams. I would like to thank my father and mother-in-law for their active cooperation in taking care of our children so that I could have a peaceful mind to study. And finally, I cannot end my appreciation without thanking my mother for supporting me spiritually throughout my Ph.D. study and my life in general.

Chapter 1

Introduction

1.1 Introduction to the Dissertation: Rationale of the Study

It is said that regular workers in Japan are more likely to be asked by their employers to work long hours and change their work location and work duties than regular workers in other countries (Tsuru 2016). In exchange for that high work commitment, Japanese regular workers have enjoyed nearly guaranteed employment through to retirement, though Japan's prolonged economic stagnation may have made the system unsustainable, and companies have started relying more and more on non-regular employees so as to adjust to changes in the business environment (MHLW 2014). As this change in employment practice was not accompanied by fundamental reform of work style of regular workers, it contributed to the persistent gender inequality in the work place and at home (i.e. most non-regular workers are women, and more housework is done by women). This study investigates the workers' acceptance and effectiveness of possible policy measures to address the issue of work-life balance and gender equality in Japan from the viewpoint of the workforce, using an online Choice Experiment (CE) and an economic experiment.

This short introductory chapter is organized as follows. Section 2 in this chapter (1.2) further discusses the sources of the work-life-imbalance in Japan. Section 3 (1.3) describes in detail the objectives and contributions of this study. Finally, section 4 (1.4) presents a roadmap of the dissertation.

1.2 Work-life-imbalance in Japan

In this section, stylized facts and previous studies on the Japanese work style and lifestyle are presented.

1.2.1 Overtime

According to OECD statistics, the average hours worked by Japanese workers annually has been decreasing rather steadily since 2000 (Figure 1.1). However, since the data includes part-time workers, whose work share has tended to increase in recent years, it may conceal persistent overtime of regular workers. According to the *Karoshi* white paper by the Ministry of Health Labour and Welfare (MHLW 2017), the decline in the annual total working hours per worker per year is due to the increase in the share of part-time workers. Similarly, although OECD cross-national comparison data shows that Japan does not stand out in terms of long working hours among G7 countries (Figure 1.2), this does not mean that long working hours are not a problem in Japan.

Kuroda (2010) analyzed a Japanese time-use survey and found that holding demographic changes constant, hours spent on market work per week has been relatively stable for the last two decades for both male and female full-time workers. Kuroda (2010) also found that, on average, Japanese males work 10 hours per week longer and Japanese females work 7 hours longer than Americans, even after adjusting for demographic differences between the two countries.

The median length of overtime work by Japanese regular full-time workers is about 15 hours per month, but there is significant variation across jobs and workers (MHLW 2016). Also, there are substantial gender differences. Long working hours are most common among males in their 30s and 40s (Figure 1.3). An international comparison also shows that

Japanese men are the most likely to work long hours (MHLW 2017) (Figure 1.4). Pervasive long-working hours among men is of concern because previous studies seem to agree that husbands' long work hours and late arrival home after work negatively affect their housework share (Nishioka and Yamauchi 2017). Men in their 30s and 40s are typically of prime child-rearing age, so their absence at home would mean that unequal housework and child-rearing burden is placed on their wives.

On the other hand, if this long work hour culture of Japan is driven by worker preference, people may not be unhappy about working late. In fact, the preferred work time of both male and female workers in Japan, on average, is longer than that of workers in the U.K. and Germany (Kuroda and Yamamoto 2014). However, the results of the same study also suggests that worker preference for long working hours in Japan is not related to earning more money for consumption or leisure. Kuroda and Yamamoto (2013) point to the existence of the peer effect, which they argue affects Japanese workers strongly as one major reason behind the long working hours. This suggests that Japanese workers would also reduce their work hours if society changed and if everyone, including their bosses and colleagues at work, started going home earlier.

1.2.2 Mandatory Relocation

Another unique feature characterizing Japanese regular employment is mandatory relocation, which makes life planning more difficult for employees and could threaten their work-life balance. In Japan, it is not common to have a specified job description in an employment contract at the time of hiring, and as priority is given to employment security, or securing long-term employment, employers are allowed to arrange workers freely (JILPT 2015). In court cases, unless an agreement to limit the work location is explicitly included

in the employment contract, it has long been accepted that employers order their employees to relocate (MHLW 2017). JILPT (2015) explains in detail the situation surrounding the mandatory relocation practice of company personnel management. The main points are summarized below:

- Among those who are currently being relocated, the average number of relocations they have made is 2.2 times for workers who relocated with their families, and it is 3.0 times for workers who relocated without their families.
- 70% of those currently on relocation are given a three-year term for a single assignment, 20% a three to five year term, and 10% a term of more than five years.
- According to a MHLW Employment Management Survey, in the year from January 1998 to December 1998, 37.0% of companies ordered work-related relocation for workers. The share of relocations was as high as 95.8% for large companies (with 5,000 or more employees).
- In 2004, about 80% of companies with more than 1000 employees had employees who were living separately from their families to engage in a work assignment. While that share is higher for larger companies, it was about 20% for all companies surveyed with 30 or more employees. The number of companies with female employees living separately from their family for work was small (only 7% for even the largest companies) (Figure 1.5).
- In 2012, about 2.5 % of all male workers were married and living separately from their family for work. The share was 0.8 % for female workers (JILPT 2017).

In summary, it seems like although not all workers experience mandatory relocation, employees of large firms in particular have a higher probability of being asked to relocate

several times, for a few years in each assignment. Moreover, male workers are more likely to live separately from family upon agreement to relocate than female workers. This could be for two reasons; married women often quit if they are ordered to relocate by their employers; and working women are usually not at risk of mandatory relocation because their contracts rule out the possibility, in exchange for less security or lower pay (e.g. the case of non-regular workers). In addition, women are more likely to accompany their husbands (quitting if they were working) when the husbands receive relocation orders, especially when they have small children (JILPT 2015). The situation regarding mandatory relocation practices suggests that unless the requirements are eased, it is difficult for both husband and wife to continue working in regular employment.

1.2.3 Mandatory Intra-firm Transfer

As in the case of mandatory job relocation, employers are allowed to freely arrange workers across different departments and divisions within the firm, in exchange for almost guaranteed job protection for regular workers. The difference between transfer and mandatory job relocation is that while job relocation requires the worker to change his/her place of residence, a transfer is a wider concept which includes cases where workers are transferred within the same building or to a location near their home. In this dissertation, I distinguish relocation and transfer by whether or not the assigned work station is so far away from the worker's current place of residence that relocation is necessary to perform the duty.

Many companies rotate employees so that they can acquire a broader set of skills, and some companies transfer employees from time to time to maintain employment when the business is struggling or to allocate business resources efficiently (Sato, Fujimura and Yashiro 2007). Some workers might welcome transfer orders, perceiving them a necessary step to climbing

the corporate ladder. Nevertheless, being available for transfer order means giving up work autonomy to a great extent, and which could make balancing work and life difficult for the worker.

Another disadvantage of rotational intra-firm transfers of workers is that it might prevent the worker from acquiring long-term experience in one field. As companies prefer hiring job changers who have longer years of experience in a given field (Kurosawa 2002), career disruption in the same field caused by transfers would lower the employability of workers who wish to switch jobs. In Japan, while male regular workers tend to stay in regular employment all their working life, about half of working women quit when they give birth to their first child (Cabinet Office 2017). The prevalence of transfer of regular workers thus makes it harder for female regular workers with career discontinuity to return to a similar position after giving birth, thus could further widen gender gap.

1.2.4 High Job security

While the high job security is not in itself bad for workers, Japanese regular workers' high job security is linked to other problematic characteristics of employment practices, as discussed above. Life-long employment, or being employed by the same employer from the first year after final schooling until retirement, is believed to be a status enjoyed by traditional regular workers in Japan (Yamada and Kawaguchi 2015). In today's economy, though, employment security varies across workers and across employment contracts. While direct measurement of employment security is hard to obtain across different employment types, average years of tenure can be a proxy. A summary of the survey conducted by JILPT (2011), Figure 1.6, shows that regular workers' average tenure is more

than double the tenure for non-regular workers, implying the fact that regular workers' employment is on average still more secure than that for other types of employment.

International comparisons of average tenure of workers and share of workers whose tenure is longer than ten years suggest that Japanese workers tend to work for the same employer for more extended periods of time (Figure 1.7 and Figure 1.8). Although this may occur in part due to demographic differences across countries (i.e. the Japanese population is one of the oldest in the world), it also reflects the low labor mobility and slow or infrequent labor turnover in Japan. The inflexible labor market is problematic for at least one reason: as many women quit when they give birth to their first child, as discussed above, the underdeveloped secondary labor market is inconvenient for them if they wish to re-enter market work after childbirth. This too could widen gender inequality.

1.2.5 Housework Share

While the above mentioned employment practices related to regular workers seem to contribute to a widening of the gender gap (for males in favor of typical regular employment and for women in favor of non-regular employment or not working), we could speculate that those gender gaps in the work place also affect housework share at home.

As seen in Figure 1.9, the average daily hours spent on housework and child care for couples with children under age 6 in Japan is almost eight hours per day for women, but only one hour per day for men. However in other developed countries such as the U.S., the U.K., France, Germany, Sweden and Norway, women spend 5-6 hours on house work and child care while men spend 2-3 hours, and the gap within a couple is not as large as Japan (Cabinet Office 2004-2015). This striking gender difference suggests that equal sharing of domestic

work and paid work responsibilities is especially important for the improvement of gender equality for married couples in Japan.

In this section, we have seen problems regarding Japanese work-style and life-style. In the following section, I will discuss how this study addresses these issues, and how the findings would contribute to the literature.

1.3 Objective and Contribution of the Study

1.3.1 Overview of Chapter 2

In the previous section, we have seen that the amount of overtime, the possibility of mandatory relocation and intra-firm transfer, and the degree of job security differ widely across workers and employment types. Regular workers tend to enjoy more stable job than non-regular workers, but they may need to work long hours and have little say in where to work or what to do at the time when their employers make relocation and transfer orders. If regular jobs which put a "limit" on these obligations (such as long overtime and mandatory relocation and transfer) were widely available, it would help both male and female workers to balance work and family responsibilities. This concept of limited-regular (in Japanese *gentei seishain*) contracts have been intensely discussed among policy makers and academia in the last few years (Imano 2012) (Tsuru 2014) (MHLW 2014), especially after the government's Regulatory Reform Council's working group on employment started discussing the concept in 2013. However, whether the contract would be accepted by the wider society is still questionable. Chapter 2 of this dissertation aims to answer the question by the mean of a Choice Experiment (CE). No other studies has ever used CE of this kind and the design of the experiments are in large part my original. The online CE offers more

than 1000 respondents different features of hypothetical jobs, aiming to analyze the preference of Japanese workers.

While the main alternative to CE is the hedonic pricing model, that approach is not fully adequate. First, there is no publicly available large-scale reliable data on pay and job characteristics. Second, Job characteristics are often standardized in reality and the variation in the characteristics found in the sample is small compared to the feasible range of characteristics. Third, it is difficult to distinguish worker preference from employer preference in an existing wage profile. CEs have shown their worth in other areas, for example, marketing, health economics, transportation, agricultural and environmental economics, to examine consumer preferences. This chapter is the first to use CE to analyze the job preference situation in Japan, and thus can be a valuable addition to the body of literature.

The main findings of this chapter suggest that there is potential for legal clarification and dissemination of limited regular contracts, which offer lower pay and security but reduce the requirements for overtime, relocation and transfers compared to regular contracts, to improve workers' welfare (under the same quality of job as regular contracts). Moreover, the empirical estimations of willingness to pay (WTP) to avoid overtime, relocation and transfer are found to be higher for women than men, suggesting limited-regular contracts are particularly useful for women. This implies that in order to attract as many male workers as female workers to accept limited regular contracts, additional factors not captured in the stated choice questions, such as promotion prospects and opportunity of career developments, would be important. When taking into account one's guilt level in situations that typical working parents encounter, it was found that women tend to feel guiltier and

have higher WTP than men for avoiding work responsibilities. This suggests that "anti-guilt" policies targeted female workers could be effective to improve gender equality.

1.3.2 Overview of Chapter 3

We have seen in the previous section that long working hours and mandatory relocation are more common among male workers than female workers. Indeed, as the norm for ideal employee behavior in firm-internal labor markets is conceptualized as an employee who puts company above family and willingly spends almost unlimited 'face time' with his/her employer and colleagues (Brinton and Mun 2016), it is often very difficult for both the husband and the wife to follow that norm, and it is usually the husband who struggles to meet that ideal worker image. The flip side of this is that the wife takes on a much larger share of housework. The importance of work-style reform and the dissemination of limited-regular contracts are discussed in Chapter 2, but that may not sufficiently address the gender gap in the division of labor within couples if the current gendered preference remains unchanged.

Indeed, after 2017, when the government announced plans to propose an overtime cap and some companies voluntarily started releasing workers earlier, NHK (Japan's national public broadcasting organization) broadcast a TV program featuring *Furari-man* (men wandering around without going home after work). This abbreviated combination of the words "*furari (*wandering, or dropping by on the way*)"* and "*sarari-man* (salaried men)" was coined by a social psychologist in 2004 (Shibuya 2004), but was not widely recognized before a number of private broadcasting stations, newspapers, magazines and Internet articles featured the topic following the TV program. According to those media reports, it seems that there is an increasing number of men who do not go straight home from work, but rather wander in the

park, coffee shops, or book stores, even if work finishes early, because they do not find a role at home. If the long-prevailing gendered division of labor within couples is causing men's unwillingness to go home to share housework, there may be a need for intervention to break the status-quo and improve gender equality. Policy, such as the introduction of paid-parental leave allocated specifically to fathers, would ensure that men get sufficient experience in housework and could be a useful measure to address gendered preferences. Chapter 3 aims to test whether additional experience of tasks affects preference for doing those tasks, through an economic experiment with Japanese married couples.

In the experiment, a total of 51 couples performed two kinds of tasks, a paid task and an unpaid task. In each couple, one person was randomly assigned to get more experience of the paid/unpaid task, and then both partners individually indicated their preference regarding division of work with their partner. Contrary to expectations, the results suggest that regardless of prior task experience or gender, subjects gave higher preference to the Traditional task pair (i.e. male breadwinner option) over the Reverse task pair (i.e. female breadwinner option) or the Mixed task pair (i.e. half-time dual earner option with equal unpaid task sharing). While the results did not give any clear evidence supporting the benefit of parental leave allocated specifically to fathers, they do suggest that the success of labor market reform might depend on whether it addresses gender stereotypes.

To the best of my knowledge, this chapter is the first to use economic experiments to analyze preference of the division of labor in married couples in Japan. Also, these findings would contribute internationally to the still under addressed body of literature, as one of the very few studies which used couple experiments regarding decisions on the division of work and

experimentally assessed the effect of a policy intervention on subject preferences. The experimental design in this chapter is in large part unique and original.

1.4 Roadmap to the Dissertation

This dissertation has three remaining chapters. Chapter 2 discusses work-life balance in Japan, and by means of a CE examines the extent to which wage workers are willing to forfeit benefits in order to improve their work style. Chapter 3 discusses the gender gap in the division of labor within couples and explores the possibility that some policy intervention might improve gender equality. Finally, Chapter 4 summarizes the main findings and some policy implications from each of the main chapters and indicates some possible extensions of research.

Chapter 2

Work-Life Balance in Japan[1]

2.1 Introduction

It has been said that Japanese labor market practices must be reformed to meet growing worker demand for flexible work arrangement (Cabinet Office 2017). Typically regular workers in Japan enjoy very stable employment (Kanemoto and MacLeod 1991) (Yamada and Kawaguchi 2015), but they are often required to work overtime and to follow their employers' orders to relocate or transfer to different work within the firm regardless of their own preferences (Sato, Fujimura and Yashiro 2007) (Tsuru 2014). Possibly this work style made more sense during economic growth period between the 1960s and 1980s, when most women quitted their jobs upon getting married so as to fulfil their responsibility to care for their families and homes while their husbands focused on their paid jobs and became the sole breadwinners in the household. However, prolonged economic stagnation and subdued wage growth after the burst of the bubble in the 1990s made it more difficult for families to continue with that life style (Imano 2012) (H. Yamada 2017), as reflected in the fact that since 1997 the number of dual earner households has exceeded the number of single earner households (with working husband and stay-home wife) (Cabinet Office 2014). The nature of the typical Japanese work style makes it difficult for both the man and the woman in a family to stay in regular employment; usually the woman ends up taking non-regular work or quitting work completely, in line with traditional gender roles (Steinberg and Nakane

[1] This work was supported by JSPS KAKENHI Grant Number JP17H02498, 'Towards a behavioural theory of the household'

2012). Perhaps there is a need for labor market reform aimed at changing this work style, which requires workers to make enormous commitments and sacrifices in their private lives. If regular jobs which put a "limit" on these obligations (such as long overtime and mandatory relocation and transfer) were widely available, it would help both male and female workers to balance work and family responsibilities. Therefore, the main component of any work style reform plan could be the dissemination of the use of limited-regular (in Japanese *gentei seishain*) contracts (Imano 2012) (Tsuru 2014).

Worker acceptance of jobs which would potentially improve their work-life balance depends on the interaction of multiple factors related to the jobs (Baum and Kabst 2013). One way of investigating workers' decision making factors is to determine the best combination of wage and non-wage characteristics of a job to allow the worker to cope with responsibilities at home and be satisfactorily rewarded. This chapter aims to estimate that combination, using choice experiment (CE) methodology. Unlike revealed choice data, stated choice data consists of choices made based on hypothetical situations. Since not many Japanese switch jobs (Kanemoto and MacLeod 1991) (Yamada and Kawaguchi 2015), and thus there is no large scale revealed choice data available, I believe the investigation of stated preferences by means of CE could contribute to knowledge of the job preference situation in Japan. The main alternative to CE is the hedonic pricing model, but it has its own problems too. First, there is no publicly available large-scale reliable data on pay and job characteristics. Second, Job characteristics are often standardized in reality and the variation in the characteristics found in the sample is small compared to the feasible range of characteristics. Third, it is difficult to distinguish worker preference from employer

preference in an existing wage profile. More discussion about the hedonic pricing model and a literature review on Japanese work-life-balance will be presented in section 15 (2.15).

The chapter is organized as follows. In the following section, I briefly describe choice experiments (2.2). Then I discuss issues of validity of CEs (2.3), present my application of CEs (2.4), and identify design dimensions of CE through a literature review (2.5), and explain about efficient designs (2.6). Then, in Section 7 (2.7), I describe the process of construction and implementation of my CE step by step, and in Section 8 (2.8) describe my final CE design. Section 9 (2.9) presents the derivation of the main model in my CE, the rank ordered logit model, and Section 10 (2.10) identifies the empirical model. Before the report of the econometrics, section 11 (2.11) presents basic information about my CE and a summary and the descriptive statistics of the survey, followed by section 12 (2.12) diagnostics of the stated choice data. Results of the CE are discussed in section 13 (2.13). In section 14 (2.14) I explain IIA assumptions and different specifications of robustness checks. In section 15 (2.15) I compare my CE results to those obtained by other means, such as hedonic pricing and other hypothetical survey methods. In section 16 (2.16), guilt level is considered as another potential factor influencing Willingness to Pay (WTP). And finally, section 17 (2.17) is a discussion of the policy implications of the results and discuss conclusions.

2.2 What is a Choice Experiment?

Stated choice experiments (CEs) are used as a means of identifying and evaluating the relative importance of factors of decision making. Stated choice data consists of choices made by subjects based on hypothetical situations, while revealed choice data consists of choices made in an actual market. In a choice experiment, each respondent is shown the

choice sets one by one and is asked to rank the hypothetical alternatives in each choice set or select the best alternative from each choice set. Alternatives are described by attributes whose levels are varied systematically.

Stated choice experiments, unlike revealed choice experiments, are particularly useful for capturing information about people's decision making on goods or services that do not yet exist or whose markets are underdeveloped. Stated choice experiments are widely used to examine consumer preferences in marketing, health economics, transportation, agricultural and environmental economics. Besides those usual areas of application, CEs could also be useful for analyzing labor market decisions in Japan.

Though the work patterns in Japan are changing, it is still common for Japanese employees to stay in one company until retirement (Yamada and Kawaguchi 2015). Since fewer Japanese switch jobs than workers in other developed countries, longitudinal data on job changing behavior are not available for analyses of the determinants of the actual labor market decisions (revealed preferences) made by Japanese workers during their careers (Kanemoto and MacLeod 1991). For that reason, the investigation of stated preferences by means of CEs could contribute to knowledge of the job preference situation in Japan.

2.3 Validity of Choice Experiments

The extent to which individuals might behave in accordance with their choice (as stated in a survey) is a major question concerning the validity of CEs. To minimize the deviation between hypothetical and actual experiments (hypothetical bias), researchers typically frame a CE in a manner that adds realism, so that the stated choices closely resemble the decisions. Few empirical studies have tested for hypothetical bias in CE, and there is no general agreement as to its existence (Hensher 2010).

The theory of planned behavior developed by psychologist Ajzen (1991) provides support for the validity of choice experiments. According to that theory, "intentions" to perform behaviors of different kinds can be predicted with high accuracy from personal evaluation of the behavior (attitudes), socially expected mode of conduct (subjective norms), and self-efficacy with respect to the behavior (perceived behavioral control). These "intentions," together with perceptions of behavioral control, account for considerable variance in actual behavior. This suggests that choices in stated preference surveys, or "intentions", should be similar to actual market choices. However, an additional element, perceived behavioral control, cannot be well explained by a CE, so an individual's actual behavior might differ from what the individual said he or she would do. Thus the validity of CEs varies across situations and across decisions.

2.4 Application of Choice Experiments

Though a CE could be a very useful tool for analyzing worker job preference, only a small number of studies have applied CE to job market research, and the majority of them examine job preferences of health/medical professionals. Lagarde and Blaauw (2009) reviewed ten studies that used discrete choice experiments to investigate the job preferences of health care providers in both advanced and developing countries. They found that the most common CE attributes besides wages are: workload, which was included in different forms such as hours worked per week, amount of after-hours work, patient list size, and staffing levels; location, which seems particularly important in developing countries; provision of housing; opportunity to benefit from further education; and improvement of facility management.

Baum and Kabst (2013) extended the above analysis to capture preferences of general job seekers by conducting a CE using graduate and undergraduate students in a German university, using the ten highest ranked job choice factors[2] as attributes in experiments (with all non-wage attributes having two levels). More recently, Yoo and Oh (2017) applied that methodology to identify policy measures to improve female labor participation and to narrow the gender wage gap. They conducted a CE using representative Korean respondents of age 19 and above, and calculated WTP (willingness to pay) for some family-friendly policies. Yoo and Oh (2017) used 5 attributes[3] including factors for essential to work-life balance, such as availability of parental leave and workplace childcare centers. Of the four non-wage attributes three have two levels and one has three levels. The above general job preference studies use relatively large numbers of attributes with small numbers of levels. It might be that the number of attributes could not be reduced to properly differentiate alternative jobs to reflect respondent preferences. On the other hand, the number of attribute levels might be set so as to simplify the model, though this would come at a cost, as we shall see in the following section.

Other more recent studies worth mentioning include Mas and Pallais (2017) and Wiswall and Zafar (2018). The former applied the CE to call center job applicants in the U.S. and

[2] The ten factors are work climate, security, work–life balance (operationalized by flexible working hours and working schedule), training, salary, person–organization fit, promotion prospects, task attractiveness and location.
[3] The five factors are annual salary, weekly working hours, and firm size, availability of parental leave and a workplace childcare center.

the latter to undergraduate students at New York University. Both studies found women's WTP for flexible work arrangement [4] is higher than men's.

To the best of my knowledge, no studies have investigated job seeker preference of traditional Japanese personnel practices through CEs.

2.5 Design Dimensions of the CE

The main design dimensions of a CE are number of choice sets, alternatives, and attributes, and number and range of attribute levels. Many authors have highlighted the influence of CE design dimensions and information structure on model outcomes. In general, while more complex experimental design would yield richer data, respondent ability to process complex information is limited, and complex experimental design could result in incomplete responses or inconsistent choices. Inconsistent results can occur if subjects develop their own decision rules and simplify complex information by considering only a portion of the information available in the choice set, or if they make more mistakes. There have been efforts to test those effects empirically, for example the former effect (respondents' developing own simplified decision rules) can be captured as biased estimates of attribute weights, or WTP, and the latter effect (increased respondent error) can be captured as larger error variance (Johnson, et al. 2013). While the problem of design complexity is understood widely, so far there is no agreement in the literature as to optimal CE task complexity, in part because those studies vary in field of study and experimental design (Table 2.1).

[4] The main flexible-work attributes of interest in those paper include flexible work scheduling, working from home, and the employer discretion over scheduling (Mas and Pallais 2017) and dismissal probability and work hours flexibility (Wiswall and Zafar 2018).

2.5.1 Number of Choice Sets

To decide the number of choice sets, we need to strike the right balance in terms of number of items that the respondents can learn from repeated choice tasks without excessive fatigue. While a larger number of choice sets might increase the number of dropouts (Meyerhoff, Oehlmann and Weller 2015) and increase status quo bias (Oehlmann, et al. 2017), too small a number of choice sets may also give rise to problems. Chung, Boyer and Han (2011) found that on average variance of error first decreases, then increases with increasing number of choice sets presented to the respondent, with the optimal number of choice sets per survey being 6. Similarly, Caussade, et al. (2005) observed a U-shaped relationship with error variance decreasing in up to 9 or 10 choice situations. Hensher, Stopher and Louviere (2001) found little evidence of fatigue effect, even for 32 choice sets.

2.5.2 Number of Alternatives

As in the case of the choice sets above, we need to take into account both positive and negative effects of increasing number of alternatives. If the number is too small, participants will not be able to find a better match to their preference, but if there are too many options available there could be a negative effect on respondents' choice consistency. As documented by DeShazo and Fermo (2002), Chung, Boyer and Han (2011) and Caussade, et al. (2005), error variance first decreases, then increases with number of alternatives, resulting in a U-shape quadratic relationship. DeShazo and Fermo (2002) concluded that precision of model estimation is maximized when the number of alternatives is three, while Chung, Boyer and Han (2011) and Caussade, et al. (2005) identified the optimal number of alternatives to be five and four respectively. On the other hand, (Meyerhoff, Oehlmann and

Weller (2015) found that designs with five alternatives result in higher drop-out rate than those with only three alternatives.

2.5.3 Number of Attributes

If the number of attributes is too small, participants may not find enough information to make choices to reflect their preferences, but if the number is too large, the cognitive load associated with greater information outweighs the potential increase in consistency induced by a more complete description of the alternative. As documented by DeShazo and Fermo (2002) and Caussade, et al. (2005), there is evidence strongly suggesting that an increase in the number of attributes results in an increase in error variance. Also, Meyerhoff, Oehlmann and Weller (2015) find that the probability of abandoning the survey significantly increases with the number of attributes.

2.5.4 Number of Attribute Levels and Level Range

As number of levels increases, number of comparisons increases and in turn experimental complexity increases. Caussade et al. (2005) found that number of levels has a significant effect on ability to choose, contributing to error variance. However, if the number of levels is too small, the model may fail to estimate more complex non-linear relationships. For example, if an attribute has only two levels, the analyst would be forced to conclude that the utility relationship for the attribute is linear (Hensher, Rose and Greene 2005). On the other hand, a small level range should minimize the variance of the error term because comparisons would be easier to assess, resulting in a more consistent process. Also, if attribute levels include extreme values, respondents may stop taking the experiment seriously, leading to inconsistent choice. Thus, a wider (narrower) level range would result

in greater (smaller) status quo bias (Oehlmann, et al. 2017) and higher (lower) error variance (Caussade, et al. 2005).

2.5.6 Structure of Information

DeShazo and Fermo (2002) claim that choice consistency is also affected by structure of information. They argue that increasing the number of attributes that differ across alternatives, increasing the mean correlation of intra-alternative attribute correlation, and increasing the dispersion of the correlation of attribute levels across alternatives all increase the variance of utility. On the other hand, Swait and Adamowicz (2001) found some evidence supporting their hypothesis that the variance of preference will be concave in their complexity measurement, entropy, which summarizes the impacts of number of alternatives, number of attributes, attribute correlation, and preference similarity among alternatives. They argue that beyond the complexity level where preference is characterized by the highest variance, increased complexity actually leads to preference consistency because of the similarity of alternatives, independent of respondent effort. Nevertheless, Oehlmann, et al. (2017) observed that respondents opt more often for the status quo when entropy, or similarity among alternatives, is high and therefore the decision is more difficult.

2.5.7 Presentation Format

Not only task design complexity but also the way in which the information is presented can affect respondent ability to correctly evaluate alternatives. However, Arentze et al. (2003) found that presentation format (visual versus non-visual representation of attributes) has little effect on the validity of stated choice data for respondents with limited literacy skills. Arentze et al. (2003) observed that adding pictorial material to a verbal description of attributes has no impact on error variance or on measurement of attribute weights. On the

other hand, other studies have found some evidence that adding visual stimuli can improve respondent information evaluability. Lack of information evaluability often occurs when information is presented only numerically, increasing response variability and leading survey respondents to resort to heuristics (such as loss aversion) to formulate responses (Bateman, et al. 2009). Bateman et al. (2009) concluded that virtual reality presentation (as opposed to only numerical presentation) enables respondents to more thoroughly evaluate the scale of gain and loss, thus narrowing preference asymmetry and reducing error.

2.6 Efficient Designs

The type of design generated should reflect the belief of analysts as to what is the most important property of the constructed design. The construction process of the designs known as *orthogonal fractional factorial designs* minimizes the correlations evidenced within a design to zero, or the attributes of the design are uncorrelated, but may not be the most statistically efficient design available. Also factorial design, in which each level of each attribute is combined with every level of all other attributes, is difficult to implement if problems (number of attributes and their levels) are too large to use complete factorials. On the other hand, *statistically efficient designs* of various kinds optimize the amount of information obtained from a design, though will likely have correlations (Hensher, Rose and Greene 2005). Efficient designs are based on the idea that minimize the size of the variance-covariance matrix given a prior for the parameters to be estimated. There are various ways of calculating the size of a matrix, which lead to different efficiency measures. The most commonly used efficiency measure is D-efficiency which minimizes the determinant of the covariance matrix. Alternatively, C-efficiency, which allows the experiment to be tailored to a specific application, can be used.

2.7 Process of Constructing and Implementing the Survey

As steps to construct and implement my stated choice survey, I conducted literature review, three focus group interviews and one pilot experiment with 100 subjects to determine the final attributes and their levels that are important to the respondents, as well as the vocabulary and language to be used in the survey.

2.7.1 Literature Review

My literature review consists on mainly three components. The first is the literature review on methodology of CEs, which includes topics such as validity of CE (covered in section 3 of this chapter (2.3)), design dimensions of CE (section 5 (2.5)), and efficient designs (section 6 (2.6)). The second is the literature review on CE on job choice decisions in general, referred in section 4 (2.4). Finally, the third components of the literature review is on Japanese work style reform which will be discussed in section 15 (2.15).

2.7.2 Focus Groups

Three focus groups were conducted during summer in 2017.

2.7.2.1 Recruitment strategy:

The three focus groups are “focused” on a given topic because they are purposive, although not necessarily representative, sampling of a specific population. Participants were selected on the criteria listed in Rabiee (2004) that they would have something to say on the topic, are within the age-range of the targeted survey respondents, have similar socio-characteristics and would be comfortable talking to the interviewer and each other: All selected participants were currently working, in their 20s-40s, and belonging to the preformed groups. In all three focus groups, I acted as the interviewer, and in two of the three focus group interviews, there was another person acting as a note-taker. More

description about those criteria and participants' work background information are following.

- **Group 1:** Four full-time staff members from the office where the interviewer is currently working. They are in their 30s and 40s. Though their employment contracts are different (one is directly hired by the office, two are seconded from other institutions, one is a dispatched worker.) and their work roles are also quite different (two are specialists and the other two are assistants) they have worked together and known each other and the interviewer for at least several months. This group interview was conducted during lunch-time on a work day in a restaurant near the office.
- **Group 2**: Two married couples whose children attend the same nursery school with the interviewer's daughter, who are in their 30s and 40s. All of them are full-time specialists and have a 3 year-old. The two men work for the IT sector (in a different companies) as an engineer and have experienced changing their jobs once. One woman works as an environmental researcher in an institute after working in a company which sells environmental monitoring equipment. Another woman is a radiation technologist in a hospital where she has worked for all 18 years of her career. This group interview is conducted over lunch in the interviewer's home on Sunday.
- **Group3:** Four students who attend the same school with the interviewer (and the note-taker). The group members are in the same program and have known each other, though they had not met the interviewer and the note-taker before. They are all currently studying full-time in a one-year Master's program financed by their

employers, who are in their 20s and 30s. Two of them are married and have a small child, one is married and does not have children, and another one is single. All of them have worked for private companies; two for construction companies, one for a meteorological company, and another for a railway company, through their entire career after their completion of undergraduate degree and never have changed jobs. The interviewer recruited them on campus when they were having lunch together. They were provided chocolates for cooperating the interview.

2.7.2.2 Strategies for Focus Group Questions:

The interviews were conducted in the order of Group 1, Group2 and then Group 3. In all three groups, prior to questions, participants were asked about their work experience.

For the main part of the interview, in Group 1, open questions were asked such as “What are your job search criteria?” and “What would enable you to better balance work and life?” followed by a question asking what they think about a list of job attributes to be included in the survey and the attribute levels.

In Group 2 and Group 3, to give them a clearer insight into the interview questions, some novel tasks were presented. The participants were handed cards and pens. They were asked to list the 3 things that they look for in a job most and 2 things they really avoid. Then the interviewer asked each participant to explain. They were also asked how the answer might change 5 or 10 years in the future. They were also asked about attribute levels.

To generate comfortable environment for participants to engage in discussions, the Group 1 interview was conducted in a casual setting, where the interviewer took a note of key statements while also facilitating discussion. On the other hand, for Group 2 and Group 3, a separate note-taker and an audio recorder were present in order not to miss any important

information and to add to formality. Findings from the three groups reported below were not sensitive to the difference in the way questions were asked and the conversations were recorded.

2.7.2.3 Findings from Focus Groups:

- The majority of participants mentioned **appropriate level of wage** as a factor for accepting a job. Tolerance of wage cut varied among participants. One participant said he would not want to tradeoff wage against "anything else" and could not accept even a very small decrease in wage, because he had recently bought a house and had to pay off the mortgage. On the other hand, another participant said she would accept a wage decrease up to the level at which she could have "the average" level of life if some other factors were improved in return. On the upside, however, none of the people interviewed had ever changed jobs purely for higher pay. Restricting wages not to exceed the limit for spousal tax and insurance exemptions was not mentioned by any participant, probably because all were currently earning above the threshold.

- Another frequently mentioned phrase was "interpersonal relationships in the workplace", by which they often mean **fairness of staff evaluation** in various aspects. Some said they appreciate bosses who "evaluate their work properly" and want to avoid bosses who give "unfairly harsh evaluations." Some other participants reported "(I get) negative reactions from coworkers when I have to leave the office early or take leave for family reasons" because their absence increased the work load of their coworkers with no feeling of being rewarded. They said "(those) tensions among coworkers must be avoided" and "sympathetic coworkers and work

environment" are things they look for in a job. On the other hand, there were opinions such as "Workers who conduct assistant tasks must not stay in that same position for a long time" and "Team members whose skill levels are low must improve their skills or leave," calling for some form of punishment to nonproductive workers.

- Work culture or policies related to **work days and times** also received a lot of interest. There were opinions that "additional paid leave" or "paid or non-paid but longer leave" to take care of family members is a desirable condition in a job. Many agreed that excess overtime should be avoided. While some reported preferring to have zero overtime, others said up to certain level (e.g. regularly working until 8 PM) is "not a big problem" but overtime above that level should be avoided. There was no discussion of part-time work or shortened work hours, probably because all participants were currently working full-time.

- Some people raised the issue of **work location**. One participant who works as a dispatched worker said she chose her work "by the convenience of the location." One participant said his wife, who was not a member of the focus group, "improved her work life balance by working from home" though the option was not available to him. Four people interviewed said "(they) have a high probability of being asked to do mandatory relocation," and all those who reported a high probability reported mandatory relocation as one of the factors they want to avoid.

- Many comments that emerged in the discussions were related to **work content**. People seemed to like jobs that offer "opportunities to improve their skills and experience" and "prospects for advancement." Challenging and productive jobs seem to give satisfaction. Also people prefer to have work autonomy and want to avoid "forced work," "menial labor" or "administrative work." Two participants who work as assistants did not mention work content as an important factor. While many people considered work content quite important, mandatory transfer to a different work assignment within the origination or across group/related organizations drew mixed reactions from the participants. Those who thought mandatory transfer is a good thing said "(it) gives (them) an opportunity to acquire a wider set of skills," "improve experience," and "built networks" given the fact that the typical personnel rotation is "on an interval of 3-5 years".

- Aside from the one dispatched worker in the focus groups, none of the participants mentioned **employment stability** as important factor. A possible explanation for that is that the other participants interviewed already had very stable employment and it was hard for them to imagine the situation where they would have to worry about losing their jobs. Even those who work in the IT sector, which is known as having relatively high share of job switchers, did not raise employment stability as a determinant for deciding jobs. Many people also stated that it is difficult to set levels of the employment stability in the survey.

- Some participants argued that it is hard for them to imagine a situation where they have to trade off one attribute for another because those matters are "beyond their

control” in their current place of employment. After employees agreed to the contract, for example, relocation and transfer of assignments were decided “solely by the employer” and employees must follow orders or “(the employees) will have to quit”. Changes in the contract cannot be negotiated by individual employees, but “negotiation is done by the labor union”. The uniqueness of choice experiments is their ability to generate hypothetical choice sets and quantify the importance of otherwise unmeasurable attributes. This argument suggests the need to ask in the survey about members’ hypothetical job choices by explicitly asking within a hypothetical situation where they had no current job (i.e. no status quo option) to make the choice easier for the respondents.

2.7.2.4 Conclusions from the Focus Group:

Important factors arising from the debate over limited- regular contracts, such as limited work time, work location, and work content were found in all focus groups, implying that the use of those attributes in the survey would be necessary. Although excess overtime as a problem was a consensus view, some people reported preferring to have zero overtime while others said they would accept overtime to some degree. The number of overtime levels and ranges should be determined as a design issue but should also take into account legal limits. Though work location seem to be important, mandatory relocation is not an issue for many people, as a large share of jobs have zero or very little probability of mandatory relocation. For that reason, level of mandatory relocation should be either zero or some possibility. Again the number of levels should be a design issue. Work content, while having many aspects, was the most commonly perceived problem, though it would be difficult to be incorporated in the CE and thus treated as constant across alternatives. The mandatory

transfer was another frequently used term in the focus groups, which seems to be an important attribute to be included in the survey.

2.7.3 Pilot Experiment

A pilot experiment was conducted in November 2017, one month before the actual CE. The two experiments differ in terms of timing, and number of respondents (effective responses in the pilot experiments were about 1/10 of the size), and in terms of one question, which was found to be necessary and was added after the pilot experiment. The added question asks what attributes respondents ignored when making their stated choice decisions (Q29 in Appendix 2.1). A total of 107 respondents completed the survey, with a dropout rate ranging from 3% to 16 % for each block of choice sets presented. The resulting stated choice data seemed to capture well the variety of respondent preference, as most choice sets yielded divided responses across the alternatives. The preliminary regression results for the data confirm that the variables of interest are worthy of further investigation as all the coefficients were significant and had the expected signs. In that light I carried the initial empirical strategy over to the actual CE. The coefficients obtained from a preliminary regression analysis of the pilot data were adopted as the prior for getting D-efficient design of the stated choice survey. However, the resulting efficiency level was not very different from that for the initial design with prior of my judgment, so the initial design was carried over to the actual CE.

2.8 Design of the CE

The CE consists of eight choice sets with three alternatives per respondent. Moreover, respondents were randomly assigned to four blocks, with each block of respondents facing a different set of choice sets. In other words, the CE has 32 different choice sets in total, but

each individual is asked to consider only eight choice sets, so as to limit cognitive load and avoid exhaustion effect.

Table 2.2 shows an example of a choice set[5]. Job A, Job B, and Job C are the three alternatives. Each alternative is described by five attributes, annual wage, overtime, employment security, transfer possibility and relocation possibility. Each respondent sees eight choice sets repeatedly, making two decisions for each choice set, i.e. choosing best job and worst job from the three alternatives.

Based on a literature review and the results of the focus groups, the five attributes which seem to play important roles in people's job choices are included in the choice sets. Attribute levels in each choice set are set using D-efficient design with the initial set of priors which was determined to reflect differences in preferences. Those priors were my initial guess for β in the baseline model, as shown in equation (5) in Section 10 (2.10) but allowing for interactions between non-wage attributes. Then with those priors, I generated D-efficient design in Stata, using the ***dcreate*** command (Hole 2016).

Table 2.3 shows the levels and ranges of the attributes. In particular, the cutoff of 15 hours of overtime per month is the median for full-time regular workers in Japan reported by the Ministry of Health Labor and Welfare in the 2018 *Karoshi* white paper. 45 hours per month is the legal limit for overtime under normal circumstances, though during busy seasons of the year, companies are allowed to exceed the limit under certain conditions. "High" employment security reflects the stability enjoyed by current regular workers, and as regular workers are traditionally said to have lifetime employment in Japan, this level provides the

[5] An example of a choice set in Japanese (Appendix table A2.1) is in Appendix 2.2.

maximum job security. While the alternative, “medium” employment security, is somewhat lower than regular workers’ job security, it is still higher than the non-regular worker’s job security level (“low”) which is not included in the choice set. As exact measurement of employment security is difficult to obtain in reality, lengths of average tenure for regular workers, limited-regular workers, and non-regular workers are used as proxies for high, medium and low security, and are assigned to the respondents as an illustration. That information, along with some description of each attribute and its levels, are given to respondents as reference information before they start considering the choice sets.

The goodness of the initial D-efficient design was confirmed through one pilot experience with 100 subjects. The D-efficiency design using its parameter estimates was almost identical in efficiency level to the original design. Therefore, the original design was adopted for the actual CE with over 1000 subjects.

2.9 Derivation of the Models

In this section, the conceptual models of the CE and derivation of the ordered logit model are described. This section closely follows (Beggs, Cardell and Hausman 1981).

In the typical choice situation of the CE, there are observations on a vector of attributes for person i , s_i, and a vector of attributes for each element in his choice set, x_{ij} for $j = 1, \dots, J$, along with an indication of the chosen alternative, say choice k .

The basic specification I will use is the random utility model

$$\boldsymbol{U_{ij} = V(s_i, x_{ij}) + \varepsilon_{ij} = V_{ij} + \varepsilon_{ij}} \qquad \textbf{(1)}$$

where V_{ij} is the utility for the representative individual and is a deterministic component of the model, while the stochastic component $\boldsymbol{\varepsilon_{ij}}$ is assumed to follow some distribution function. For computational simplification, CE analyses often rely on logit model.

We now drop the person index, and as the basis of the logit specification, $\boldsymbol{\varepsilon_j}$ s are assumed to be independently and identically distributed extreme value random variates. The conditional distribution of the extreme value distribution assume the convenient properties, the independence of irrelevant alternatives (IIA) assumptions. The usual logit derivation that the probability that $U_j > U_k$ can be computed as

$$\boldsymbol{Pr(U_j > U_k,\ j \neq k) = e^{Vj}/(e^{Vj} + e^{Vk})} \qquad \textbf{(2)}$$

While the Multinomial logit (MNL) is the most often used choice model, in the following empirical analysis I employ **rank-ordered logit** specification. The distinction from the MNL is that dependent variable in rank-ordered logit records the rankings of the alternatives, whereas for MNL, dependent variable marks only the best alternative. Therefore, rank-ordered logit uses richer information and thus adopted as the main model in the empirical analysis explained in the following sections.

The ordered case can be computed as the extension of equation (2):

$$\boldsymbol{Pr(U_1 > U_2 \ldots > U_H,\ \text{for } H \leq J) = \prod_{h=1}^{H} [e^{Vh} / \sum_{m=h}^{H} e^{Vm}]} \qquad \textbf{(3)}$$

Where J=number of alternatives and H=number of alternatives ranked. In particular, J=H=3 for the CE in this chapter. As specified a particular linear in parameter form for $\mathbf{V}_{ij}$ in equation (1), $\mathbf{V}_{ij}=\mathbf{Z}_{ij}\boldsymbol{\beta}$ where $\mathbf{z}_{ij}$s are combinations of the s_i and x_{ij}s. Then for a particular person i the ranking of his/her J choices can be written as $\mathbf{R}_i=(\mathbf{r}_1,\ldots,\mathbf{r}_j)$ so that the probability of the observed ranking is

$$\boldsymbol{\pi}(\boldsymbol{R}_{\mathbf{i}}) = \boldsymbol{Pr}(\boldsymbol{U}_{\mathbf{r1}}>\boldsymbol{U}_{\mathbf{r2}}\ldots>\boldsymbol{U}_{\mathbf{rj}})=\prod_{\boldsymbol{h=1}}^{\boldsymbol{J-1}}[\boldsymbol{exp}(\boldsymbol{Z}_{\mathbf{irh}}\boldsymbol{\beta})/\sum_{\boldsymbol{m=h}}^{\boldsymbol{J}}\boldsymbol{exp}(\boldsymbol{Z}_{\mathbf{irm}}\boldsymbol{\beta})] \qquad \textbf{(4)}$$

Since for h=J the numerator and denominator cancel. For example, when h=j=3, this simplifies as

$$\boldsymbol{\pi}(\boldsymbol{R}_{\mathbf{i}}) = \left(\exp(z_{i1}\beta)/\sum_{rm=1}^{3}\exp(z_{irm}\beta)\right)\left(\exp(z_{i2}\beta)/\sum_{rm=2}^{3}\exp(z_{irm}\beta)\right)$$

= (probability the best alternative is chosen from full choice set) (probability runner-up is chosen from choice set without the best alternative)

For an independent sample of N individual the log likelihood is

$$\mathbf{L}(\boldsymbol{\beta}) = \sum_{i=1}^{N}\log\pi(\mathbf{R}_i)$$

$$=\sum_{i=1}^{N}\sum_{h=1}^{J-1}Z_{irh}\beta-\sum_{i=1}^{N}\sum_{h=1}^{J-1}\left[\log\sum_{m=h}^{J}\exp(Z_{irm}\beta)\right]$$

The log likelihood function is globally concave in β so that a unique maximum exists (Beggs, Cardell and Hausman 1981).

2.10 Empirical Model Identification

In this section, the empirical model that forms the basis of the analysis in the chapter will be explained.

Five attributes which seem to play important roles in people's job choices are included in the model: wage, overtime, relocation, transfer, and security. The deterministic part of an individual's utility in equation (1) can be specified as:

$$\begin{aligned} V = {} & \beta_0 + \beta_1 \boldsymbol{Wage} + \beta_2 \boldsymbol{Wage^2} + \beta_3 \boldsymbol{Overtime1} \\ & + \beta_4 \boldsymbol{Overtime2} + \beta_5 \boldsymbol{Overtime3} \\ & + \beta_6 \boldsymbol{Security2} + \beta_7 \boldsymbol{Transfer2} \\ & + \beta_8 \boldsymbol{Relocation2} \end{aligned} \qquad \textbf{(5)}^{6}$$

Where ***wage*** refers to annual wage (including overtime and bonuses) in million yen which we expect to have quadratic effects; ***overtime 2-4*** are dummy variables that represent the amount of overtime: ***overtime 2***=1 if overtime is between 0 and 15 hours (the average) per month, 0 otherwise; ***overtime 3*** =1 if overtime is between 15 and 45 hours (legal maximum at regular time), 0 otherwise; ***overtime 4*** =1 if overtime is more than 45 hours, 0 otherwise. The baseline is zero overtime. ***Security*** refers to the level of employment security; ***security2*** =1 if employment security is as strong as that of traditional regular employees, 0 otherwise. Assuming that traditional regular employees have had so called "life-long employment,"

[6] The baseline model as in the equation (5) is the linear model except for wage. I have tried interaction model allowing for interaction between non-wage attributes but it did not add much to this baseline model.

meaning workers are expected to stay in the same company until retirement, which gives maximum job security, the baseline job security is employment less secure than life-long employment (but more secure than non-regular employment). ***Transfer2*** is a dummy variable that represents the likelihood of being transferred by the employer to a different department/section within the firm which requires a major change in job contents: ***Transfer2*** =1 if there is possibility of intra-firm transfers, 0 otherwise. The baseline is zero possibility of intra-firm transfers. Finally, ***Relocation2*** is a dummy variable that represents the likelihood of being asked by the employer to relocate to a different branch which is too far away for commuting from the individual's current residence: ***Relocation2*** =1 if there is some possibility of mandatory relocation, 0 otherwise. That means there is no possibility of mandatory relocation in the baseline. We expect positive sign for coefficients of wage and security, and negative sign for overtime, transfer and relocation. We also expect negative sign for coefficient of wage^2, because we expect diminishing marginal utility from higher wage.

Note that the βs are not identified in the logit model. Rather the multiplication $\lambda\beta$ (where λ is the scale parameter defined as $\lambda = \sqrt{\frac{\pi^2}{6\sigma^2}}$, where π^2 is a constant and σ^2 is the variance of $\boldsymbol{\varepsilon}$) is identified (Hensher, Rose and Greene 2005).

The significance and magnitude of the β coefficients indicate the relative importance of those attributes that statistically influence respondent job preference. The marginal rate of substitution of any two variables represents the trade-offs made between the two attributes, and in particular, trade-offs between wage and another attribute provide estimates of willingness to pay (WTP) for the particular characteristic. WTP is a measure to determine

the amount of money individuals are willing to forfeit in order to obtain the particular characteristic of the job. In simple linear models, WTP measures are calculated as the (negative of the) ratio of wage and another parameter estimates, holding all else constant. However, as the equation (5) has $\boldsymbol{Wage^2}$ in the model, the WTP depends on the level of wage. As an example, below the simplified model with only ***Wage*** and ***Overtime*** is shown. WTP for Overtime measures the change in wage that compensates a change in Overtime such that utility remain constant. As shown in the last term in the third equation below, the WTP depends on wage level.

(Simplified Example)

$$V = \beta_1 \boldsymbol{Wage} + \beta_2 \boldsymbol{Wage^2} + \beta_3 \boldsymbol{Overtime}$$

$$dV = \left[\beta_1 + 2\beta_2 \boldsymbol{Wage}\right] d\boldsymbol{Wage} + \beta_3 \, d\boldsymbol{Overtime} = 0$$

$$WTP = d\,\boldsymbol{Wage} / d\,\boldsymbol{Overtime} = -\beta_3 / \left[\beta_1 + 2\beta_2 \boldsymbol{Wage}\right]$$

Though β is not identified, because $\lambda\beta$ is identified, WTP measures can be identified, as λ in the numerator and the denominator cancels out.

(Simplified Example)

$$WTP = -\lambda\beta_3 / \lambda\left[\beta_1 + 2\beta_2 Wage\right]$$

Furthermore, in addition to the variables related to the five attributes in the CEs, some socio-demographic characteristics of the decision maker, such as gender, age, marital status, education, having children, and income; might be important to an explanation of an

individual's job choice. However, if I just add those socio-demographic variables, they drop out of equation (5). Thus, those variables are only meaningful in the model to be included as interaction to the attribute variables, to identify differences in preferences arising from the decision maker's characteristics. In the result section, I will show some specification with interaction terms in which some important socio-demographic variables are interacted with other non-wage variables. The marginal utility of the individual with the particular characteristics q then becomes

$$\beta_k = \beta_{ko} + \beta_k' \boldsymbol{q} \quad for\ k > 2 \qquad \textbf{(6)}$$

Where q is a vector of socio-demographic variables and β_k is a vector of associated parameters.

For example, if q= (***Female, No_children***), where ***Female***=1 if the individual is female and ***Female***=0 if the individual is male, and ***No_children***=1 if the individual has no children and ***No_children***=0 if the individual has at least one child, then

$$\beta_k = \beta_{ko} + \beta_{k1}\boldsymbol{Female} + \beta_{k2}\boldsymbol{No\_children} \quad for\ k > 2 \qquad \textbf{(7)}$$

2.11 Basic Information about the CE, Summary and Descriptive Statistics

2.11.1 Basic Information

The CE was conducted by an online survey company from late December 2017 to early January 2018. The survey company used a panel of signed up potential respondents, and an

invitation to participate in the survey was sent to the targeted subjects, Japanese aged 20 to 59 with at least one job experience. Effective answers were collected from a total of 1,046 respondents.

2.11.2 Summary and Descriptive Statistics

The panel is diverse in terms of age, gender and other factors, but is not necessarily representative (see Table 2.4). In particular, as is the often the case with online surveys, the sample underweights young cohort. In an effort to balance young and old in the sample, the company collected responses so as to balance the 20s-30s group and the 40s-50s group. The mean age was 41.67. Though the sample may not be representative, it still covers the current and potential labor force, which is the focus of this study. More details about the characteristics of the respondents are provided below.

2.11.2.1 Age

As can be seen in Figure 2.1, there seems to be a linear positive relationship between respondent age and survey participation for the younger cohorts but that trend is not clear for the older cohorts. This is because the online survey company collected responses from the 20s-30s group and the 40s-50s group separately, so as to balance the number of responses from the two groups. As it is commonly the case for online surveys in Japan, it is difficult to collect responses from young people. In an effort to be sure to collect the information of the younger cohorts, the survey company took extra measures (e.g. more survey invitations were sent to signed up potential respondents) to secure a sufficient number of responses, thus probably led to differences in patterns of age-participation relationships between the younger and the older.

2.11.2.2 Education

As can be seen in Figure 2.2, my sample had a higher level of education achievement than the national distribution. There were 48 percent 4 year university graduates whereas the share for the national distribution is 26 percent (2012).

2.11.2.3 Prefecture of residence

Figure 2.3 shows high frequency of participation in the Kanto region (Tokyo, Kanagawa, and Saitama), and in Aichi and Osaka. While those prefectures have a high concentration of business activities and have large populations, my sample slightly overweighs subjects from those populous prefectures.

2.11.2.4 Income

My panel covers diversified income groups. Equivalent measurements of individual annual income nationwide were not available, so instead, Statistical Survey of Actual Status for Salary in the Private Sector (2016) was used for comparison. Though the share of high earners in my panel is slightly higher than the national comparison, my panel has a lower mean income than the national mean because many subjects reported zero income (Figure 2.4).

2.11.2.5 Current job, employment status

My panel has a higher share of regular employees than non-regular employees (sum of part-time employees, *arbeit* (non-student), and dispatched worker), and the ratio is about the same as the national distribution (2012), 7:3 (Figure 2.5). My panel has more housewives/househusbands and fewer students than the national distribution, but the ratio of working people to non-working people (including both housewives/husbands and students) is almost the same as in the national distribution, 8:2 (Figure 2.5).

2.11.2.6 Current job, company size and industry

Respondents who are currently working work for various sizes of employers and in various industries (Figure 2.6, Figure 2.7). The most populous industry is manufacturing (22% of working respondents in my sample, and 18% in the national sample).

2.12 Diagnostics of the Stated Choice Data

To confirm that my stated choice data was trustworthy and meaningful as a source of information, various checks were run to give diagnostics. Overall, the results of the checks are supportive for the use of the data in the interested analysis.

2.12.1 Drop-out Rates

After receiving invitation, the potential respondents came to the webpage and started answering the online survey, but some of them drop out without completing the whole survey. The number of drop-outs for the four choice set groups was 35, 29, 36, and 46, resulting the ratio of drop-out to initial respondents to be 10%, 9%, 10%, and 13% respectively. The drop-out rates seem to be reasonably low.

2.12.2 Coherence

One choice set had one alternative that clearly dominated the other alternatives, assuming that utility is increasing with wage, and decreasing with overtime, relocation and transfer. We can use that choice set as a check for response coherence. It was found that 89% of subjects who saw the choice set correctly chose the dominating alternative as the best. This indicates that the respondents actually considered each alternative carefully to make their decisions, rather than following some unreasonable simplified rules.

2.12.3 Ignored Variables

Some variables are more relevant for some individuals. It is possible that when making decisions, respondents completely ignore some of the variables that are less relevant to them. One question asked whether the participant had ignored any variables was included in the survey following the stated choice questions. 45% of the respondents stated that they did not ignore any of the variables when making decisions, and that they considered all variables in all the choice sets they were assigned. This is another indication that the included variables were important for job choice decisions and that the respondents took the CE seriously and responded with care. Respondents who ignored at least one variable when making decisions, most often ignored security and transfer, and ignored wage, overtime and relocation less frequently.

2.12.4 Monetary Evaluation of Extreme Overtime

In another question, respondents were asked, “How much additional wage would you like in order for you to switch to a job with 45 hours or more overtime per month?” Most respondents asked for a considerable increase. 46% respondents said they need to get at least 3 million yen more in annual wages than in their current job to accept 45 hours or more overtime per month, while 35% said they would never accept such long hours for a pay increase of 3 million yen or less. This result is consistent with the high WTP estimates obtained in the empirical analysis reported in the following section.

2.12.5 Reality and Applicability

In the survey, after the stated choice questions, there are questions asking how realistic and applicable the hypothetical job choice situations were for the participants. On average,

respondents stated that the hypothetical job choice situations were realistic and the situations applied to them. This suggests that the design of the CE successfully addressed the issue of hypothetical bias.

2.13 Results of the CE

In this section, I use rank-ordered logit regression to assess the importance of the factors influencing workers' job choice decisions. The dependent variable is the ranking of the particular job in the choice set (3 indicating the best job, 2 indicating the second best job, and 1 indicting the worst job). The independent variables include the five attributes described by equation (5) above.

2.13.1 Main Effects and the WTP

Table 2.5 shows the simplest model with no interaction and no socio-demographic variables. All coefficients have the expected sign and are significant at 1%, except the low level of overtime (overtime2, which captures overtime of less than 15 hours per month) with 5% significance level. WTP is calculated in million yen based on these estimated parameters (Table 2.6). Because the benchmark model has wage square, WTPs depend on level of wage. For example, the WTP for Overtime2 for wage 3 million yen can be interpreted to mean that people are willing to pay 0.08 million yen to avoid overtime2 (15 hours or less overtime per month). WTPs for overtime in particular could be also given a wage elasticity interpretation since elasticity of labor supply is proportional to the inverse of the WTP. For example, the WTP for Overtime2 for wage 3 million yen can be alternatively interpreted to mean that people are willing to be compensated 0.08 million yen more to do 15 hours or less overtime

per month. Basically, high WTP means low elasticity. It should be also noted that "WTP" is "WTA (willingness to avoid)" with the high numbers basically showing that individuals are very reluctant to accept high overtime. The sign of security WTP is negative, because unlike other attributes, having security is a good thing, so people are willing to receive (not pay) for losing high job security. All WTPs in Table 2.6 are significant at 1% for all wage levels, except WTPs for Overtime2 at the 5-10% significance level depending on level of wage[7].

Table 2.6 shows that people expressed strong dislike for overtime4 (i.e. 45 hours or more overtime per month). For example, people with an annual wage of 3 million yen are willing to pay 1.5 million yen, half of their annual wage, in order to avoid such overtime, and people in the highest wage group (8 million yen) are willing to pay more than their wage (11.7 million yen) to avoid 45 hours or more overtime[8]. Giving this the wage elasticity interpretation, the results mean that people with an annual wage of 3 million yen would seek for 50 % wage increase while people with an annual wage of 8 million yen would require doubling their wage to do 45 hours or more overtime. This would imply that labor supply of people in the highest wage group is much less elastic than that of people in the lowest

[7] To compute the standard errors, the delta method is used instead of bootstrap because Dowd et al. (2014)has shown that the two methods produce almost identical results.

[8] The derivative approach adopted in this chapter can overestimate WTP compared to calculating it directly by solving a quadratic equation. However, I am using the derivative approach for simplicity and using the direct approach would give the same relative ranking of the WTPs.

wage group. People are also willing to pay a high amount in order to avoid job transfer, ranging from 1.02 to 7.99 million yen, depending on their wage level.

It might be counterintuitive that WTP for transfer rather than relocation is so high, as relocation can be considered as an extreme version of transfer. One possible reason is that the share of workers who are under high pressure of relocation is lower than those under high transfer possibility, thus respondents do not evaluate avoiding relocation as worth of large portion of wage. In the sample, about 10 % of subjects stated that they have "high possibility" to be given relocation order within next five years, compared to 15 % of those with high transfer possibility within the same time frame.

2.13.2 WTP Based on Different Socio-demographic Characteristics

We can also calculate WTPs based on the extended model with interaction terms to see the differences in WTPs arising from specific socio-demographic characteristics of respondents. In the extended model (Model 1 in Appendix Table A2.2 in Appendix 2.3), I interact ***Female*** dummy, which takes value 1 if the respondent is female and 0 otherwise; and ***No_children*** dummy, which takes value 1 if the respondent has no children and 0 otherwise; and female and no_children interaction (***Female * No_children***) , which takes value 1 if the respondent is female and has no children and zero otherwise; and ***Age_cohort*** dummies (30s, 40s and 50s) which takes the value 1 if the respondent is in the specific age cohort and zero otherwise; with all variables except wages. In the model, therefore, the baseline is male in their 20s. Table 2.7 shows the WTP estimation using the parameters from the extended model. Hereafter, for simplicity, I fix wage at 3 million yen and age cohort at the 30s, but technically, it is possible to calculate WTPs for all six wage levels and each age cohort (20s, 30s, 40s and 50s).

The leftmost column in Table 2.7, "base" refers to the case "male, with children, in their 30s, with wage 3 million yen". The "female" column refers to the case, "female, with children, in their 30s, with wage 3 million yen." The "no children" column represents the case, "male, without children, in their 30s with wage 3 million yen." And finally, the "female* no_children" refers to the case "female, without children, in their 30s with wage 3 million yen."

The table is informative in several aspects. Firstly, when included interaction terms, some WTP becomes insignificant. For example, men with children do not seem to value security or dislike relocation significantly.

Secondly, the sign of WTPs for transfer may be counterintuitive for some people, as WTPs for transfer are all positive, meaning respondents dislike transfer, at any given respondent characteristics. Some people may like to experience different works for their career development or simply as a taste for variety. While that explanation may be right for some people, this empirical results (positive WTP) hold true for all those specified individual characteristics. And moreover, the distaste for transfer was observed across different education levels in another specification (see Model 3 in Appendix Table A2.3, and its WTP measures in Appendix Table A2.4 and Appendix Table A2.5[9] in Appendix 2.4.)

[9] The WTPs for non-educated people (with two-year college education or lower education levels) are found to be statistically higher in absolute value for Overtime4, Security2 and Transfer2 than those for educated people (with four year university education or higher education levels).

Finally, from Table 2.7 we can speculate that gender and presence of children might influence the level of WTPs for each attribute. We will examine whether the differences in WTPs across socio-demographic groups are statistically significant in the following.

2.13.3 Statistical Differences in WTP across Different Socio-demographic Characteristics

Table 2.8 and Table 2.9 report the comparison of WTP measures across gender and across different family structure.

Table 2.8 shows gender differences in WTP, calculated as male minus female. Positive (negative) sign indicates higher WTP for male (female) for all attributes but Security2. For WTP for Security2, which has either zero or negative value for all social groups, positive (negative) sign indicates higher absolute WTP for female (male). It shows that women tend to have higher WTPs in absolute value than men for most of the attributes, for both with children groups and with without children groups. The high female WTPs, especially for overtime, imply that labor supply of women are less elastic than that of men. These findings can add to the observations found in existing study, where women, in general, have higher WTP for flexible work arrangements than men. (Mas and Pallais 2017) (Wiswall and Zafar 2018)

Table 2.9 shows differences in WTP by the presence of children, calculated as with children minus without children. It suggests that when comparing the two male groups, men without children have higher absolute WTPs than men with children. On the other hand, when comparing between the two female groups, there are no statistically significant differences.

This finding may seem counter-intuitive, as we expect parents tend to value flexible work arrangement more to take care of their children. One possible explanation can be that fathers work hard for high pay (as reflected in smaller absolute WTPs for the which-children group than for without-children counterparts) to support his family, while mothers do not feel increased pressure to earn more to meet family demand. Another possible explanation is that men with higher tolerance to overtime etc. have the higher possibility of having children than men with lower tolerance, while whether a woman has children does not depend on her job preferences. There might also be a factor specific to Japanese value which resulted in the different responses to presence of children in female preferences from other countries: in Japan, many women consider working hard lowers her value in marriage market thus do not prefer to work like typical regular workers even before marrying or having children (Okuda 2018).

2.13.4 WTP and Existing Income

The above specification does not allow the WTP to depend on respondents' actual income level, however existing income might also affect workers' job choice. Here, I add an interactive dummy for existing high income and see how it affects WTP. In the sample, mean wage group was the annual wage of 2-3 million yen. The added ***High_income*** dummy captures income level higher than the mean and takes the value =1 if the existing wage is more than 3 million yen and 0 otherwise. ***High_income*** dummy is then interacted with non-wage attributes. The interaction tem is significant for most attributes except Overtime2 and Overtime3 (Model 4 in Appendix Table A2.6 in Appendix 2.5). Table 2.10 and Table 2.11 report the resulting WTP measures by existing income, for those who have higher than mean income and those who have lower income respectively. It seems that those who have lower income tend to have significantly higher WTP in absolute value, except WTPs for the Overtime2 with insignificant differences. This result implies

that people with lower existing income have less elastic labor supply, probably because those lower wages are resulting from their need to put limitations to work to better cope with personal life.

2.14 IIA Assumptions and Robustness Check

2.14.1 IIA Test

One common objection to the model adopted in the main analysis of this chapter is that the model carries the independence of irrelevant alternatives (IIA) assumption, which may be too restrictive. IIA means that, all else being equal, a person's choice between two alternative outcomes does not depend on the availability or attributes of the other alternatives. To see if the IIA assumption holds, I conducted commonly used IIA test, the Hausman and McFadden (HM) test, following the steps suggested by Cheng and Long (2007), comparing the estimated coefficients of the full model with the ones from restricted models (either the one with the best choice being dropped (Model (2) in Table 2.12) or the worst choice being dropped (Model (3) in Table 2.12)). To be precise, I used conditional logit equivalent to the rank ordered logit of my main analysis for the IIA tests. The idea is to treat the decision problem as a pair of problems: 1) select best from three options and then 2) select best from the two rejected options in the first choice. The significant values of HM in Table 2.13 indicate that the IIA assumption has been violated, thus by using either of the two restricted models, the test suggests IIA does not hold.

IIA assumption is less likely to be met under labeled experiments than under unlabeled experiments, because the label attached to an alternative acts like an attribute for the alternative and may be correlated with the attributes used within the experiment, or the

decision makers may use assumptions surrounding the labels attached to the alternatives as proxies for omitted attributes (Hensher, Rose and Greene 2005). Unlabeled experiment, in contrast to labeled experiment, is an experiment in which the heading or title of each alternative is generic or uninformative to the decision maker. Unlabeled experiments are believed to be more appropriate for studies for establishing WTP for specific attributes (Hensher, Rose and Greene 2005). Because the main purpose of this chapter is to establish WTP for job attributes, I adopted unlabeled experiment with each job alternative in a given choice set being named Job A, Job B, and Job C, thus it came as a surprise for me to find the rejection of IIA assumption in the HM test. Having said that, some studies have questioned the underlying IIA assumption which typically accompanies unlabeled experiments using rank-ordered or best-worst data. Those studies argue that in rank-ordered data or best-worst data, the obtained estimates are not stable across stages because an individual will pay more careful attention to their top choice rather than carefully ranking all alternatives (Hausman and Ruud 1987) or the respondents may exhibit different preferences in positive (i.e. best) choice situations from those in negative (i.e. worst) choice situations and thus affected by the framing effects (Rose 2014) (Giergiczny, et al. 2013). On the other hand, it is noteworthy that other studies found evidence that preferences are stable across ranking stages and symmetric between best and worst choice (Collins and Rose 2011) (Scarpa, et al. 2011). Moreover, when using the same experimental design, some studies have found that results from choice-based conjoint analysis and those of ranking-based conjoint analysis are not different (Alejandro, Oviedo and Campos 2008), especially in a small choice set setting (Akaichi, Nayga and Gil 2013). Based on those arguments both against and supporting best-worst data, I conducted another set of HM test using data with only the best alternative being recorded as 1 while other two alternatives as 0 (Model (4) in

Table 2.10) to be compared with the full model. Similar to the previous test results, the significant value of HM using this restricted model in Table 2.11 indicates that IIA does not hold.

However, as pointed by Cheng and Long (2007), the IIA tests often reject the assumption when the alternatives seem distinct, and thus they conclude that the tests of the IIA assumption that are based on the estimation of a restricted choice set are unsatisfactory for applied work. Moreover, the scale of the utility function is not identified, and thus these coefficients are not of direct interest to economists. What matters to the economic interpretations of the results is the consistency of WTP estimates. Table 2.14 reports the WTP estimation for wage 3 million yen in their 30s (in million yen) by using the restricted model (4) above. The comparison with the WTP measures with those calculated from full-model (as reported as Table 2.7 in section 13 (2.13.2)) implies that WTP measures are consistent across rank stages, thus supporting my use of logit model under IIA assumption (Table 2.15).

There are some alternative specifications. For example, a mixed logit model fully relaxes IIA assumption by allowing parameter estimates to vary across individuals, and a nested logit model partially relaxes the IIA assumption by maintaining IIA for choices within the same nest but relaxing it for choices across nests. However, findings from some studies imply that the coefficients estimation are likely to be similar to my estimation even if mixed logit model or nested logit model were used instead (Christiadi and Cushing 2007) (Dahlberg and Eklöf 2003).

2.14.2 Marital Status

Another specification as reported as Model 2 in Appendix Table A2.2 in Appendix 2.3 can be used to see if marital status influences WTP differently for male and female, like the way the presence of children does. In that extended model, I interact ***female*** dummy, which takes value 1 if the respondent is female and 0 otherwise; and ***married*** dummy, which takes value 1 if the respondent is married and 0 otherwise; and female and married interaction (***female * married***) , which takes value 1 if the respondent is female and married and zero otherwise; and ***age_cohort*** dummies (30s, 40s and 50s) which takes the value 1 if the respondent is in the specific age cohort and zero otherwise; with all variables except wages. Table 2.16 shows the WTP estimation using the parameters from the extended model for the annual wage at 3 million yen and age cohort at the 30s.

The comparison between the left two columns (“base” and “married”) suggests effects of marriage for men while the comparison between the right two columns (“female” and “female*married”) gives an idea for female. It seems like that male WTPs are higher in absolute value when they are not married than when they are married, while female WTPs do not differ very much between the two groups. This finding is in line with what we have seen in Table 2.7 about the presence of children effects. Just like the presence of children case, here we cannot say if the marriage changes men or there had been pre-existing differences between the two male groups before marriage. However, combined the two results, we can speculate that it is not likely the appearance of children which changes male

WTPs, given the fact that most people get married first and have children in Japan[10] (MHLW 2010).

2.14.3 Number of Children and Age of the Youngest Children

While the No_children dummy in the model in chapter 13 only captures the binary effect of the presence of children, it might be the case that raising children have both fixed-cost and variable-cost for parents and it might be worth separating the two effects. By fixed-cost, I mean the cost arising once having a first child compared to zero children, and it is assumed to remain the same regardless of the number of children or age of children. Once having a child, the life is never the same as before, with a lot more to worry about and a lot more to enjoy. On the other hand, variable-cost is defined as a marginal cost arising from having more children or having older children. It is often said that having two children does not cost as much as the double the cost of raising one child. Also, it can be expected that costs of raising one more infant and costs of raising one more pre-school child must be different. To see these differences, I run two other extended models with categorical dummy variables of 1) the total number of children (Model 5 in Appendix Table A2.7 in Appendix 2.6) and 2) age group of the youngest children in the family (Model 6 in Appendix Table A2.7 in Appendix 2.6). Those children dummies are interacted with non-wage attribute variables and with gender to see differences between men and women. Table 2.17-2.18 and 2.19-2.20 reeport the resulting WTPs from the models.

[10] Only 10% of birth given by parents older than the 30s are with a marriage period shorter than the pregnancy period. As this number includes the cases that parents get married soon after they discover the pregnancy, the occurrence of giving birth before marriage is expected to be even rarer.

Table 2.17 shows the difference in WTPs by the total number of children for men in their 30s with annual wage of 3 million yen, and Table 2.18 shows that for women counterpart. There seem to be no clear patterns across the different number of children for both genders. It is probably because our sample is limited in number when categorized by the different number of children as seen in Table 2.19. For example, there are no men who have four children, and there are only two women who have four children.

Table 2.20 shows the difference in WTPs by the age of the youngest children for men in their 30s with annual wage of 3 million yen, and Table 2.21 shows that for women counterpart. In the model, the age-of-the-youngest-children dummy is categorized as ***infant_toddler***=1 if the youngest children within the family are 0-3 years old and ***infant_toddler*** =0 otherwise, ***preschool***=1 if 4-6 years old and ***preschool*** =0 otherwise, ***elementary_school***=1 if 7-12 years old and ***elementary_school***=0 otherwise, and finally ***teenager***=1 if 13-18 years old and ***teenager***=0 otherwise. Again, there seem to be no clear patterns across different age of the youngest children for both genders. It is also probably because our sample is limited in number when categorized by the age of the youngest children as seen in Table 2.22. For example, there are only two men and two women whose youngest children are teenagers.

2.15 Comparison with Alternative Means to WTP

The main alternative for calculating WTP is hedonic pricing model, but it has its own problems. First of all, There is no publicly available large-scale and reliable data on pay and job characteristics. If we have information on the overtime, for example, and on wage for a category of employees, such as women with children, with comparison with other workers

in a same establishment, we can infer an overtime–wage trade-off for the category of employees. However, in general, official statistics on firms and establishments do not contain detailed information about employee characteristics (Morikawa 2017). Second, Job characteristics are often standardized in reality, and that means the variation in the characteristics, such as working hours, terms of overtime and so on found in the sample is small compared to the feasible range of characteristics. Third, existing wage profile reflect both demand side and supply side interest and it is difficult to distinguish worker preference from employer preference. In Japan, it is not common to have a specified job description in an employment contract at the time of hiring, so it is doubtful that workers accurately consider tradeoffs between wage and non-wage characteristics of a job. Despite those limitations, there are some studies which used this methodology to look into Japanese work-life balance issues.

It seems like there is no agreement with regard to the existence of actual wage differences between workers with traditional regular work style and workers with more worker-friendly work style with otherwise similar characteristics through the hedonic pricing analysis. Toda (2015) concluded the wage differences exist. The author used the 2012 Working Person Survey conducted by a private research firm, Recruit Works Institute, covering nearly 1,000 people aged 18-59 who lived in Tokyo and surrounding prefectures. After controlling for variables in standard wage regressions, including years of tenure and educational levels of the worker, the author found that there is a 10% hourly wage discount to have limit on work location for employees in big companies and a 10% hourly wage discount for having limit on work hours for female workers. Kuroda and Yamamoto (2013) also found negative wage implications for having access to company's work-life-balance policies especially for male

workers. The authors used the 2009 International Comparative Survey on Work-Life Balance conducted by RIETI covering full-time white-collar workers in Japan. After controlling for the selection bias and individual fixed effects, they found for male workers, having used a flex-time working option resulted in 5-9% lower wages. In contrast to the above two studies, (Yasui, et al. 2016) concluded there are no such wage discriminations against limited-regular workers. Using the Blinder-Oaxaca decomposition to analyze an online survey, covering about 2000 people aged 15 and above who work in big firms (RIETI 2015), they found that although workers limited by their work location and work content receive statistically lower "monthly" wage than unlimited workers, an 80% and a 90% of their wage differences respectively can be explained by the difference in observed characteristics which are included in standard wage regressions[11]. Moreover, when it comes to "hourly wages", surprisingly, their results show all types of limited-regular workers receive statistically higher (not lower) wages than unlimited regular workers after controlling for the observed characteristics.

The different findings across studies suggest that the extent of the possible wage cut for a worker to adopt more worker-friendly working style could vary depending on work location, company size, and whether it is a white-collar job or not. To understand whether the work style reform can be successful nationwide, it is important to also measure how workers who currently do not have any limit in work commitments evaluate the proposed change in work-

[11] The controlled characteristics are gender, education, age, age squared, yeas of tenure, years of tenure squared, industry, occupation, marital status, number of children, prefecture of residence, and hours of work.

style. So this lack of agreement through hedonic pricing model is a support for the use of CEs to analyze the potential of yet to exist limited-regular contracts nationwide.

While the above hedonic pricing analyses are based on the current wage cases of various types of workers, how much it can be applied to the potential cases under the reform scenario is questionable. There are some studies which used hypothetical scenario surveys to workers, but those studies assess workers' monetary evaluation of each aspect of a job separately. In reality, though, people would try to gather more information about the job and evaluate the whole package of the variables of their interest before making a job choice. Therefore, CEs give more accurate analysis as they enable researchers to look at people's preferences under more realistic conditions. Nevertheless, those recent surveys based on hypothetical questions can be useful as a comparison to the WTP estimation carried out in this chapter.

Overall, studies using hypothetical survey questions conclude that workers believe the monthly wage should be increased by 10-30% for them to take an unstable job or to commit to some work-related obligations. Morikawa (2010) used an individual survey conducted by Ministry of Economy Trade and Industry (2006), which asked people aged 20-60 how much wage premium they would request (ranging from +0% to +50% of a typical regular worker's wage) in order to accept unstable employment and some possibility of mandatory relocations or intra-firm transfers. The result suggests on average, people ask for 10% of wage-premium for both accepting unstable employment and some possibility of mandatory relocations and intra-firm transfers. Kume, Otake and Tsuru (2014) confirmed the findings by Morikawa (2010) using different data set. The authors used an online survey which asked hypothetical questions to both regular and non-regular workers about acceptable level of wage differentials (ranging from +0% to +50% of the wage of an otherwise equivalent

regular worker) to accept unstable employment and some possibility of mandatory relocations/intra-firm transfers (RIETI 2013). After excluding some invalid responses[12], the respondents on average asked for 21% premium for having an unstable job and 19% for having been exposed to the risk of mandatory relocations/intra-firm transfers. Similarly, Tsuru, Kume, et al. (2013) analyzed a panel survey of non-regular workers, which asked respondents to consider trading-off wage and employment stability and possibility of mandatory relocations/transfers under the hypothetical situation (RIETI 2009-2011). The authors found the respondents on average asked for 20% premium for switching to an unstable job and 27% for switching to a job with a risk of mandatory relocations/intra-firm transfers.

Calculated from the numbers in Table 2.7 in Section 13 (2.13.2) , our estimation of WTPs for people with the lowest wage level (annual 3 million yen) suggest that they are willing to pay the amount equivalent to 3% of their annual wage to avoid small overtime, 19% to avoid moderate overtime, and 50% to avoid extreme overtime. The WTP for relocation shows that in order to avoid such possibility, they are willing to pay the amount equivalent to 12% of their annual wage, and they are also willing to receive additional 12% of their annual wage to give up high job security. And to avoid transfer possibility, they are willing to pay as much as 34% of their annual wage. Overall, it seems that the WTP estimations in the previous section are not so far off from the estimations derived from other means with hypothetical scenario approach.

[12] About 30% of respondents chose “I don’t know” and did not select any of the listed wage levels.

Moreover, there is an investigation on the monetary evaluation of companies' other work-life balance (WLB) policies under hypothetical scenarios, and the perceived price (in terms of the wages) for having those policies seems to be comparable to having a stable job or having a limit in work-related obligations. Kuroda and Yamamoto (2013) analyzed a survey which asked hypothetical questions to both employers and employees, to see how much wage cut is justifiable to introduce WLB policies in a company (RIETI 2012). The questions for employers asked what percentage of the wage should be cut to newly introduce 1) longer than mandatory parental leaves, 2) short-time working options, and 3) flexible work arrangements (flex-time, work from home etc.). The questions for employees asked that in order for employees (not only themselves but also their colleagues) in their companies to use WLB policy 1) 2) and 3) above, how much the wage should be set compared to the current wage. According to the hypothetical questions, they found that there are gaps in the right amount of wage discounts perceived by companies and workers. After excluding invalid responses[13], on average, workers seem to accept lower wage (about 20% lower than the current wage) than companies would offer (about 10% lower), when they see the necessity of having the WLB policies available at their company. These results show that there is a friction in the market and there is room for policy to improve efficiency.

2.16 Guilt Effects

Not only the socio-demographic characters of a respondent but also his /her emotional state might affect their job choice decisions. Moreover, if there are gender differences in the

[13] 40-60% of workers/companies responded they want 0%/100% wage cut for having WLB policies, indicating they would not want WLB policies to be implemented at any wage rate.

emotional state, then it might be the actual factor of gender differences in WTPs. Some studies found women feel more guilt than men (Etxebarria, et al. 2009) while other studies found men and women feel about the same level of guilt but on the different situations (Martínez, et al. 2011). I included some questions in the survey to explore the feelings of guilt about working and parenting. I drafted seven items to describe the situations faced by working mothers and fathers that potentially generate guilt, and asked respondents to rate each situation on the five-point scale ranging from 1 (very guilty) to 5 (not at all guilty) as in Table 2.21. In case the situations do not apply to them, they were asked to answer as they imagine what they would feel in the situation. Working men and women tend to feel stress when 1) work can interfere with family life (i.e., work-to-family conflict: WFC) and when 2) family life can interfere with work (i.e., family-to-work conflict: FWC) (Shimazu, et al. 2013). Item (1) and (2) in Table 2.23 describe typical FWC situations, while item (3)-(5) describe typical WFC situations. The item (6) and (7) are taken from a previous study (Martínez, et al. 2011), which they found the largest gender differences in responses, where men significantly feel more guilty when facing the situation in item (6) and women feel more guilty when facing the situation in item (7). It should be noted that item (1) and (2) are different to the others as they depend on how the work is managed. The results of my survey are summarized in Table 2.24. It shows that women tend to feel guiltier than men in general and the differences are statistically significant at 1% for all the guilt questions.

Now, to see if the level of guilt in each item affects WTP measures, an extended model with dummy variables of each question are added in the model together with interactions between the guilt dummy variables and non-wage attribute variables. Table 2.25 shows WTP measures for a person only differ by guilt levels, with annual wage of 3 million yen and in

the 30s. The **Not_guilty** is a person rating all guilt questions as 5 (not at all guilty) and the **Guilty** is for a person rating all guilt questions as 1 (very guilty). It seems that guilty person tends to have higher WTPs. As our sample women tend to have higher guilt level than men, it might be the case that higher female WTPs than men's are partly because of their higher guilt level. However, although Table 2.25 shows the extreme cases of opposite guilt levels, the differences are not significant for most variables but Overtime4 and Transfer2 (see Table 2.26).

One may then wonder if the guilt effects differ across gender. We can test it by another model which include guilt-gender interactions. Table 2.27 shows the resulting WTPs. It can be read that female WTPs are higher when they are guiltier, but the pattern is less obvious for male WTPs. Next, we take a close look at Overtime4 WTP for female which has the largest differences for the two female groups with different level of guilt, to see what kind of guilt has the largest influence.

Table 2.28 shows WTP for Overtime 4 for female, with annual wage of 3 million yen and in the 30s, by varying the rating scores for each guilt question. For example, if the person only feels very guilty (rating 1) for q1 about taking paid-leave and not at all guilty (rating 5) in all other six situations, her WTP for Overtime 4 is 0.83 million yen. From Table 2.26, we can see female WTP for overtime 4 is the highest when she feels guilty for canceling a kids' event for work (1.73 million yen). On the other hand, if she feels very guilty about leaving the office early but not guilty about other items, her WTP for Overtime 4 (0.73 million) is actually lower than not feeling guilty at all for leaving the office early or anything else (0.80 million yen).

The above findings about guilt questions are interesting in two ways. Firstly, it suggests that the higher female WTPs than male WTPs might be actually from the higher likelihood of women feeling guilty than men and the greater guilt sensitivity to WTPs for women than men. Secondly, though as discussed in earlier sections, female WTPs are mostly indifferent across family structure (i.e. having or not having children) or marital status, women are more prone to be guilty and thus have higher WTPs as working mothers.

2.17 Policy Implication and Conclusions

CEs are a useful tool to analyze labor market in Japan and they give interesting policy implications. While the main alternative to CE is the hedonic pricing model, that approach has its own problems because of lack of or inadequate data availability of existing wage cases. CEs have shown their worth in other areas, for example, marketing, health economics, transportation, agricultural and environmental economics, to examine consumer preferences, and as the first study to use CE to analyze the job preference situation in Japan, this chapter can be a valuable addition to the body of literature.

For the CE in this chapter, I conducted the literature review, three focus group interviews and one pilot experiment with 100 subjects to determine final attributes and their levels that are important to the respondents, as well as the vocabulary and language to be used in the survey. More than 1000 subjects participated in the CE, in which they made a series of best-worst choices from a set of three jobs described by five attributes such as annual wage, overtime, employment security, transfer possibility, and possibility of relocation. The main findings from the empirical analysis are the following:

- The benchmark model suggests that when choosing a job, people significantly consider the amount of overtime (when it is more than the average amount), job security, and the possibility of intra-firm job transfer and relocation.
- The WTP calculation suggests that people are willing to forfeit a large portion of their wage to avoid extreme overtime and job transfer in particular.
- Job choice preferences seem to differ significantly between males and females. Women, in general, have higher WTPs than men to avoid overtime, relocation, and transfer, and to have more secure jobs (i.e. labor supply of women are less elastic than that of men).
- WTPs differ by the presence of children for male, while the difference isn't significant for female. Male WTPs are higher in absolute value for those attributes when not having children than having children.
- When taking into account one's guilt level in situations that typical working parents encounter, it was found that women tend to feel guiltier and have higher WTPs for avoiding work responsibilities than men.

The above results imply that limited-regular contracts can be an appealing option for workers with various characteristics, especially women. As workers would accept such offer with lower pay, employers should also see economic interest in offering such contracts. However, in reality, limited-regular contracts are not widespread, and thus there is room for policies to remove frictions in the market and improve market efficiency. For example, policies to clarify legal framework around limited regular contracts could be useful.

It should be noted that limited regular contracts should be designed to maintain the job quality of otherwise the same regular contracts. In practice, limited regular jobs differ in other dimensions from regular jobs and often bring with them less opportunities for growth and promotion. However, in the state preference questions in this chapter, it was emphasized that other aspects of the job stay constant, because promotion prospects and other job quality aspects are often subjective and difficult to measure. It is important for policy makers to ensure that job qualities would be the same as corresponding regular jobs in order for many workers to accept the alternative.

Another important issue to be concerned about wider use of limited regular contracts is that it could widen the wage gap between men and women. The results of the empirical analysis in this chapter could be interpreted as an invitation for Japanese employers to lower wages for women relative to men –or more subtly- to price limited regular jobs so that they are taken only by women. Dealing with this problem represents a challenge for policy.

Probably, as a complement to the offering of limited regular contracts, “anti-guilt” policies can be useful, since women tend to feel guiltier in situations typical working parents encounter and have higher WTPs for avoiding work responsibilities than men. The government can promote images to praise working mothers through campaigns so they can be proud of what they are doing and feel okay about not being able to do everything perfectly. The government can also give recognitions to those who are juggling with work and child-rearing, so they would know there are many other women who are struggling and they are not alone.

Chapter 3

A Married Couple Experiment[14]

3.1 Introduction

Japanese women of child-rearing age tend to spend more time doing housework and less time working than their male counterparts, and the gender gap is remarkably wide in Japan compared to that in other developed countries. For example, an international comparison of average daily hours spent on housework and child-rearing for couples with children under age 6 finds that in Japan, women spend almost eight hours per day while men spend only one hour. However in other developed countries such as the U.S., the U.K., France, Germany, Sweden and Norway, women spend 5-6 hours while men spend 2-3 hours and the average gap within a couple is not as large as in Japan (Cabinet Office 2004-2015). In a 34-country comparison, Japanese men were clear outliers; their average 2.5 hours of housework per week (childcare was not included) was less than one tenth their wives' 27 hours per week. This exceptionally low proportion of family work performed by Japanese men is a common phenomenon regardless of age of children or of whether their wives work full time, part time or not at all (North 2009). In a comparative study, (Leah and Huffman 2014) find that Japan is the only case among the 31 countries studied where women report experiencing significantly more of both life-work conflict and work-life conflict than men, suggesting

[14] This work was supported by JSPS KAKENHI Grant Number JP17H02498, 'Towards a behavioural theory of the household.'

that there is the male-breadwinner norm in Japan and Japanese men experience substantial separation of the spheres of work and family life. This striking gender difference suggests that equal sharing of domestic work and paid work responsibilities is especially important for the improvement of gender equality for married couples in Japan.

While the highly gendered division of work in Japan compared to the situation in other developed countries may reflect differences in cultural norms (men as earners and women as caretakers) (Fahlen 2014), in some part, this may be the result of friction in the economy, and policy may be able to narrow that gender gap. For example, shortage of child care facilities and tax deductions for spouses which place an income cap on the second earner in a Japanese household (typically the woman), are said to discourage Japanese female labor force participation (Aoyagi, Ganelli and Tawk 2016) (Steinberg and Nakane 2012).

Another potential source of persisting gender inequality in the division of work is the social barrier to men taking parental leave; because taking parental leave would be an excellent way to get men to start being caring and active fathers (Haaas and Hwang 2008). While Japanese parental leave policy is relatively generous[15]compared to that in other developed countries and while it allows both fathers and mothers to take paid-leave, fewer than three percent of fathers make use of such leave compared to 70 percent in Sweden (Steinberg and

[15] Firms are legally required to offer parental leave until the child becomes one year old (one and a half if childcare service is not available), and one of the following until the child reaches the age of three: (1) reduced work hours; (2) flextime; (3) changes of start or ending time of work; (4) policy ensuring no overtime work; and (5) provision of childcare service in the workplace. Many large Japanese firms offer more generous policies such as parental leave until the child is three years old, or reduced work hours for the parents of preschool age children (Kato, Kawaguchi and Owan 2013).

Nakane 2012). A qualitative study using interviews with human resource managers in large Japanese firms (Brinton and Mun 2016) find that managers implicitly assume that their companies' parental leave policy pertains only to female employees and is designed to enable recruitment and retention of talented women. Brinton and Mun (2016) claim that development of Japanese parental leave policy, which is marked by long leave period, leave with only partial wage replacement and norms that support only mothers as leave-takers, reinforces a traditional gendered division of labor. As a result of such gendered corporate culture and policy constructions, the idea of taking parental leave for men is highly non-normative in Japan, and very few men even consider the possibility of taking parental leave (Takahashi, et al. 2014).

To improve gender equality, family policy could increase support for working mothers, not only surrounding childbirth but also during the period of child-rearing, and changing the cultural norm of the male breadwinner might become a primary objective. Past reforms of parental leave policy in Japan seem to have been insufficient to achieve these goals. Asai (2015) found little evidence that the labor supply pattern for new mothers changed in response to the increased cash benefit during maternity leave, a reform implemented in Japan in 2001. The above mentioned interviews with human resource managers by Brinton and Mun (2016) revealed that the managers' implementation of parental leave and their evaluation of leave-takers occur within the context of norms about ideal employee behavior in firm-internal labor markets (which they conceptualize as an employee who puts company above family and willingly spends almost unlimited 'face time' with his/her employer and colleagues) and about gendered division of care work at home. Failing to meet those employers' expectations of the ideal worker behavior upon return to work for mothers

results in lower wage and promotion prospects in Japan (Kato, Kawaguchi and Owan 2013), and as a result, widening gender inequality. Unless policy is reformed to increase support for working mothers beyond the period of one year after childbirth, and to change deep-rooted traditional family roles and lack of support from husband, it is unlikely that parental leave will narrow the gender gap in Japan.

One way to improve the parental leave policy scheme might be to assign a father's quota, or days of non-transferable paid-parental leave specifically allocated to the father, as in the case in Swedish policy. A comparative study of five Nordic countries shows that while there are number of important features for the implementation of gender equality in the division of take-up of parental leave, such as universal coverage, a father's quota, a relatively long period of available leave, wage-based compensation, flexibility and other equality incentives that offer bonuses to couples who share leave, the father's quota seems to be the overall most effective policy instrument for encouraging fathers to take leave (Haas and Rostgaard 2011). Haas and Hwang (2008) analyzed 356 Swedish fathers working in large private companies and found that fathers who took more days of parental leave were more likely to take full responsibility for the children when the mothers worked, to spend more time with the children on a workday, and to be engaged in specific childcare tasks, especially those related to physical caregiving (2008). These findings suggest that if Japan introduced paid-parental leave allocated specifically to fathers so as to ensure that they get sufficient experience in housework and child-rearing, men might become more favorable to taking part in domestic work and women might be more comfortable working outside the home.

Chapter 3 of my dissertation reports a testing of this hypothesis that additional experience of tasks affects preference for doing those tasks, through an economic experiment with Japanese married couples.

In my experiment, a total of 51 couples performed two kinds of tasks, one paid task and one unpaid task. In each couple, one person was randomly assigned to get more experience of the paid/unpaid task. Then both the husband and the wife separately indicated their preference regarding the division of work with their partner. Contrary to our expectations, subjects expressed higher preference for the Traditional task pair (i.e. male breadwinner option) over the Reverse task pair (i.e. female breadwinner option) or the Mixed task pair (i.e. half-time dual earner option with equal task sharing). Obviously, this intervention is a long way from forced parental leave for men, but the intention is to see if forced experience can change task preferences. While the results did not give any clear evidence supporting the benefit of parental leave allocated specifically to fathers, that may be the result of the limited time of the experiment or the absence of rewards of child care from the task in the experiment. In reality, fathers might develop more emotional attachment to their children while taking parental leave and that might increase the likelihood of their level of contribution in child-rearing later. Though the nonpaid task in this experiment is designed to evoke emotional reward from imagining happy children solving the cleaned maze, it should be noted that the task in this chapter is more like general housework (but not really a child rearing task), and that could change the results and the policy implications. Also, further study is required to determine whether longer policy intervention could result in different and positive outcomes by extending the treatment time and varying the fixed cost for the dual-earner option.

The Chapter is organized as follows. In section 2) the literature on division of household labor in Japan and the literature on couple experiments are reviewed. In section 3), my experiment design is described. In section 4), research questions and hypotheses are stated to motivate the analysis. In section 5), the experiment implementation is described. In section 6), sample characteristics and task performance of the participants are reported which overall seem to support the validity of the experiment. In section 7), analytical results from both inferential statistics and regression analysis are discussed. And finally, section 8) discusses policy implications and concludes.

3.2 Literature

3.2.1 Division of Household Labor in Japan

The inequitable household division of labor among Japanese married couples as described above has been widely known and many researchers empirically investigated the factors behind such phenomenon. A recent study by Nishioka and Yamauchi (2017) analyzed previous studies and categorized the identified determinants of husband's contribution to housework and child-rearing in Japan into six major factors; 1) time constraints, measured as husbands' working hours and time to arrive at home after work, 2) level of necessity of housework/childcare, measured as wives' employment status, working hours and time to arrive at home after work, and the age of their youngest children, 3) power balance within the couple, measured as their income difference, age difference, and education difference, 4) availability of an alternative resource, such as whether or not to be living with their parents, 5) husbands' social status, measured by husbands' education, career status, and income and finally 6) the couples' ideology measured by degree of agreement to statements such as " husbands should work outside, and wives should stay home" and "the husband

and the wife should equally divide household work." According to Nishioka and Yamauchi (2017), besides husbands' long work hours and late arrival home after work which seemingly have been agreed to negatively affect husbands housework share, most factors were found to be significantly influencing husbands' contribution to housework and child-rearing in one study while other studies found them insignificant, thus there has been no uniform consensus of the most influential determinants.

Another active topic of discussion both among researchers and general public surrounding couples' division of household labor in Japan is their sense of fairness and its influence on marital satisfaction. It is said that married women in Japan often perceive their share of housework to be fair even when they assume a disproportionately larger share than their husbands (Fuwa and Tsutsui 2010) (Nakamura and Akiyoshi 2015). Nevertheless, studies found the larger share of husbands' housework and childcare responsibilities is associated with the higher evaluations for husbands and greater marital satisfaction by wives (Lee 2008) (Yamaguchi 2006), especially in dual-earner couples (Kobayashi, et al. 2016) (Kubo 2016). The perception of fairness about the division of labor at home is not static but rather changing over time and influenced by the surrounding society. Nakamura and Akiyoshi (2015) find that social comparison with others is a key mechanism that explains women's perception of fairness of household division of labor, especially with reference to people with similar life circumstances rather than non-specific others. Even though the average length of time when fathers are engaged in taking care of or playing with children under three years old has not been increased much since 1998 (Nishioka and Yamauchi 2017), it seems like making a complaint about the absence of husbands' support for wives has become more acceptable and common. In 2017, the term "*wanope ikuji* (Solo child-rearing

and completing all household chores alone)" was selected as a candidate for the buzzword of the year (Japan Times 2017). This abbreviated combination of the words "one" and "operation" was originally used to describe the harsh environment at fast-food restaurants and convenience stores where employees take responsibility and perform all types of work alone. Recently it has seen as surge in usage, in reference to women juggling work with child-rearing and all household chores without support from others, especially the husband. Some people say that the selection of the term further increase social awareness about the difficulty Japanese women are facing, and thus let women express their hardship about the shortage of husbands' support in housework and seek help.

While many studies have tried to investigate the mechanisms of the inequitable household division of labor among Japanese married couples and its potential problems, studies which incorporated experimental approach was not many till today. To the best of my knowledge, this is the first to use economic experiments to analyze preference of the division of labor in married couples in Japan, in an attempt to draw policy implication to improve gender equality.

3.2.2 Couple Experiments

There is relatively large literature on couple experiments which study how couples cooperate or not cooperate to maximize income (Iversen, et al. 2011) (Kebede, et al. 2014) (Cochard and Couprie 2016). The experimental approach could also be useful to study division of labor within the couple, because neither non-market input nor outputs can be observed by widely available statistics, while experimental approach could overcome these difficulties by allowing a direct measure of individual domestic productivities and the

control of the market wage. However, the number of studies which conducted couple experiments regarding their decisions on the division of work is very few (Gorges 2015).

Gorges (Gorges 2015) was probably the first to experimentally analyze couple decisions on how to divide labor explicitly. In her experiment, 20 real couples and 20 pairs of strangers were invited to play a paid task and/or an unpaid task, paired up either with their partner or a stranger of the opposite sex. Participants were then asked to decide on how to divide labor, when they have options either both players complete the paid task or have one of the players perform an unpaid task, thereby tripling the pay-rate for the partner playing the task. After completing their tasks, participants decided individually what portion of their income to invest in a common pool, where it is increased by 20% and distributed equally between the two players. The author found that women were significantly more likely to give up their income autonomy and perform the unpaid task when playing with their partner rather than with an unfamiliar man, representing the often observed phenomenon across time and countries, the female specialization on household work and male specialization on labor-market work within a couples.

While the experiment in Gorges (2015) could not distinguish possible factors causing gendered specialization within a couple, in an experiment by Cochard, Couprie and Hopfensitz (2017), productivity (and wage rate) was controlled to investigate preference for work-division by cohabiting couples. In their experiment, 64 true couples and 55 unrelated pairs of strangers of opposite sex participated, and were divided into a group where men get an advantaged return from doing a paid task and women get a disadvantaged return from the same task, and the other group where women are advantaged and men are disadvantaged. They compare behavior when men (women) are in the advantaged position, then found no

gender differences in contribution to the public goods (i.e. housework) in all conditions. These findings support the assumption that labor specialization by spouses is driven by differences in net benefits from labor market activity.

The results from an experiment by Couprie, Cudeville and Sofer (2017) suggest different mechanisms in the allocation of time among men and women between market and household work, namely the preference to follow gender roles and the stereotypes that men (women) are better at certain kinds of tasks. To assess the effects of gender roles and stereotypes in couples' division of labor, the decision on sharing of highly-gender stereotyped household tasks was compared with the decision on the sharing of neutral tasks, with the sample of total 81 established couples. They found that having to perform gendered tasks, compared to neutral tasks, induce couples to deviate more often from the division of labor to maximize household income i.e. women on average overspecialize in the "feminine task" and men in "masculine" task.

While above-mentioned couple experiments regarding the division of labor in couples were focusing on factors which would cause unequal sharing of market work or household work, experiments to assess policy effects on couple's work choice are lacking. Schröder, et al., (2013) conducted an experiment motivated by the assumption that income tax system (either joint tax or individual tax) affects work effort for couples. Among the few related studies including the above three, my experiment is most close to this paper by Schröder, et al., (2013) in design. As in my experiment, real couples are invited to join the experiment and each member of the couple is assigned to a paid task. In their experiment, the paid task was solving mazes on two different piece-rates. Couples then are asked to make a joint decision on who (the husband or the wife) will do non-pained compulsory work thus have a shorter

time for the paid work. Players face different tax scheme in stage 1 and stage 2, and thus the effects of the tax system on work effort, which is measured as the number of correctly solved mazes, are estimated. Furthermore, in the second round, participants are asked whether they would have liked to switch the mazes with their partner as a test of satisfaction. While Schröder, et al. could not find strong evidence of existence of tax system effects on work effort, what they found was the gender identity effects that affects female and male work effort differently: Only male secondary earners (i.e. men with lower piece-rate) increase his work effort when he is unsatisfied with his role, in an effort to keep his male identity as the breadwinner, while dissatisfaction alone is expected to discourage both male and female workers.

Although there are a number of similarities between Schröder, et al. (2013) and my experiment in design, their study does not explicitly discuss policy effects on couple's preferences of work division, which is the main focus of my study. Their results seem to imply that both men and women equally prefer to have higher piece rate and avoid non-paid compulsory work (a proxy for housework) regardless of the tax system. My study aims to look into the case closely to assess whether or not a possible policy intervention (giving the extra experience of tasks) would affect couples' preference for work division of the tasks.

3.3 Experiment Design

Participants perform two kinds of tasks, Task A and Task B, and only Task A will be paid. That is Task A is a proxy for market work while Task B is a proxy for housework. In each couple, one person is randomly assigned to get greater experience in the paid task, and the other gets more of the unpaid task.

For the economic analysis to be discussed in later sections, we divide couples into two groups to see if the extra experience of tasks will affect subjects' preference for doing the task. In Group 1, couples' division of work is what we commonly observe in real life today, i.e. the husband plays more of the paid task while the wife plays more of the unpaid task. In Group 2, the work division is reversed, i.e. the wife plays more of the paid task while the husband plays more of the unpaid task.

3.3.1 Experiment Order

There are three rounds preceded by a ***practice session*** in which participants have an opportunity to experience both Task A and Task B for 1 minute each (Figure 3.1). The practice session does not generate any pay. After the practice session, Round 1 begins. After Round 1, Round 2 and Round 3, participants are asked to fill out some questionnaire and finally, payment to each participant will be made.

In ***Round 1*** **(Single Tasks)**, everyone separately performs each task sequentially. Everyone performs Task A first, and then Task B next for 4 minutes each (8 minutes in total). Participants will be paid individually, and the money will be given at the end of the experiment.

In ***Round 2*** participants perform assigned tasks to create different roles within each couple **(Assigned Roles).** In Round 2, either the husband or the wife will perform Task A (and thus be the breadwinner) while the other perform Task B for four minutes. The amount of compensation from Task A that the one person earned will be paid to both the husband and the wife regardless of who performed Task A at the end of this experiment.

After participants completed two rounds, I gave extra instruction for ***Round 3*** because it is more complex. In short, couples are asked to indicate their preferences of work divisions right before the Round 3, and actual tasks to be played is selected by lottery. Then, the compensations from Round 3 will be paid to the husband and the wife equally.

By making sure to pay the same amount of compensations for both the wife and the husband in Round 2 and Round 3, I made sure that the couple's monetary interest is always the same and there is no conflict because my focus is not about joint income maximization hypothesis.

The decision making about the division of work in Round 3, and the performance of the Round 3 are the most critical part of this experiment. Hereafter, the design of Round 3 will be described. The detailed information about the two tasks, and the work division choice and the randomization mechanisms for actual selection of the Round 3 task pair will be introduced later in the experiment implementation section.

3.3.2 Round 3

In Round 3, each couple's task division depends on the individual preference, their partner's preference, and luck. That is, before starting the tasks in Round 3, participants individually rank their choices of task-pair to be played by the couple, and actual task-pair is determined randomly by lottery from weighted choices from theirs and their partner's recorded rankings for incentive compatibility.

There are four kinds of possible task pairs to choose from. Two of them are *specialization* task pairs, for which either one in the couple specializes in the paid-task while the other specializes in the unpaid-task for four minutes. For the task pair **Traditional**, the husband plays task A, and the wife plays task B. And for the **Reverse** pair, the wife plays task A, and

the husband plays task B. For those specialization pairs, the individual pay is whatever the Task A player generated regardless of their roles.

The other two task pairs are *equality* task pairs, for which both members in the couple perform the same task(s). For the **Power** pair, both the husband and the wife play Task A for four minutes but a fixed cost will be deducted from their total earnings. The fixed cost represents the cost of childcare or outsourcing housework for a dual-worker couple, though it was not explicitly told as such to the participants. What participants were explained was that the amount of fixed cost equals average earnings from our pilot experiment, and the exact amount of the cost would be announced after they finish Round 3. We set the fixed cost this way in a hope to test the Hypothesis 3, gender differences in confidence and preference for competition. Because the exact amount of cost is unknown to the participants when they make a choice and they have to think whether they can over-perform the average from some unknown subject group from the pilot experiment, it requires confidence and appetite for competition for them to favor this task-pair. The other equality task pair is the **Mixed** pair. For this task pair, both the husband and the wife play Task A and Task B for only a half of the time (i.e. 2 minutes) for each task. For both equality task pairs, the final pay the husband and the wife receive will be the same and calculated by their combined earnings from Task A with the fixed-cost (for task pair Power) or without the fixed-cost (for task pair Mixed).

In summary, the final payoff for person *i*, or π_i, depends on her own payoff in each round and her partner *j*'*s* payoffs, such that:

$$\pi_i = \pi_{1i} + (\pi_{2i} + \pi_{2j}) + (\pi_{3i} + \pi_{3j} - COST) = \pi_{1i} + \pi_2 + \pi_3$$

Where π_1 is the payoff from Round 1 and differ from one individual to the other. π_2 is the payoff from Round 2 which is the same amount within the couple, and $\pi_{2i} = 0 \; and \; \pi_{2j} > 0$ for person *i* and her partner *j* if they are in **Group 1**, and $\pi_{2i} > 0 \; and \; \pi_{2j} = 0$ for person *i* and her partner *j* if they are in **Group 2**. And finally, π_3 is the payoff from Round 3 which is again the same amount within the couple, and

for task pair **Traditional** couple, $\pi_{3i} = 0, \; \pi_{3j} > 0, \; COST = 0,$

for task pair **Reverse** couple, $\pi_{3i} > 0, \; \pi_{3j} = 0, \; COST = 0,$

for task pair **Power** couple, $\pi_{3i} > 0, \; \pi_{3j} > 0, \; COST > 0$, and

for task pair **Mixed** couple $\pi_{3i} > 0, \; \pi_{3j} > 0, \; COST = 0.$

Furthermore, to incentivize participants to truthfully refract their task pair preferences to their choice ranking, we employed randomization process with weights based on their order of the preferences to determine the actual task-pair to be played in Round 3. This is to make sure that participants carefully consider filling in all levels of the ranking (from top rank to the bottom rank). The randomization made sure that the higher they rank a task pair, the more likely they will get to play it, so that it becomes incentive compatible. Nevertheless, the final task-pair decision will depend partly on luck and partly on their partner's answers. The detail on the implementation of the randomization is described in the experiment implementation section.

3.4 Research Questions and Hypotheses

While the general hypothesis of this chapter is that additional experience of a task and the gender of participants affect their preference for doing the task, more detailed hypotheses can be described in this section.

3.4.1 (Hypothesis 1) Experience Effects

The first hypothesis is that greater experience of tasks will lead to greater preference for doing the tasks. People may expect positive learning effects by continuing in the same role and thus see not-switching as the efficient choice, or it might be caused by the status quo bias (Samuelson and Zeckhauser 1988), which leads people to rationally choose to stay in the current role even when switching may seem like the efficient choice in terms of income maximization. Deviation from this hypothesis can happen if people show more preference for variety and choose to conduct different tasks from the previous round, or partly by gender identity effects (Akerlof and Kranton 2000), which leads men and women to prefer conducting the manly/womanly tasks regardless of the previous experience in those tasks.

3.4.2 (Hypothesis 2) Equality

The second hypothesis is that women prefer equality option to efficiency option more than men. Couprie, et al. (2012) observed that in both French and German data, women in general select equality option over efficient option more often than men when deciding on the distribution of the household income.

3.4.3 (Hypothesis 3) Competition, Confidence

The third hypothesis is that compared to women, men prefer an option which requires competition with unknown others and confidence for doing better than them. Based on the speedboat races in Japan, Booth and Yamamura (2016) show that men are overconfident

than women; thus they take more aggressive strategy and are better at competing. There are other experimental studies which support those findings, for example, Kamas and Preston (2012) also found that men are more confident than women, and Booth and Nolen (2012) observed that in general, men prefer competition more than women.

3.5 Experiment Implementation

3.5.1 General Setting

We ran a total of 10 sessions (5 morning sessions and 5 afternoon sessions) in four different public conference rooms in Koto-ku, Tokyo, Japan.[16] Sessions were held on Saturdays, January to February 2018. A total of 51 couples (102 people) participated. The number of participants per session ranged from 6 (3 couples) to 14 (7 couples). Participants were recruited through job advertisements on classified apps, [17]printed flyers (either handed out or posted in mailboxes) and word of mouth. They were informed that each couple could earn on average 8,000 yen (or 4,000 yen per person) for doing easy tasks that required no special skills or knowledge, in a session of approximately 2 hours. However, no information was provided about the purpose or the content of the experiment prior to their arrival at the venue. To participate, partners had to be married (either legally or de-facto marriage with

[16] Venues and dates are as follows: 1) January 6 and February 17 at Pal City Koto, 2) January 20 at Sunamachi Culture Center, 3) February 10 at Tiara Koto 4) February 24 at Furuishiba Culture Center. An image of one venue is in Appendix 3.6.
[17] Mercari Atte and Jimoty were two apps used.

some kind of documentation), aged 20-59, and Japanese national.[18] To lower the cost of participation, child-care was provided in a separate room for some sessions.

Each experimental session lasted approximately 90 minutes. To avoid communication and interaction between partners during the experiment, we seated the partners apart. All participants were seated facing forward, and we ensured there was enough space between each participant to produce privacy and minimize peer effects. [19]The author read out the instructions at the front, and a team of experimenters consisting of a professor (as a supervisor) and one or two university students (as assistants) was present at all sessions. The instructions were given orally, and a copy of the supplementary visual guide was provided to participants.[20] Participants were encouraged to ask questions whenever an instruction was unclear to them, and all participants were required to pass a test of understanding to complete the experiment.

Upon arrival of each participating couple, we handed each person an identification card with a couple ID number, and with one shape (either Circle or Triangle) indicating their role in some part of the experiment. In each couple, the person who had a circle shape was randomly assigned to get greater experience in the paid task, while the other (with a triangle) got more of unpaid task experience. In Group 1, the husband performed more of the paid task (i.e. the husband had the circle card) while the wife performed more of the unpaid task

[18] We restricted the sample to Japanese national for two reasons; first to avoid any language communication problem and second to reflect demography of Japanese society, as the share of foreigners to Japanese population is less than 2 % (Ministry of Internal Affairs and Communications 2017).
[19] An image of a typical session can be found in Appendix 3.6.
[20] A copy of the materials shown to participants can be found in Appendix 3.1-3.3. A written instruction (not presented to participants) is in Appendix 3.4.

(i.e. the wife had the triangle card). In Group 2, the work division was reversed, (i.e. the wife had the circle card, and the husband had the triangle card). Groups were randomly assigned as in Table 3.1 so as to have the balanced number of observations for each group. Group allocations were not explicitly announced to the participants, though participants could guess them by their allocated shape and task assignment.

3.5.2 Tasks in the Experiment[21]

3.5.2.1 Task A

Task A is the paid task, a proxy for market work in this experiment. Task A is marking the correct answers printed on an exam paper into bubble sheets. The exam questions were taken from a Microeconomics workbook. The correct answers are already indicated by stars on the exam paper so the participants do not need to read the questions. Bubble sheets have a set of blank circles that correspond to the questions, and participants are asked to fill the circle which corresponds to the choice with a star. Participants gain 20 yen for each correctly marked answer and lose 5 yen for each incorrectly marked answer.

When determining the task as the proxy for market work, a number of conditions were taken into consideration. The main conditions considered are as follows:

First, for the purpose of making the implementation smoother, the task performance should be countable in short time, and should be suitable for piece-rate calculation based on effort.

[21] Images of the tasks are in Appendix 3.6.

By adopting bubble sheets, we could scan the sheets and instantly calculate the number of correctly or incorrectly marked answers.

Second, the task should be doable for anyone and should require no special skill or knowledge. This condition is aimed at recruiting a wide range of workers, and to make sure that the task is gender-neutral, i.e. both men and women can do the task equally well. The latter assumption is important in my experiment because we have only one piece-rate, while Schröder, et al. (2013) used two piece-rates to artificially create disadvantaged/advantaged players within a couple. By already indicating the correct answers on the exam paper, we satisfy this second condition.

Thirdly, the task should not be fun in itself and should require effort. The task should also at least looks like meaningful work, so people will take it seriously. This condition is in an effort to mimic real-world market work, which is not always fun and requires effort though the workers understand there is some kind of meaning in the work. It is also because, as documented by Chandler and Kapelner (2013), people seem to make more effort when the tasks appear meaningful. I believe that by using actual questions from a micro-economics workbook and adopting the style of an examination, this third condition is met in this experiment.

3.5.2.2 Task B

Task B is the unpaid task, which is a proxy for housework in this experiment. It involves erasing pencil marks from mazes. Participants do not get penalized for doing too little or even doing nothing. However, to induce effort, participants are told that their cleaned mazes will be reused by children. In real-world housework also, people do not get a monetary reward but get emotionally rewarded by, for example, the satisfaction of cooking a delicious

dinner or tidying a room. The image of children solving the cleaned mazes is expected to serve as such an emotional reward. Though this task is designed to evoke emotional reward from imagining happy children solving the cleaned maze, it should be noted that some rewards of childcare are absent from the task, as the task does not involve actual interaction with children. For that reason, the task in this chapter is more like general housework (but not really a child rearing task) and this could change the results and the policy implications.

As for the paid task, a number of conditions were taken into consideration when deciding the unpaid task. The most important requirement was that the task should not be very exciting or intellectual. In a pilot experiment with GRIPS students, we tried a different task, sorting poker chips by color, but some participants found it entertaining because the task felt like a game to them. As a proxy for housework, I wanted the task to be a little bit boring and tiring yet doable for anyone. Also, in order to mimic real household production closely possible, it needed to be a manual work task. While some people enjoy solving mazes, erasing mazes is not very much fun, and rubbing pencil marks off of pages after pages requires real effort, so I chose this as the unpaid task in this experiment.

3.5.3 *Round 3 Choice and Randomization*

To mask the purpose of the experiment (work division within a couple) and to make the later lottery implementation easier, the four task pairs were given color names (green, blue, red and yellow) in the experiment. The names used in the experiments and the corresponding ones I used in the analysis to reflect what the task pairs represent in the couple context are shown in Table 3.2.

I provided a visual guide and recording sheets to support the participants' understanding of each task pair and the recording of their preferences (see Figure 3.2 and Figure 3.3 for the translated versions).

Participants were told to take as much time as they need to fill in their task-pair ranking table in the recording sheets provided, where they were told to write down the name of the task-pair color into each box corresponding to their preference order. Before making their decisions, participants were given careful instructions on how the compensation from each task pair would be calculated and how the actual task-pair to be played would be determined. We gave them a test of understanding which they were all required to pass before moving to the next steps.

After participants ranked the four task pairs, we used a pile of 16 poker chips for randomization (four of each of the four task pair colors), that we had placed in front of each participant. For the task pair the person ranked at the top we took four chips of that same color from the pile and put them into a bag. We put three chips into the bag for his/her second option, two chips for the third option and one chip for the least preferred option. The partner then carried out the same exercise, putting their chips into the same bag. Then, we let one of them draw out 1 chip. The color of that chip determined which task pair the couple would do. This randomization process is to make sure people's decision making is incentive compatible because the higher they rank a task pair, the more likely they will get to play it.

3.6 Sample Characteristics and Task Performance Distributions

Depending on the sample selection or the level of participants' seriousness in doing the tasks, the external validity of the experiment might be questionable. Overall, the sample

characteristics and the task performance distributions seem to be satisfactory to ensure the external validity of this experiment.

3.6.1 Sample Characteristics

Since the participation was voluntary, a participation bias in the recruitment process cannot be excluded. With couples in particular, we need to be aware of the fact that participation might be harder for couples with high opportunity cost of time for getting the joint attention of parents, particularly those from nuclear family with young children, and also that samples will be biased in favor of couples with relatively healthy and stable relationship (Munro 2018). By providing childcare in some sessions, we dealt with the first issue. However, it seems like the second issue could not be ruled out, since our identical survey questions for Ch1 and Ch2 experiments, which asked about their satisfaction on task division with their partner for doing housework, taking care of children, taking care of elderly or disabled relatives, and contributing to the household income, revealed that the participants of this couple experiment are more satisfied on average in all of those categories. While the differences in the mean satisfaction indicators between the two samples are not very large (i.e. at most 0.6 point in 5-point scales ranging from 1=very satisfied to 5= very dissatisfied) it remains to be seen whether it is an important factor in external validity.

Still, our total of 102 subjects displays substantial heterogeneity suggesting that the findings from this experiment could be generalized to Japanese couples at large. Several charts are provided below to show the distribution of the sample from the socioeconomic questionnaire taken at the end of the experiment.

3.6.1.1 Age

As can be seen in Figure 3.4, while the participation was restricted to the core working-age subjects (i.e. 20-59 years old), close to half of participants (44 people) were in their 30s, followed by 33 people in their 40s. We had smaller participants from people in their 20s (17 people) and 50s (8 people). Since mothers' mean age at first birth in Japan is 30 years old (Demographic Survey 2016), the high participants of people in 30s and 40s reflect that our provision of child-care services was somewhat helpful in lowering the hurdles for their participants[22].

3.6.1.2 Education

More than a half (60 people) of participants had graduated from 4-year universities and 5 of participants had Master's degree (Figure 3.5). As it is typically the case for behavioral economics experiments, our sample was drawn from the more educated sample than entire population (Henrich, Heine and Norenzayan 2010).

3.6.1.3 Income

There is a significant variation in the annual income of my samples, ranging from zero to 10-14 million yen. The highest proportion was of the 6-8 million yen group which consists of 18 participants, but there are in total 24 people who earn either zero or less than 2 million yen annually (Figure 3.6).

[22] 8 couples used the child-care while in experiments.

3.6.1.4 Living with Parents

Close to 80 of the participants are living neither with their parents nor their partner's parents[23] (Figure 3.7). As this implies that most participants are living in a household with only two adults, it can be imagined that sharing housework with the spouse is a situation they face every day.

3.6.1.5 Living with children

About half of the participants were living with their children (Figure 3.8). And for those who have children, 24 have one child, 19 have two children, and 6 have three children (Figure 3.9). Four couples answered the question about whether or not living with children differently, and of the four couples, three couples, as well as another one couple, had answered number of children differently. While there was no question to ask about detailed family structure, possible reasons for those different answers within the couple might be that the husband and the wife are living in separate houses or they have children with different partners. For the former, it is not uncommon in Japan that a married couple in a healthy relationship lives separately for a while upon work assignment[24] or childbirth[25].

---

[23] Because a small number of couples gave conflicting answers (i.e the husband (wife) says "yes," and the wife (husband) says "no"), the number of responses in each category is an odd number. See footnote 8 and 9 for cases about living separately from spouses in Japan.

[24] In 2004, about 80% of companies with more than 1000 employees had employees who were living separately from their family to engage in a work assignment. While the share is higher for larger companies, it was about 20% of total companies surveyed with 30 or more employees (JILPT, Yu-sufuru rodo tokei [Useful labor Statistics 2015] 2015).

[25] According to a survey (Benesse Educational Research & Development Institute 2015) of 1,500 mothers with children between 4 months and 11 months old, 54.7% stayed at her parents' house at/around giving birth to the child.

3.6.2 Task Performance

Figure 3.10 shows the distribution of Task A performance for each Round, as calculated by the number of correctly filled answers. . In Round 3, some players (those who played task pair Mixed) only performed the task for a half of the time, so for comparison, the chart shows performance adjusted for those time differences (i.e. Task B performance for the Mixed pair was multiplied by two). However, we are still unable to make direct comparison across different rounds because in Round 2 Task A performers are a half the size of Round 1, and number of performers are also not the same in Round 3. Nevertheless, we can see some signs of learning effects, as the peak of the curve shifting slightly towards the right, but also the variance seems to be increasing.

Figure 3.11 shows the Task B performance distribution in each round, as the number of mazes cleaned. The same caveats as the Task A performance distributions apply here. However, from the below charts, it seems many people quickly master Task B, and gave full effort to complete as many mazes as possible within the allocated time.

3.7 Results

As discussed in previous sections, there are various motivations that drive particular task division choices but if the individual acted in an *efficient* way to maximize household income, the decision process would be the following. If the person thought both he/she and his/her partner could earn more than the average from the pilot experiment, then they should choose the Power option. If they thought at least one of them had below average earning ability, they should decide which one of them is better and let the better one specialize in Task A. (i.e. if they think the husband can earn more than the wife, they should choose

Traditional, and if they think the wife can earn more than the husband, they should choose Reverse). Also, if they assumed that the wife and the husband were equally bad at Task A and below average, they could choose the Mixed option (i.e. they would be indifferent between the two specialization task pairs and the Mixed option). Figure 3.12 shows the decision tree of what a household income maximizing couple would do, as an illustration of the motivations. However, it should be emphasized that the decision tree might be modified. For example, risk aversion would promote choice of Mixed, or desire for equality would promote choice of Mixed and Power.

Although, people choose a task pair for many other reasons, but the household income maximization motive seems to be the most efficient choice. Men and women seem to have equal ability to choose wisely. As can be seen in Table 3.3, 25 men and 21 women correctly choosing the efficient task pair (given Task A performance in Round 1). For most people, the efficient pair was either Power or Traditional; both men and women chose these task pair options with good reliability. We have to note, though, that the calculation of the efficient option was ex-post, as participants had no information about their partners' performance or the exact cost to be deducted when they made decisions. Below, I discuss the actual preferences of the participants, as revealed by their task pair rankings in the experiment.[26]

Eyeballing the results of the distribution of best-choice (Figure 3.13) and worst-choice (Figure 3.14) by group suggests that the general hypothesis, that experience of a task affects

[26] The complete ranking distribution by group and gender is in Appendix 3.5 (Appendix Table A.3.1).

preference for doing the task, does not seem to hold. For example, regardless of assigned group, more than a quarter of subjects preferred the Power option the most. This may reflect the fact that, in the absence of real-world obstacles, such as shortage of child-care facilities or tax system bias against dual-earner couples, more people would choose to outsource the housework and both the husband and the wife would participate in market work. This points to a need for policy intervention. On the other hand, while the general trend of worst choice was also similar across Groups, the choice between Mixed and Reverse might have been affected by group. This points to a need for further investigation. In this section, we examine whether the more specific hypotheses (Hypothesis 1- Hypothesis 3) hold, first in the results of inferential statistics and then in the regression results.

3.7.1 Inferential Statistics

First, to see if Hypothesis 1 holds, I test the null hypothesis (Hypothesis 1*) that there is no experience effect:

Hypothesis 1: Proportion of Traditional is higher in Group 1 than in Group 2.

Hypothesis 1: There is no difference between Group 1 and Group 2.*

It is expected that when choosing within the specialization options, Group 1 will choose Traditional more often, and Group 2 will choose Reverse more often. The results of the test shown in Table 3.4, shows that 65 % of Group 1 participants and 76 % of Group 2 participants chose Traditional over Reverse. Although we designed our paid task to be a gender-neutral task, both groups designated the husband to specialize in the paid task. The p-value of the test shows the result fails to reject the null hypothesis of no difference in choices between the two groups, but even if there were a difference, it would be the opposite

of the expressed expectation, since subjects actually chose the less experienced task over the more experienced task. This might be a result of preference for variety or some other reasons.

Second, to see if Hypothesis 2 holds, I test the null hypothesis (Hypothesis 2*) of no equality effects:

Hypothesis 2: Higher portion of women choose Power and Mixed over Reverse and Traditional.

Hypothesis 2: There is no difference between men and women.*

It is expected that women are more likely to prefer equality in their roles than men, when we compare choices between specialization task pairs and equality task pairs. There are 2x2 comparisons: Power and Reverse, Power and Traditional, Mixed and Reverse, and Mixed and Traditional. The results of the test, shown in Table 3.5, show that in all four cases, equal or slightly greater number of women chose equality pairs (Power and Mixed) than men. On the other hand, both men and women similarly preferred Power option to the two specialization task pairs, they are similarly equally divided for the choice between Mixed and Reverse, and they similarly preferred Traditional to Mixed options. The p-value indicates that the test fails to reject the null hypothesis of no difference in choice between men and women[27].

[27] Failure to reject a null hypothesis does not provide unequivocal evidence that there are no gender differences, since the failure to reject may actually be the result of low statistical power, the probability that it will correctly lead to the rejection of the null hypothesis (List, Sadoff and Wagner 2011).

Third, to see if Hypothesis 3 holds, I test the null hypothesis (Hypothesis 3*) that there is no competition and confidence effect:

Hypothesis 3: Higher portion of men choose Power over Mixed.

Hypothesis 3: There is no difference between men and women.*

It is expected that men are more likely to choose the option which requires competition and confidence than women when we compare Power and Mixed options. The results of the test, shown in Table 3.6, show that while both men and women were more likely to choose Power option over Mixed option, the share of men who chose Power was slightly higher than that of women. However, the p-value indicates that the statistical test fails to reject the null hypothesis of no difference in choices between men and women.

3.7.2 Regression Results

The simple tests above do not control for confounds and sample variation in characteristics that might affect choices. In this section, regression analysis with those controls will be given.

Individual *i*'s utility function is specified as

$$U_i = \beta_{0i} + \beta_{1i}\,\mathbf{Reverse} + \beta_{2i}\,\mathbf{Traditional} + \beta_{3i}\,\mathbf{Balance} + \boldsymbol{\varepsilon}_i \quad \textbf{(5)}$$

The independent variables in the equation (5) are dummy variables which take value 1 for the corresponding task pair option and 0 otherwise. The benchmark is when the option is the Power task pair, thus all the dependent variables take value zero and $U_i = \beta_{0i} + \boldsymbol{\varepsilon}_i$. As the basis of the logit specification, $\boldsymbol{\varepsilon}$ is assumed to be independently and identically distributed

extreme value random variates. The rank-ordered logit estimation is used to utilize the maximum amount of information that can obtained from the individual ranking of subject task-pair preference order. To assess the effects of gender and assigned group on preferences, interaction terms are added to equation (5). That is, a Group 2 dummy and a Female dummy are interacted with each task pair type to create a three way interaction. Model 1 in Appendix 3.7 reports the benchmark results of the rank-ordered logit regression. It can be seen that coefficients of two-way interactions **Traditional#Female** and **Mixed#Group2** are negative and significant at 10%. Coefficients of three-way interactions **Traditional#Female#Group2** and **Mixed#Female#Group2** are positive and significant at 5%. All other interactive terms have insignificant coefficients. The result suggest that gender and task experience (i.e. assigned group) may affect the choice of Traditional or Mixed options in some cases, though other coefficients being insignificant implies that gender and task experience do not affect the task division preference overall. These results are robust even after controlling for some variables such as age, not having children, and whether the wife is housewife (Model 2-4 in Appendix Table A3.2 in Appendix 3.7) The combined overall effects of gender and task experience can be assessed by calculating the marginal utility that can be achieved from selecting each task pair and comparing it across gender and group.

The coefficients on Table 3.7 indicate the marginal utility that can be achieved from selecting each task pair, instead of the benchmark Power task pair, calculated based on the Model 1in Appendix Table A3.2 in Appendix 3.7. The all negative coefficients indicate that regardless of gender and group, people on average prefer the Power task pair the most among the four possible task-pair choices. For any category of subjects, either male or female, in Group 1 or Group 2, the smallest negative coefficients are for the Traditional

option, suggesting that subjects give second preference to the Traditional task pair. The third and the fourth preferred choices vary by category, but no consistent pattern by gender or by groups alone was visible. The least favored option was the Reverse option for males in Group 1 and for females in Group 2, and the Mixed option for females in Group 1 and males in Group 2. This could be because experience effects are different for men and women in terms of least preferred choice. For example, while the difference between the coefficients for Reverse and Mixed are smaller for men, those differences are larger for women, which affords a clearer distinction between the two task-pair options in the opposite direction (women in Group 1 favored Reverse, while women in Group 2 favored Mixed.) It is still a puzzle why Group 2 women dislike the Reverse option more than Group 1 women. It contradicts the initial assumption about experience effects.

Next, I included another variable, Task A performance (measured by the number of correctly filled answers in Round 1), to see if it would affect people's task division preferences. Model 5 in Table 3.8 reports the rank-ordered logit estimates. As the table shows, this specification has something to say only for the Traditional option. It can be read that while the positive coefficient for **Traditional-female** interaction indicates that women, in general, are in favor of Traditional compared to men, the coefficient for the three-way interaction of **Traditional-female-Round1_correct** (the measure of Task A performance) is negative, indicating that the higher her performance is, the more she would dislike Traditional option. This is contrary to the male case, where the positive coefficient of the **Traditional-Round1_correct** interaction indicates that the higher the performance is, the more he would like the Traditional option. The opposing preferences of Traditional option between male and female high performers are observed even after controlling for their

partner's task performance (as measured by the partners' number of correctly marked questions in Round1, the dummy variable **P_Round1_Correct** in Model 5' in Appendix Table A3.3 in Appendix 3.8)

If what might actually help women from deviating from gendered traditional work division was improving her performance in the paid task, the question is whether the extra experience of the task would improve the task performance, thus could have an indirect or second-round effect on task division preferences. To test this, I run a regression where the dependent variable is the Task A performance in Round 3 (measured by number of correctly filled answers per minute), and the explanatory variables are: 1) Round 1 performance, 2) Grouping in Round 2, 3) female dummy, 4) Group and Female interaction, 5) Task pair performed in Round 3, 6) Task pair preference (option ranked as best) for Round 3, and 7) Interaction between 5) and 6) to capture satisfaction effects. The variable 2) is my main interest. The variable 7) was included based on the argument by (Schröder, et al. 2013) that task satisfaction affects work effort.

Table 3.9 reports the result. It shows only the coefficient for the Round 1 performance (as measured by the number of correctly filled answers in Round 1) was significant. Signs of the coefficient for Group 2 dummy (capturing the effect of Group 2 on men) and for the coefficient for Group 2-female interaction are in line with my expectation, because in Group 2 men get less experience of Task A while women get more experience of Task A, so the effect on performance was expected to be negative and positive respectively. However, those coefficients were not statistically significant. Coefficients of the satisfaction measurements, which takes the value 1 if the actual task pair assigned matched with their

most preferred task pair, were negative for all the task pairs included, contrary to the expectations, but not statistically significant. While the results from this exercise are not satisfactory to draw any policy implications, there is an important caveat that the number of observation in this regression analysis is small, as the number of subjects who played Task A in Round 3 is limited.

3.8 Conclusions

Contrary to our expectations, there was no evidence of a relationship between prior task experience or gender and couple's task division choice. What we did find was that regardless of task experience or gender, the Traditional task division (i.e. male breadwinner option) was more preferred to the Reverse task division (i.e. female bread winner option) and the Mixed task division (i.e. half-time dual earner option with equal unpaid task sharing). While most people gave the highest preference to the power couple option (i.e. full-time dual earner option with fees to avoid unpaid work), this may suggest that the power couple decision is sensitive to expectations about fees.

While Mixed and Reverse options were unpopular by both gender and groups, the relative preference between the two low ranked task-pairs might be influenced by assigned groups, especially for women. That is, women in Group 1 seem to favor the Reverse option more, while women in Group 2 dislike the Mixed option less. Because the hypothesis was that a Group 2 female would favor Reverse, as that was her role in the previous round, the results are counterintuitive. Possibly some people have (a) a taste for variety, which would lead them to act against our expectation (e.g. I tried task x in Round 2 so now I want to try task y) or (b) a taste for responsibility (e.g. my partner did task x in round 2 so now it is my turn to take on this boring task). However, as this is a choice between low ranked options, we

cannot rule out the possibility that people just paid less careful attention to their lower choice (Hausman and Ruud 1987).

A relatively popular choice, Traditional option, is affected differently by male and female performance. While more productive males favor Traditional more than less productive males, more productive females favor Traditional less than less productive females. While this suggests that improvement of female performance could be a channel to greater female labor participation, experience of tasks and task satisfaction do not seem to affect task performance significantly (though it should be noted that this analysis was drawn from a small sample).

As the time for extra experience of tasks was short in this experiment, the failure to find significant effects of task experience on task division preference did not rule out the possibility of the existence of such effects. There is a need for further studies to examine the effectiveness of policies such as parental leave for fathers. For example, Haas (1992) and Haas and Hwang (2008) found that among fathers in Sweden who took parental leave, those who took more days of leave were more likely to report sharing responsibility for childcare, doing specific childcare tasks and spending more time in childcare than men who took fewer days. While this may reflect the fact that fathers who took longer parental leave developed tighter emotional ties with their children and thus became more willing to be involved in childcare, what the experimental results suggest is that temporary intervention on division of work between husband and wife does not have a clear impact on their role preferences (especially in the absence of sufficient rewards of childcare), so policy measures need to prescribe a sufficiently long period (e.g. setting a mandatory minimum requirement for

paternity leave days for fathers). Further studies with larger samples, longer experiment time (especially time for treatment), as well as more complex designs that vary cost of avoiding unpaid task, are required.

Chapter 4

Conclusion

There is wide ongoing discussion among policymakers and researchers of the necessity of work-style reform in Japan. Regular workers in Japan are more likely to be asked by their employers to work long hours and change their work location and work duties than regular workers in other countries. These traditional labor practices seem to negatively affect work-life balance of workers, as they often require workers to put work above their private life. Moreover, this traditional Japanese work style contributes to persistent gender inequality in the work place and at home. In the Japanese labor market, most non-regular workers, especially part-time workers, are women. It appears that women cut back on work so as to take care of home and family more frequently than men. In fact, Japanese men's contribution to housework and child-rearing is much smaller than that of men from other developed countries. It is possible that the success of labor market reform would depend at least in part on whether the reform addresses the stereotypical gendered preferences on division of work. This dissertation has two main chapters, Chapter 2, which reports an online Choice Experiment (CE) and Chapter 3, reporting an economic experiment, to investigate worker acceptance of and effectiveness of potential policy measures to address the issues of work-life balance and gender equality in Japan from the viewpoint of the workforce.

4.1 Main Findings and Policy Implications

Chapter 2 explores worker acceptance of limited-regular contracts, i.e. regular employment contracts with a limitation on one or more of the traditional regular worker obligations such as overtime, relocation, and intra-firm transfer, in exchange for lower pay. This is the first

CE study on the Japanese labor market. While the main alternative to CE is the hedonic pricing model, the CE approach successfully overcome the limitation of the hedonic pricing model which is the lack of or inadequate data availability of existing wage cases. CEs have shown their worth in other areas, for example, marketing, health economics, transportation, agricultural and environmental economics, to examine consumer preferences, and this chapter extends the CE approach to studying labor reform by examining the tradeoffs workers are willing to accept. It is found that limited regular contracts would be widely accepted by workers (if the job quality is maintained), suggesting the need for policymakers to support employers who adopt such employment contracts. For example, reform clarifying the legal framework for the employment and dismissal of limited regular workers could encourage more employers to offer such contracts.

More than 1000 subjects participated in the CE, in which they made a series of best-worst choices from a set of three jobs described by five attributes such as annual wage, overtime, employment security, possibility of transfer, and possibility of relocation.

The main findings from the empirical analysis suggest that people significantly consider amount of overtime (when it is more than the average), job security, and possibility of intra-firm job transfer and relocation when choosing a job. In particular, people are willing to forfeit a significant portion of their wage to avoid extreme overtime and job transfer. For example, people with an annual wage of 3 million yen are willing to forfeit half of their annual wage in order to avoid 45 hours or longer overtime per month, and people in the highest wage group are willing to pay more than their wage to avoid such extreme overtime. As an alternative and probably more natural interpretation, the results can mean that people with an annual wage of 3 million yen would seek for 50 % wage increase while people with

an annual wage of 8 million yen would require doubling their wage to do 45 hours or more overtime.

The findings from Chapter 2 also suggest that job choice preferences seem to differ significantly between males and females. Women, in general, have higher willingness to pay (WTP) than men to avoid overtime, relocation, and transfer, and to have more secure jobs. Especially, the largest gender difference in amount of WTP for subjects with an annual wage of 3 million yen is found to be that of WTP to avoid extreme overtime, which accounts for almost one-third of wages for subjects who have children. The high female WTPs, especially for overtime, imply that labor supply of women are less elastic than that of men.

Chapter 2 also investigates a soft factor which might be related to gender difference regarding WTP. When taking into account one's guilt level in situations that typical working parents encounter, it was found that women tend to feel guiltier and have higher WTP than men for avoiding work responsibilities. Women also seem to feel more stress than men in both situations when (a) work can interfere with family life and (b) family life can interfere with work. Moreover, the stronger the guilt for women, the larger the portion of wages they are willing to forfeit to avoid overtime, intra-firm transfer, and relocation. For example, when women feel "very guilty" about canceling a kids' event for work, their WTP for avoiding extreme overtime reaches as high as 1.73 million yen for those who earn 3.0 million yen per year.

These results from the CE in Chapter 2 imply that limited-regular contracts can be an appealing option for workers with various characteristics, especially for women with children. While this might have been predicted by common sense, the approach in Chapter 2 advances the argument by quantifying the tradeoff between wage and the non-wage job

attributes. This is important because the WTP estimates reported in Chapter 2 could be used as a benchmark for justifying differences in wages between regular workers and limited-regular workers and could provide legal grounds for contract reform. For the success of the promotion of wider use of limited regular contracts, policy makers should ensure the job quality of the limited regular contracts to be as high as regular one, and it does not widen wage gap between men and women.

Chapter 3 explores the effectiveness of prior task allocation for improving equality in task division among couples, by testing whether additional experience of a task affects preference for doing the task. This chapter also uses an original approach to testing the hypothesis: an economic experiment with Japanese married couples. The results suggest that gender stereotypes are persistent and policy measures would need to be in place for a sufficiently long period to take effect, though further study is needed to test effectiveness under extended time allocation.

In the experiment reported in Chapter 3, a total of 51 couples performed two kinds of tasks, one paid task and one unpaid task. In each couple, one person was randomly assigned to get more experience of the paid/unpaid task. Partners separately indicated their preference regarding the division of work within the couple. The main findings of both inferential statistics and regression analysis suggest that, contrary to our expectations, there was no evidence of a relationship between prior task experience or gender and a couple's task division choice. In particular, the results of the regression analysis indicate that regardless of gender and prior experience, subjects expressed higher preference for the Traditional task pair (i.e. male breadwinner option) over the Reverse task pair (i.e. female breadwinner option) or the Mixed task pair (i.e. half-time dual earner option with equal task sharing).

The results of regression analysis in Chapter 3 also suggest that while more productive males favor traditional task division more than less productive males, more productive females prefer the traditional option less than less productive females. This could be a reminder that the success of labor market reform depends on whether it addresses gender stereotypes, or this may suggest that improvement of female performance could be a channel to greater female labor participation, another possible approach for policymakers to improve gender equality.

4.2 Possible Extension of Research

As a possible extension of the work in Chapter 2, I would like to propose three directions for further research. First, it would be valuable to examine in detail what drives male-female gap in job choice preferences and in the associated WTP. One possible factor suggested in Chapter 2 was gender difference in guilt level and guilt sensitivity to WTP, but other factors might also impact gender gap in preferences. For example, current and previous employment status and work satisfaction, and current housework share with spouse and satisfaction with work division could influence worker job choice. Because those soft factors are difficult to take into account in conventional research methodology based on actual wage cases (i.e. hedonic pricing model), extension of the model to incorporate such factors within the CE framework could contribute to the knowledge of worker preference in job choice.

Second, the effect of children on WTP should be examined with a larger sample. While the results here did suggest that presence of children affect workers' WTP, especially for men, we could say more with a larger sample. Because there was not much variation in the age of youngest children and the total number of children in the sample used for Chapter 2 CE,

the analysis did not satisfactorily determine whether those factors also affect parents' WTP. A larger sample might reveal a clear pattern in WTP, for example, whether younger children or larger total number of children increase parents' WTP to avoid overtime, relocation or transfer. It would be also possible to determine with more observations if the current situation of child care (e.g. nursery school, nannies, and support from grandparents) impacts parents' WTP.

Finally, an employer side experiment or interview would complement results of Chapter 2 in terms of feasibility of the policy presented. While the results of Chapter 2 suggest that workers are likely to accept limited-regular employment options with lower pay, there is a lack of employer side information. A similar CE to quantify employers' WTP for keeping traditional human resource management practices would determine whether there is a gap between labor supply side and demand side. Even an interview with employers regarding barriers against offering of limited-regular contracts would be useful for finding specific areas of focus for policy makers. Nevertheless, this chapter has shown that CE analysis of the Japanese labor market is of value.

As a possible extension to Chapter 3, I identify here three main areas of improvement. First, an experiment which allocates longer experiment time (especially time for treatment) would be necessary to investigate policy effectiveness since we could not rule out the possibility that failure to find effect of task experience was due to limited time allocation. Since existing studies say very little about optimal experiment time for doing tasks, a number of trials would be needed.

Second, an experiment with a larger sample would be beneficial since like many other experimental studies involving real couples, recruitment of participants was challenging too. If more resources were available, a larger sample would yield more convincing policy implications.

Lastly, more complex designs that vary the cost of avoiding unpaid task are needed. Although I carefully set the cost for the Power option to test a hypothesis about gender differences in preferences, there might have been some hidden signal in the way cost was calculated which was attracting both male and female participants indifferently. It would be worth examining whether varying cost would affect the results. Nevertheless, this chapter has shown that it is possible to conduct economics experiments with married couples in Japan.

Appendix 2.1: Survey Questionnaire (English Language Version[28])

Introduction

We would be grateful if you could spare the time to take part in our survey – it should only take around 15-20 minutes to complete. The survey concerns the working life of employees in Japan. The research is being conducted by researchers from the National Graduate Institute for Policy Studies

We would like to ask following questions to people who are 20-59 years old and have had at least one job (excluding *part-time work* while in school or college, self-employed, and family employee.).

Please click on the appropriate answer to each question or type in your answer where required.

Section 1

Q1-1: What is your current employment status? [tick one]

Self-employed

Family employee (in self-employed business)

Executive of company or corporation

Regular Employee of private company or organization

Government employee

---

[28] Note that the actual survey was in Japanese.

Part-time employee

Arbeit

Dispatched worker from temporary labor agency

Contract worker

Housewives/Househusbands

Student

Retired (excluding housewives/household husbands)

Unemployed (excluding housewives/household husbands)

Other (Specify): _____

Q1-2: Please select all employment status you have had before. [tick all that apply]

Self-employed

Family employee (in self-employed business)

Executive of company or corporation

Regular Employee of private company or organization

Government employee

Part-time employee

Arbeit

Dispatched worker from temporary labor agency

Contract worker

Housewives/Househusbands

Student

Retired (excluding housewives/household husbands)

Unemployed (excluding housewives/household husbands)

Other (Specify): _____

＜The following questions are about your current work＞

Q2: Approximately how many employees are working for the organization that employs you? Indicate the approximate numbers including the head office, all branch offices, branch stores, sales offices and factories. **[tick one]**

1 to 5 people

6 to 29 people

30 to 99 people

100 to 299 people

300 to 499 people

500 to 999 people

1,000 to 4,999 people

5,000 or more people

Q3: What is the main business of your organization? [tick one]

Agriculture, Forestry, or Fisheries

Mining

Construction

Manufacturing

Electricity, Gas, Heat supply and Water

Information and communications

Transport and Postal activities

Wholesale and Retail trade

Finance and Insurance

Real estate and Goods rental and Leasing

Accommodations, Eating and Drinking services

Education, Learning support

Medical, Health care and Welfare

Other services

Government

Other

Q4: In a typical working week, about how many hours do you work including overtime work?

_____ hours per week

Q5: In a typical working week, about how many hours of overtime do you work? If you do not work overtime, please write zero.

_____ hours per week

Q6: In a typical working week, about how many days do you work?

_____days per week

Q7: For how many years have you been working for your present employer? [tick one]

Less than a year

A year to less than 5 years

5 years to less than 10 years

10 years to less than 20 years

20 years to less than 30 years

30 years to less than 40 years

More than 40 years

Q8: To what extent is your salary or wage based on your seniority (age, length of tenure)? [tick one]

The wage is based almost entirely on seniority

The wage is based mostly on seniority

The wage is based slightly on seniority

The wage is not at all based on seniority

Q9: What is the possibility that you will be ordered job relocation [29] which requires a move from your home by your current employer in the next five years? Choose the answer that most closely matches your current situation. [tick one]

I might be asked to move to another city

There is little possibility of job relocation for me which requires a move from my home.

There are no branches/plants far away from the head-quarters

Don't know

Q10: What is the possibility that you will be transferred to different division or department [30] within next five years by your current employer? [tick one]

Strong possibility

Some possibility

Little possibility

Don't know

[29] Job relocation is where you are asked by your employer to move your place of work. This question asks about the case when you are asked to move another city or region far enough away that you won't be able to commute on a daily basis from your current home. The move might be temporary or permanent.

[30] This question asks about the chance of intra-firm transfer across divisions or departments which lead to a major change in job contents.

Q11: What is the possibility that you will be unemployed within the next two years? [tick one]

Strong possibility

Some possibility

Little possibility

Don't know

Q12: How long does it take for you to commute? [write zero if you work from home]

_____hours_____minutes

Q13: How true for you is each of the following statements about your current work? True, Basically true, I cannot say either way, Not really true, Untrue

I am satisfied with my work overall: 1 2 3 4 5

I am satisfied with my salary/wage: 1 2 3 4 5

I wish to work longer hours than I do now: 1 2 3 4 5

The location of my work is convenient to me: 1 2 3 4 5

I like the contents of my work: 1 2 3 4 5

There is good communication between colleagues in the department: 1 2 3 4 5

There is good communication between my boss and his/her subordinates: 1 2 3 4 5

I wish to work more days: 1 2 3 4 5

I get enough opportunity for training: 1 2 3 4 5

I have prospects for promotion or career advancement: 1 2 3 4 5

Q14: Are you currently job seeking? *(Please answer regardless of whether you presently have a job.)* If you are job seeking, please indicate the duration of your search.

Seeking a job (___year/s ___month/s)

Not seeking a job

＜The following questions are about your work experience＞

Q15: How many years in total of work experience do you have? _____years

Q16: Have you ever switched jobs (excluding *part-time work* while in school or college)? If you have, please indicate how many times you have changed your jobs.

Yes, ____times

NO, GO TO Q20

＜The following questions are about your previous work＞

Q17: What was the period of employment (from yyyy/mm to yyyy/mm)

From _____/_____ to _____/_____

Q18: What was your employment status?

Regular employee

Part-time employee

Arbeit

Dispatched worker from temporary labor agency

Contract worker

Other (Specify): _____

Q19: What are the reasons for disengagement? [select all that apply]

Didn’t like the job

Lack of recognition or reward

Low salary

Overworked

To start own business

Commute was too long

Too many relocation orders

Temporary job, Unstable employment

Bad relationships with coworkers/bosses

Dissatisfied with the company's management policy

No confidence in company's future

Got married

Gave birth, To take care of children

Health Problem

To look after family member

Family member got a relocation order

Parental leave was not accessible

Unwanted order of employment transfer

Voluntary retirement

Employer went bankrupt

Fired

Contract term was expired

Intended to work for a short term from the beginning

Found a new job

To study

Other (Specify): _____

We want to find more about what aspects of a job are important for you. We want you to think about what makes a particular job attractive for you or unattractive.

Q20: What are the factors you think is important for you when deciding on a job. [tick all that apply.]

The level of the wage

Number of overtime hours

Possibility of mandatory relocation

Possibility of intra-firm job transfers

Job security

Number of employees in the organization

Name recognition of the employer

Flexible hours

Employment close to home

Base working hours per day (excluding overtime)

Number of work days per week

Number of paid leave days per year, Ability to take the agreed leave

Availability of telecommuting/working from home

Other benefit (e.g. housing allowance)

Amount of retirement allowance

Education and training opportunities

Prospect for promotion

Expertise in duties

Social contributions of the work

Good communication with colleagues

Good communication with bosses

<Please read the following information before you answer next questions.>

Information 1

The median (mid-point) length of overtime work for Japanese regular full-time workers is about 15 hours per month. In some jobs there is a lot of overtime in some jobs there is very little overtime.

ABOUT44%
<10Hours

MEDIAN
About 15 Hours

ABOUT 1.6%
>45 Hours

Source: Ministry of Health labor and Welfare, White Paper on "*karoshi*" 2016

Information 2

Many companies (especially big companies) have branches which are far away from the head-quarters, and may order employees to relocate. Depending on employment contract and type of assigned job, the possibility of relocation order may vary across workers in a same company. Once the order is made, employees are expected to follow the order regardless of their preference.

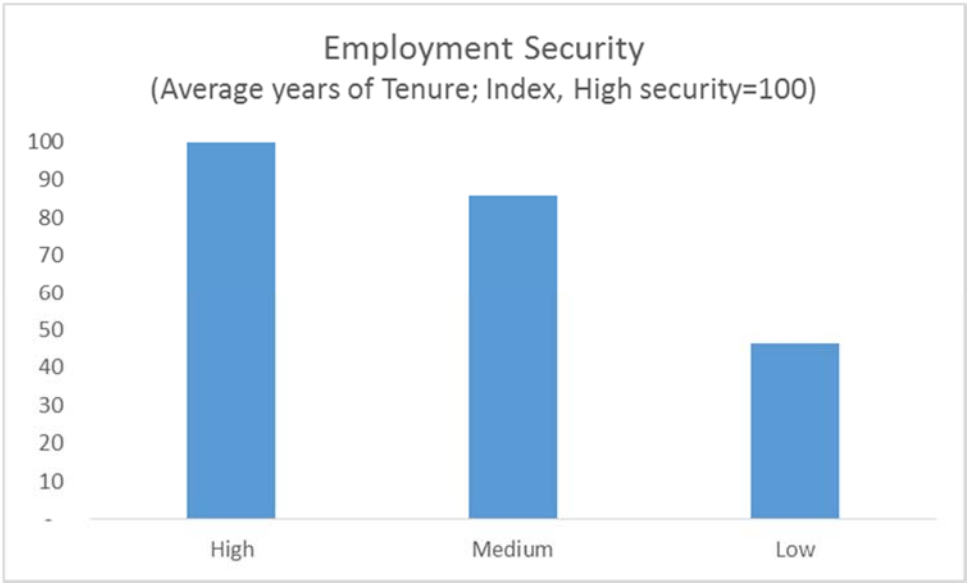
<u>Information 3</u>

Many companies rotate employees across different departments/sections within the firm, to let them acquire a wider set of skills. On the other hand, some companies also transfer employees from time to time to maintain employment when the business is going bad or to allocate business resources efficiently. In either case, employees are expected to follow the order and start the new assignment regardless of their preference.

Information 4

Life-long employment, or being employed by the same employer from the first year after the final schooling till the retirement date, is believed to be a status enjoyed by traditional regular workers in Japan. In today's economy, employment security differs across workers.

In the survey questions we shall see below, we categorize employment security into three levels. The lower the employment security is, the shorter the average tenure tend to be for the job. In what follows, we suppose Medium employment security and Low employment security to have the average tenure shorter compared to High employment security by the degree shown in the chart below.


Employment Security
(Average years of Tenure; Index, High security=100)
100
90
80
70
60
50
40
30
20
10
-
High
Medium
Low

<We are now going to show you a series of hypothetical job choices. Each job description will have some information about the working conditions. In each question you must select the job you like the most and the job you like the least. There are total 8 questions.

In the options we will give you information about the total annual wage (including overtime pay and bonuses), the amount of overtime, the possibility of relocation, the possibility of changes in job contents and employment stability. The jobs included are all regular jobs, meaning they are not temporary jobs with low employment stability. Please study the information carefully, then make your choice.

What about other features of the job, such as the type of work, the size of the company and so on? Here, we want you to suppose that the other features of the job are almost the same as your current job or the last full-time job you held.>

Q21: Please choose the job you like the MOST and the job you like the LEAST from Job A, Job B, and Job C below.

Example	Job A	Job B	Job C
Total annual wage	4 million yen	3 million yen	3.5 million yen
Overtime	More than 45 Hours/month	zero	zero
Employment Stability	Medium	High	Medium
Possibility of relocation	Zero	Zero	Some
Possibility of intra-firm transfer	Some	Zero	Zero
Which Job do you **LIKE the MOST**?	○	○	○
Which Job do you **LIKE the LEAST**?	○	○	○

Choice question like above example * eight times (Q21-Q28)

There are 4 blocks of eight choice sets, so there are in total 32 choice sets

＜The following questions ask you about your assessment of the above questions ＞

Q29: Did you ignore any of the below attributes when you make the hypothetical job choice in Q21-Q28? [check all that apply]

Total annual wage

Overtime

Employment Stability

Possibility of relocation

Possibility of intra-firm transfer

None of above

Q.30: How true for you is each of the following statements?

True, Basically true, I cannot say either way, Not really true, Untrue

The choice sets presented were realistic: 12345

The hypothetical situation applied to me: 12345

Q.31. Now, suppose you were offered a job with the following features:

Overtime: More than 45 hours per month

Employment stability: High

Possibility of intra-firm transfer: Some

Possibility of relocation: Some

Would you switch to this job if [please choose all that apply],

The wage was 1m Yen per year lower than your current salary

The wage was the same as your current salary

The wage was 1m Yen per year higher than your current salary

The wage was 3m Yen per year higher than your current salary?

None of above

<The following questions are about you, your family and your home-life>

Q32: What is your highest educational qualification? [tick one]

Secondary School (until age 16)

Graduated from High School

Graduated from College - Associate's Degree (2 year)

Graduated from College - Bachelor's Degree (4 year)

Master's Degree - MS, MA, MBA, etc.

Doctoral Degree - DVM, Ph.D, DDS, etc.

Q33: 'Anyone supported' is someone claimed as dependent on the last tax return. Are you ...[tick one]

Supporting someone in the family.

Supported by someone in your family.

Neither supporting nor being supported. (including single-person household)

Q34: Approximately how much is your total annual income in 2016 before tax but including bonuses and overtime pay? [tick one]

None

Less than 1 million yen

1 million yen to less than 2 million yen

2 million yen to less than 3 million yen

3 million yen to less than 4 million yen

4 million yen to less than 5 million yen

5 million yen to less than 6 million yen

6 million yen to less than 8 million yen

8 million yen to less than 10 million yen

10 million yen to less than 14 million yen

More than 14 million yen

Q35: Do you have a spouse or a partner with whom you live?

Yes

No GO TO Q43

Q36: What is your spouses or partner's highest educational qualification? [tick one]

Secondary School (until age 16)

Graduated from High School

Graduated from College - Associate's Degree (2 year)

Graduated from College - Bachelor's Degree (4 year)

Master's Degree - MS, MA, MBA, etc.

Doctoral Degree - DVM, Ph.D, DDS, etc.

Q37: Is your spouse or partner in paid employment?

Yes

No GO TO Q41

Q38: Which of the following best describes your spouse or partner's employment status? [tick one]

Self-employed

Family employee (in self-employed business)

Executive of company or corporation

Regular Employee of private company or organization

Government employee

Part-time employee

Arbeit

Dispatched worker from temporary labor agency

Contract worker

Other (Specify): _____

Q39: How many hours a week does your spouse/partner work including overtime work?

_____ hours per week

Q40: What is your spouse or partner's approximate annual income from the job in 2016 before tax but including bonuses and overtime pay? [tick one]

None

Less than 1 million yen

1 million yen to less than 2 million yen

2 million yen to less than 3 million yen

3 million yen to less than 4 million yen

4 million yen to less than 5 million yen

5 million yen to less than 6 million yen

6 million yen to less than 8 million yen

8 million yen to less than 10 million yen

10 million yen to less than 14 million yen

More than 14 million yen

Q41: Do you live with your parent(s) or your spouse/partner's parent(s)?

Yes

No

Q42: Do you have any children with whom you live?

Yes

No GO TO Q45

Q43: How many children do you have?

0-3 years old: _____Child/ren, zero

4-6years old: _____Child/ren, zero

6-12years old: _____Child/ren, zero

13-18 years old: _____Child/ren, zero

19 years old or older: _____Child/ren, zero

Q44: Which of the following systems, services or support have you used when bringing up your children? [tick all that apply]

I have taken parental leave

I switched to working part time/short time

My partner has taken parental leave

My partner switched to working part time/short time

Baby sitter or nanny

Nursery School

Kindergarten

Support from parents, siblings or other close relatives of you or your partner

Support from friends or other more distant relatives

Other (please specify)_____

I was a stay-home mom/dad when I was raising my children

Q45: How much time do you spend on average doing following activities every day? [tick one]

Zero; 0-1hour; 1-2 hours; 2-3 hours; 3-4 hours; 4-5 hours; 5 hours or more

Housework 1234567

Caring for and helping your children 1234567

Caring for and helping other relatives 1234567

Community services, volunteering 1234567

Educational activities 1234567

Q46: In your household, who is responsible for the following? (choose the option that is closest to the situation in your household)

Always me, Mostly me, Mostly my spouse, Always my spouse, Someone else

Separating garbage for recycling 12345

Remembering the garbage day and putting out the garbage 12345

Cleaning the house 12345

Turning off the lights 12345

Paying household bills 12345

Q47: (SKIP THIS QUESTION IF NO SPOUSE) **How satisfied are you with your spouse/partner on the division of work in the following?**

Satisfied, Basically Satisfied, I cannot say either way, Not really satisfied, Unsatisfied

Home making 12345

Child rearing 12345 DO NOT HAVE CHILDREN

Taking care of other relatives 12345 DO NOT HAVE OTHER RELATIVES WHO NEED CARE

Paid-work 12345 NEITHER ME OR MY SPOUSE CURRENTLY HAVE PAID WORK

Q48: How much guilty would you feel by acting as each of the statement below. If the situation does not apply to you or your spouse/partner, please answer as you imagine what you would feel in the situation.

Very Guilty, Somewhat Guilty, I cannot say either way, Not really guilty, Not Guilty at all

I took paid leave when my managers and colleagues are working a lot of overtime.

1234

I left the office on time for a family event when my managers and colleagues are working.

1234

I did not prepare healthy dinner for me and for my family for the entire week.

1234

Because I was working I missed my child(ren)'s event which I had promised to go.

1234

I did not see my elderly parents or other relatives who needs care for the last one month.

1234

Not earning enough income to satisfy the demands of my child (extra-academic activities, clothes, games…)

1234

Not being able to spend time with my child when we are at home because I must perform tasks that do not concern the family.

1234

Q49: (SKIP THIS QUESTION IF NO SPOUSE) **How much guilty do you think your spouse/partner would feel by acting as each of the statement below? If the situation does not apply to you or your spouse/partner, please answer as you imagine what you would feel in the situation.**

Very Guilty, Somewhat Guilty, I cannot say either way, Not really guilty, Not Guilty at all

I took paid leave when my managers and colleagues are working a lot of overtime.

1234

I left the office on time for a family event when my managers and colleagues are working.

1234

I did not prepare healthy dinner for me and for my family for the entire week.

1234

Because I was working I missed my child(ren)'s event which I had promised to go.

1234

I did not see my elderly parents or other relatives who needs care for the last one month.

1234

Not earning enough income to satisfy the demands of my child (extra-academic activities, clothes, games…)

1234

Not being able to spend time with my child when we are at home because I must perform tasks that do not concern the family.

1234

Thank you for your help with this study!

Appendix 2.2: An Example of a Choice Question in Japanese

Appendix Table A2. 1 An example of a choice question in Japanese

	仕事A	仕事B	仕事C
年間の総賃金	４００万円	３００万円	３５０万円
残業時間	45時間以上/月	0時間/月	0時間/月
雇用の安定性	中くらい	高い	中くらい
異動の可能性	なし	なし	あり
転勤の可能性	あり	あり	なし
もっとも好ましい	○	○	○
もっとも好ましくない	○	○	○

Appendix 2.3: Extended Model 1 & 2

Appendix Table A2. 2 Extended Models: Model 1 (Presence of Children), Model 2 (Marital Status)

VARIABLES	Model 1 (Presence of Children)	Model 2 (Marital Status)
Wage_cont	1.362***	1.369***
	(0.0678)	(0.0679)
Wage_sq	-0.0774***	-0.0779***
	(0.00516)	(0.00516)
Overtime2	-0.0417	-0.0864
	(0.142)	(0.129)
Overtime3	-0.521***	-0.772***
	(0.170)	(0.162)
Overtime4	-1.409***	-1.802***
	(0.235)	(0.222)
Security2	0.0220	0.460***
	(0.124)	(0.109)
Transfer2	-0.376***	-0.538***
	(0.127)	(0.115)
Relocation2	0.000350	-0.208**
	(0.0926)	(0.0815)
Interactions with Female Dummy		
Overtime2	-0.123	0.0417
	(0.0993)	(0.104)
Overtime3	-0.444***	-0.164
	(0.0972)	(0.116)
Overtime4	-0.845***	-0.387**

	(0.137)	(0.157)
Security2	0.381***	0.0205
	(0.0920)	(0.0889)
Transfer2	-0.604***	-0.306***
	(0.0978)	(0.101)
Relocation2	-0.317***	-0.102
	(0.0684)	(0.0705)
Interactions with	No_Children	Married
Overtime2	-0.0570	0.0180
	(0.0952)	(0.0955)
Overtime3	-0.220**	0.229**
	(0.0937)	(0.0996)
Overtime4	-0.403***	0.286**
	(0.129)	(0.135)
Security2	0.418***	-0.360***
	(0.0889)	(0.0893)
Transfer2	-0.120	0.103
	(0.0843)	(0.0867)
Relocation2	-0.225***	0.138**
	(0.0595)	(0.0608)
Interactions with	Female*No_Children	Female*Married
Overtime2	0.208	-0.0774
	(0.132)	(0.134)
Overtime3	0.234*	-0.257*
	(0.138)	(0.146)
Overtime4	0.485**	-0.313
	(0.189)	(0.198)
Security2	-0.342***	0.294**
	(0.118)	(0.118)
Transfer2	0.156	-0.345***
	(0.131)	(0.132)

Relocation2	0.268***	-0.109
	(0.0907)	(0.0921)
Interactions with age cohorts		
30s # Overtime2	-0.0170	-0.0202
	(0.127)	(0.127)
40s # Overtime2	0.0641	0.0649
	(0.137)	(0.138)
50s # Overtime2	0.0332	0.0327
	(0.132)	(0.137)
30s # Overtime3	0.276*	0.278*
	(0.162)	(0.162)
40s # Overtime3	0.392**	0.380**
	(0.170)	(0.171)
50s # Overtime3	0.284*	0.262
	(0.167)	(0.173)
30s # Overtime4	0.610***	0.612***
	(0.223)	(0.223)
40s # Overtime4	0.681***	0.675***
	(0.232)	(0.234)
50s # Overtime4	0.589**	0.573**
	(0.231)	(0.238)
30s # Security2	-0.0222	-0.0274
	(0.109)	(0.108)
40s # Security2	0.0213	0.0310
	(0.115)	(0.116)
50s # Security2	-0.103	-0.0762
	(0.117)	(0.120)

30s # Transfer2	-0.265**	-0.229*
	(0.119)	(0.119)
40s # Transfer2	-0.217*	-0.173
	(0.126)	(0.128)
50s # Transfer2	-0.269**	-0.218*
	(0.127)	(0.131)
30s # Relocation2	-0.0558	-0.0562
	(0.0843)	(0.0849)
40s # Relocation2	-0.162*	-0.166*
	(0.0901)	(0.0914)
50s # Relocation2	-0.138	-0.142
	(0.0895)	(0.0922)
Observations	25,104	25,104
Number of groups	8,368	8,368

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix 2.4 : Extended Models 3 and its WTP measures

Appendix Table A2. 3 Extended Model: Model 3 (Education)

VARIABLES	Model 3 (Education)
Wage_cont	1.375***
	(0.0678)
Wage_sq	-0.0781***
	(0.00516)
Overtime2	-0.0720
	(0.149)
Overtime3	-0.578***
	(0.176)
Overtime4	-1.550***
	(0.239)
Security2	0.233*
	(0.131)
Transfer2	-0.543***
	(0.135)
Relocation2	-0.0337
	(0.0989)
Interactions with Female Dummy	
Overtime2	-0.116
	(0.100)
Overtime3	-0.422***
	(0.0982)
Overtime4	-0.801***
	(0.138)

Security2	0.321***
	(0.0919)
Transfer2	-0.555***
	(0.0997)
Relocation2	-0.308***
	(0.0701)
Interactions with No-Children	
Overtime2	-0.0527
	(0.0948)
Overtime3	-0.222**
	(0.0934)
Overtime4	-0.395***
	(0.129)
Security2	0.401***
	(0.0885)
Transfer2	-0.104
	(0.0838)
Relocation2	-0.217***
	(0.0596)
Interactions with Female*No_Children	
Overtime2	0.206
	(0.132)
Overtime3	0.224
	(0.138)
Overtime4	0.465**
	(0.190)
Security2	-0.312***
	(0.117)
Transfer2	0.130
	(0.132)
Relocation2	0.265***

	(0.0916)
Interactions with age cohorts	
30s # Overtime2	-0.0118
	(0.127)
40s # Overtime2	0.0700
	(0.137)
50s # Overtime2	0.0346
	(0.133)
30s # Overtime3	0.270*
	(0.161)
40s # Overtime3	0.390**
	(0.169)
50s # Overtime3	0.284*
	(0.166)
30s # Overtime4	0.613***
	(0.221)
40s # Overtime4	0.704***
	(0.231)
50s # Overtime4	0.606***
	(0.229)
30s # Security2	-0.0415
	(0.108)
40s # Security2	-0.0279
	(0.115)
50s # Security2	-0.147
	(0.116)
30s # Transfer2	-0.253**

	(0.117)
40s # Transfer2	-0.183
	(0.124)
50s # Transfer2	-0.237*
	(0.125)
30s # Relocation2	-0.0534
	(0.0845)
40s # Relocation2	-0.146
	(0.0907)
50s # Relocation2	-0.131
	(0.0900)
Interaction with Educated Dummy	
Overtime2	0.0394
	(0.0667)
Overtime3	0.0895
	(0.0706)
Overtime4	0.185*
	(0.0950)
Security2	-0.255***
	(0.0581)
Transfer2	0.208***
	(0.0671)
Relocation2	0.0336
	(0.0458)
Observations	25,104
Number of groups	8,368

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Educated=1 if graduated from 4-year universities or graduate schools, and Educated =0 otherwise.

Appendix Table A2. 4 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, calculated from Model 3 (Educated Subjects)

WTP	Base	Female	No_children	Female*No_children
Overtime2	0.05	0.18**	0.11	0.01
Overtime3	0.24***	0.71***	0.49***	0.70
Overtime4	0.83***	1.71***	1.26***	1.64***
Security 2	0.07	-0.28***	-0.37***	-0.38***
Transfer 2	0.65***	1.26***	0.76***	1.23***
Relocation2	0.06	0.40***	0.30***	0.35***

***p<0.01 , **p<0.05, *p<0.1

Appendix Table A2. 5 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, calculated from Model 3 (Uneducated Subjects)

WTP	Base	Female	No_children	Female*No_children
Overtime2	0.09	0.22***	0.15	0.05
Overtime3	0.34***	0.81***	0.58***	0.80***
Overtime4	1.03***	1.92***	1.47***	1.84***
Security 2	-0.21**	-0.57***	-0.65***	-0.66***
Transfer 2	0.88***	1.49***	0.99***	1.46***
Relocation2	0.10	0.44***	0.34***	0.38***

***p<0.01 , **p<0.05, *p<0.1

Appendix 2.5: Extended Models 4

Appendix Table A2. 6 Extended Model4 (Existing Income)

VARIABLES	Model 4 (Existing Income)
Wage_cont	1.369***
	(0.0681)
Wage_sq	-0.0776***
	(0.00517)
Overtime2	-0.0293
	(0.158)
Overtime3	-0.546***
	(0.183)
Overtime4	-1.596***
	(0.253)
Security2	0.329**
	(0.138)
Transfer2	-0.579***
	(0.141)
Relocation2	-0.0851
	(0.105)
Interactions with Female Dummy	
Overtime2	-0.138
	(0.118)
Overtime3	-0.418***
	(0.115)
Overtime4	-0.690***
	(0.160)
Security2	0.130
	(0.103)
Transfer2	-0.439***

	(0.114)
Relocation2	-0.245***
	(0.0811)
Interactions with No-Children Dummy	
Overtime2	-0.0658
	(0.0977)
Overtime3	-0.216**
	(0.0959)
Overtime4	-0.356***
	(0.133)
Security2	0.329***
	(0.0889)
Transfer2	-0.0638
	(0.0864)
Relocation2	-0.197***
	(0.0601)
Interactions with Female*No-Children	
Overtime2	0.222
	(0.143)
Overtime3	0.395**
	(0.196)
Overtime4	-0.193
	(0.120)
Security2	0.0586
	(0.136)
Transfer2	0.223**
	(0.0949)
Interactions with age cohorts	
30s # Overtime2	-0.00516
	(0.127)
40s # Overtime2	0.0754
	(0.136)

50s # Overtime2	0.0396
	(0.132)
30s # Overtime3	0.267*
	(0.162)
40s # Overtime3	0.382**
	(0.170)
50s # Overtime3	0.272
	(0.168)
30s # Overtime4	0.606***
	(0.223)
40s # Overtime4	0.678***
	(0.233)
50s # Overtime4	0.577**
	(0.231)
30s # Security2	-0.0195
	(0.109)
40s # Security2	0.0215
	(0.115)
50s # Security2	-0.0951
	(0.117)
30s # Transfer2	-0.265**
	(0.118)
40s # Transfer2	-0.219*
	(0.125)
50s # Transfer2	-0.273**
	(0.126)

30s # Relocation2	-0.0559
	(0.0852)
40s # Relocation2	-0.161*
	(0.0909)
50s # Relocation2	-0.144
	(0.0904)
Interaction with High_Income dummy	
Overtime2	-0.0241
	(0.0790)
Overtime3	0.0420
	(0.0832)
Overtime4	0.220**
	(0.112)
Security2	-0.345***
	(0.0669)
Transfer2	0.228***
	(0.0770)
Relocation2	0.0938*
	(0.0543)
Observations	25,104
Number of groups	8,368

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2.6: Extended Models 5 & 6

Appendix Table A2. 7 Extended Models: Model 5 (Total Number of Children), Model 6 (Age of the Youngest Children)

VARIABLES	Model 5 (Total Number of Children)	Model 6 (Age of the Youngest Children)
Wage_cont	1.376***	1.366***
	(0.0681)	(0.0682)
Wage_sq	-0.0782***	-0.0776***
	(0.00518)	(0.00520)
Overtime2	-0.0897	-0.118
	(0.124)	(0.125)
Overtime3	-0.732***	-0.764***
	(0.156)	(0.157)
Overtime4	-1.798***	-1.838***
	(0.216)	(0.216)
Security2	0.413***	0.456***
	(0.107)	(0.107)
Transfer2	-0.483***	-0.484***
	(0.113)	(0.114)
Relocation2	-0.199**	-0.229***
	(0.0798)	(0.0796)
Interactions with Female Dummy		
Overtime2	0.0637	0.0885
	(0.0805)	(0.0869)
Overtime3	-0.256***	-0.202**
	(0.0892)	(0.0971)
Overtime4	-0.422***	-0.352***
	(0.119)	(0.130)

Security2	0.113	0.0367
	(0.0697)	(0.0743)
Transfer2	-0.460***	-0.452***
	(0.0814)	(0.0887)
Relocation2	-0.107*	-0.0458
	(0.0555)	(0.0596)
Interaction with Youngest_chidren dummy		
Infant_toddler # Overtime2		0.0966
		(0.162)
Preschool # Overtime2		0.331*
		(0.201)
Elementary_school# Overtime2		-0.142
		(0.189)
Teenager # Overtime2		0.107
		(0.165)
Over18 # Overtime2		-0.105
		(0.179)
Infant_toddler # Overtime3		0.177
		(0.160)
Preschool # Overtime3		0.648***
		(0.194)
Elementary_school# Overtime3		0.106
		(0.157)
Teenager # Overtime3		0.166
		(0.154)
Over18 # Overtime3		0.0678
		(0.168)
Infant_toddler # Overtime4		0.454**
		(0.221)

Preschool # Overtime4	1.178***
	(0.236)
Elementary_school# Overtime4	0.232
	(0.231)
Teenager # Overtime4	0.373*
	(0.209)
Over18 # Overtime4	-0.0302
	(0.231)
Infant_toddler # Security2	-0.264*
	(0.158)
Preschool # Security2	-0.219
	(0.242)
Elementary_school# Security2	-0.492***
	(0.153)
Teenager # Security2	-0.618***
	(0.165)
Over18 # Security2	-0.411***
	(0.151)
Infant_toddler # Transfer2	0.0645
	(0.148)
Preschool # Transfer2	-0.100
	(0.180)
Elementary_school# Transfer2	-0.0175
	(0.154)
Teenager # Transfer2	0.423***
	(0.125)
Over18 # Transfer2	0.160
	(0.165)

Infant_toddler # Relocation2	0.309***
	(0.102)
Preschool # Relocation2	0.270**
	(0.136)
Elementary_school# Relocation2	0.170*
	(0.101)
Teenager # Relocation2	0.114
	(0.106)
Over18 # Relocation2	0.189*
	(0.115)
Interactions with Youngest_children*Female	
Infant_toddler # female # Overtime2	-0.138
	(0.208)
Preschool # female # Overtime2	-0.571**
	(0.243)
Elementary_school# female # Overtime2	-0.0375
	(0.251)
Teenager # female # Overtime2	-0.434
	(0.293)
Over18 # female # Overtime2	-0.0777
	(0.233)
Infant_toddler # female # Overtime3	-0.0762
	(0.210)
Preschool # female # Overtime3	-0.511*
	(0.264)
Elementary_school# female # Overtime3	-0.00119
	(0.248)
Teenager # female # Overtime3	-0.508**
	(0.242)
Over18 # female # Overtime3	-0.337

	(0.236)
Infant_toddler # female # Overtime4	-0.429
	(0.306)
Preschool # female # Overtime4	-1.176***
	(0.336)
Elementary_school# female # Overtime4	-0.0831
	(0.330)
Teenager # female # Overtime4	-0.595*
	(0.334)
Over18 # female # Overtime4	-0.414
	(0.320)
Infant_toddler # female # Security2	0.0313
	(0.191)
Preschool # female # Security2	0.100
	(0.282)
Elementary_school# female # Security2	0.342
	(0.211)
Teenager # female # Security2	0.740***
	(0.249)
Over18 # female # Security2	0.561***
	(0.199)
Infant_toddler # female # Transfer2	-0.170
	(0.207)
Preschool # female # Transfer2	0.0741
	(0.286)
Elementary_school# female # Transfer2	-0.0642
	(0.239)
Teenager # female # Transfer2	-0.522*

		(0.278)
Over18 # female # Transfer2		-0.0526
		(0.225)
Infant_toddler # female # Relocation2		-0.343**
		(0.142)
Preschool # female # Relocation2		-0.142
		(0.188)
Elementary_school# female # Relocation2		-0.212
		(0.160)
Teenager # female # Relocation2		-0.0797
		(0.209)
Over18 # female # Relocation2		-0.410**
		(0.162)
Interactions with Age Cohort		
30s # Overtime2	-0.0184	-0.0158
	(0.128)	(0.127)
40s # Overtime2	0.0528	0.0867
	(0.137)	(0.138)
50s # Overtime2	0.00695	0.0839
	(0.130)	(0.137)
30s # Overtime3	0.263	0.257
	(0.162)	(0.161)
40s # Overtime3	0.381**	0.411**
	(0.170)	(0.170)
50s # Overtime3	0.321**	0.386**
	(0.163)	(0.171)
30s # Overtime4	0.580***	0.575***
	(0.223)	(0.222)

40s # Overtime4	0.641***	0.679***
	(0.232)	(0.231)
50s # Overtime4	0.634***	0.748***
	(0.224)	(0.232)
30s # Security2	-0.00353	-0.0236
	(0.110)	(0.108)
40s # Security2	0.0230	0.0102
	(0.116)	(0.116)
50s # Security2	-0.179	-0.145
	(0.116)	(0.124)
30s # Transfer2	-0.264**	-0.254**
	(0.120)	(0.119)
40s # Transfer2	-0.213*	-0.222*
	(0.126)	(0.127)
50s # Transfer2	-0.256**	-0.334**
	(0.124)	(0.136)
30s # Relocation2	-0.0694	-0.0744
	(0.0847)	(0.0844)
40s # Relocation2	-0.169*	-0.165*
	(0.0906)	(0.0914)
50s # Relocation2	-0.107	-0.0893
	(0.0885)	(0.0937)
Interactions with Children_total		
1 child # Overtime2	0.178	
	(0.125)	
2 children # Overtime2	0.00452	
	(0.148)	
3 children # Overtime2	-0.00445	

	(0.310)
4 children # Overtime2	-18.25***
	(0.0773)
1 child # Overtime3	0.150
	(0.120)
2 children # Overtime3	0.346**
	(0.136)
3 children # Overtime3	0.183
	(0.277)
4 children # Overtime3	-1.002***
	(0.0928)
1 child # Overtime4	
	0.541***
2 children # Overtime4	(0.158)
	0.477**
3 children # Overtime4	(0.195)
	0.421
4 children # Overtime4	(0.407)
	-55.54
1 child # Security2	-0.406***
	(0.127)
2 children # Security2	-0.341**
	(0.133)
3 children # Security2	-0.291
	(0.303)
4 children # Security2	-17.84***
	(0.0943)
1 child # Transfer2	0.115

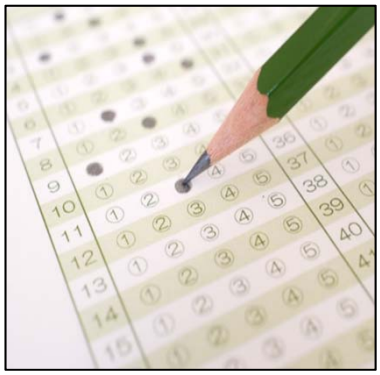
	(0.101)
2 children # Transfer2	0.0508
	(0.131)
3 children # Transfer2	0.291
	(0.297)
4 children # Transfer2	0.314***
	(0.0762)
1 child # Relocation2	0.188**
	(0.0787)
2 children # Relocation 2	0.220**
	(0.0871)
3 children # Relocation 2	0.160
	(0.141)
4 children # Relocation 2	-18.18***
	(0.0731)
Interactions with Children_total*Female	
1 child # Female # Overtime2	-0.296*
	(0.174)
2 children # Female # Overtime2	-0.0929
	(0.199)
3 children # Female # Overtime2	-0.498
	(0.389)
4 children # Female # Overtime2	18.31***
	(0.558)
1 child # Female # Overtime3	-0.136
	(0.175)
2 children # Female # Overtime3	-0.178
	(0.197)
3 children # Female # Overtime3	-0.269

	(0.375)
4 children # Female # Overtime3	1.046**
	(0.479)
1 child # Female # Overtime4	
	-0.635***
2 children # Female # Overtime4	(0.240)
	-0.226
3 children # Female # Overtime4	(0.278)
	-0.556
4 children # Female # Overtime4	(0.535)
	55.25***
	(0.441)
1 child # Female # Security2	
	0.311*
2 children # Female # Security2	(0.165)
	0.00489
3 children # Female # Security2	(0.169)
	-0.00230
4 children # Female # Security2	(0.377)
	18.33***
	(0.464)
1 child # Female # Transfer2	
	-0.0712
2 children # Female # Transfer2	(0.161)
	-0.0967
3 children # Female # Transfer2	(0.206)
	-0.948**
4 children # Female # Transfer2	(0.445)
	-3.137***
	(0.486)
1 child # Female # Relocation 2	

	-0.131	
2 children # Female # Relocation 2	(0.123)	
	-0.142	
3 children # Female # Relocation 2	(0.126)	
	-0.408	
4 children # Female # Relocation 2	(0.346)	
	17.53***	
	(0.561)	
Observations	25,104	25,104
Number of groups	8,368	8,368

Appendix 3.1: Visual Guide (in Japanese)

ビジュアルガイド



タスク A

○　正しくマークした一問あたり２０円の報酬が出ます

×　間違えてマークした一問あたり５円減額します

問題用紙に書かれた正しい回答をマークシートにえんぴつで書き写す作業です。（問題文を読む必要はありません）

タスク B

報酬は出ません。

迷路に書かれたえんぴつの線を消しゴムで消す作業です。

ペナルティーもありませんが、子供たちが再利用するので、なるべく多くきれいに仕上げてください。

実験のタイムスケジュール

開始前の説明

練習	全員	タスク A (1 分) タスク B (1 分)		
本番１回目 (個人プレー)	全員	タスク A (4 分) タスク B (4 分)		
本番２回目 (夫婦ペア： 指定されたタスク)	●の人	タスク A (4 分)	▲の人	タスク B (4 分)

夫婦間の３回目の作業分担についての選択

本番３回目 (夫 婦 ペ ア ： 選択したタスク)	(緑) または (青) または (赤) または (黄) …説明は次ページ (4 分)

アンケート回答

謝金のお渡し

第三回のタスクの組み合わせ

	旦那さん	奥さん	一人当たりの報酬額
(緑)	タスク A (4 分)	タスク B (4 分)	旦那さんが稼いだ金額
(青)	タスク B (4 分)	タスク A (4 分)	奥さんが稼いだ金額
(赤)	タスク A (4 分)	タスク A (4 分)	夫婦それぞれが稼いだ金額の合計から費用を引いた金額
(黄)	タスク A と B (2 分×2)	タスク A と B (2 分×2)	夫婦それぞれが半分の時間で稼いだ合計金額

（報酬の計算例）

(緑)：旦那さんが稼いだ金額

- ３０問正しくマークし、間違いがゼロの場合。

一人当たり報酬：３０問×２０円＝６００円

- 同じように３０問マークしていても、１つ不正確なマークがあった場合

一人当たりの報酬：３０問×２０円－１問×５円＝５９５円

(青)：**奥さん**が稼いだ金額

報酬の計算方法は緑と同じ。

(赤)：夫婦それぞれが稼いだ金額の合計から**費用を引いた**金額

旦那さん（２５問）、奥さん（３０問）、トライアル実験の平均（１５問）で、間違いがゼロの場合。

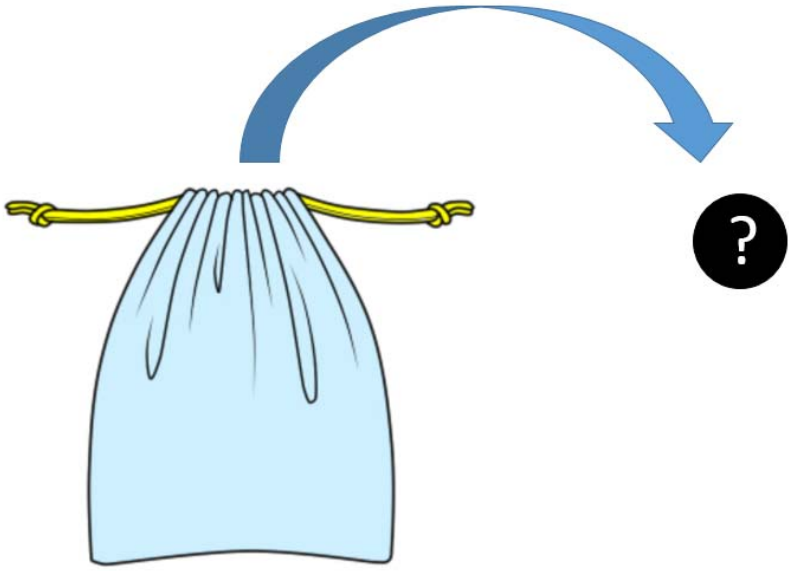
一人当たりの報酬：（２５＋３０－１５）×２０円＝８００円

(黄)：夫婦それぞれが**半分の時間で**稼いだ合計金額

旦那さん（１５問）、奥さん（１０問）、間違いがゼロの場合。

一人当たりの報酬：（１５＋１０）×２０円＝５００円

第三回の作業の組み合わせの決定方法：くじ引き



- 記録シートの２ページ目、<u>タスクの組み合わせの希望</u>に回答する。
- 順位が高い色のチップが多くなるように、袋の中に入れる。（１位の色が４枚、２位の色が３枚、３位の色が２枚、４位の色が１枚）。

  ※アシスタントが一緒に行いますので、しばらくお待ちください。
- パートナーのチップも同じ袋に入れる。
- ●の人が袋の中から一枚チップをひき、選ばれた色のタスクの組み合わせが第三回の作業に決定する。

Appendix 3.2: Recording Sheet (in Japanese)

理解度テスト

1. どちらのタスクに対して報酬が支払われますか？

□タスク A　　□タスク B

2. タスク A では、正しくマークシートを塗りつぶした一問あたりの報酬はいくらですか?

_____ 円

3. タスク A では、不正確にマークシートを塗りつぶした場合、報酬がいくら減額されますか？

_____ 円

4. タスク B では、迷路が少ししか消し終わらなかったり、完全に消せていない場合にペナルティはありますか？

□ある　　□ない

タスクの組み合わせの希望

第三回であなたのご夫婦が行っていただくタスクの組み合わせとして、好ましいものから順に１から４まで色をご記入ください。（例えば緑が一番好ましい場合、順位１の横に緑とご記入ください）。順位付けはご夫婦で相談することなく、各自で行ってください。

順位	タスクの組み合わせ
1	
2	
3	
4	

※実際に第三回で行っていただく作業の組み合わせは、あなたとあなたのパートナーの希望を考慮して、くじ引きで決定します。

Appendix 3.3: Questionnaire (in Japanese)

Q1. あなたの性別をお答えください。

男性

女性

Q2. あなたの年齢をお答えください。

２０代

３０代

４０代

５０代

Q3.あなたの最後に卒業された学校をお答えください。

中学校 卒業（盲/聾/養護学校の中学部を含む）

高等学校 卒業（盲/聾/養護学校の高等部を含む）

短期大学 卒業（専門学校、高等専門学校等を含む）

大学 卒業（防衛大学などの省庁大学校を含む）

大学院修士課程 修了

大学院博士課程 修了

Q4.ここで扶養家族とは、健康保険上の扶養家族のことをいいます。当てはまるものを 1 つ選んでください。

自分は家族の誰かを扶養している

自分は家族の誰かに扶養されている

どちらにもあてはまらない

Q5.あなたの 2016 年のボーナス・残業代を含めた税引き前の総収入はおよそいくらでしたか。

なし

100 万円未満

100～200 万円未満

200～300 万円未満

300～400 万円未満

400～ 500 万円未満

500～600 万円未満

600～800 万円未満

800～1,000 万円未満

1,000～1,400 万円未満

1,400 万円以上

Q6.あなたかあなたの配偶者（夫または妻）の親と同居していますか。

同居をしている

同居をしていない

Q7.あなたは一緒に暮らしている１８歳以下のお子さんがいますか。

いる

いない　Q.10 へ

Q8.一緒に暮らしているお子さんの年齢をお答えください。また、年齢ごとにお子さんの人数をお答えください。【複数選択可】

０－３歳：_____ 人

４－６歳：_____人

７－１２歳：_____人

１３－１８歳：_____人

Q9.育児の方法であてはまるものをすべてお答えください。お子さんが小学生以上の場合は、小学生になる前の育児方法について、あてはまるものをすべてお答えください。【複数選択可】

あなたが育児休暇を取得した

あなたが時短勤務またはパートタイムで働くことにした

あなたの配偶者（夫または妻）が育児休暇を取得した

あなたの配偶者（夫または妻）が時短勤務またはパートタイムで働くことにした

ベビーシッターを利用した

保育園

幼稚園

両親、兄弟姉妹、その他の近い親戚に手伝ってもらった

友人や遠い親戚に手伝ってもらった

その他（具体的に　　）

あなたは育児期間中、専業主婦／主夫だった

Q.10.あなたはふだん一日のうち平均してどれくらいの時間を次のことに使っていますか？あてはまるものを１つ選んでください。

	使っていない	0～1時間未満	1～2時間未満	2～3時間未満	3～4時間未満	4～5時間未満	5時間以上
Q10-1. 家事	1	2	3	4	5	6	7
Q10-2. 育児	1	2	3	4	5	6	7
Q10-3. 両親や親戚の介護や手伝い	1	2	3	4	5	6	7
Q10-4. ボランティアなどの社会活動	1	2	3	4	5	6	7
Q10-5. 勉強	1	2	3	4	5	6	7

Q.11.ご夫婦の間での**理想の**仕事と家事・育児分担について、あなたの考えにもっとも近いものを１つ選んでください。

夫がフルタイムで働き、妻は専業主婦として全ての家事・育児を担う。

夫はフルタイムで働き、妻はパートタイムで働き大部分の家事・育児を担う。

夫も妻もフルタイムで働き、家事・育児は半分ずつ担う。

夫がパートタイムで働き大部分の家事・育児を担い、妻はフルタイムで働く

夫は専業主夫として全ての家事・育児を担い、妻はフルタイムで働く

トータルでの負荷（仕事と家事・育児）は平等でなくても、夫と妻がそれぞれ比較的得意なことを担当する。

Q12. あなたはあなたの配偶者（夫または妻）との間の**実際の**役割分担について、どのくらい満足していますか。もっとも当てはまるものを１つ選んでください。

Q12-1.　家事

とても満足

満足

どちらでもない

不満

とても不満

Q12-2. 育児

とても満足

満足

どちらでもない

不満

とても不満

子供はいない

Q12-3.両親や親戚の介護

とても満足

満足

どちらでもない

不満

とても不満

介護の必要がある親戚はいない

Q12-4.家計への貢献

とても満足

満足

どちらでもない

不満

とても不満

私も私の配偶者も現在働いていない

Q.13.あなたは次の状況のとき、どれくらい罪悪感を感じますか？該当しない場合は、仮にその状況にあると想定して回答してください。

Q13-1. 同僚や上司がたくさん残業しているのに、有給休暇をとった。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-2.　同僚や上司が残業する必要があるのに、家族の用事のため定時に退社した。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-3. この一週間、自分や家族のために健康な夕食を用意しなかった。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-4. 子供の行事に参加する約束をしたのに、仕事のため行けなくなった。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-5. この一ヶ月、手伝いが必要な高齢の親（または親戚）のところに行っていない。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-6. 習い事、ゲーム、服など、子供の要求に応じられるほどの所得がない。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q13-7. 家には居るものの、他の用事があるため、子供にかまってあげられない。

とても罪悪感を感じる

少し罪悪感を感じる

どちらでもない

罪悪感はあまりない

罪悪感は全然ない

Q14.あなたの仕事は次のどれにあたりますか。当てはまるものを 1 つ選んでください。

自営業主

家族従業者（自営業主の家族）

会社経営者・役員

正規の職員・従業員（民間企業）

公務員

パート

アルバイト

労働者派遣事業所の派遣社員

契約社員・嘱託

専業主婦／主夫　　Q.19 へ

学生　　Q.19 へ

定年退職者（専業主婦／主夫は除く）　　Q.19 へ

無職（専業主婦／主夫は除く）Q.19 へ

その他（具体的に　　）

Q.15. あなたは、ふだん一週間のうち何時間ぐらい仕事をしていますか（残業時間を含みます）。

週 _____時間

Q16. あなたは、ふだん一週間のうち何時間くらい残業していますか。

週_____ 時間

Q17.あなたは、ふだん一週間のうち何日くらい仕事をしていますか。

週_____日

Q18.あなたの通勤時間（家からあなたの職場まで）は片道どのくらいですか。在宅勤務（テレワーク）の場合は、0 時間とご記入ください。

_____時間_____分

Q19.次の各項目は、あなたに当てはまりますか。「ぴったり当てはまる」を「1」、「全く当てはまらない」を「5」として、当てはまる番号に〇をつけてください。

	ぴったり当てはまる	当てはまる	どちらでもない	当てはまらない	全く当てはまらない
人と競争することは楽しい。	1	2	3	4	5
夫婦は対等・平等であるべきだと思う。	1	2	3	4	5
夫は外で働き、妻は家庭を守るべきだと思う。	1	2	3	4	5

Q20.50%の確率で 10 万円が当たる「スピードくじ」があります。当たれば、賞金は今日すぐに支払われます。外れた場合、賞金はゼロです。あなたは「スピードくじ」をいくらなら買いますか。下の各値段について、それぞれ買う場合は「1」を、買わない場合は「2」に〇をつけてください。

「スピードくじ」が 10 円なら	1 買う	2 買わない
2,000 円なら	1 買う	2 買わない
4,000 円なら	1 買う	2 買わない
8,000 円なら	1 買う	2 買わない
15,000 円なら	1 買う	2 買わない
25,000 円なら	1 買う	2 買わない
35,000 円なら	1 買う	2 買わない
50,000 円なら	1 買う	2 買わない

Q21.実験で行った作業は、どれくらい男らしい/女らしいと思いますか？それぞれのタスクにつき、１０段階で評価してください。

男らしい ◀――――――――――――――▶ 女らしい

タスク A	1	2	3	4	5	6	7	8	9	10
タスク B	1	2	3	4	5	6	7	8	9	10

Q22.実験の本番第三回で行ってもらった作業の組み合わせの選択の際に考慮したことは、以下のどれですか？当てはまるものを全て選んでください。

両方のタスクをできること。

男性らしいタスクを夫が、女性らしいタスクを妻が行うこと。

夫婦の作業分担が平等であること。

すでに経験したタスクであること。

フルで４分間一つのタスクに集中できること。

自分とパートナーの成績の予想。

自分の成績とトライアル実験の平均点の予想。

パートナーの成績とトライアル実験の平均点の予想。

その他（具体的に：　　）

ご回答いただきましてありがとうございました。

Appendix 3.4: Instruction (in Japanese)

夫婦の意思決定に関する実験の進行

[参加者が会場に来たら、事前の申し込み者のリストを見て出欠をとる。参加者に ID 番号のカード（カップル ID 番号と●か▲のどちらかが書かれたカード）を渡し、夫婦リストに記録する。夫と妻が離れた席になるように、座席に誘導する（●と▲のそれぞれにつき、夫と妻の数が同数になるように）。全員が前を向いて着席し、なるべく隣の人と間を空ける。]

始まりの挨拶と連絡事項

このたびは、夫婦の意思決定に関する実験にご参加いただきありがとうございます。皆様のご協力は私どもの研究に大変重要なものです。あらためて感謝申し上げます。

実験の最後には、謝金をお渡しします。謝金の一部は成果報酬になっておりますので、ご自身とパートナーの成績によって変動します。最低でも一人当たり２０００円で平均的には４０００円になるように設定しています。作業内容はこれから詳しく説明しますので、注意して聞いてください。質問があれば、そちらに伺いますので挙手してください。

まず最初にお願いですが、個人作業が主ですので、他の参加者の方とご相談したり、お話されないようにお願い致します。また、携帯電話やその他の電気機器の電源はお切りください。

個人情報のお取り扱いですが、実験で得られた情報は研究の目的のみで利用し、外部に提供することはありません。また、実験中はお配りした ID カードの情報で皆様をご確認

し、皆様のお名前や作業の成績などが他の方に知られないようにいたします。謝金は、実験の最後に手渡しし、金額は他のご夫婦にはお教えしません。

では作業についてお話します。まずお手元の資料がそろっているか確認してください。

- ビジュアルガイドが１セット（３ページ）
- 記録シートが１枚（２ページ）、
- 問題用紙とマークシート４枚（●の人）、または３枚（▲の人）
- アンケート（６ページ）、
- 鉛筆と消しゴム、
- 使用済の迷路が４セット。使い方は後で説明します。
- ４色のチップがそれぞれ４枚

ありますでしょうか？ない方は手をあげてください。

ありますでしょうか。次に、記録シートの上にあるカップル ID 番号と●か▲が、お渡ししている **ID カードの情報**と合っているか確認をお願いします。[見本を手にとって見せる]次に、机の上の４枚のマークシートの「氏名」の欄とその下の枠内に、カップル ID と●か▲と、練習・本番１、２、３回目（●の人）１、３回（▲の人）が書かれていることをご確認ください。

では実験のメインとなる２つの作業（タスク A とタスク B）について説明します。タスク A のみ出来高に応じて報酬が払われます。お手元のビジュアルガイドをごらん下さい。

[手にとって見せる]

タスク A

報酬ありのタスク A は、マークシートを塗りつぶす作業です。お手元にマークシート４枚と試験問題のプリントが４セットあると思いますので、まず最初の練習のときに１セット、その後の本番で人によって合計１～３セット、使います。試験問題を読む必要はありません。プリントにはすでに正しい答えが星印で記されていますので、皆さんの作業はプリントの星印の番号をマークシートにえんぴつで写すだけです。

※マークシートは１２３４なので、試験問題の abcd に対応させてください。

マークシートをはじめて使う方はいらっしゃいますか？記入のよい例と悪い例がシートに書いてあります。円の中をはみ出さないように、しっかり塗ってください。[手にとって見せる]間違いは消しゴムで完全に消して、他の部分には書かないようにしてください。

タスク A では正しくマークされた問題一つにつき２０円が報酬として払われます。間違えてマークされた問題は５円ペナルティーとして引かれますので、なるべく間違えないように多くの問題を塗りつぶしてください。

タスク B

次に報酬がでないタスク B は、迷路に書かれた鉛筆の線を消しゴムで消す作業です。[手にとって見せる]。最低何枚は消してください、という決まりはありません。消し終わった枚数が少なくても、消し残しがあっても、謝金にペナルティーは発生しません。ですが皆さんが消し終わった迷路は、子供たちに配られて再利用されるので、なるべく多い枚数をきれいに消すようにして下さい。

実験の最後にお渡しする謝金の額は、参加者ご自身とパートナーのタスク A の成績によって変動します。謝金は一人ひとり別の封筒に入れてお渡しします。

実験のタイムスケジュール

ビジュアルガイドの実験のタイムスケジュールをみてください。まず作業の練習の後、本番が３回（各４分）あります。**練習**では、タスク A とタスク B をそれぞれ１分試していただきます。練習では報酬は発生しません。練習が終わったら本番の 1 回目が始まります。本番が３回終わった後、アンケートにお答えいただき、謝金をお渡しして終了です。

本場の３回目については、２回目の作業が終わったあとに説明しますので、まずは１回目と２回目について説明します。

本番一回目

本番一回目は**個人プレー**です。全員が個別にタスク A とタスク B の順番で４分ずつ（合計８分）行っていただきます。一回目のタスク A の報酬は、個人の成績により個別に計算されます。

本番２回目

２回目は**夫婦ペア**として報酬が計算されます。ID カードで**指定されたタスク**を行っていただきます。カードで●をもらっている人はタスク A、▲の人はタスク B を、それぞれ４分間行います。ご夫婦のどちらか（●をもっている方）が行ったタスク A の報酬金額を奥様と旦那様にお支払いします。

本番の全３回の報酬は一人ひとり合計を計算して、実験の最後に支払われます。

ここまででご質問のある方はいらっしゃいますか？

ご質問がなければ、さっそく練習を始めたいと思います。用意はいいですか？「よーい、はじめ！」

[練習開始]

はい、終わってください。練習が終わりましたが、作業の方法でご不明な点がある方はいらっしゃいますか？

質問がないようでしたら、本番をはじめます。ここからはタスク A に対して報酬が出ます。まず１回目は個人プレーですので、皆さんタスク A、タスク B の順番で作業をして、報酬は後で個別に支払われます。ではタスク A を４分間はじめます。「よーい、はじめ！」

次にタスク B を４分間行います。もし４分以内に全てのページが消し終わってしまった場合は、そのまま待っていてください。「よーい、はじめ！」

[第一回目開始]

はい、終わってください。マークシートをいったん回収します。

【本番第一回目のマークシートを回収し、平均点を計算。】

では次に第二回目をはじめましょう。再度申し上げますが、２回目は夫婦ペアで報酬が計算されます。ご夫婦のどちらかがタスク A でもう一方がタスク B を行います。ID カー

ドにある形が●の人はタスク A、▲の人はタスク B を始める用意をしてください。では第二回目、４分間、「よーい、はじめ！」。

[第二回開始]

はい、終わってください。マークシートをいったん回収します。

第三回に移る前に、ご夫婦で相談なしでそれぞれ、第三回で行うタスクの希望を確認します。記録シートの２ページ目、とビジュアルガイドの２ページ目をご覧下さい。

第三回も**夫婦ペア**として報酬が計算されますので、ご夫婦のタスクの組み合わせを４通り用意しています。４種類の組み合わせは、ぞれぞれ色の名前がつけられています。

では４つのタスクの組み合わせについて説明します。ビジュアルガイドの２ページ目をご覧ください。一つずつ説明します。

第３回のタスクの組み合わせ

(緑)

旦那さんがタスクＡ、**奥さん**がタスクＢを、同時にスタートしてそれぞれ４分間行っていただきます。ご自身とパートナーの報酬はそれぞれ、**旦那さんが**タスクＡで稼いだ金額と同額になります。

例えば、もし旦那さんが３０問正しくマークシートを塗りつぶしたとして、間違いがゼロだとすると、旦那さんも奥さんも３０問×２０円＝６００円の報酬が受け取れることになります。もし同じように３０問マークしていても、１つ不正確なマークがあった場合は、一人当たりの報酬が３０問×２０円―１問×５円＝５５５円となります。

(青)

奥さんがタスクＡ、**旦那さん**がタスクＢを、同時にスタートしてそれぞれ４分間行っていただきます。ご自身とパートナーの報酬はそれぞれ、**奥さんが**タスクＡで稼いだ金額と同額になります。報酬の計算方法は緑と同じです。

(赤)

旦那さんと奥さんの両方がタスクＡを、４分間行っていただきます。ご夫婦のどちらもタスクＢをしない代わりに、**一定の費用**が引かれます。費用は、事前のトライアル実験で計算された４分間のタスクＡの平均報酬額です。費用の金額は、第三回の作業の終了後に発表します。最終的なご自身とパートナーの報酬はそれぞれ、旦那さんと奥さんがタスクＡで稼いだ報酬の合計から費用を引いたものになります。

例えば、間違いがゼロのシンプルな例ですが、もし旦那さんが２５問正しくマークし、奥さんが３０問正しくマークしたとして、トライアル実験の平均が１５問の正しいマークだとします。一人当たりの報酬は、（２５＋３０－１５）×２０円＝８００円になります。

(黄色)

旦那さんと奥さんの両方がタスクＡとタスクＢを、それぞれのタスクにつき**２分間**行っていただきます。一人当たりの報酬は旦那さんと奥さんがタスクＡで稼いだ金額の合計になります。

例えば、間違いがゼロのシンプルな例ですが、もし旦那さんが１５問正しくマークし、奥さんが１０問正しくマークした場合、一人当たり報酬は、（１５＋１０）×２０円＝５００円になります。

４種類のタスクの組み合わせについてと、それぞれの報酬の計算方法について、ご質問はありますか？

４種類のタスクの組み合わせについて、質問がなければ、記録シート１ページ目の**理解度テスト**にこたえてください。［例を見せる］私かアシスタントが確認しに行きますので、皆さんが全問正解した後に次にすすみます。できた人から手を挙げてください。その間に第二回までのマークシートをお返しします。

【マークシートを返す。アシスタントの人は正解をチェックする。間違えた箇所は個別に理解してもらうまで説明する。】

皆さん理解度テストをクリアしましたので、本番３回目で行っていただく組み合わせの希望と、実際に行っていただく組み合わせの選択方法についてご説明します。

記録シートの２ページ目とビジュアルガイドの３ページ目を開いてください。まず、私が回答スタートしてください、と言ってから皆さんに一斉に、記録シートに**タスクの組み合わせの希望**の順位付けをしていただきます。好ましいものから順に１から４まで、先ほど説明した４種類の組み合わせの色をご記入ください。例えば、緑が一番好ましい場合、順位１の横に「緑」とご記入ください。実際に行っていただく作業の組み合わせは、くじ引きで決まります。私かアシスタントが皆様とパートナーの方を順番にまわりますので、順位が高い色のチップが多くなるように袋の中に入れます。（１位の色が４枚、２位の色が３枚、３位の色が２枚、４位の色が１枚）このとき、チップを入れる作

業は参加者の皆さんと私かアシスタントが一緒に行いますので、待っていてください。パートナーのチップも同じ袋に入れたら、●の人が袋の中から一枚チップを引き、選ばれた色のタスクの組み合わせが、ご夫婦が実際に行っていただくものとなります。

組み合わせの選び方についてご質問はありますか？この回答はとても重要なものですので、お時間を長めに取りたいと思いますので、慎重に回答してください。では、記録シートに回答をスタートしてください。

皆さん、順位付けはなさいましたか？では実際に行っていただくタスクを決めるためのくじ引きの袋をもってきますのでしばらくお待ちください。どの色に決まったかは、私たちから個別にお伝えします。

[くじ引きの実施。研究者とアシスタントがそれぞれの夫婦の代表者にくじ引きの袋をもっていく。●の人にチップを一枚ひいてもらう。出た色を記録シートに記入するとともに、それぞれ参加者に伝える。]

皆さん、ご自分の夫婦が行うタスクの組み合わせは決まりましたか？第三回は４分間です。黄色 (もしいらっしゃればですが)の組み合わせの方には、２分経過したところでお知らせするので、タスク B に移ってください。

[第三回開始]

はい、終わってください。

これで作業の部分は終わりです。先ほど、第三回目の後に赤の報酬の計算に使う費用の金額を発表すると言いましたが、費用は １２００円でした.

では、報酬の合計金額と、お渡しする謝金の額を計算しますので、その間に机の上のアンケートのご記入をお願いいたします。［見本を見せる］。謝金の計算に少々時間がか

かりますので、アンケートが書き終わった方は、着席のまましばらくお待ちください。計算が終わった方から謝金をお渡しします。受け取られましたら、すぐに金額を確認して、**レシート**にご記入ください。さらに**データの使用に関する同意書**に署名をお願いします。署名が終わりましたら、お帰りいただいて結構です。

本日はどうもありがとうございました。

Appendix 3.5: Complete Ranking Distributions

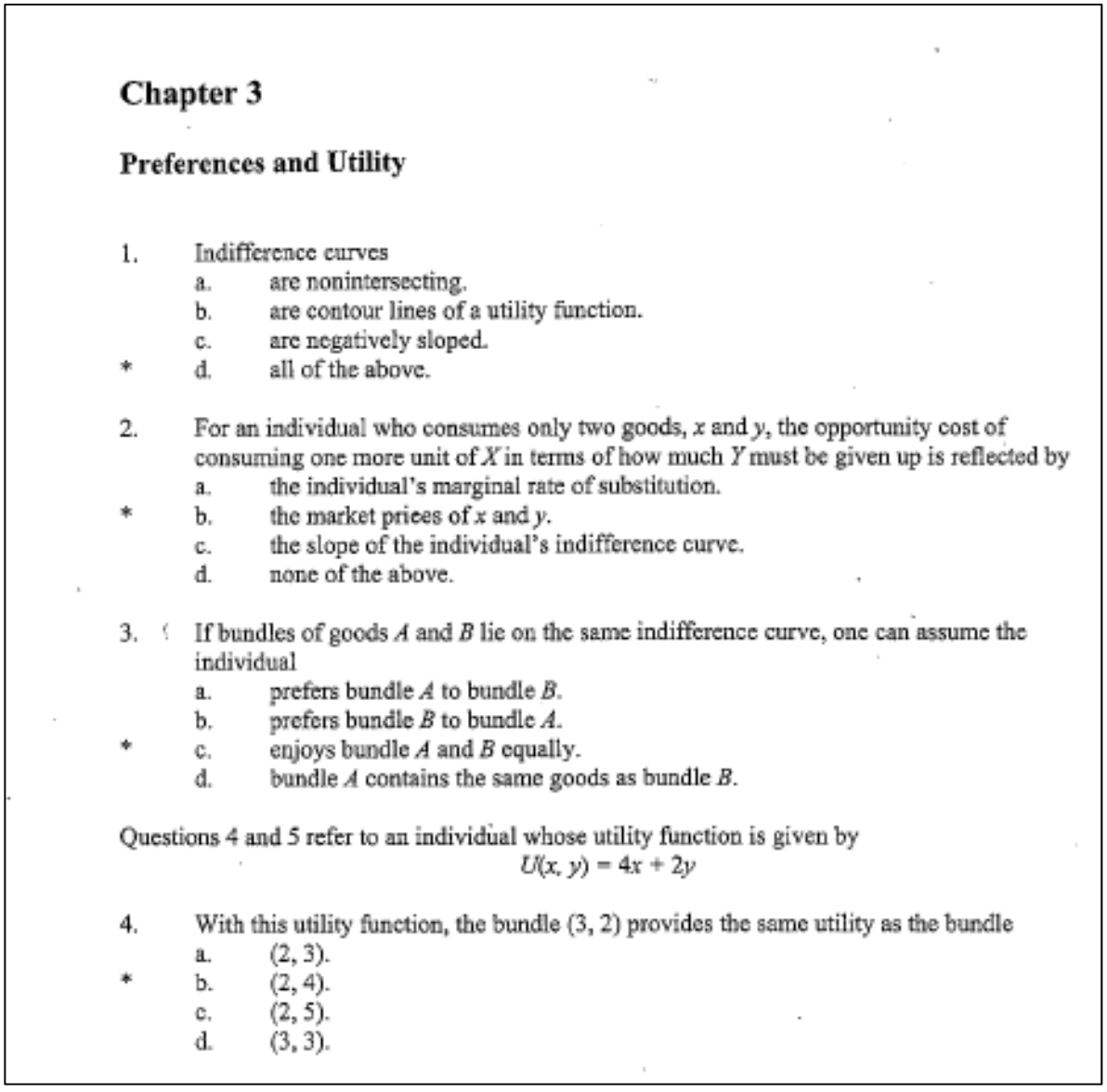
Appendix Table A 3. 1 Complete Ranking Distributions

Ranked 1st	Ranked 2nd	Ranked 3rd	Group 1 (Male, Female)	Group 2 (Male, Female)
Traditional	Reverse	Power	(1, 0)	(1, 0)
Traditional	Reverse	Mixed	(2, 0)	(0, 2)
Traditional	Power	Reverse	(2, 1)	(4, 0)
Traditional	Power	Mixed	(5, 1)	(1, 3)
Traditional	Mixed	Reverse	(0, 1)	(0, 1)
Traditional	Mixed	Power	(0, 2)	(1, 1)
Reverse	Traditional	Power	(0, 0)	(0, 1)
Reverse	Traditional	Mixed	(0, 0)	(0, 0)
Reverse	Power	Traditional	(1, 1)	(0, 0)
Reverse	Power	Mixed	(1, 1)	(0, 0)
Reverse	Mixed	Traditional	(0, 0)	(0, 0)
Reverse	Mixed	Power	(1, 0)	(0, 0)
Power	Traditional	Reverse	(3, 5)	(3, 4)
Power	Traditional	Mixed	(2, 1)	(5, 3)
Power	Reverse	Traditional	(1, 6)	(4, 1)
Power	Reverse	Mixed	(1, 0)	(1, 0)
Power	Mixed	Traditional	(3, 3)	(1, 3)
Power	Mixed	Reverse	(0, 1)	(0, 2)
Mixed	Traditional	Reverse	(0, 0)	(0, 0)
Mixed	Traditional	Power	(0, 0)	(0, 0)
Mixed	Reverse	Traditional	(1, 0)	(0, 1)

Mixed	Reverse	Power	(0, 0)	(0, 0)
Mixed	Power	Traditional	(2, 0)	(3, 2)
Mixed	Power	Reverse	(0, 3)	(1, 1)

Appendix 3.6: Images of Task A, Task B, venue, and typical session

Image of Task A



Chapter 3

Preferences and Utility

1. Indifference curves
   a. are nonintersecting.
   b. are contour lines of a utility function.
   c. are negatively sloped.
   * d. all of the above.

2. For an individual who consumes only two goods, x and y, the opportunity cost of consuming one more unit of X in terms of how much Y must be given up is reflected by
   a. the individual's marginal rate of substitution.
   * b. the market priees of x and y.
   c. the slope of the individual's indifference curve.
   d. none of the above.

3. If bundles of goods A and B lie on the same indifference curve, one can assume the individual
   a. prefers bundle A to bundle B.
   b. prefers bundle B to bundle A.
   * c. enjoys bundle A and B equally.
   d. bundle A contains the same goods as bundle B.

Questions 4 and 5 refer to an individual whose utility function is given by

$$U(x, y) = 4x + 2y$$

4. With this utility function, the bundle (3, 2) provides the same utility as the bundle
   a. (2, 3).
   * b. (2, 4).
   c. (2, 5).
   d. (3, 3).

Image of Task B

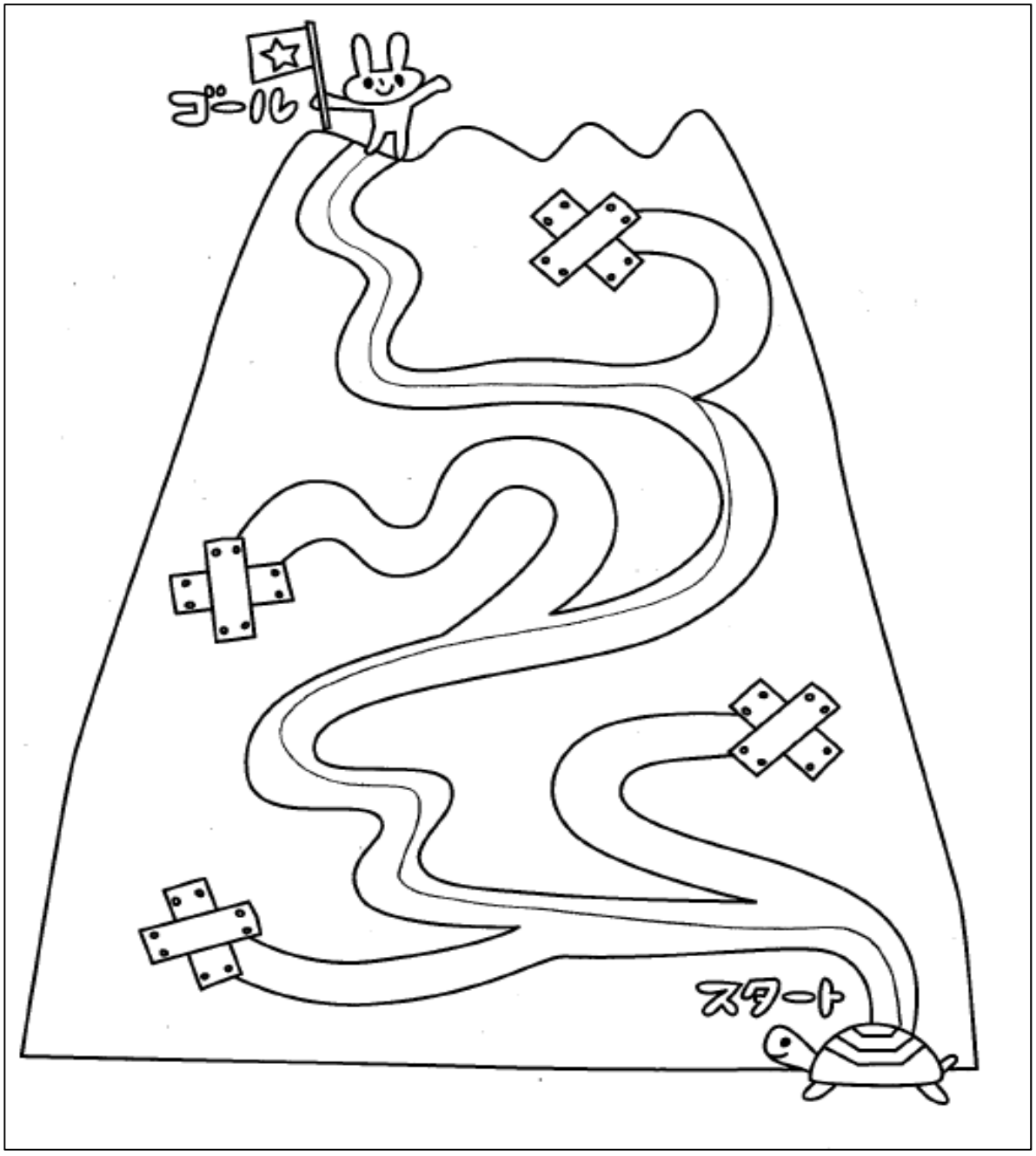


Image of a Venue

Image of a Typical Session

Appendix 3.7: Benchmark Model and Models with Control Variables

Appendix Table A 3. 2 Benchmark Rank-ordered Logit Model and Models with Control Variables such as Age Cohort, Not Having Children and whether Wife is Housewife

VARIABLES	Model 1 (Benchmark)	Model 2 (Age)	Model 3 (Children)	Model 4 (Housewife)
Reverse	-1.241***	-1.438**	-1.044**	-1.089***
	(0.390)	(0.637)	(0.434)	(0.417)
Traditional	-0.269	-0.0469	-0.200	-0.186
	(0.348)	(0.549)	(0.394)	(0.380)
Mixed	-1.009***	-1.114*	-0.604	-0.919**
	(0.363)	(0.619)	(0.399)	(0.389)
Interaction with Female Dummy				
Reverse#Female	-0.220	-0.0622	-0.191	-0.242
	(0.548)	(0.559)	(0.552)	(0.550)
Traditional#Female	-1.003*	-0.906*	-0.979*	-1.016*
	(0.522)	(0.529)	(0.525)	(0.523)
Mixed#Female	-0.851	-0.683	-0.817	-0.864
	(0.558)	(0.567)	(0.560)	(0.563)
Interaction with Group2 Dummy				
Reverse#Group2	-0.770	-0.770	-0.673	-0.842
	(0.583)	(0.609)	(0.587)	(0.589)
Traditional#Group2	-0.654	-0.707	-0.600	-0.715
	(0.520)	(0.545)	(0.525)	(0.527)
Mixed#Group2	-1.052*	-1.044*	-0.892	-1.101*
	(0.573)	(0.594)	(0.579)	(0.579)
Interaction with Female*Group2				
Reverse#Female#Group2	0.907	0.908	0.813	0.916
	(0.805)	(0.820)	(0.811)	(0.805)
Traditional#Female#Group2	1.564**	1.676**	1.564**	1.563**
	(0.746)	(0.762)	(0.750)	(0.747)
Mixed#Female#Group2	2.161***	2.140***	1.991**	2.162***
	(0.803)	(0.818)	(0.810)	(0.806)
Control Variables (as an interaction term)	None	Age Cohort	Not Having Children	Whether Wife is Housewife
Reverse#30s		-0.264		
		(0.600)		
Reverse#40s		0.279		
		(0.623)		
Reverse#50s		1.069		
		(0.859)		

Traditional#30s		-0.872		
		(0.536)		
Traditional#40s		0.0514		
		(0.556)		
Traditional#50s		0.602		
		(0.805)		
Mixed#30s		-0.476		
		(0.606)		
Mixed#40s		0.236		
		(0.615)		
Mixed#50s		0.991		
		(0.824)		
Reverse#No_Children			-0.499	
			(0.405)	
Traditional#No_Children			-0.235	
			(0.377)	
Mixed#No_Children			-0.928**	
			(0.403)	
Reverse#Housewife				-0.549
				(0.520)
Traditional#Housewife				-0.214
				(0.456)
Mixed#Housewife				-0.331
				(0.511)
Observations	408	408	408	408
Number of groups	102	102	102	102

Appendix 3.8: Model with Control for Partner Productivity

Appendix Table A 3. 3 Rank-ordered Logit Model with Task A Performance with Control for Partner Productivity

VARIABLES	Model 5' With Control for Partner Productivity
Reverse	-2.212
	(1.537)
Traditional	-3.465**
	(1.428)
Mixed	-2.757*
	(1.472)
Interactions with Task A Performance	
Reverse#Round1_correct	0.00114
	(0.0187)
Traditional#Round1_correct	0.0368**
	(0.0172)
Mixed#Round1_correct	0.00469
	(0.0178)
Interactions with Female Dummy	
Reverse#Female	-0.561
	(1.791)
Traditional#Female	4.625***
	(1.700)
Mixed#Female	-0.348
	(1.778)
Interactions with Task A Performance* Female Dummy	
Reverse#Female#Round1_correct	0.0107
	(0.0253)
Traditional#Female#Round1_correct	-0.0721***
	(0.0244)
Mixed#Female#Round1_correct	0.00753
	(0.0250)
Interaction with Partner Task A Performance	
Reverse#P_Round1_correct	0.00792
	(0.0128)
Traditional#P_Round1_correct	0.00555
	(0.0117)
Mixed#P_Round1_correct	0.0139
	(0.0128)

Observations	408
Number of groups	102

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Task A Performance is measured by the number of correctly filled answers in Round 1.

References

Ajzen, Icek. 1991. "The theory of planned behavior." *organizational behavior and human decision process* 179-211.

Akaichi, Faical, Rodolfo M. Nayga, and Jose M. Gil. 2013. "Are results from non-hypothetical choice-based conjoint analyses and non-hypothetical recorded-ranking conjoint analyses similar?" *American Journal of Agricultural Economics.*

Akerlof, George, and Rachel Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* .

Alejandro, Caparrós, Jose L.. Oviedo, and Pablo Campos. 2008. "Would you choose your preferred option? Comparing choice and recorded ranking experiments." *American Journal of Agricultural Economics.*

Aoyagi, Chie, Giovanni Ganelli, and Nour Tawk. 2016. "Minimum Wage as a Wage Policy Tool in Japan." *IMF Working Paper 16/232.*

Arentze, Theo, Aloys Borgers, Harry Timmermans, and Romano DelMistro. 2003. "Transport stated choice responses: effects of task complexity, presentation format and literacy." *Transportation Research* Part E 39 229-244.

Asai, Yukiko. 2015. "Parental leave reforms and the employment of new mothers: quasi-experimental evidence from Japan." *Labor Economics.*

Bateman, Ian J, Brett H. Day, Andrew P. Jones, and Simon Jude. 2009. "Reducing gain-loss asymmetry: A virtual reality choice experiment valuing land use change." *Journal of Environmental Economics and Management* 106-118.

Baum, M., and R. Kabst. 2013. "Conjoint implications on job preferences: The moderating role of involvement." *Int. J. Hum. Res. Manag.* 24, 1393–1417.

Beggs, S, S Cardell, and J Hausman. 1981. "Assessing the potential demand for electric cars." *Journal of Econometrics* 1-19.

Benesse Educational Research & Development Institute. 2015. *Sanzen sango no seikatsu to sapo-to nitsuiteno chosa [Survey on life and support after childbirth]*. https://berd.benesse.jp/jisedai/research/detail1.php?id=4704.

Booth, Alison, and Eiji Yamamura. 2016. "Performance in Mixed-Sex and Single-Sex Tournaments: What We Can Learn from Speedboat Races in Japan." *IZA Discussion Paper No. 10384.*

Booth, Alison, and Patrick Nolen. 2012. "Choosing to compete: How different are girls and boys?" *Journal of Economic Behavior & Organization.*

Brinton, Mary, and Eunmi Mun. 2016. "Between state and family: managers' implementation and evaluation of parental leave policies in Japan." *Socio-Economic Review.*

Cabinet Office. 2017. *"Action Plan for the Realization of Work Style Reform" issued by the Council for the Realization of Work style Reform.* March 28.

—. 2017. *Danjo kyodo sankaku hakusho heisei 29 nenban [Gender equality white paper 2017]*. http://www.gender.go.jp/about_danjo/whitepaper/h29/zentai/index.html.

—. 2004-2015. *Otto no Kyoryoku [Husband's cooperation]*. Accessed May 11, 2018. http://www8.cao.go.jp/shoushi/shoushika/data/ottonokyouryoku.html.

—. 2014. "Tomobatarakito Setaisu no suii [Changes in the number of dual-income households]." *Gender Equality Bureau Cabinet Office.* Accessed April 11, 2018.

http://www.gender.go.jp/about_danjo/whitepaper/h26/zentai/html/zuhyo/zuhyo01-02-08.html.

Caussade, S, JD Ortuza, LI Rizzi, and DA Hensher. 2005. "Assessing the influence of design dimensions on stated choice experiment estimates." *Transportation Research Part B 39* 621-640.

Chandler, Dana, and Adam Kapelner. 2013. "Breaking monotony with meaning: Motivation in crowdsourcing markets." *Journal of Economic Behavior & Organization.*

Cheng, Simon, and J. Scott Long. 2007. "Testing for IIA in the Multinominal Logit Model."

Christiadi, and Brian Cushing. 2007. "Conditional Logit, IIA, and Alternatives for Estimating Models of Interstate Migration." *Paper presented at the 46th annual meeting of the Southern Regional Science Association.*

Chung, Chanjin, Tracy Boyer, and Sungill Han. 2011. "How Many Choice Sets and Alternatives are Optimal? Consistency in Choice Experiments." *Agribusiness, Vol.27* 114-125.

Cochard, Franc¸ois, and He´le`ne Couprie. 2016. "Do spouses cooperate? An experimental investigation." *Review of Economics of the Household.*

Cochard, Franc¸ois, He´le`ne Couprie, and Astrid Hopfensitz. 2017. "What if women earned more than their spouses? An experimental investigation of work-division in couples." *Experimental Economics.*

Collins, A.T., and J.M. Rose. 2011. "Estimation of stochastic scale with best-worst data." *International Choice Modelling Conference.* Leeds, United Kingdom.

Couprie, Hélène Couprie, Miriam Beblo, Denis Beninger, François Cochard, and Astrid Hopfensitz. 2012. "Equality-Efficiency Trade-off within French and German Couples- A Comparative Experimental Study." *Thema Working Paper.*

Couprie, Hélène, Elisabeth Cudeville, and Catherine Sofer. 2017. "Efficiency versus Gender Roles and Stereotypes: An Experiment in Domestic Production." *THEMA Working Paper No.2017-10, Université de Cergy-Pontoise, France.*

Dahlberg, Matz, and Matias Eklöf. 2003. "Relaxing the IIA Assumption in Locational Choice Models: A Comparison Between Conditional Logit, Mixed Logit, and Multinomial Probit Models." *Working Paper Series, Department of Economics, Uppsala University.*

DeShazo, J.R., and German Fermo. 2002. "Designing Choice Sets for Stated Preference Moethods: The Effects of Complexity on Choice Consistency." *Journal of Environmental Economics and Management 44* 123-143.

Dowd, Bryan, William Greene, and Edward Norton. 2014. "Computation of Standard Errors." *Health Services Research* 731–750.

Etxebarria, Itziar, M. José Ortiz, Susana Conejero, and Aitziber Pascual. 2009. "Intensity of Habitual Guilt in Men and Women: Differences in Interpersonal Sensitivity and the Tendency towards Anxious-Aggressive Guilt." *The Spanish Journal of Psychology.*

Fahlen, Susanne. 2014. "The agency gap: policies, norms, and working time capabilities across welfare states." *Oxford University Press.*

Fuwa, Makiko, and Junya Tsutsui. 2010. "A Cross-National Comparison of the Perceived Fairness of the Division of Household Labor." *Kazoku syakaigaku kenkyu.*

Giergiczny, Marek, Stephane Hess, Thijs Dekker, and Phani Kumar Chintakayala. 2013. "Testing the consistency (or lack thereof) between choices in best-worst surveys)." *3rd International Choice Modeling Conference.* Sydney.

Gorges, Luise. 2015. "The power of love: A subtle driving force for unegalitarian labor division?" *Review of Economics of the Household.*

Haaas, Linda, and Philip Hwang. 2008. "The Impact of Taking Parental Leave on Fathers' Participation in Childcare and Relationships with Children: Lessons from Sweden." *Community, Work and Family.*

Haas, Linda. 1992. " Equal parenthood and social policy." *State University of New York Press.*

Haas, Linda, and Tine Rostgaard. 2011. "Father's rights to paid parental leave in the Nordic countries: consequences for the gendered division of leave." *Community, Work & Family.*

Hausman, Jerry, and Paul Ruud. 1987. "Specifying and Testing Econometric Models for Rank-Ordered Data." *Journal of Econometrics.*

Henrich, Joseph, Steven Heine, and Ara Norenzayan. 2010. "The weirdest people in the world?" *BEHAVIORAL AND BRAIN SCIENCES.*

Hensher. 2010. "Hypothetical bias, choice experiments and willingness to pay." *Transportation Research Part B: Methodological* Volume 44, Issue 6, July 2010, Pages 735-752.

Hensher, David A., John M. Rose, and William H. Greene. 2005. *Applied Choice Analysis A Primer.* New York: Cambridge University Press.

Hensher, David A., Peter R. Stopher, and Jordan J. Louviere. 2001. "An exploratory analysis of the effect of numbers of choice sets in designed choice experiments, an airline choice application." *Journal of Air Transport Management* 373-379.

Hole, Arne Risa. 2016. "Creating efficient designs for discrete choice experiments." *Nordic and Baltic Stata Users Group meeting.* September. https://www.stata.com/meeting/nordic-and-baltic16/slides/norway16_hole.pdf.

Imano, Koichiro. 2012. *Seishain shometsu jidai no jinji kaikaku [Personnel reform in the era of extinction of regular employees].* Tokyo, Japan: Nikkei publishing.

Iversen, Vegard, Cecile Jackson, Bereket Kebede, Alistair Munro, and Arjan Verschoor. 2011. "Do Spouses Realise Cooperative Gains? Experimental Evidence from Rural Uganda." *World Development.*

Japan Times. 2017. *Japan's buzzwords of 2017 cover everything from politics to poop.* Nov. 9. https://www.japantimes.co.jp/news/2017/11/09/national/japans-buzzwords-of-2017-include-poop-kanji-workbook-aufheben-hifumin-plus-other-top-candidates/#.WxAKDlIcST4.

JILPT. 2011. *Kigyo kara mita genteiseishain no katsuyo jittai [Actual utilization of limited regular employees from the perspective of companies]*. www.jil.go.jp/institute/reports/2013/documents/0158_01.pdf.

JILPT. 2015. "Yu-sufuru rodo tokei [Useful labor Statistics 2015] ." *The Japan Institute for Labour Policy and Training* http://www.jil.go.jp/kokunai/statistics/kako/2015/index.html.

—. 2017. *Yu-sufuru rodo tokei 2017 [Useful labor statistics 2017]*. http://www.jil.go.jp/kokunai/statistics/kako/2017/index.html.

Johnson, F. Reed, Emily Lansar, Marshall. Deborah, Vikram Kilambi, Axel Muhlbacher, Dean A Regier, Brian W. Bresnahan, Barbara Kanninen, and John F.P. Bridges. 2013. "Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force." *Value in health* 3-13.

Kamas, Linda, and Anne Preston. 2012. "The importance of being confident: gender, career choice, and willingness to compete." *Journal of Economic Behavior & Organization.*

Kanemoto, Yoshitsugu, and Bentley MacLeod. 1991. "The theory of contracts and labor practice in Japan and the United States." *Managerial and decision economics* 159-170.

Kato, Takao, Daiji Kawaguchi, and Hideo Owan. 2013. "Dynamics of the Gender Gap in the Workplace: An econometric case study of a large Japanese firm (Revised)." *RIETI Working Paper.*

Kebede, Bereket, Marcela Tarazona, Alistair Munro, and Arjan Verschoor. 2014. "Intra-household efficiency: An experimental study from Ethiopia." *Journal of African Economies.*

Kobayashi, Mayu, Miki Kobayashi, Tsunao Okumura, and Emiko Usuki. 2016. "Sharing housework between husbands an wives: how to improve marital satisfaction for working wives in Japan." *Journal of Labor Policy.*

Kubo, Keiko. 2016. "The Evaluation of Wives about Husbands' Participationin Housework and Childcare in Dual-Career Couples." *The Japan Society of Home Economics.*

Kume, Koichi, Fumio Otake, and Kotaro Tsuru. 2014. "Tayoka suru Seiki Hiseiki rodosha no shugyo kodo to ishiki RIETI Web anke-to chosa no gaiyo [Employment behavior and consciousness of diversifying regular and irregular workers - Overview of RIETI Web survey]." *RIETI Policy Discussion Paper Seriese* 14-P-003.

Kuroda, Sachiko. 2010. "Do Japanese Work Shorter Hours than before? Measuring trends in market work and leisure using 1976–2006 Japanese time-use survey." *Journal of Japanese and International Economies.*

Kuroda, Sachiko, and Isamu Yamamoto. 2013. "Do Peers Affect Determination of Work Hours? Evidence Based on Unique Employee Data from Global Japanese Firms in Europe." *Journal of Labor Research.*

Kuroda, Sachiko, and Isamu Yamamoto. 2014. "Kibo rodo jikan no kokusai hikaku; Kaso sitsumon ni yoru rodokyokyu danseichi no keisoku [International comparison of desired labor hours: Measurement of labor supply elasticity by virtual questions]." *Japan Economic Research.*

Kuroda, Sachiko, and Isamu Yamamoto. 2013. "Wage Premiums for Firms' Work-Life Balance Practice: Evidence from Japanese matched firm-worker data." *RIETI Discussion Paper Seriese* 13-J-004.

Kurosawa, Masako. 2002. "Chuto saiyo shijyo no macchingu: manzokudo, chingin, kunren, seisansei [Satisfaction, Wages, Training and Productivity of New Hires]." (The Japanese Journal of Labour Studies).

Lagarde, M, and D Blaauw. 2009. "A review of the application and contribution of discrete choice experiments to inform human resources policy interventions." *Human Resources for Health.*

Leah, Ruppanner, and Matt Huffman. 2014. "Blurred Boundaries: Gender and Work-Family Interference in Cross-National Context." *Work and Occupations.*

Lee, Kipyong. 2008. "The Effect of Husband's Participation in Household Work on Wife's Marital Satisfaction: Introducing the Concept of the Expectation Su$ciency." *Japanese Journal of Family Sociology.*

List, J.A., S. Sadoff, and M, Wagner. 2011. "So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design." *Experimental Economics.*

Martínez, Pilar, J María Carrasco, Gonzalo Aza, Angeles Blanco, and Isabel Espinar. 2011. "Family Gender Role and Guilt in Spanish Dual-earner Families." *Sex Roles* 813–826.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 3722-3759.

Meyerhoff, Jürgen, Malte Oehlmann, and Priska Weller. 2015. "The Influence of Design Dimensions on Stated Choices in an Environmental Context." *Environ Resource Econ* 61:385–407.

MHLW. 2016. *Heisei 28 nen ban karoshi tou boshi taisaku hakusho [2016 White papers to prevent death from overwork]*. http://www.mhlw.go.jp/wp/hakusyo/karoushi/16/.

—. 2017. *Heisei 29 nen karoshi boshi taisaku hakusho (honbun) [2017 White papers to prevent death from overwork]*. http://www.mhlw.go.jp/wp/hakusyo/karoushi/17/index.html.

—. 2010. *Shusseini kansuru tokei [Birth statistics]*. http://www.mhlw.go.jp/toukei/saikin/hw/jinkou/tokusyu/syussyo06/index.html.

—. 2014. *Tayona seishain no fukyu kakudai no tameno yushikisha kaigi hokokusho [Report of discussion meeting for experts to disseminate and expand "various kinds of regular employment"]*. http://www.mhlw.go.jp/stf/shingi/0000052525.html.

—. 2017. *Tenkin ni kansuru koyokanri no hinto to shuho [Hints and Methods of Employment Management Regarding Relocation]*. https://www.kantei.go.jp/jp/singi/sousei/meeting/kpi_kenshouteam/h29-11-29-shiryou1.pdf.

Ministry of Economy Trade and Industry, Japan. 2006. "Survey for Basic Policy Committee of Industrial Structure Council."

Ministry of Internal Affairs and Communications. 2017. "the Basic Resident Registration System for Foreign Residents."

Morikawa, Masayuki. 2017. "Are Part-time Employees Underpaid or Overpaid? Productivity–wage gaps in Japan." *RIETI Discussion Paper Series 17-E-077*.

Morikawa, Masayuki. 2010. "Insecurity of Employment and Work-Life Balance: From the viewpoint of compensating wage differentials." *RIETI Discussion Paper Series 10-E-052*.

Munro, Alistair. 2018. "Intra-household Experiments: a survey ." *Journal of Economic Surveys* .

Nakamura, Mayumi, and Mito Akiyoshi. 2015. "What Determines the Perception of Fairness Regarding Household Division of Labor between Spouses?" *PLoS ONEW 10(7).*

Nishioka, Hachiro, and Masakazu Yamauchi. 2017. "Otto no kaji ya ikuji no suikou hindo ha takamatta noka? [Has the frequency of husbands' housework and child care increase?]." *Journal of PopulationProblems.*

North, Scott. 2009. "Negotiating What's ‘Natural’: Persistent Domestic Gender Inequality in Japan." *Social Science Japan Journal.*

Oehlmann, Malte, Jürgen Meyerhoff, Petr Mariel, and Priska Weller. 2017. "Uncovering context-induced status quo effects in choice experiments." *Journal of Environmental Economics and Management Volume 81* 59-73.

Okuda, Shoko. 2018. *Josei katsuyaku ni honro sareru hitobito [People who are at the mercy of female success]*. Tokyo: Kobunsha.

Rabiee, Fatemeh. 2004. "Focus-group interview and data analysis." *Proceedings of the Nutrition Society* 655–660.

RIETI. 2009-2011. *Hakenrodosha no seikatsu to kyushoku kodo ni kansuru anke-to chosa [A questionnaire survey on life and job seeking behavior of dispatched workers]*. https://www.rieti.go.jp/jp/projects/research_activity/temporary-worker/result.pdf.

—. 2015. *Heisei 26 nendo seishain hiseishain no tayou na hatarakikata to ishiki ni kansuru Web chosa [FY2014 Web survey on various work style and consciousness of regular employees and non-regular employees]*. https://www.rieti.go.jp/jp/projects/research_activity/workers2014/.

—. 2012. *Jinteki shihon keisei to wa-ku raihu baransu ni kansuru kigyo jugyoin chosa [Survey of Companies and Employees on Human Capital Development and Work-Life Balance]*. https://www.rieti.go.jp/jp/projects/research_activity/wlb2011/.

—. 2013. *Tayoka suru seiki hiseiki rodosha no shugyo kodo to ishiki ni kansuru Web chosa [Survey on employment behavior and awareness of diversifying regular and irregular workers]*. https://www.rieti.go.jp/jp/projects/research_activity/workers2012/.

Rose, John. 2014. "Interpreting discrete choice models based on Best-Worst data: A matter of framing." *Transportation Research Board 93rd Annual Meeting.* Washington DC.

Samuelson, William, and Richard Zeckhauser. 1988. "Status quo bias in decision making." *Journal of Risk and Uncertainty.*

Sato, Hiroki, Hiroyuki Fujimura, and Atsushi Yashiro. 2007. *Atarashii jinji romu kanri 3rd edition [Contemporary Human Resource Management]*. Tokyo, Japan: Yuhikaku Publishing.

Scarpa, R., S. Notaro, J. Louviere, and R. Raffelli. 2011. "Exploring scale effects of best/worst rank ordered choice data to estimate benefits of tourism in alpine grazing commons." *American Journal of Agricultural Economics.*

Schröder, Melanie, Norma Schmitt, Britta Heynemann, and Claudia Brünn. 2013. "Income Taxation and Labor Supply: An Experiment on Couple's Work Effort." *Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Tax Evasion and Labor Supply.*

Shibuya, Shozo. 2004. *"Joshi ga yomeru" to omoshiroi [It will be interesting if you can "read your boss"]*. Tokyo: Shinkosha.

Shimazu, Akihito, Kazumi Kubota, Arnold Bakker, Eva Demerouti, Kyoko Shimada, and Norito Kawakami. 2013. "Work-to-family Conflict and Family-to-work Conflict among Japanese

Dual-earner Couples with Preschool Children: A Spillover-Crossover Perspective." *Journal of Occupational Health.*

Steinberg, Chad, and Masato Nakane. 2012. "Can Women Save Japan?" *IMF Working Paper 12/248.*

Swait, Joffre, and Wiktor Adamowicz. 2001. "Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice." *Organizational Behavior and Human Decision Processes* 141-167.

Takahashi, Mieko, Saori Kamano, Tomoko Matsuda, Setsuko Onode, and Kyoko Yoshizumi. 2014. "Worklife balance in Japan: new policies, old practices." *Oxford University Press.*

Toda, Akihito. 2015. "Gentei seishain no jittai kigyo kibo betsu ni okeru chingin manzokudo no chigai [Actual condition of limited regular employees-Difference in wages and satisfaction levels by company size]." *Monthly Journal of The Japan Institute for Labour Policy and Training.*

Tsuru, Kotaro. 2014. *Gentei seishain kara nihonjin no hatarakikata wo kaeru [Change Japanese way of working from "limited regular employee"]*. December 11. Accessed April 11, 2018. https://www.rieti.go.jp/jp/papers/contribution/tsuru/26.html.

—. 2016. *Jinzai kakusei keizai [Human resources arousal economy]*. Tokyo: Nikkei Business Publications.

Tsuru, Kotaro, Koichi Kume, Fumio Otake, and Hiroko Okudaira. 2013. "Hiseikirodosha kara mita hosho chingin, fuanteikoyo, anmokutekina seishain kosoku to chingin puremiamu no bunseki [Compensating wage for non-regular workers, analysis of wage premium for unstable employment]." *RIETI Discussion Paper* 13-J-003.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the workplace, investment in human capital, and gender." *The Quarterly Journal of Economics* 457-507.

Yamada, Hisashi. 2017. *Doitsurodo doitsuchingin no shogeki [Shock of equal wage for equal work]*. Tokyo, Japan: Nikkei publishing.

Yamada, Ken, and Daiji Kawaguchi. 2015. "The changing and unchanged nature of inequality." *The Journal of Economic Inequality.*

Yamaguchi, Kazuo. 2006. "Fuufu kankei manzokudo to wa-ku raifu baransu [Marital satisfaction and work-life balance]." *RIETI Discussion Paper Series 06-J-054.*

Yasui, kengo, Shinpei Sano, Koichi Kume, and Kotaro Tsuru. 2016. "Mugentei seishain to gentei seishain no chingin kakusa [Wage gap between limited and unlimited regular workers]." *RIETI Discussion Paper Seriese* 16-J-061.

Yoo, W, and I Oh. 2017. "Effect of work-family balance policy on job selection and social sustainability: The Case of South Korea." *Sustainability* 9(5), 849.

Tables

Table 2. 1 Studies on Design Dimensions of the CE

	Meyerhoff et al. (2015)	Oehlmann et al. (2017)	Chung et al. (2011)	Hensher et al. (2001)	DeShazo and Fermo (2002)	Caussade, et al. (2005)
Application	Environmental valuation	Environmental valuation	Marketing	Transportation	Environmental valuation	Transportation
Number of choice sets	6,12,18,24	6,12,18,24	1-20	4,8,12,24,32	-	6,9,12,15
Number of alternatives	3-5	3-5	3-12	11	6-9, 2-7	3-6
Number of attributes	4-7	4-7	3	4,9	4-7,9	3-6
Number of levels	2-4	2-4	11	2-4	3	2-4
Experimental design	C-efficient	C-efficient	Orthogonal	fractional factorial	Random	D-efficient
Area of interest	Germany	Germany	Korea	Australia, New Zealand	Guatemala, Costa Rica	Chile
Main focus of interest	Rate of dropout	Status quo effect	Error variance	Fatigue effects	Error variance	Error variance
Number of observations	90,354	23,118	10,000	2,616	3,900	8,020

Table 2. 2 Choice Set Example

	Job A	Job B	Job C
Annual Wage	6 million yen	7 million yen	8 million yen
Overtime	0 hours/month	15-45 hours/month	0 hours/month
Employment Security	Medium	High	Medium
Transfer Possibility	Some	None	None
Relocation Possibility	None	None	Some
Best Job	○	○	○
Worst Job	○	○	○

Table 2. 3 Levels and Ranges of the Attributes

Attribute Name	Level
Wage1 (Baseline)	3 million yen
Wage 2	3.5 million yen
Wage 3	4 million yen
Wage 4	6 million yen
Wage 5	7 million yen
Wage 6	8 million yen
Overtime 1 (Baseline)	0 hours per month
Overtime 2	0-15 hours per month
Overtime 3	15-45 hours per month
Overtime 4	45- hours per month
Relocation 1 (Baseline)	Zero possibility of relocation
Relocation 2	Some possibility of relocation
Transfer 1 (Baseline)	Zero possibility of transfer
Transfer 2	Some possibility of transfer
Security 1 (Baseline)	Less secure than regular contracts
Security 2	As secure as regular contracts

Table 2. 4 Mean Comparisons of Selected Socio Demographic Variables

	Panel	Representative
Gender (Male=2, Female=1)	1.49	1.50
Age***	41.67	40.46
Marriage*** (Not Married=2, Married=1)	1.58	1.63
Education*** (Middle School=1, High school=2, 2-year College=3, 4-year College=4, Graduate School=5)	3.35	2.74
Having Children*** (Having Children=1, Not Having Children=2)	1.58	1.24
Annual Income (in million yen)*** (less than 1 =1, 1-2 =2, 2-3=3, 3-4=4, 4-5=5, 5-6=6, 6-8=7, 8-10 =8, more than 10=9)	4.26	4.32

Asterisks indicate the panel mean and the representative mean are significantly different. ***p<0.01, **p<0.05, *p<0.1

Table 2. 5 Simplest Model with No Interaction (Main Effects)

VARIABLES	Benchmark Model
Wage_cont	1.350***
	(0.0670)
Wage_sq	-0.0773***
	(0.00510)
Overtime2	-0.0665**
	(0.0323)
Overtime3	-0.514***
	(0.0351)
Overtime4	-1.327***
	(0.0500)
Security2	0.324***
	(0.0313)
Transfer2	-0.905***
	(0.0325)
Relocation2	-0.307***
	(0.0221)
Observations	25,104
Number of groups	8,368

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2. 6 WTP estimations (in million yen) by Different Amount of Annual Wage (Incl. Overtime Pay)

WTP	3 million yen	3.5million yen	4 million yen	6 million yen	7 million yen	8 million yen
Overtime2	0.08**	0.08**	0.09**	0.16**	0.25**	0.59*
Overtime3	0.58***	0.64***	0.70***	1.22***	1.92***	4.54***
Overtime4	1.50***	1.64***	1.81***	3.14***	4.96***	11.72***
Security 2	-0.37***	-0.40***	-0.44***	-0.77***	-1.21***	-2.86***
Transfer 2	1.02***	1.12***	1.24***	2.14***	3.38***	7.99***
Relocation2	0.35***	0.38***	0.42***	0.73***	1.15***	2.71***

***p<0.01 , **p<0.05, *p<0.1

Table 2. 7 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Presence of Children

WTP	Base	Female	No_children	Female*No_children
Overtime2	0.07	0.20***	0.13*	0.03
Overtime3	0.27***	0.77***	0.52***	0.75***
Overtime4	0.89***	1.83***	1.34***	1.74***
Security 2	0.00	-0.42***	-0.47***	-0.51***
Transfer 2	0.71***	1.39***	0.85***	1.35***
Relocation2	0.06	0.41***	0.31***	0.37***

***p<0.01 , **p<0.05, *p<0.1

Note: The base case is males with children.

Table 2. 8 Difference in WTP Estimations (in million yen) by Gender

WTP	With children	No children
Overtime2	-0.14	0.09
Overtime3	-0.49***	-0.23**
Overtime4	-0.94***	-0.40***
Security 2	0.42***	0.04
Transfer 2	-0.67***	-0.50***
Relocation2	-0.35***	-0.05

***p<0.01 , **p<0.05, *p<0.1

Notes: Differences in WTP Estimates = Male WTP Estimates – Female WTP Estimates.

Table 2. 9 Difference in WTP Estimations (in million yen) by Presence of Children

	Male	Female
Overtime2	-0.06	0.17
Overtime3	-0.25**	0.02
Overtime4	-0.45***	0.09
Security 2	0.47***	0.08
Transfer 2	-0.13	0.04
Relocation2	-0.25***	0.05

***p<0.01 , **p<0.05, *p<0.1

Notes: Differences in WTP Estimates = With_children WTP Estimates – No_children WTP Estimates.

Table 2. 10 WTP Estimations (in million yen) for High_Income sample

WTP	Base	Female	No_children	Female*No_children
Overtime2	0.06	0.22**	0.14*	0.05
Overtime3	0.26***	0.73***	0.50***	0.72***
Overtime4	0.85***	1.62***	1.25***	1.57***
Security 2	0.04	-0.10	-0.33***	-0.26***
Transfer 2	0.68***	1.17***	0.75***	1.17***
Relocation2	0.05	0.32***	0.27***	0.29***

***p<0.01 , **p<0.05, *p<0.1

Note: The base case is males with children. For simplicity, the wage level of the hypothetical job is fixed at annual 3 million yen, and the age of the subjects is in their 30s.

Table 2. 11 WTP Estimations (in million yen) for not High_Income sample

WTP	Base	Female	No_children	Female*No_children
Overtime2	0.04	0.19**	0.11	0.02
Overtime3	0.31***	0.77***	0.55***	0.77***
Overtime4	1.10***	1.86***	1.49***	1.82***
Security 2	-0.34***	-0.49***	-0.71***	-0.64***
Transfer 2	0.93***	1.42***	1.00***	1.43***
Relocation2	0.16*	0.43***	0.37***	0.40***

***p<0.01 , **p<0.05, *p<0.1

Note: The base case is males with children. For simplicity, the wage level of the hypothetical job is fixed at annual 3 million yen, and the age of the subjects is in their 30s.

Table 2. 12 Conditional Logit Models for IIA Test

VARIABLES	(1) Full Model	(2) Restricted Model (without best choice)	(3) Restricted Model (without worst choice)	(4) Restricted Model (with only best choice)
Wage_cont	1.362***	0.979***	1.282***	1.526***
	(0.0586)	(0.0948)	(0.0874)	(0.0754)
Wage_sq	-0.0774***	-0.0535***	-0.0730***	-0.0861***
	(0.00477)	(0.00754)	(0.00719)	(0.00625)
Overtime2	-0.0417	0.108	-0.130	-0.123
	(0.137)	(0.216)	(0.207)	(0.180)
Overtime3	-0.521***	-0.402**	-0.481**	-0.534***
	(0.130)	(0.190)	(0.207)	(0.179)
Overtime4	-1.409***	-1.074***	-1.330***	-1.655***
	(0.147)	(0.213)	(0.242)	(0.209)
Security2	0.0220	-0.133	0.250*	0.150
	(0.0919)	(0.140)	(0.140)	(0.124)
Transfer2	-0.376***	-0.169	-0.436***	-0.550***
	(0.0866)	(0.127)	(0.135)	(0.120)
Relocation2	0.000350	0.0308	-0.0156	-0.0342
	(0.0870)	(0.140)	(0.123)	(0.112)
Interactions with Female Dummy				
Overtime2	-0.123	-0.227	0.0536	-0.0468
	(0.102)	(0.153)	(0.156)	(0.138)
Overtime3	-0.444***	-0.342**	-0.398***	-0.553***
	(0.0951)	(0.138)	(0.151)	(0.133)
Overtime4	-0.845***	-0.692***	-0.871***	-1.096***
	(0.103)	(0.149)	(0.167)	(0.148)
Security2	0.381***	0.205*	0.384***	0.465***
	(0.0666)	(0.105)	(0.0998)	(0.0882)

Transfer2	-0.604***	-0.451***	-0.656***	-0.717***
	(0.0652)	(0.0960)	(0.103)	(0.0897)
Relocation2	-0.317***	-0.288***	-0.302***	-0.356***
	(0.0649)	(0.106)	(0.0910)	(0.0831)
Interactions with No-Children Dummy				
Overtime2	-0.0570	0.0117	-0.166	-0.102
	(0.0940)	(0.142)	(0.146)	(0.126)
Overtime3	-0.220**	0.0933	-0.563***	-0.512***
	(0.0874)	(0.126)	(0.141)	(0.122)
Overtime4	-0.403***	-0.138	-0.693***	-0.694***
	(0.0937)	(0.136)	(0.151)	(0.133)
Security2	0.418***	0.356***	0.294***	0.426***
	(0.0615)	(0.0968)	(0.0936)	(0.0813)
Transfer2	-0.120**	-0.120	-0.106	-0.122
	(0.0585)	(0.0872)	(0.0911)	(0.0797)
Relocation2	-0.225***	-0.248***	-0.134	-0.229***
	(0.0591)	(0.0945)	(0.0844)	(0.0764)
Interactions with Female*No-Children				
Overtime2	0.208	0.227	0.126	0.167
	(0.132)	(0.202)	(0.200)	(0.177)
Overtime3	0.234*	-0.0143	0.452**	0.472***
	(0.124)	(0.182)	(0.196)	(0.173)
Overtime4	0.485***	0.209	0.800***	0.823***
	(0.135)	(0.195)	(0.220)	(0.194)
Security2	-0.342***	-0.281**	-0.241*	-0.359***
	(0.0862)	(0.135)	(0.130)	(0.115)
Transfer2	0.156*	0.0990	0.281**	0.224*
	(0.0853)	(0.126)	(0.133)	(0.117)
Relocation2	0.268***	0.238*	0.219*	0.294***
	(0.0843)	(0.137)	(0.119)	(0.108)
Interactions with age cohorts				

30s # Overtime2	-0.0170	-0.215	0.189	0.134
	(0.122)	(0.195)	(0.180)	(0.158)
40s # Overtime2	0.0641	-0.152	0.286	0.230
	(0.131)	(0.206)	(0.194)	(0.170)
50s # Overtime2	0.0332	-0.221	0.260	0.198
	(0.130)	(0.206)	(0.191)	(0.168)
30s # Overtime3	0.276**	-0.00324	0.515***	0.456***
	(0.117)	(0.173)	(0.183)	(0.159)
40s # Overtime3	0.392***	0.177	0.534***	0.516***
	(0.124)	(0.184)	(0.194)	(0.170)
50s # Overtime3	0.284**	0.154	0.315	0.325*
	(0.123)	(0.182)	(0.194)	(0.169)
30s # Overtime4	0.610***	0.376*	0.797***	0.819***
	(0.133)	(0.193)	(0.223)	(0.190)
40s # Overtime4	0.681***	0.500**	0.796***	0.840***
	(0.141)	(0.203)	(0.235)	(0.202)
50s # Overtime4	0.589***	0.293	0.843***	0.854***
	(0.140)	(0.203)	(0.233)	(0.199)
30s # Security2	-0.0222	0.157	-0.227*	-0.150
	(0.0790)	(0.120)	(0.120)	(0.107)
40s # Security2	0.0213	0.274**	-0.328**	-0.184
	(0.0841)	(0.128)	(0.129)	(0.114)
50s # Security2	-0.103	0.289**	-0.532***	-0.389***
	(0.0840)	(0.128)	(0.128)	(0.114)
30s # Transfer2	-0.265***	-0.342***	-0.0622	-0.200*
	(0.0781)	(0.114)	(0.121)	(0.108)
40s # Transfer2	-0.217***	-0.265**	-0.112	-0.194*

	(0.0831)	(0.121)	(0.130)	(0.115)
50s # Transfer2	-0.269***	-0.455***	0.0104	-0.140
	(0.0831)	(0.122)	(0.129)	(0.115)
30s # Relocation2	-0.0558	-0.0423	-0.0553	-0.0262
	(0.0779)	(0.125)	(0.110)	(0.0997)
40s # Relocation2	-0.162*	-0.228*	-0.0373	-0.0789
	(0.0829)	(0.133)	(0.117)	(0.106)
50s # Relocation2	-0.138*	-0.175	-0.0478	-0.0573
	(0.0826)	(0.133)	(0.117)	(0.106)
Observations	41,840	16,736	16,736	25,104

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2. 13 Hausman and McFadden (HM) IIA test

Comparing Model(2)-Model(1)	
	Chi2(44)=275.57
	Prb>Chi2=0.0000
Comparing Model(3)-Model(1)	
	Chi2(44)=275.57
	Prb>Chi2=0.0000
Comparing Model(4)-Model(1)	
	Chi2(44)= 190.28
	Prb>Chi2=0.0000

Table 2. 14 WTP Estimations (in million yen) from the Restricted Model (4), for Annual Wage 3 million Yen in their 30s

WTP	Base	Female	No_children	Female* No_children
Overtime2	-0.01	0.04	0.09	-0.03
Overtime3	0.08	0.62***	0.58***	0.66***
Overtime4	0.83***	1.91***	1.52***	1.79***
Security 2	0.00	-0.46***	-0.42***	-0.53***
Transfer 2	0.74***	1.45***	0.86***	1.35***
Relocation2	0.06	0.41***	0.29***	0.35***

***p<0.01, **p<0.05, *p<0.1

Note: The base case is males with children.

Table 2. 15 Differences in WTP Estimations (in million yen) by Models

WTP differences	Base	Female	No_children	Female* No_children
Overtime2	-0.08	-0.17	-0.04	-0.06
Overtime3	-0.19	-0.14	0.07	-0.09
Overtime4	-0.06	0.08	0.18	0.05
Security 2	0.00	-0.04	0.04	-0.02
Transfer 2	0.03	0.07	0.02	0.01
Relocation2	0.00	0.00	-0.03	-0.02

Notes: Differences in WTP Estimates = WTP Estimates from the full model (1) – WTP Estimates from the restricted model (4).

Table 2. 16 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Marital Status

WTP	Base	Married	Female	Female*Married
Overtime2	0.12	0.10	0.07	0.14*
Overtime3	0.55***	0.29***	0.73***	0.76***
Overtime4	1.32***	1.00***	1.75***	1.78***
Security 2	-0.48***	-0.08	-0.50***	-0.43***
Transfer 2	0.85***	0.74***	1.19***	1.46***
Relocation2	0.29***	0.14**	0.41***	0.37***

***p<0.01 , **p<0.05, *p<0.1

Note: The base case is males who are not married.

Table 2. 17 Male WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Number of Children

WTP	0 children	1 child	2 children	3 children	4 children
Overtime2	0.12	-0.08	0.11	0.12	20.25***
Overtime3	0.52***	0.35***	0.14	0.32	1.62***
Overtime4	1.34***	0.75***	0.82***	0.88**	62.59***
Security 2	-0.45***	0.00	-0.08	-0.13	19.23***
Transfer 2	0.82***	0.70***	0.77***	0.50	0.48***
Relocation2	0.30***	0.09	0.05	0.12	20.34***

***p<0.01, **p<0.05, *p<0.1

Table 2. 18 Female WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Number of Children

WTP	0 children	1 child	2 children	3 children	4 children
Overtime2	0.05	0.18	0.15	0.60**	-0.02
Overtime3	0.80***	0.78***	0.61***	0.89***	0.75
Overtime4	1.81***	1.91***	1.53***	1.96***	2.13***
Security 2	-0.58***	-0.47***	-0.21**	-0.25	-1.12**
Transfer 2	1.33***	1.28***	1.38***	2.06***	4.44***
Relocation2	0.41***	0.35***	0.33***	0.69**	1.13*

***p<0.01, **p<0.05, *p<0.1

Table 2. 19 Summary of Total Number of Children in the Sample Panel in Their 30s

Total Number of Children	Male	Female	Total
0	135	124	259
1	21	56	77
2	29	45	74
3	6	11	17
4	0	2	2
Total	191	238	429

Table 2. 20 Male WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Age Group of the Youngest Children

WTP	0 children	Infant_ toddler	Preschool	Elementary_ school	Teenager
Overtime2	0.15*	0.04	-0.22	0.31	0.03
Overtime3	0.56***	0.37**	-0.16	0.45***	0.38**
Overtime4	1.40***	0.90***	0.09	1.15***	0.99***
Security 2	-0.48***	-0.19	-0.24	0.07	0.21
Transfer 2	0.82***	0.75***	0.93***	0.84***	0.35**
Relocation2	0.34***	-0.01	0.04	0.15	0.21*

***p<0.01 , **p<0.05, *p<0.1

Note: Infant_toddler: 0-3 years old, Preschool: 4-6 years old, Elementary_school: 7-12 years old, Teenager: 13-18 years old

Table 2. 21 Female WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Age Group of the Youngest Children

WTP	0 children	Infant_toddler	Preschool	Elementary_ school	Teenager
Overtime2	0.05	0.10	0.32**	0.25	0.41
Overtime3	0.79***	0.68***	0.64***	0.67***	1.17***
Overtime4	1.79***	1.77***	1.79***	1.63***	2.04***
Security 2	-0.52***	-0.26**	-0.39***	-0.35**	-0.66***
Transfer 2	1.32***	1.44***	1.35***	1.41***	1.43***
Relocation2	0.39***	0.43***	0.25*	0.43***	0.35*

***p<0.01 , **p<0.05, *p<0.1

Note: Infant_toddler: 0-3 years old, Preschool: 4-6 years old, Elementary_school: 7-12 years old, Teenager: 13-18 years old

Table 2. 22 Summary of Age Group of the Youngest Children in the Sample Panel in Their 30s

Age of the Youngest Children	Male	Female	Total
No children	135	124	259
Infant_toddler	31	63	94
Preschool	13	29	42
Elementary_school	10	20	30
Teenager	2	2	4
Total	191	238	429

Note: Infant_toddler: 0-3 years old, Preschool: 4-6 years old, Elementary_school: 7-12 years old, Teenager: 13-18 years old

Table 2. 23 Guilt Questions; "How much guilty would you feel by acting as each of the statement below?"

Guilt Questions	Situations faced by working parents that potentially generate guilt
q1	I took paid leave when my managers and colleagues are working a lot of overtime.
q2	I left the office on time for a family event when my managers and colleagues are working.
q3	I did not prepare healthy dinner for me and for my family for the entire week.
q4	Because I was working I missed my child(ren)'s event which I had promised to go.
q5	I did not see my elderly parents or other relatives who needs care for the last one month.
q6	Not earning enough income to satisfy the demands of my child (extra-academic activities, clothes, games…)
q7	Not being able to spend time with my child when we are at home because I must perform tasks that do not concern the family.

Note: Respondents rated each situation on the five-point scale ranging from 1 (very guilty) to 5 (not at all guilty)

Table 2. 24 Mean Scores for Guilt Questions

	Male		Female	
Guilt Questions	Mean	Std.Dev.	Mean	Std.Dev.
q1	2.98	1.16	2.66	1.17
q2	3.13	1.10	2.78	1.12
q3	2.95	1.01	2.28	1.07
q4	2.46	1.10	2.09	1.11
q5	2.75	1.04	2.36	1.04
q6	2.64	1.04	2.45	1.00
q7	2.72	0.97	2.40	1.02

Note: Five-point scale ranging from 1 (very guilty) to 5 (not at all guilty)

Table 2. 25 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Guilt Level

	Not_guilty	Guilty
WTP	all guilt questions=5	all guilt questions=1
Overtime2	0.03	0.13
Overtime3	0.50***	0.66***
Overtime4	0.89***	1.82***
Security 2	-0.57***	-0.34***
Transfer 2	0.64***	1.35***
Relocation2	0.22***	0.37***

***p<0.01, **p<0.05, *p<0.1

Table 2. 26 Differences in WTP Estimations (in million yen) by Guilt Level

WTP Differences	(Not_guilty-Guilty)
Overtime2	-0.10
Overtime3	-0.16
Overtime4	-0.93***
Security 2	-0.24
Transfer 2	-0.71***
Relocation2	-0.15

***p<0.01, **p<0.05, *p<0.1

Note: Differences in WTP Estimates = WTP Estimates for Not_Guilty – WTP Estimates for Guilty.

Table 2. 27 WTP Estimations (in million yen) for Annual Wage 3 Million yen in Their 30s, by Guilt Level and Gender

WTP	Not_guilty	Guilty	Not_guilty_Female	Guilty_Female
	all guilt questions=5	all guilt questions=1	all guilt questions=5	all guilt questions=1
Overtime2	0.10	0.03	-0.04	0.20
Overtime3	0.49***	0.35**	0.68***	0.74***
Overtime4	1.20***	0.97***	0.80***	2.23***
Security 2	-0.47***	-0.30**	-0.82***	-0.30***
Transfer 2	0.61***	0.86***	0.93***	1.52***
Relocation2	0.20*	0.20**	0.32**	0.44***

***p<0.01, **p<0.05, *p<0.1

Table 2. 28 WTP Estimations (in million yen) for Overtime 4, Females with Annual Wage 3 Million Yen in Their 30s

Value for guilt questions	Guilty situations	WTP for Overtime4
all guilt questions=5	Not_guilty	0.80***
q1=1, other guilt =5	Guilty_Taking paid-leave	0.83*
q2=1, other guilt =5	Guilty_Leaving the office early	0.73*
q3=1, other guilt =5	Guilty_Not cooking	1.03***
q4=1, other guilt =5	Guilty_Canceling a Kids' Event	1.73***
q5=1, other guilt =5	Guilty_Not caring elderly parents	1.52***
q6=1, other guilt =5	Guilty_Not earning enough for kids	0.38
q7=1, other guilt =5	Guilty_Not caring kids while at home	0.80

***p<0.01, **p<0.05, *p<0.1

Table 3. 1 Grouping and Number of Participants

Conditions	Male	Female	Total
Circle	Group 1 (26)	Group 2 (25)	(51)
Triangle	Group 2 (25)	Group 1 (26)	(51)
Total	(51)	(51)	(102)

Table 3. 2 Task Pair Names

Task pair names used in the experiment	Task pair names in the analysis
Green (緑 in Japanese)	Traditional
Blue (青 in Japanese)	Reverse
Red (赤 in Japanese)	Power
Yellow (黄 in Japanese)	Mixed

Table 3. 3 Correctly Chosen Efficient Pair

	Male	Female
Traditional	8	6
Reverse	2	1
Power	15	14
Mixed	0	0
Correct Total	25	21
Wrong Total	26	30

Table 3. 4 Choice between Reverse and Traditional

	Group1	Group2
Reverse	35%	24%
Traditional	65%	76%

Pearson chi2(1) = 1.3835 Pr = 0.239

Table 3. 5 Choice between (1) Power and Reverse, (2) Power and Traditional, (3) Mixed and Reverse, and (4) Mixed and Traditional

(1)	Male	Female
Power	84%	84%
Reverse	16%	16%
Pearson chi2(1) = 0.0000 Pr = 1.000		

(2)	Male	Female
Power	65%	73%
Traditional	35%	27%
Pearson chi2(1) = 0.7286 Pr = 0.393		

(3)	Male	Female
Mixed	49%	57%
Reverse	51%	43%
Pearson chi2(1) = 0.6296 Pr = 0.427		

(4)	Male	Female
Mixed	29%	33%
Traditional	71%	67%
Pearson chi2(1) = 0.1821 Pr = 0.670		

Table 3. 6 Choice between Mixed and Power

	Male	Female
Mixed	22%	27%
Power	78%	73%
Pearson chi2(1) = 0.4769 Pr = 0.490		

Table 3. 7 Marginal Utility from Selecting Each Task Pair, Instead of the Benchmark Power Task Pair

	Group1		Group2	
	Male	Female	Male	Female
Reverse	-1.24***	-1.46***	-2.01***	-1.32***
Traditional	-0.27	-1.27***	-0.92***	-0.36
Balance	-1.01***	-1.86***	-2.06***	-0.75**

Table 3. 8 Rank-ordered Logit Model with Task A Performance

VARIABLES	Model 5
Reverse	-1.749
	(1.336)
Traditional	-3.076**
	(1.242)
Mixed	-1.971
	(1.276)
Interactions with Task A Performance	
Reverse#Round1_correct	0.00205
	(0.0186)
Traditional#Round1_correct	0.0364**
	(0.0172)
Mixed#Round1_correct	0.00661
	(0.0176)
Interactions with Female Dummy	
Reverse#Female	-0.539
	(1.789)
Traditional#Female	4.592***
	(1.698)
Mixed#Female	-0.237
	(1.771)
Interactions with Task A Performance* Female Dummy	
Reverse#Female#Round1_correct	0.0110
	(0.0252)
Traditional#Female#Round1_correct	-0.0710***
	(0.0244)
Mixed#Female#Round1_correct	0.00702
	(0.0250)
Observations	408
Number of groups	102

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Task A Performance is measured by the number of correctly filled answers in Round 1.

Table 3. 9 Determinants of Task A Performance in Round 3

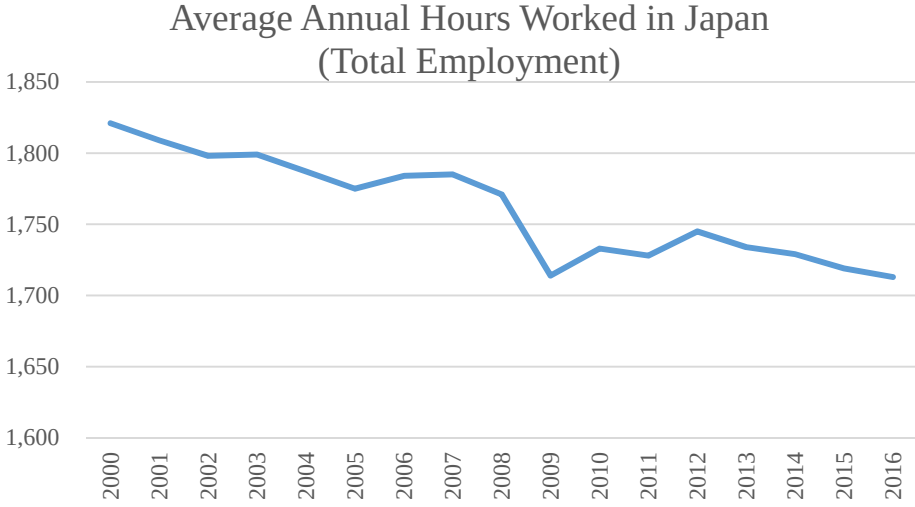
VARIABLES	
Round1_correct	0.238***
	(0.0271)
Group 2	-1.161
	(1.031)
Female	-1.427
	(1.290)
Group 2#Female	1.402
	(1.549)
Task Pair Performed in Round 3	
Traditional	0.682
	(1.126)
Reverse	1.278
	(1.315)
Mixed	1.652
	(1.309)
Task Pair Preference	
P_Traditional	-0.783
	(1.076)
P_Reverse	-1.178
	(2.401)
P_Mixed	-1.049
	(1.157)
Satisfaction Effect	
Traditional#P_Traditional	-0.0728
	(1.441)
Reverse#P_Reverse	-0.0118
	(2.421)
Mixed#P_Mixed	-0.475
	(3.204)
Constant	4.684**
	(2.248)
Observations	71
R-squared	0.641

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the Task A performance in Round 3 (measured by number of correctly filled answers per minute). Round1_correct is the Task A performance in Round 1. Satisfaction Effect captures the effect of having been able to play the task pair which the subject had expressed the highest preference for.

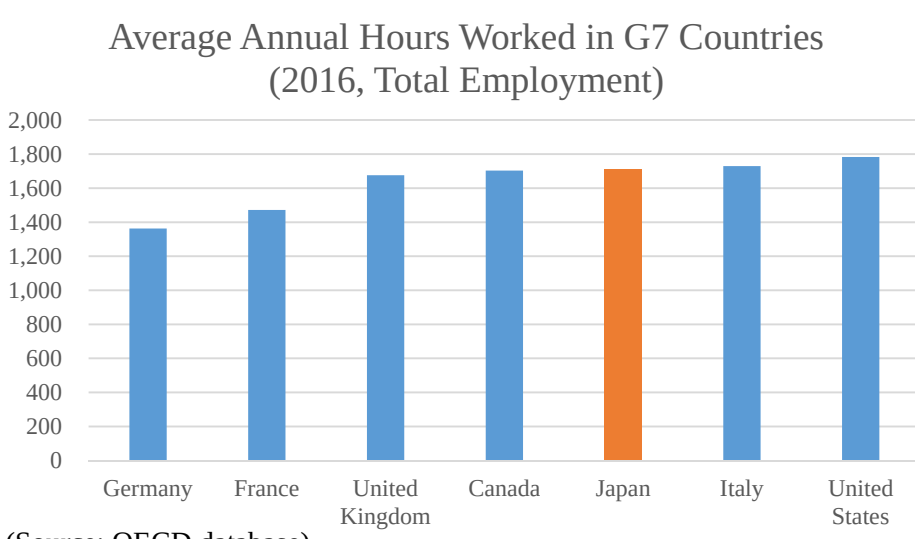
Figures

Figure 1. 1 Average Annual Hours Worked in Japan



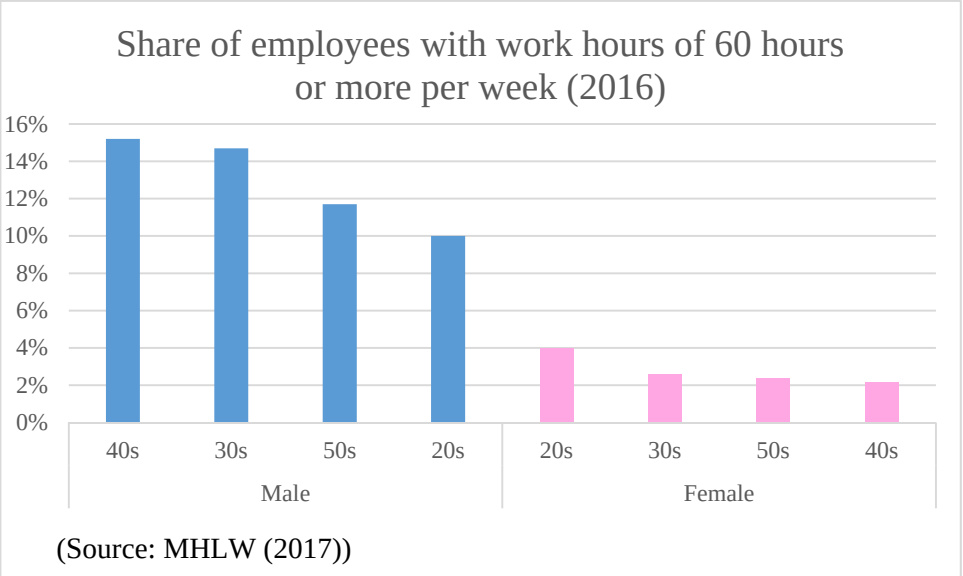
(Source: OECD database)

Figure 1. 2 Average Annual Hours Worked in G7 Countries



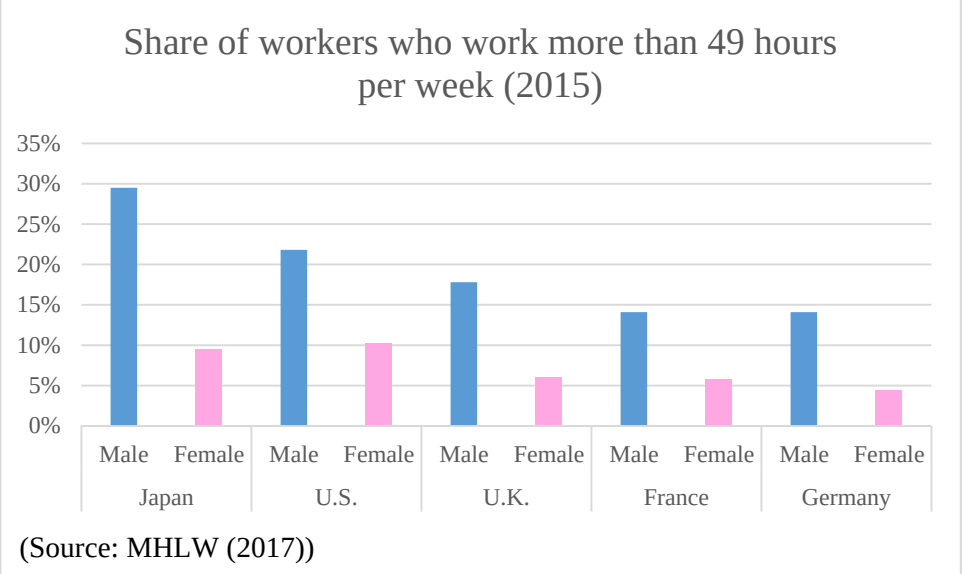
(Source: OECD database)

Figure 1. 3 Share of Employees with Work Hours of 60 Hours or More Per Week in Japan



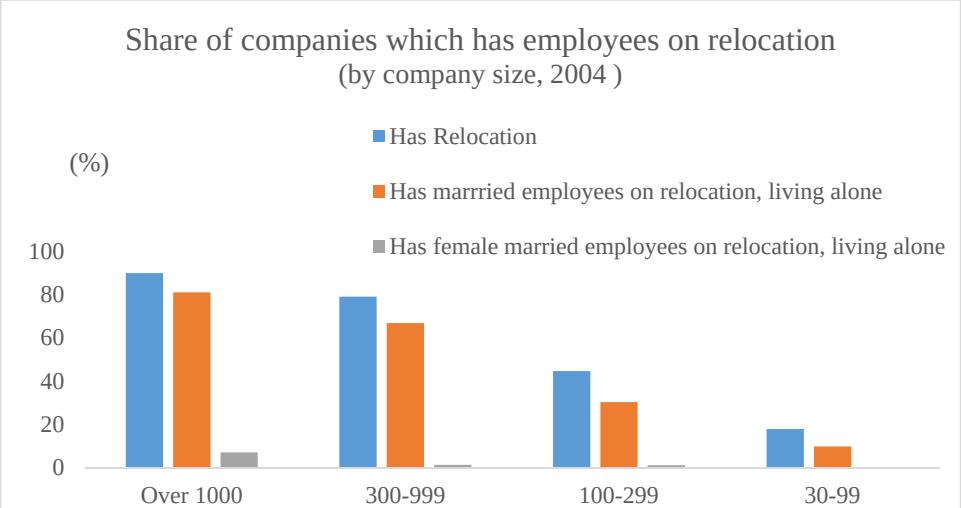
(Source: MHLW (2017))

Figure 1. 4 Share of Workers Who Work More Than 49 Hours per Week by Country



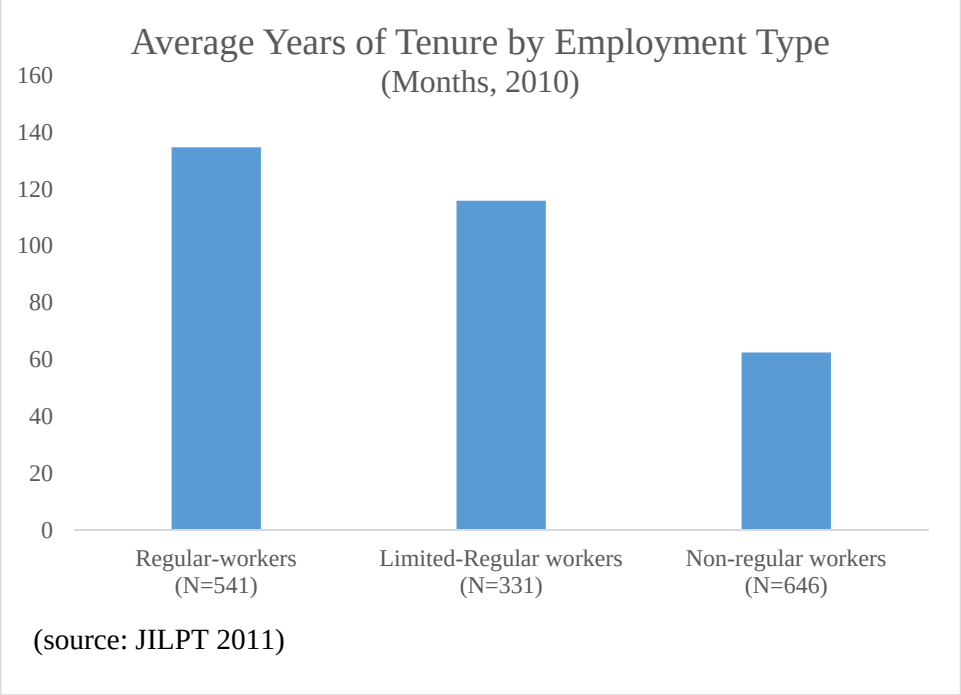
(Source: MHLW (2017))

Figure 1. 5 Share of Companies Which Has Employees on Relocation



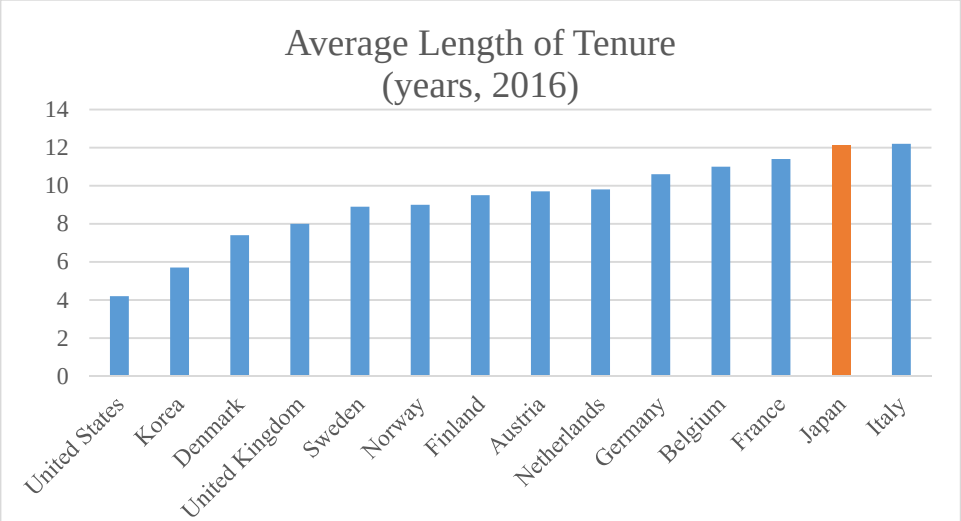
(source: JULPT (2015))

Figure 1. 6 Average Years of Tenure by Employment Type



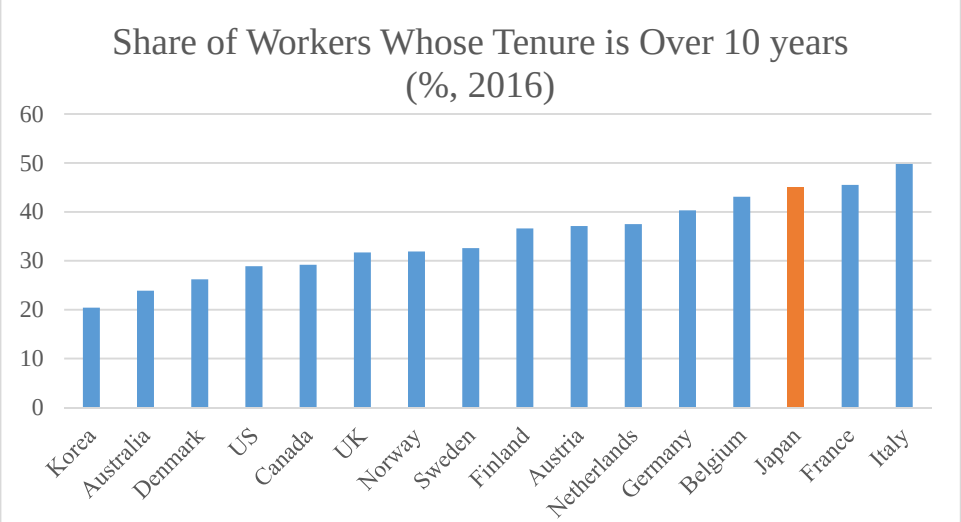
(source: JILPT 2011)

Figure 1. 7 Average Length of Tenure



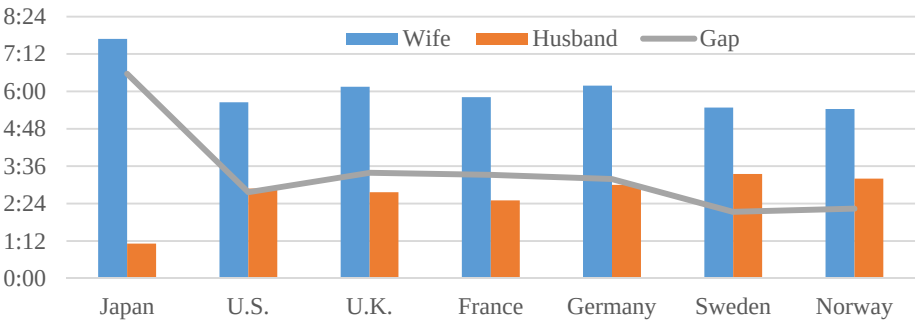
(source: JILPT (2017))

Figure 1. 8 Share of Workers Whose Tenure is over 10 Years



(source: JILPT (2017))

Figure 1. 9 Daily Hours Spent for Housework and Child Care



*2004 for European countries, 2015 for the U.S., and 2011 for Japan.
(source: Cabinet Office)

Figure 2. 1 Distribution of Age in the Sample Panel

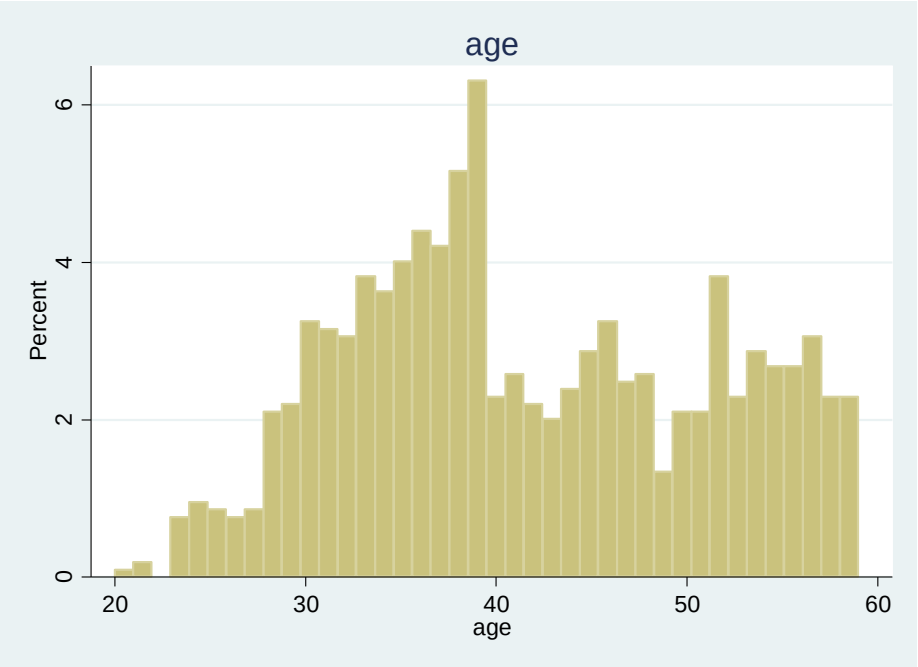


Figure 2. 2 Distribution of the Highest Education Received in the Sample Panel

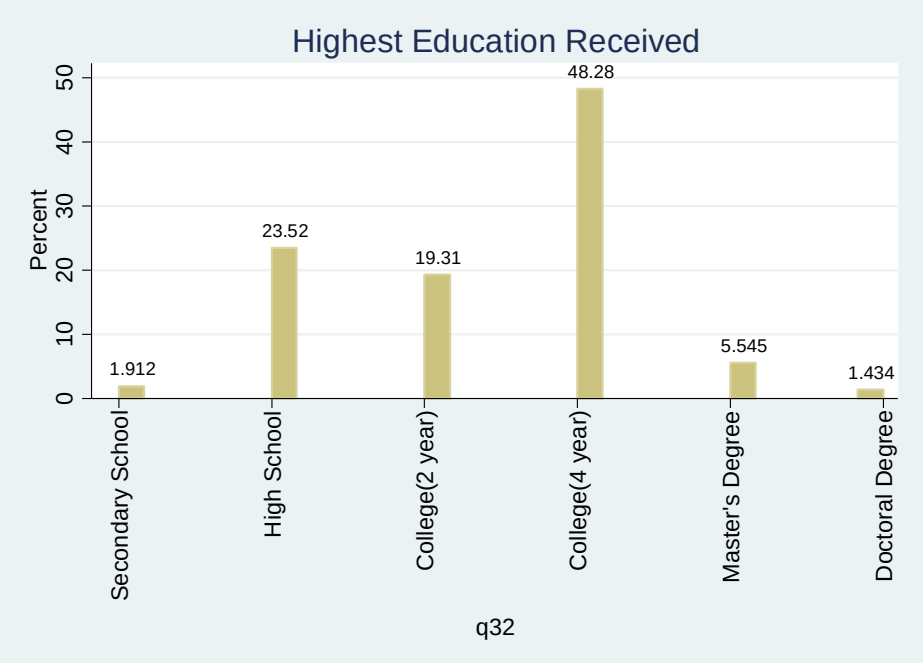


Figure 2. 3 Prefecture Distribution in the Sample Panel

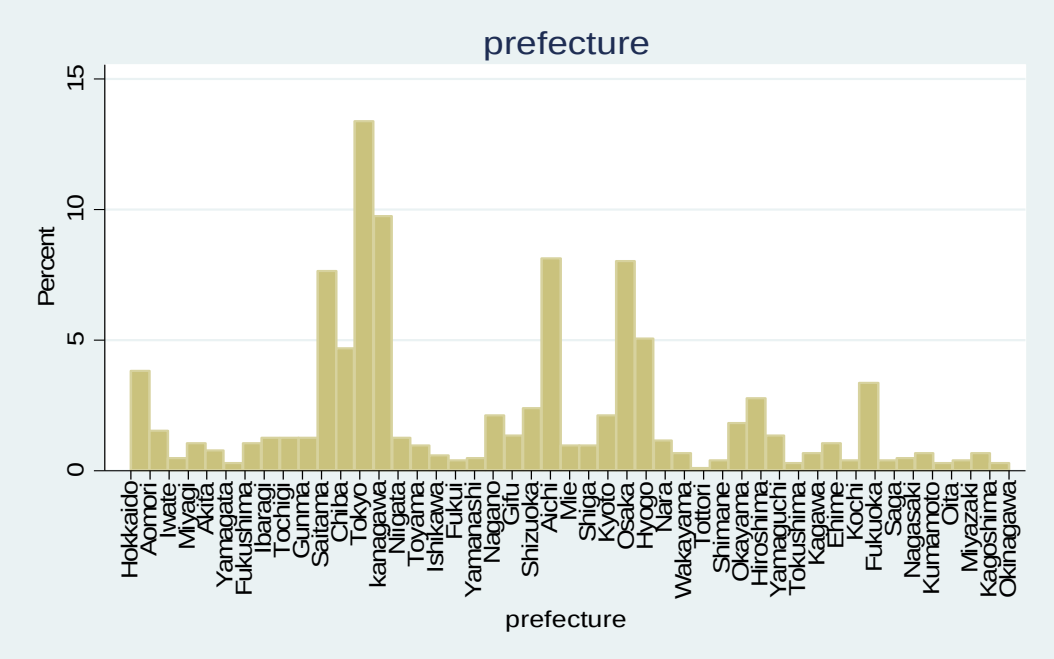


Figure 2. 4 Distribution of Annual Income in the Sample Panel

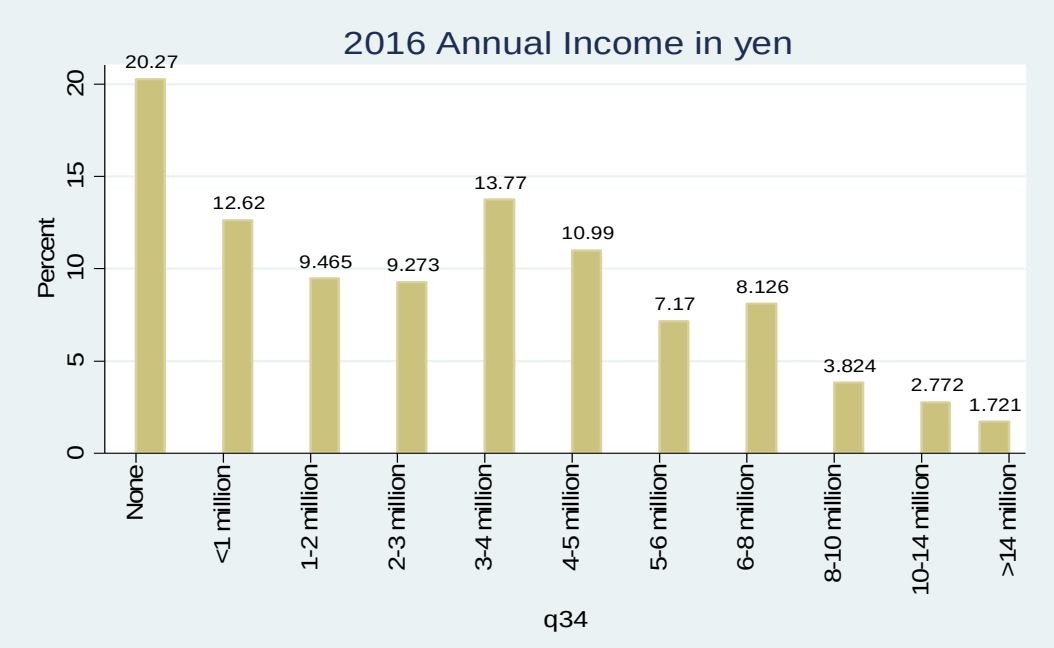


Figure 2. 5 Distribution of Current Job in the Sample Panel

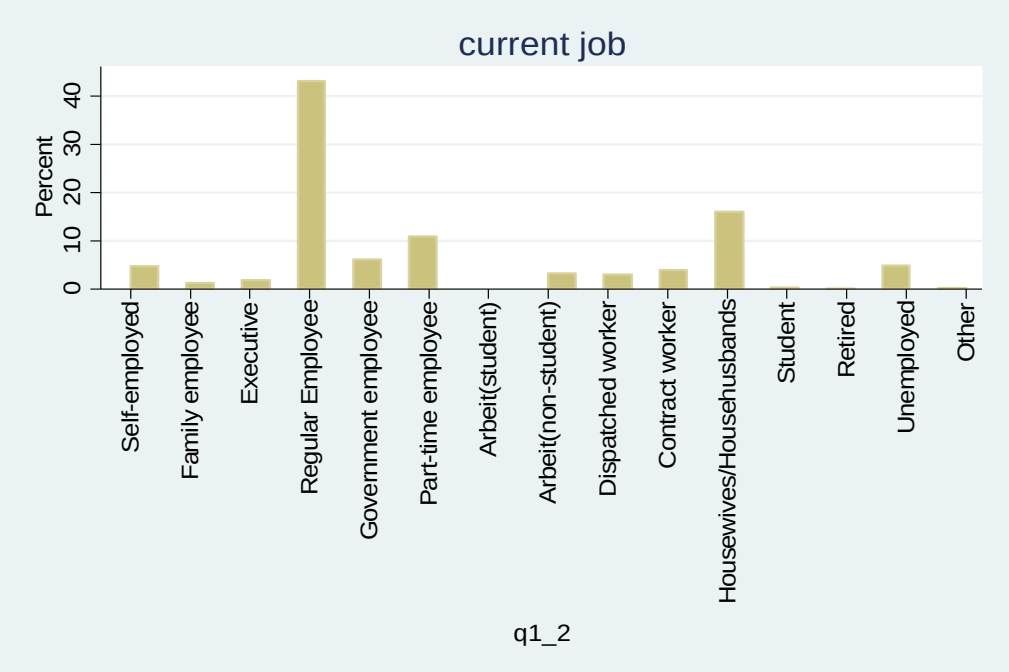


Figure 2. 6 Distribution of Size of Employers in the Sample Panel

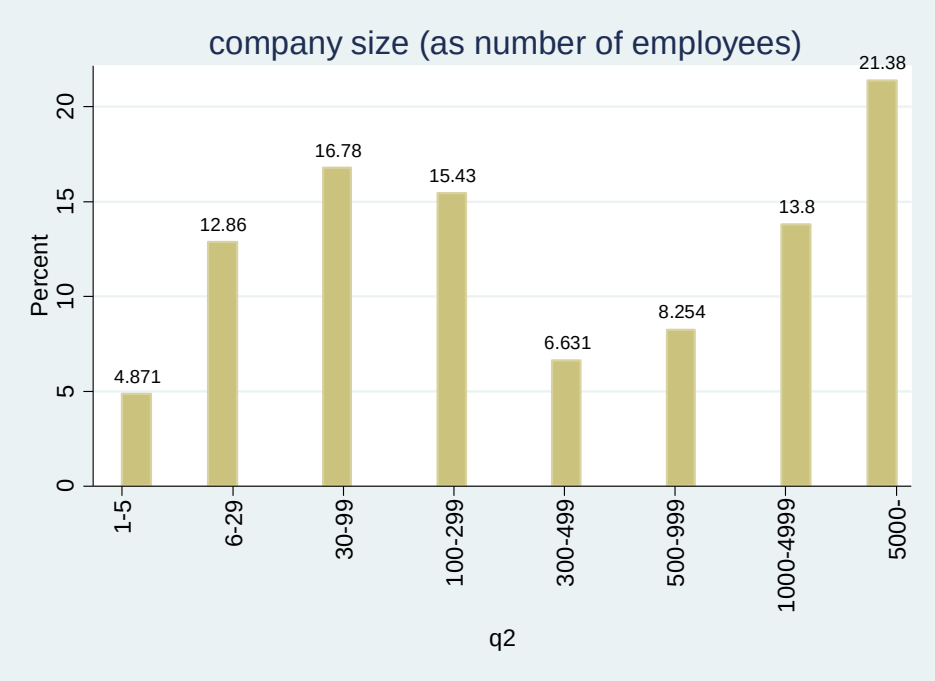


Figure 2. 7 Distribution of Industry of Current Employer in the Sample Panel

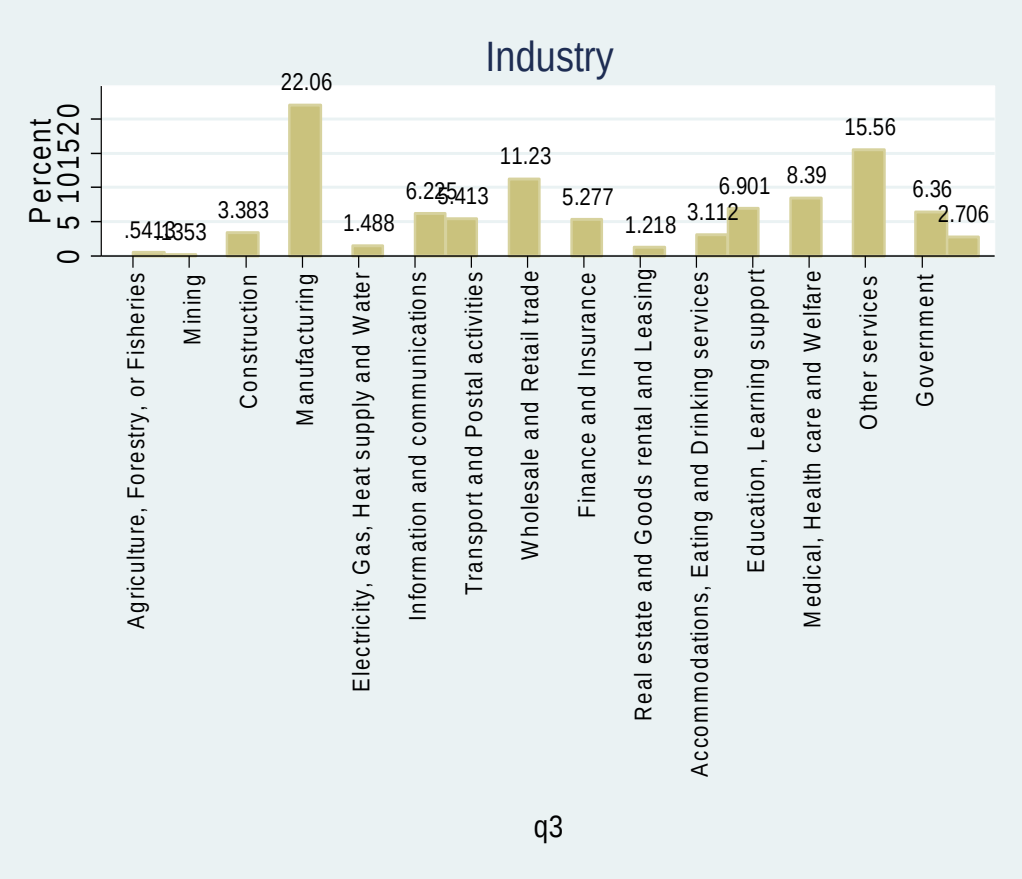


Figure 3. 1 Time Order of the Experiment

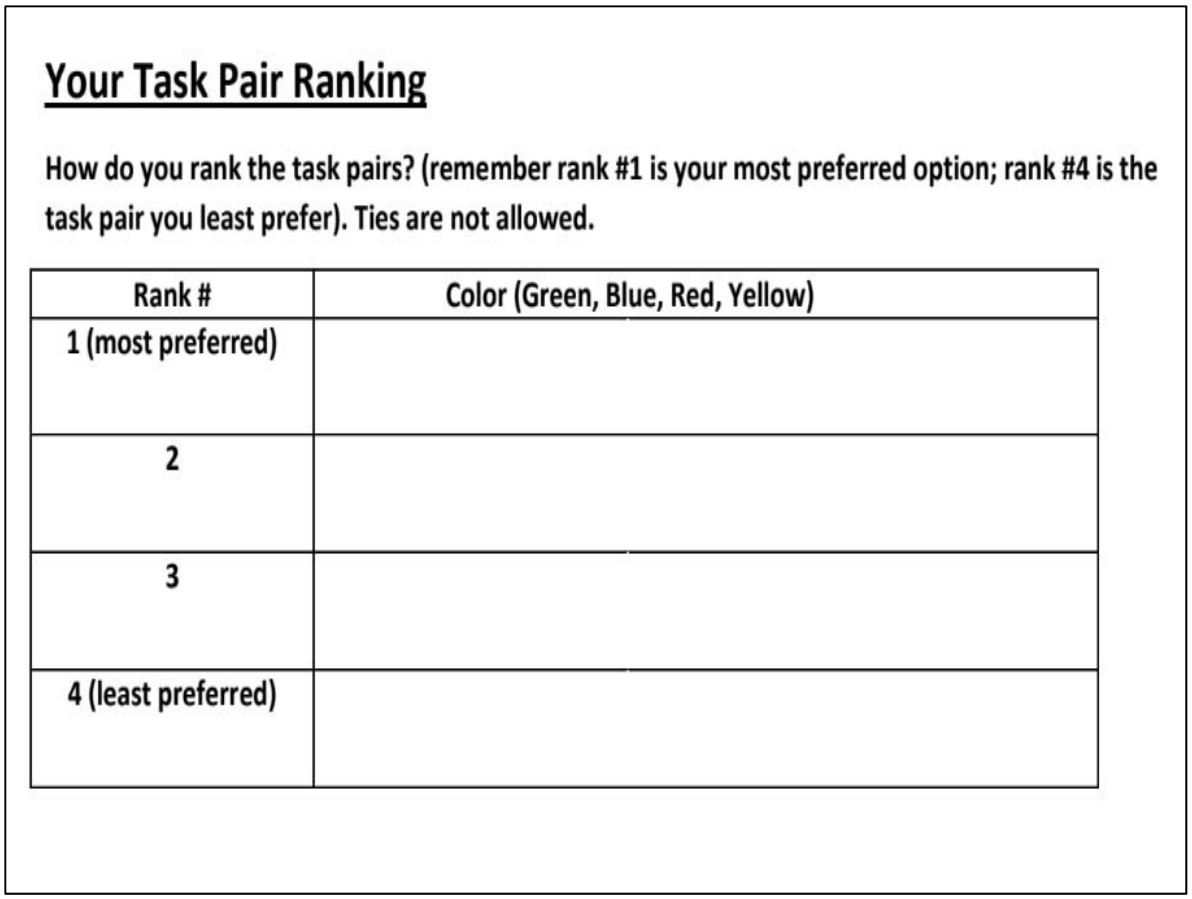
<table>
<tr><td colspan="5">Instruction</td></tr>
<tr><td>Practice Session</td><td>Everyone</td><td colspan="3">Task A (1 minute)
Task B (1 minute)</td></tr>
<tr><td>Round 1
(Single Tasks)</td><td>Everyone</td><td colspan="3">Task A (4 minute)
Task B (4 minute)</td></tr>
<tr><td>Round 2
(Couple Tasks: Assigned Roles)</td><td>●</td><td>Task A
(4 minutes)</td><td>▲</td><td>Task B
(4 minutes)</td></tr>
<tr><td colspan="5">Decision making about Division of Work in Round 3</td></tr>
<tr><td>Round 3
(Couple Tasks: Selected Roles)</td><td colspan="4">(Green) ,
(Blue) ,
(Red) ,
or (Yellow)
(4 minutes)</td></tr>
<tr><td colspan="5">Questionnaire</td></tr>
<tr><td colspan="5">Payment to each participant</td></tr>
</table>

Figure 3. 2 Description of Task Pairs for Round 3 in the Visual Guide (Translated in English)

Task Pairs for Round 3

Pair	Role for Husband	Role for Wife	Pay for each of you will be calculated as
(Green)	A (4 minutes)	B (4 minutes)	The earnings the husband generated
(Blue)	B (4 minutes)	A (4 minutes)	The earnings the wife generated.
(Red)	A (4 minutes)	A (4 minutes)	The sum of earnings you and your partner generated minus the cost.
(Yellow)	A&B (2 minutes*2)	A&B (2 minutes*2)	The sum of earnings you and your partner generated by only using a half of the time.

Figure 3. 3 How to Report Task Pair Preferences for Round 3 in the Recording Sheet (Translated in English)



Your Task Pair Ranking

How do you rank the task pairs? (remember rank #1 is your most preferred option; rank #4 is the task pair you least prefer). Ties are not allowed.

Rank #	Color (Green, Blue, Red, Yellow)
1 (most preferred)	
2	
3	
4 (least preferred)	

Figure 3. 4 Distribution of Age in the Sample Panel

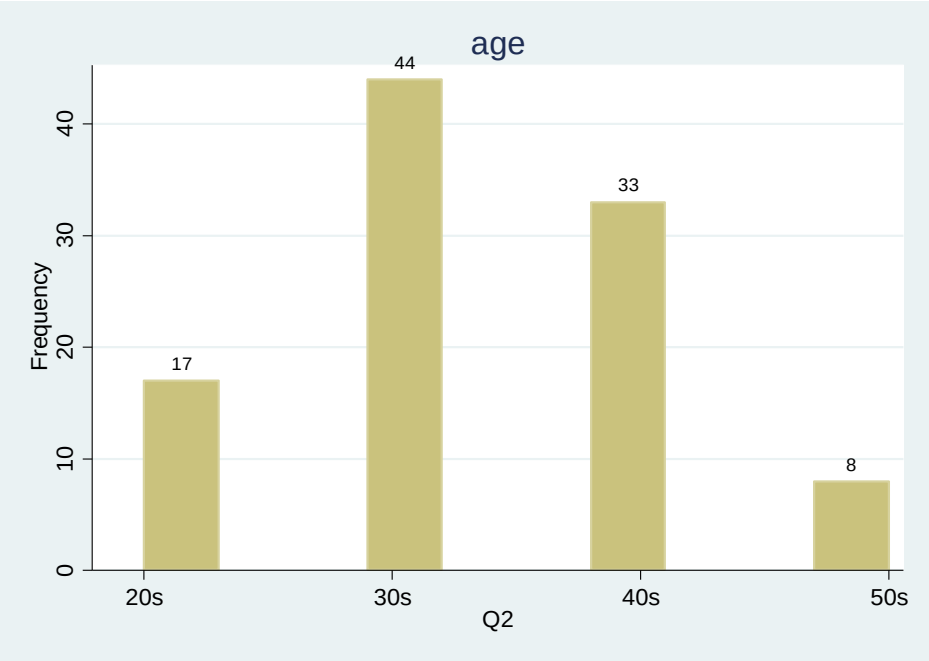


Figure 3. 5 Distribution of the Highest Education Received in the Sample Panel

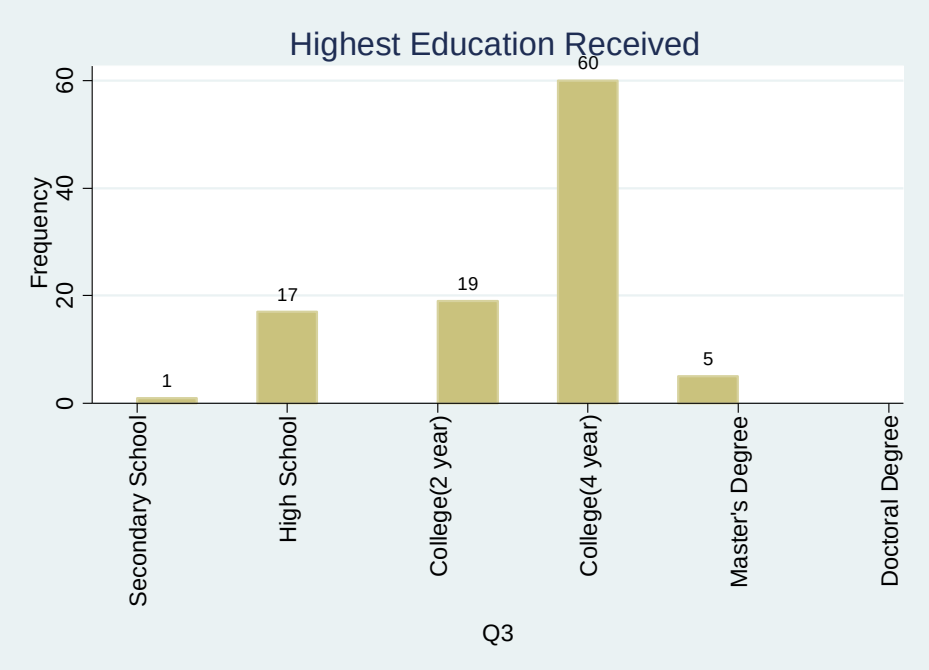


Figure 3. 6 Distribution of Annual Income in the Sample Panel

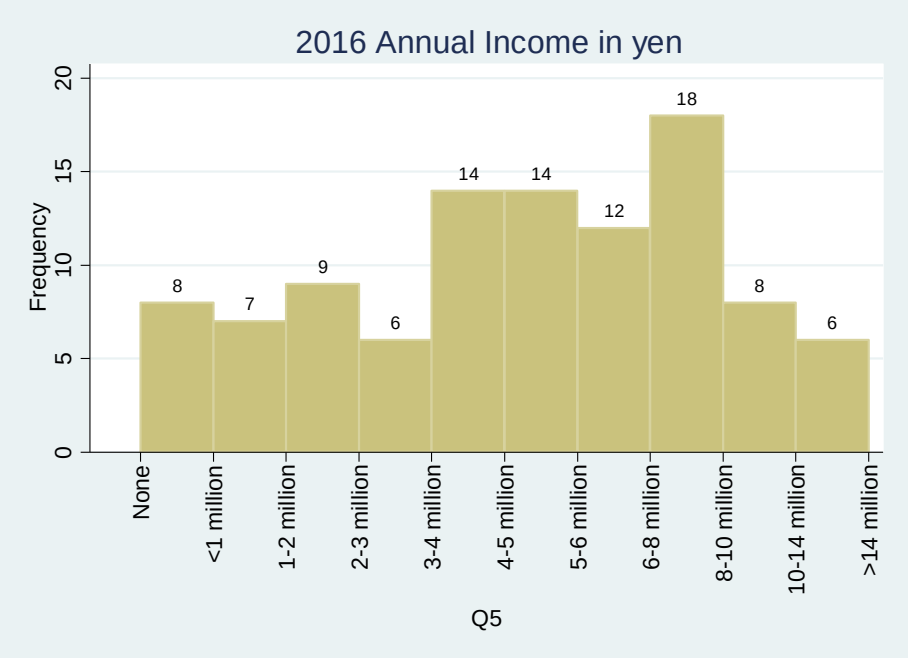


Figure 3. 7 Distribution of Subjects Living With Parents/ Parents-in-law in the Sample Panel

Figure 3. 8 Distribution of Subjects Living with Children in the Sample Panel

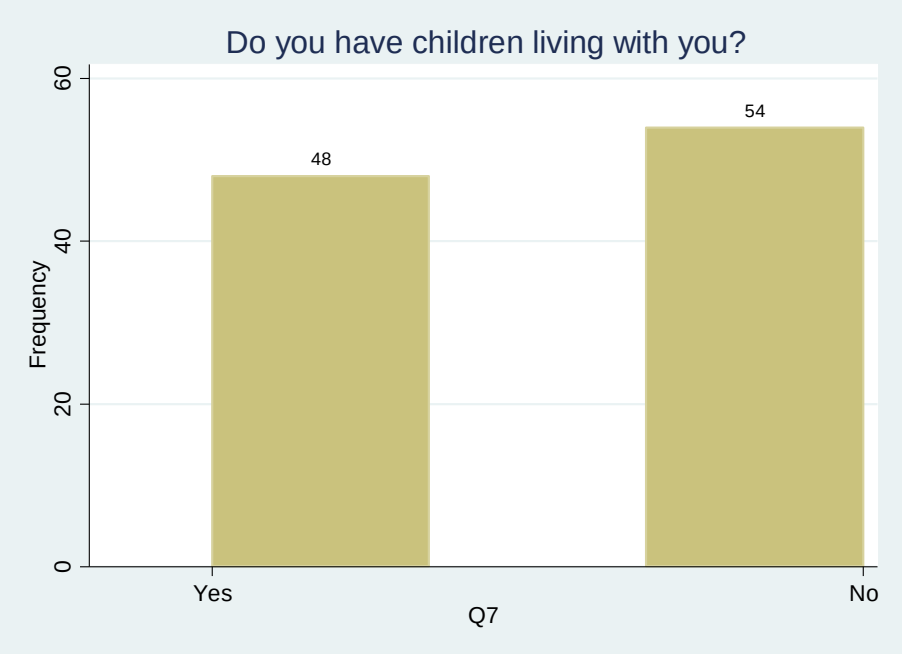


Figure 3. 9 Distribution of Number of Children for Subjects Who are Living with Children in the Sample Panel

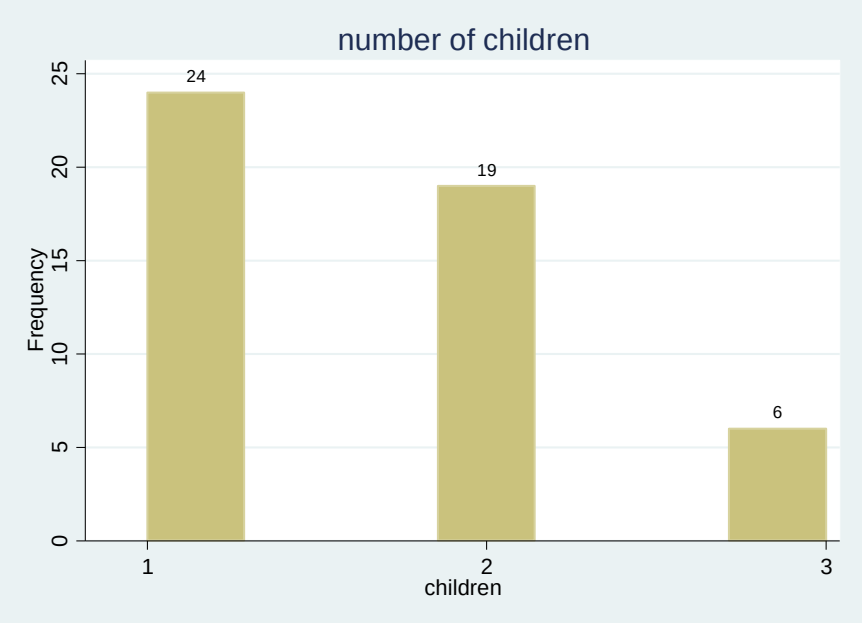


Figure 3. 10 Task A performance in Each Round

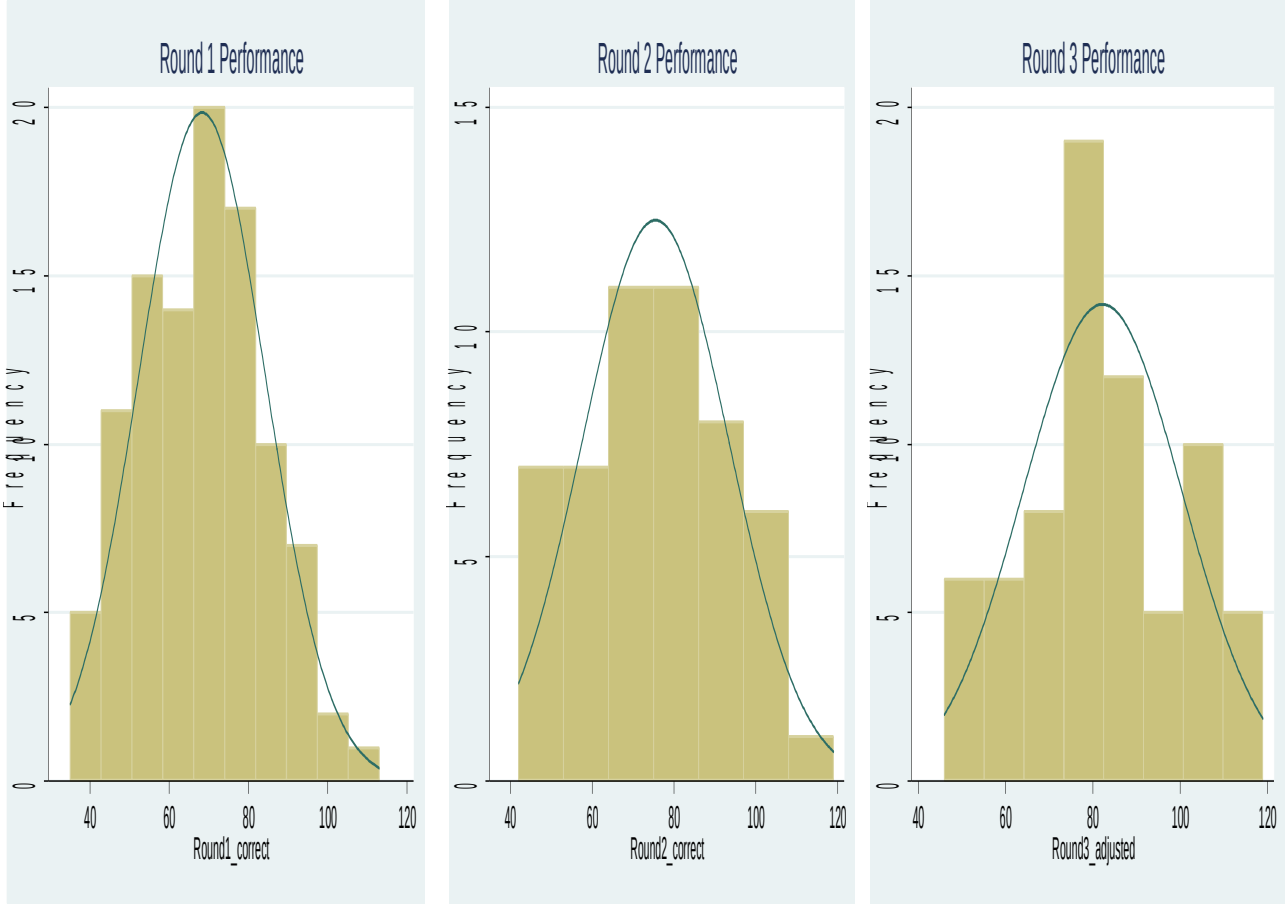


Figure 3. 11 Task B Performance in Each Round

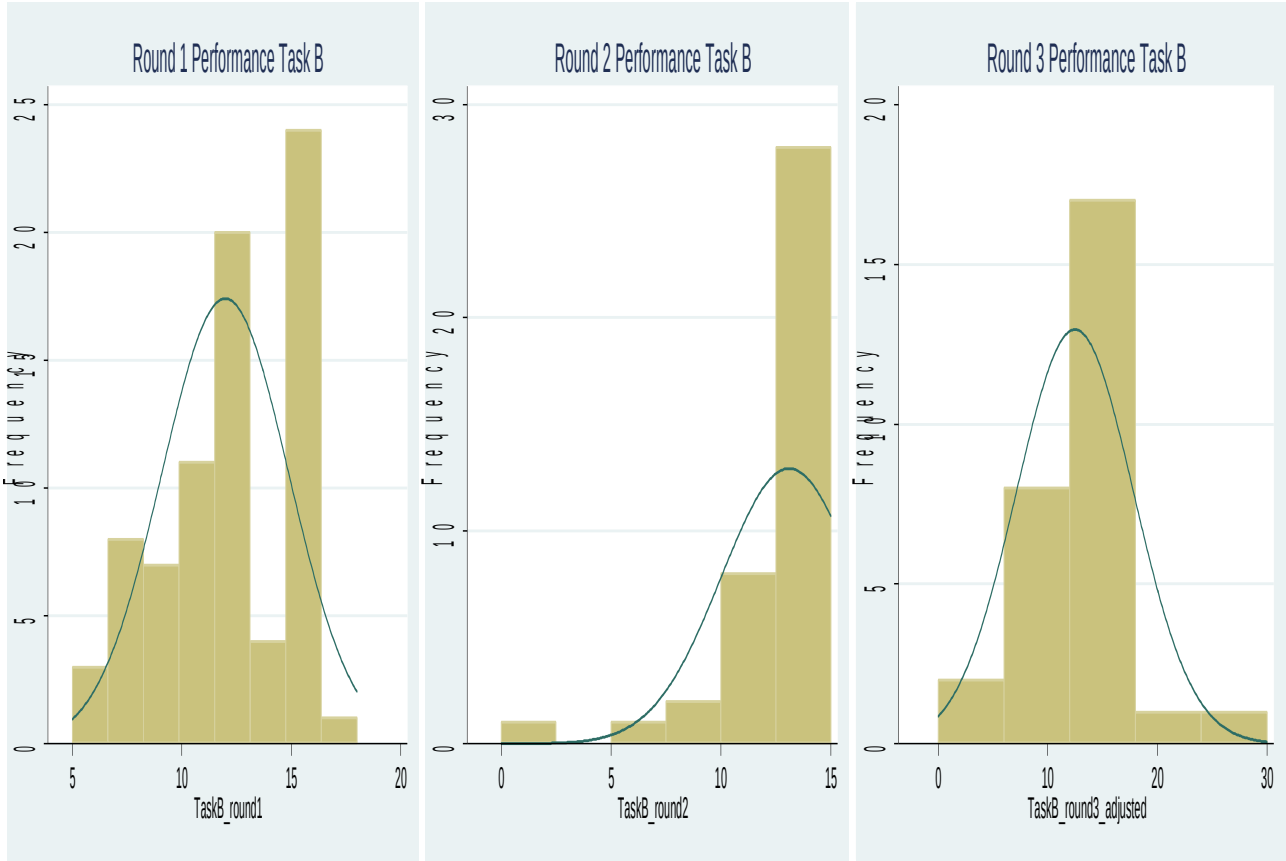


Figure 3. 12 An Example of Decision Tree (Household Income Maximization)

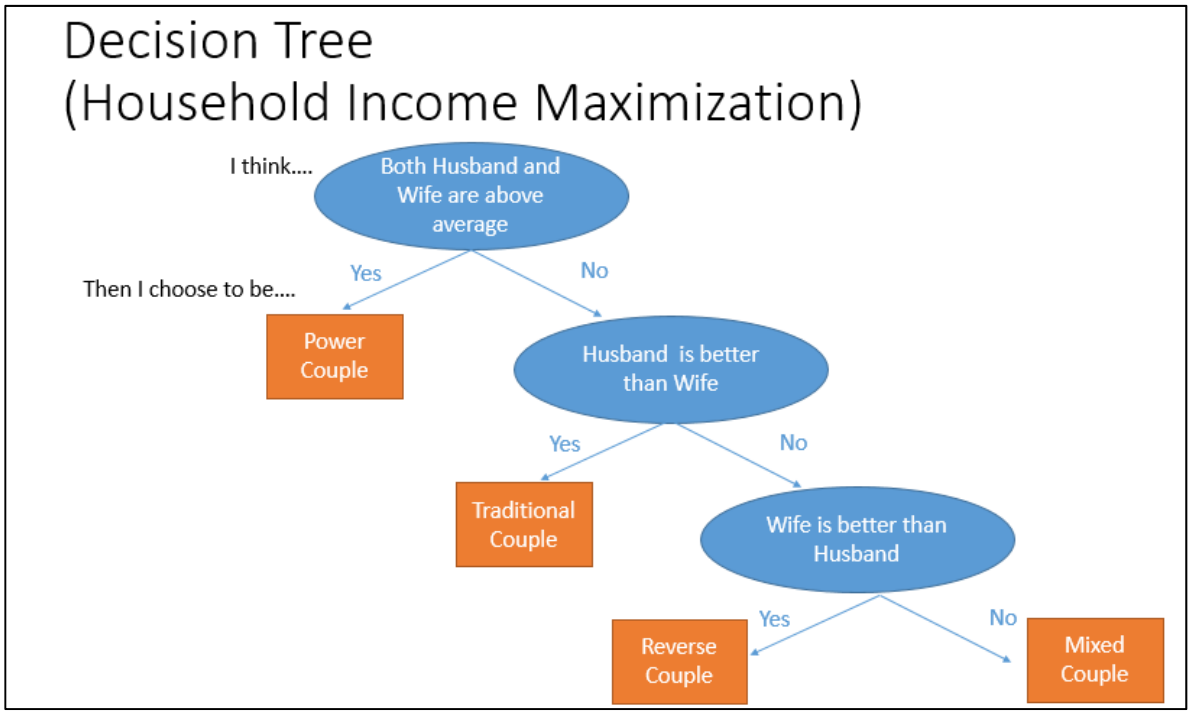


Figure 3. 13 Distribution of the Best Choice

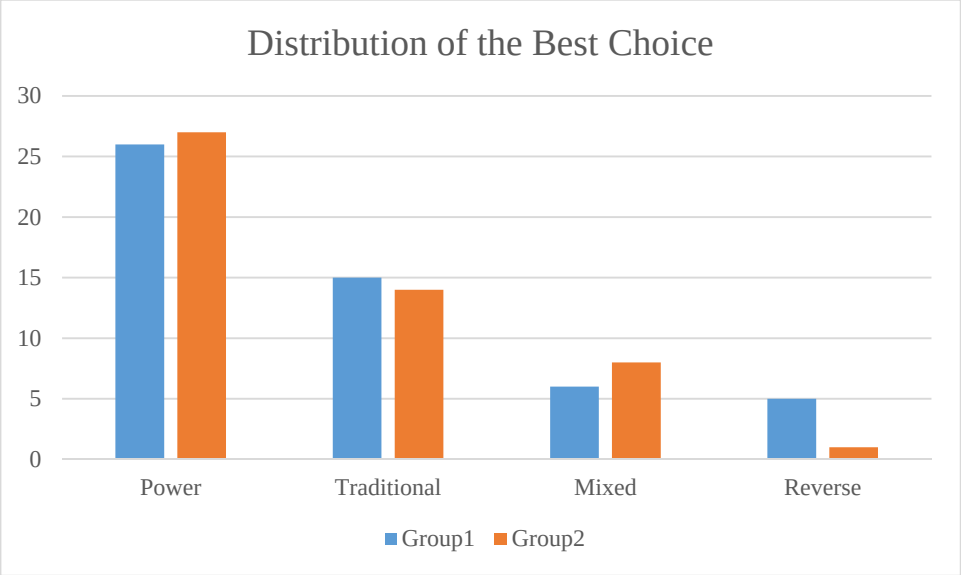


Figure 3. 14 Distribution of the Worst Choice