ベストワースト尺度法におけるチェックボックスの位置効果
-企業が取り組む子育て支援に関するウェブ調査を事例として-

Checkbox Positioning Effect on Best–Worst Scaling
- Evidence from Online Survey Data on Corporate Support for Child Care and Upbringing in Japan-

大床太郎*
Taro Ohdoko

Email: ohdoko@dokkyo.ac.jp

選択モデリング研究において多くの研究が蓄積されている順序ないし位置効果は、同様に表明選好法の一手法であるベストワースト尺度法について十分な研究蓄積があるとは言い難い。とりわけ、ベストワースト尺度法の質問形式には、選択セットにおいて回答者に提示するチェックボックスの位置効果が存在している可能性がある。そこで、日本の企業が取り組む子育て支援に関するウェブ調査データを用いてチェックボックスの位置効果を分析した。その結果、ベストワースト尺度法において採用したアイテムの重要度・多様性パラメータについて、推定値の順序関係は変わらなかった一方で、推定値の大きさには位置効果の存在が確認され、その効果を相殺するために選択セットのチェックボックス位置をランダムに変更して回答者に提示すべきことが示唆された。

Although several ordering or positional effects have been considered in the context of choice modeling research, no studies have investigated the application of best–worst scaling (BWS). In particular, there may be a checkbox positioning effect associated with the use of the BWS format. Using survey data on corporate support of child care and upbringing for employees in Japan, we demonstrate that checkbox positioning affect the absolute, rather than relative, value and heterogeneity of importance of several items. Our results indicate that the checkbox positioning of BWS should be rotated.

*: 獨協大学経済学部

1. Introduction

Best–worst scaling (BWS), or maximum difference scaling, which was developed as one of the stated preference (SP) techniques (Louviere and Woodworth 1990, Finn and Louviere 1992), has been increasingly applied to estimate preferences in many contexts. In addition, there have been many studies conducted to compare BWS with other SP techniques.

BWS operates as follows: respondents are presented with a set of survey items such as those used in this paper, as outlined in the Appendix. They are then asked to choose each item as either most important or least important. Respondents are then provided several sets, which items are varied across questions. The responses indicate the value or importance of the various items.

It has been demonstrated that BWS is superior methodologically. According to Cohen (2003), BWS works as "a rating method that does not experience scale use bias, forces trade-offs, and allows each scale point to be used once and only once." BWS has been tested against various rating methods (Hein et al. 2008; Jaeger et al. 2008; Jaeger and Cardello 2009; Mielby et al. 2012), and BWS was proved to be either as good or better than the other methodologies with regards to the discrimination of the value of items[1]. Furthermore, in choice modeling (CM) studies, which is one of the SP techniques, Carlsson and Martinsson (2001) and Lusk and Schroeder (2004) found that there are few or no hypothetical biases with regard to marginal-willingness-to-pay, which has been one of the major concerns with SP techniques. As BWS concentrates on marginal values, we can expect there to be fewer hypothetical biases compared with other SP techniques (see also Lusk and Briggeman 2009; footnote 2).

However, because BWS involves social surveys, it may or may not suffer from ordering or positional effects. In the CM context, Chrzan (1994) suggested that there are three positional effects: choice set order; profile order within choice sets; and attribute order within profiles. In the psychological context, Dobel et al. (2007) indicated that there are certain influences on writing systems with locational biases. As noted below, BWS can suffer from checkbox positioning, thus we should clarify whether or not there exists bias in estimating preference structures. To the extent of our knowledge, there are no such studies such issues related to BWS.

The objective of this paper is to confirm whether a checkbox positioning effect exists with regard to BWS. As a case study, we use BWS data on corporate support for child care and upbringing in Japan, where respondents uniformly use the Japanese left-to-right lateral writing system. If our results are unaffected by such bias, then BWS is a rigorous procedure with regard to checkbox positioning. If not, we should randomly assign two formats to respondents: a best–worst and a worst–best checkbox positioning, as described below.

The remainder of the article is organized as follows. Section 2 briefly summarizes the previous BWS applications, and ordering or positional effects on CM studies. Section 3 presents our survey design and econometric model. We present and discuss the estimated results in Section 4. Finally, we present concluding remarks and topics for future research in Section 5.

2. Literature Review

2.1 Best–Worst Scaling

BWS was first introduced by Louviere and Woodworth (1990), while Finn and Louviere (1992) was the first published application of the technique. Two formats have been used in previous studies: the BWS format, whose probability properties were outlined by Marley and Louviere (2005) and the best–worst conjoint format, whose probability properties were described by Marley and Pihlens (2012). The latter format, which is a hybrid of BWS and CM, is beyond the scope of this paper, thus we concentrate on the former format.

Regarding BWS applications to food-related issues, wine marketing studies are the most prevalent (Cohen 2009; Goodman 2009; Jaeger et al. 2009; Mueller and Rungie 2009; Sirieix et al. 2011; Loose and Lockshin 2012). Some applications focused on specific food: pork meat (Jaeger et al. 2008), fruit juice and pizza (Louviere and Islam 2008), fruit juice and various foods (Jaeger and Cardello 2009), amounts and types of fat in ground beef (Lusk and Parker 2009), purchase decision for beef (Sawada et al. 2010), conceptual profiles of dark chocolates (Thomson et al. 2010) olive oil (Dekhili et al. 2011) and snacks (Mielby et al.

[1] It has been said that ease of use is strength of BWS; however, previous studies have presented mixed results. Indeed, most studies (Hein et al. 2008; Jaeger et al. 2008; Jaeger and Cardello 2009) found that BWS was easier to adopt. In particular, Lee et al. (2008) suggested that BWS required less respondent time than a rating scale, used by the traditional Schwarz's Value Survey. However, Mielby et al. indicated the opposite result in the context of rating snacks by adolescent respondents.

2012). Other studies have considered food safety in relative to general social issues (Finn and Louviere 1992), breakfast bars (Hein et al. 2008), new food technologies and attributes (Lusk and Briggeman 2009), and general food quality (Lagerkvist et al. 2012).

On health-related issues, studies have considered dermatology consultations (Coast et al. 2006; Flynn et al. 2008), treatment of cardiac arrest occurring in a public place (Lancsar et al. 2007), quality of life in general (Flynn et al. 2007; Coast et al. 2008), young physicians' preference for practice establishment (Gunther et al. 2010), general social care (Potoglou et al. 2011), reasons to escalate antirheumatic treatment (van Hulst et al. 2011), standard health insurance coverage (van der Wulp et al. 2012), experts' assessment of the relative practicality and effectiveness of measures to reduce human exposure to Escherichia coli O157 (Cross et al. 2012), adolescents' concern for health and nonhealth consequences of smoking (Marti 2012), and the attributes in the World Anti-Doping Code's Spirit of Sport statement (Mazanov et al. 2012).

Other studies have focused on psychological issues: Schwarz's value and travel benefits (Lee et al. 2006), Schwarz's value (Lee et al. 2008), ethical beliefs on social issues (Auger et al. 2007), list of values (Lee et al. 2007), and bland equity constructs (Menictas et al. 2010). Given the popularity of the BWS approach, a rigorous procedure for the application of BWS is required.

2.2 Ordering or Positional Effects

Ordering or positional effects can threaten the reliability of preference estimates. As noted earlier, Chrzan (1994) demonstrated that three ordering effects exist in CM studies, which he acknowledged were a function of sequential and positional effects: (i) choice set order, (ii) profile order within choice sets, and (iii) attribute order within profiles. He concluded that choice set order and attribute order occur in unpredictable ways, profile order affects the significance of parameters, and practitioners should rotate those orders to offset such effects.

After Chrzan (1994), several CM studies investigated ordering effects (Farrar and Ryan (1999), Scott and Vick (1999), Kjær et al. (2006), among others). Then, a comprehensive investigation was undertaken by Day et al. (2012). They first reviewed previous studies on positional or ordering effects in CM exercises, and found there were two patterns of ordering effects: "position-dependent," which relates to the position in the sequence of tasks; "precedent-dependent," which relates to the repeating nature of CM exercises. In addition, they presented hypotheses that could explain position-dependent and precedent-dependent effects. Second, they designed their research framework so as to investigate both effects simultaneously, by splitting their respondents into two subsamples: respondents facing advance disclosure, where every task was shown to respondents in advance of conducting choice exercises, and respondents facing a stepwise disclosure of tasks, where each task was sequentially disclosed. Finally, they suggested that the position-dependent ordering effect is primarily related to the stepwise revelation of choice tasks and that advanced revelation appeared to mitigate the effect, and that the precedent-dependent ordering effect existed in both samples[2].

Apart from the previous studies on the ordering effects of CM, there seems to be certain influence of the left–right position of checkboxes on BWS format because of the appearance of the format (see Appendix). In a psychological context, Dobel et al. (2007) conducted experiments consisting of tasks of drawing and arrangement of transparencies with regard to simple sentences, which was tested in both German, which has a left-to-right writing system, and Israeli, which has a right-to-left writing system, for both preschooler and adult samples. They found significant spatial bias in action representation with writing system. Positioning checkboxes under the BWS format may or may not be influential in accordance with writing system[3].

However, to the extent of our knowledge, there are no studies that have focused on the ordering or positional effects of the BWS format. As BWS has been increasingly employed in many contexts, as shown in the previous section, we require a more rigorous procedure for applications of BWS. In particular, because there are no studies on the checkbox positioning effect in the BWS framework, our study considers the best way to structure BWS surveys.

3. Data and Methodology

3.1 Online Survey

Support for child care and upbringing of employees' children is a major political issue in Japan. The Ministry of Health, Labour and Welfare in Japan

[2] We design our research as stepwise disclosure format, which remains the topic for future research.

[3] We should have arranged our survey to include different writing systems. This is a limitation that should be addressed in future research.

(MHLW) indicated a significant difference between Japan's future population projection and what the Japanese people actually desire, which can be mainly responsible to issues concerning childbirth, parenting and the work style[4]. The Child Care and Family Care Leave Act consists of several measures to overcome this situation, including the child care leave system, measures to reduce working hours, limits on overtime, bans on overtime work, limited late-night work, sick/injured child care leave system and so on. It appears that most of these measures focus on work-time-related measures (MHLW 2012).

In contrast, several commendations have been given to workplaces where work-life balance is well established in Japan: "Kobe Danjo Iki-iki Jigyo-sho Hyosho (Commendation of Office with Lively Men and Women in Kobe);"[5] "Hyogo Shigoto to Seikatsu no Balance Kigyo Hyosho (Commendation of Corporate with Work-Life Balance in Hyogo),"[6] and so on. Furthermore, MHLW in Japan has established the social label, "Kurumin," to encourage companies to support the child care and upbringing of their employees' children[7]. Thus, support of child care and upbringing of employees' children has also been a major concern of Japanese companies.

Although employees' working hours is a major concern of companies and employees, the first step in addressing this issue has not been researched and clarified. While the particular circumstances of each company affects what measures can be implemented, it is also necessary to evaluate the preferences of employees regarding such measures. Therefore, we decided to survey employees' references toward corporate support of child care and upbringing, focusing only on work-time-related measures.

We conducted an online survey of workers with an occupation listed in the web research panel of Nikkei Research Inc. in several cities around Kobe, Hyogo, Japan: Kobe, Amagasaki, Akashi, Nishinomiya, Ashiya, Itami, Takarazuka, Miki, Ono, Sanda, Kawanishi, Ikeda, Minoh, Suita, Toyonaka and Osaka. We selected cities around Kobe as the study sites where the path-breaking commendation as discussed above affects. We organized our questionnaire as follows[8]: first, we asked respondents about their work situation; second, about the availability of corporate support for child care and upbringing, to insure respondents were familiar with all elements of the available support; third, BWS about work-time-related support as described below; fourth, their views on their workplace; and their socioeconomic characteristics. We administered our survey during February 23–27, 2013. We distributed 17,986 surveys and received 2,004 responses (response rate of $11.2\%\sim100\times2{,}004/17{,}896$). A summary of the basic demographics are provided in Table 1.

A balanced incomplete block design (BIBD) has been suggested for organizing the items to be analyzed in BWS choice sets (see e.g. Cohen 2009). BIBD enables us to analyze large numbers of items to obtain a full ranking from a relatively small number of subsets. If a BIBD has v items with r replications in b blocks of size k and the pair frequency λ, it can be denoted by $B(v,k,b,r,\lambda)$. To organize the BWS choice set, we select seven items with reference to "Ryoritsu Shien no Hiroba"[9]: child care leave (CCL), sick and/or injured child care leave (SIL), reducing working hours (RWH), flexible working hours (FWH), advancing or delaying starting or finishing hours (AD), working from home (WH), exemption from overtime work (EOW) (see Table 2). When assigning items with choice sets, we employed R 3.0.0 and the function "bibd" in the R package OPDOE 1.0-7 (Rasch et al. 2011) with $B(7,3,7,3,1)$ (see Table 3 and Appendix). Consequently, we created seven choice sets for respondents along with the choice set ordering in Table 3.

4 MHLW Web site (URL: http://www.mhlw.go.jp/english/policy/affairs/dl/05.pdf, retrieved on Sep 12th 2013).

5 Kobe city Web site (URL: http://www.city.kobe.lg.jp/life/community/cooperation/ikiikijimusho/, retrieved on Sep 12th 2013) [Japanese only].

6 Hyogo Work and Life Center Web site (URL: http://www.hyogo-wlb.jp/modtreepage01_1774/, retrieved on Sep 12th 2013) [Japanese only].

7 Kurumin is "Next-generation authorization mark" in accordance with the Act on Advancement of Measures to Support Raising Next-Generation Children in Japan (MHLW 2012, P.172).

8 Before implementation, we conducted a preliminary qualitative assessment using 14 undergraduate students in the Faculty of Economics, Dokkyo University, Japan, to improve the design of the questionnaire.

9 A comprehensive Web site on work–life balance in Japan is run by MHLW (URL: http://www.ryouritsu.jp/index.html, retrieved on Sep 12th 2013).

Table 1: Demographics

Item	Sub-Item	BW	WB	P value	Item	Sub-Item	BW	WB	P value
		Ans.	Ans.				Ans.	Ans.	
Gender	Male	595	542	0.060	No. of	1	12	6	0.326
	Female	417	450		family members	2	174	146	
Age	30	74	88	0.021*		3	208	201	
	40	284	301			4	269	253	
	50	385	396			5	259	285	
	60	269	207			6	71	75	
Marital	Married	661	656	0.702		7+	19	26	
status	Unmarried	351	336		No. of children	0	753	710	0.301
Income	200	37	38	0.005**	(6 to 14 years old)	1	155	160	
(JPY	400	170	175			2	92	102	
million)	600	258	224			3+	12	20	
	800	173	235		No. of children	0	874	847	0.453
	1,000	127	119		(5 years old or	1	105	102	
	1,200	77	63		younger)	2+	33	43	
	1,400	41	34						
	1,600	24	13						
	1,800	5	12						
	1,800+	14	5						
	No answer	86	74						

Note: Ans. denotes number of answers; P value is estimated using the χ^2 statistical test; ** and * indicate null is rejected at 1% and 5% significance level, respectively (null: proportion ratios are identical between samples).

Table 2: Work-Time-Related Corporate Support for Employees' Child Care and

Attribute	Item (Variable name)	Description
1	Child care leave (CCL)	In principle, employees can obtain child care leave from the child's birth to the day before the child's first birthday.
2	Sick and/or injured child care leave (SIL)	Employees can obtain this child care leave if the child suffers from an acute disease.
3	Reduced working hours (RWH)	Reduction in the prescribed daily, weekly or monthly number of works hour.
4	Flexible working hours (FWH)	There exists a core period of the day when employees are expected to be at work, however, employees can choose when they work during the remaining hours, subject to a total prescribed number of working hours.
5	Advance or delay starting or finishing times (AD)	For example, suppose those who plan to work between 9:00 and 17:00 (with a one-hour lunch break) change their working hour to 9:30 to 17:30 (with a one-hour lunch break) with total working hours unchanged.
6	Working from home (WH)	A work arrangement in which employees do not commute to a central place of work. Instead, they work from home with total working hours unchanged.
7	Exemption of overtime work (EOW)	Working beyond the prescribed number of hours is not allowed for employees who have children who are less than three years old.

Table 3: Balanced Incomplete Block Design for Seven Attributes: B(7,3,7,3,1)

Choice Set No.	Attribute No.		
1	1	2	3
2	1	4	5
3	1	6	7
4	2	4	6
5	2	5	7
6	3	4	7
7	3	5	6

To investigate the checkbox positioning effect of BWS, we created two subsamples: the BW sample, in which the respondents were assigned the format in Appendix, where the most important option is located on the left side and the least important one on the right of the choice set, and the WB sample, in which the respondents were assigned the opposite format, where the least important option is located on the left and the most important one on the right. The χ^2 tests on whether the two subsamples are statistically identical indicate that most of the socio-economic characteristics in Table 1 are identical, apart from age and income. As the proportion of samples in age and income was qualitatively similar between the subsamples, we assume they have identical demographics distributions.

3.2 Econometric Model

Each response to a best–worst question can be seen as choosing the two items that maximize the difference between two items on an underlying scale of importance. If a choice set has J items, there are $J(J-1)$ possible best–worst combinations a respondent n ($n = 1,\cdots,N$) could choose. When a particular pair of items is chosen as best and worst, it denotes a choice out of all $J(J-1)$ possible pairs that maximizes the difference in importance.

Let us assume γ_i is the location of the value of item i, chosen as the best item, and γ_j is that of item j, chosen as the worst item, on the underlying scale of importance. Then, we can assume the random utility model $U_i = \gamma_i + \varepsilon_i$ and $U_j = \gamma_j + \varepsilon_j$, where ε_i and ε_j are relevant random error terms. When a respondent n chooses item i and item j as the best and worst, respectively, the choice probability out of a choice set with J items is equal to the probability that the difference in U_i and U_j is greater than all other $J(J-1)-1$ possible differences in the choice set. When ε_i and ε_j are distributed i.i.d. type I extreme values, this probability takes the familiar multinomial logit form, as follows:

$$P_{ij} = \exp(\gamma_i - \gamma_j)/\left(\sum_{l=1}^{J}\sum_{m=1}^{J}\exp(\gamma_l - \gamma_m) - J\right).$$

The parameter γ_i or γ_j can be estimated by the maximum likelihood estimation procedure. The estimated γ_i denotes the importance of item i relative to an item that was normalized to zero[10].

Revelt and Train (1998) demonstrated that RPL with the use of repeat data could relax the assumptions of a multinomial logit, namely, preference homogeneity and the independence of irrelevant alternatives. The choice probability of respondent n is given as follows within the parameter space Ω:

$$\pi_n = \int \prod_{t=1}^{T} P_{nijt} \cdot f(\gamma|\Omega)d\gamma,$$

where $t\,(t = 1,\cdots,T)$ denotes the number of replications of BWS questions, P_{nijt} is the form of Equation 1, and $f(\gamma|\Omega)$ is known as a mixing distribution. Previous studies have frequently employed the normal distribution for $f(\gamma|\Omega)$, which we also utilized. The parameters were estimated by maximizing a simulated log-likelihood function, evaluated at 100 pseudorandom Halton draws in this article[11].

In particular, we can specify the estimated importance parameter for a respondent n and item i can be specified as $\hat{\gamma}_{ni} = \bar{\gamma}_i + \sigma_i\mu_{ni}$, where $\bar{\gamma}_i$ and σ_i are the mean and standard deviation parameter of γ_i in the

10 In estimating a multinomial logit model, parameters are estimated confounding with the scale parameter that is inversely proportionate to the variance of the error term (Swait and Louviere 1993; Louviere et al. 2000). For simplicity, we assume the scale parameter is set to one.

11 Train (2009) provides more computational details for the RPL.

population, and μ_{ni} is a random term normally distributed with mean zero and unit standard deviation. As we can obtain the coefficient of variation (CV) of γ_i by utilizing the mean and standard deviation parameter, which denotes the degree of heterogeneity in preferences and can be compared between different populations, we focused on both the mean and CV estimates when evaluating the checkbox positioning effect on preferences for simplicity.

We employed Limdep 9.0 + NLOGIT 4.0.1 (Econometric Software, Inc., NY, USA) to estimate RPL, with 1,000 Monte Carlo simulations of the mean and the variance matrix of the mean parameters to estimate confidence intervals (CI) for each λ and CV (Krinsky and Robb 1986). In searching for the best-fit model, we give high priority to the significance of the standard deviation parameters in order to grasp the structure of the preference heterogeneities. Then, we employed the complete combinatorial (CC) procedure (Poe et al. 2005) to observe whether the checkbox positioning effect exists in order to check each individual item in the choice set.

4. Results and Discussion

We provide estimated RPL results of the two subsamples, BW sample and WB sample, in Table 4. Every standard deviation parameter is significant in both subsamples, which indicates that all the items were heterogeneously preferred, while most of the mean parameters were significant with the exception of FWH in the WB sample, which means that the parameter was not significant. We normalized EOW to zero, thus, every mean parameter measures importance relative to EOW. Then, the estimated values and CV are provided in Tables 5 and 6, respectively.

For the BW sample, respondents ranked the items as follows in accordance with the value estimates (Table 5): CCL, SIL, RWH, AD, FWH, EOW, and WH. The CV estimates indicated that the order of the degree of preference heterogeneity was as follows in terms of absolute value: FWH, WH, SIL, AD, CCL, and RWH (Table 5). On the other hand, for the WB sample, respondents ranked the items as follows in terms of the value estimates (Table 5): CCL, SIL, RWH, AD, FWH, and EOW, which were all equally ranked, and WH. The CV estimates indicated that the order of the degree of preference heterogeneity is as follows in terms of absolute value: WH, AD, SIL, RWH, and CCL (Table 5), where FWH could not be included (Table 5).

According to the CC statistics in Table 5, the mean value estimates of CCL and AD in the BW sample were significantly larger than in the WB sample, while that of WH in the BW sample was significantly smaller than in the WB sample. On the other hand in Table 6, the CV estimates of AD and RWH in the BW sample were significantly smaller than in the WB sample, while that of WH in the BW sample was significantly larger than in the WB sample.

With regard to the mean value estimates, the respondents in both subsamples analogously prefer all items relating to child care leave, then all items relating to flexible working hours. Each rank also appears to be similar between subsamples, while the item assigned the middle rank, FWH, is equivalently preferred to EOW in the WB sample only. Thus, it indicates that relative importance is not affected by checkbox positioning. However, there exist statistical differences between subsamples on several items: CCL is the best; AD is the middle; and WH is the worst. It suggests that checkbox positioning is influential on the absolute degree of importance, and is assigned extreme and modest importance.

With regard to the CV estimates, the relative degree of heterogeneity was almost similar between subsamples. In both samples, the item assigned the higher value tends to be preferred less heterogeneously. However, there exist statistical differences on several items: AD, RWH, and WH. Although we cannot make any firm conclusions, the results do indicate that the item assigned the highest degree of heterogeneity, WH, is affected by checkbox positioning along with several other items.

Above all, if the concern of BWS practitioners is the “relative importance” of items, they do not have to care about checkbox positioning very much. However, to estimate the “absolute value” significantly, and to understand heterogeneous preference structures fully, the results suggest that we should rotate checkbox positioning as much as possible to offset its positional effect[12].

12 This procedure depends on the nature of each survey, such as the research budget.

Table 4: RPL Results of Two Samples

Sample	BW		WB	
Parameter estimates	Mean	SD	Mean	SD
CCL	2.195**	2.532**	1.926**	2.249**
	(22.851)	(23.415)	(22.789)	(22.544)
SIL	0.756**	1.424**	0.862**	1.529**
	(12.748)	(21.392)	(14.026)	(21.653)
RWH	0.749**	0.737**	0.667**	0.817**
	(16.488)	(12.185)	(14.260)	(13.997)
FWH	0.336**	1.459**	–0.031	1.425**
	(5.735)	(20.563)	(–0.514)	(20.262)
AD	0.525**	0.920**	0.366**	0.908**
	(11.134)	(14.961)	(7.771)	(15.375)
WH	–0.943**	2.204**	–0.723**	2.115**
	(–10.952)	(22.700)	(–9.439)	(21.913)
No. of observations	7,084		6,944	
No. of samples	1,012		992	
Halton draw replications	100		100	
Log likelihood	–10,329.550		–10,255.570	
McFadden's ρ				
No coefficients	0.186		0.175	
Constants only	0.129		0.116	

Note: ** indicates significance at the 1% level; t values in parentheses; SD is standard deviation.

Table 5: Estimated Value of Items and CC Results

	BW sample	WB sample	CC statistics
CCL	2.195	1.926	0.017*
	[1.998; 2.365]	[1.752; 2.076]	
SIL	0.756	0.862	0.110
	[0.641; 0.874]	[0.743; 0.985]	
RWH	0.749	0.667	0.100
	[0.662; 0.832]	[0.578; 0.752]	
FWH	0.336	0	n.a.
	[0.217; 0.450]	[n.a.]	
AD	0.525	0.366	0.008**
	[0.438; 0.621]	[0.280; 0.463]	
WH	–0.943	–0.723	0.028*
	[–1.118; –0.769]	[–0.880; –0.570]	

Note: ** and * indicate significance at the 1 and 5% levels, respectively; 95 % confidence intervals in brackets. CC denotes complete combinatorial (null: parameters are identical across samples).

Table 6: CV Estimates of Value of Items and CC Results

	BW sample	WB sample	CC statistics
CCL	1.154	1.168	0.449
	[1.050; 1.264]	[1.017; 1.353]	
SIL	1.884	1.774	0.290
	[1.602; 2.244]	[1.545; 2.083]	
RWH	0.984	1.225	0.060*
	[0.814; 1.173]	[0.994; 1.509]	
FWH	4.333	n.a.	n.a.
	[3.160; 6.666]	[n.a.]	
AD	1.752	2.481	0.028**
	[1.397; 2.197]	[1.915; 3.368]	
WH	–2.337	–2.925	0.045*
	[–2.841; –1.986]	[–3.713; –2.429]	

Note: ** and * indicate significance at the 1 and 5% levels, respectively; 95 % confidence intervals in brackets. CC denotes complete combinatorial (null: parameters are identical across samples).

5. Concluding Remarks

Using survey data on corporate support for child care and upbringing of employees' children in Japan, we demonstrated that a checkbox positioning effect exists for BWS. This type of positional effect can be influential on the absolute value and heterogeneity of importance of several items, not on the relative ones. Although we could not conclude why the effect operates in this way, it indicates that the BWS format should be used carefully, including checkbox positioning in accordance with the scope of each research.

Several issues remain unaddressed. First, we omitted other positional effects of the BWS format to obtain a large sample size. Day et al. (2012) suggested that it would be fruitful to investigate several CM ordering effects, which will have the same effects as BWS ordering effects. Second, because we collected data on many socio-economic characteristics, the RPL with cross-term model is a promising approach to the estimation of robust values of the support for work-time-related corporate support for employees' child care and upbringing, where the BW and WB samples should be pooled to cancel the checkbox positioning effect. Third, because we gathered data on employees' views on workplace diversity, latent class or latent clustering analysis (see Greene and Hensher 2003 for further details) is a promising approach for the inclusion of those attitudinal covariates, which has been employed in previous BWS applications (Mueller and Rungie 2009; Dekhili et al. 2011; Sirieix et al. 2011; Lagerkvist et al. 2012; Loose and Lockshin 2012). We intend to consider the latter two issues in future research.

Acknowledgements

This research was supported by a Ministry of Education, Culture, Sports, Science, and Technology Grant-in-Aid for Scientific Research (No. 23730238), and Information Science Research Institute and personal research grant at Dokkyo University. The author gratefully acknowledges the comments and advice of anonymous reviewers, and the cooperation of Nikkei Research Inc. and the kind survey respondents.

References

(1) Auger P, TM Devinney, JJ Louviere, "Using Best-Worst Scaling Methodology to Investigate Consumer Ethical Beliefs Across Countries," Journal of Business Ethics, Vol.70, pp.299–326 (2007).

(2) Carlsson F, P Martinsson, "Do Hypothetical and Actual Marginal Willingness to

Pay Differ in Choice Experiments?," Journal of Environmental Economics and Management, Vol.41, pp.179–192 (2001).
(3) Chrzan K, "Three Kinds of Order Effects in Choice-Based Conjoint Analysis," Marketing Letters, Vol.5, No.2, pp.165–172 (1994).
(4) Coast J, TN Flynn, L Natarajan, K Sproston, J Lewis, JJ Louviere, TJ Peters, "Valuing the ICECAP Capability Index for Older People," Social Science and Medicine, Vol.67, pp.874–882 (2008).
(5) Coast J, C Salisbury, D de Berker, A Noble, S Horrocks, TJ Peters, TN Flynn, "Preferences for Aspects of a Dermatology Consultation," Epidemiology and Health Services Research, Vol.155, pp.387–392 (2006).
(6) Cohen E, "Applying Best-Worst Scaling to Wine Marketing," International Journal of Wine Business Research, Vol.21, No.1, pp.8–23 (2009).
(7) Cohen SH, "Maximum Difference Scaling: Improved Measures of Importance and Preference for Segmentation," Sawtooth Software Research Paper Series, URL: http://www.sawtoothsoftware.com/download/techpap/maxdiff.pdf, retrieved on Sep 12th (2013).
(8) Cross P, D Rigby, G Edwards-Jones, "Eliciting Expert Opinion on the Effectiveness and Practicality of Interventions in the Farm and Rural Environment to Reduce Human Exposure to Escherichia Coli O157," Epidemiology and Infection, Vol.140, No.4, pp.643–654 (2012).
(9) Day B, IJ Bateman, RT Carson, D Dupont, JJ Louviere, S Morimoto, R Scarpa, P Wang, "Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies," Journal of Environmental Economics and Management, Vol.63, pp.73–91 (2012).
(10) Dekhili S, L Sirieix, E Cohen, "How Consumers Choose Olive Oil: The Importance of Origin Cues," Food Quality and Preference, Vol.22, pp.757–762 (2011).
(11) Dobel C, G Diesendruck, J Bölte, "How Writing System and Age Influence Spatial Representations of Actions: A Developmental, Cross-Linguistic Study," Psychological Science, Vol.18, No.6, pp.487–491 (2007).
(12) Farrar S, M Ryan, "Response-Ordering Effects: A Methodological Issue in Conjoint Analysis," Health Economics Letters, Vol.8, pp.75–79 (1999).
(13) Finn A, JJ Louviere, "Determining the Appropriate Response to Evidence of Public Concern : The Case of Food Safety," Journal of Public Policy and Marketing, Vol.11, No.1, pp.12–25 (1992).
(14) Flynn TN, JJ Louviere, TJ Peters, J Coast, "Estimating Preferences for a Dermatology Consultation Using Best-Worst Scaling: Comparison of Various Methods of Analysis," BMC Medical Research Methodology, Vol.8, No.76 (2008).
(15) Flynn TN, JJ Louviere, TJ Peters, J Coast, "Best-Worst Scaling: What it Can Do for Health Care Research and How to Do it," Journal of Health Economics, Vol.26, pp.171–189 (2007).
(16) Goodman S, "A Multi-Country Comparison of Consumer Choice for Wine in Retail Stores," International Journal of Wine Business Research, Vol.21, No.1, pp.41–49 (2009).
(17) Greene WH, DA Hensher, "A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit," Transportation Research Part B, Vol.37, pp.681–698 (2003).
(18) Günther OH, B Kürstein, SG Riedel-Heller, HH Künig, "The Role of Monetary and Nonmonetary Incentives on the Choice of Practice Establishment: A Stated Preference Study of Young Physicians in Germany," Health Services Research, Vol.45, No.1, pp.212–229 (2010).
(19) Hein KA, SR Jaeger, BT Carr, CM Delahunty, "Comparison of Five Common Acceptance and Preference Methods," Food Quality and Preference, Vol.19, pp.651–661 (2008).
(20) Jaeger SR, AV Cardello, "Direct and Indirect Hedonic Scaling Methods: A Comparison of the Labeled Affective Magnitude (LAM) Scale and Best-Worst Scaling," Food Quality and Preference, Vol.20, pp.249–258 (2009).
(21) Jaeger SR, PJ Danaher, RJ Brodie, "Wine Purchase Decisions and Consumption Behaviours: Insights from a Probability Sample Drawn in Auckland, New Zealand," Food Quality and Preference, Vol.20, pp.312–319 (2009).
(22) Jaeger SR, AS Jørgensen, MD Aaslyng, WLP Bredie, "Best-Worst Scaling: An Introduction and Initial Comparison with Monadic Rating for Preference Elicitation with Food Products," Food Quality Preference, Vol.19, pp.579–588 (2008).
(23) Kjær T, Bech M, Gyrd-Hansen D, Hart-Hansen K, "Ordering Effect and Price Sensitivity in Discrete Choice Experiments: Need We Worry?," Health Economics, Vol.15, pp.1217–1228 (2006).
(24) Krinsky I, AL Robb, "On Approximating the Statistical Properties of Elasticities," Review of Economics and Statistics, Vol.68, pp.715–719 (1986).
(25) Lagerkvist CJ, J Okello, N Karanja, "Anchored vs. Relative Best-Worst Scaling and Latent Class vs. Hierarchical Bayesian Analysis of Best-Worst Choice Data: Investigating the Importance of Food Quality Attributes in a Developing Country," Food Quality and

Preference, Vol.25, pp.29–40 (2012).
(26) Lancsar E, JJ Louviere, TN Flynn, “Several Methods to Investigate Relative Attribute Impact in Stated Preference Experiments,” Social Science and Medicine, Vol.64, pp.1738–1753 (2007).
(27) Lee JA, GN Soutar, JJ Louviere, TM Daly, “An Examination of the Relationship Between Values and Holiday Benefits across Cultures using Ratings Scales and Best-Worst Scaling,” Publication Advancing Theory, Maintaining Relevance, Brisbane, Australia, URL: http://hdl.handle.net/10453/3163, retrieved on Sep 12th 2013, (2006).
(28) Lee JA, GN Soutar, JJ Louviere, “The Best-Worst Scaling Approach: An Alternative to Schwartz’s Values Survey,” Journal of Personality Assessment, Vol.90, No.4, pp.335–347 (2008).
(29) Lee JA, GN Soutar, JJ Louviere, “Measuring Values Using Best-Worst Scaling: The LOV Example,” Psychology and Marketing, Vol.24, No.12, pp.1043–1058 (2007).
(30) Loose SM, L Lockshin, “Testing the Robustness of Best-Worst Scaling for Cross-National Segmentation with Different Numbers of Choice Sets,” Food Quality and Preference, Vol.27, No.2, pp.230–242 (2013).
(31) Louviere JJ, DA Hensher, JD Swait, “Stated Choice Methods: Analysis and Application,” Cambridge University Press, United Kingdom (2000).
(32) Louviere JJ, T Islam, “A Comparison of Importance Weights and Willingness-to-Pay Measures Derived from Choice-Based Conjoint, Constant Sum Scales and Best–Worst Scaling,” Journal of Business Research, Vol.61, pp.903–911 (2008).
(33) Louviere JJ, GG Woodworth (1990), “Best-Worst Scaling: A Model for Largest Difference Judgments,” Working Paper, Faculty of Business, University of Alberta, Edmonton (1990).
(34) Lusk JL, BC Briggeman, “Food Values,” American Journal of Agricultural Economics, Vol.91, No.1, pp.184–196 (2009).
(35) Lusk JL, N Parker, “Consumer Preferences for Amount and Type of Fat in Ground Beef,” Journal of Agricultural and Applied Economics, Vol.41, No.1, pp.75–90 (2009).
(36) Lusk JL, TC Schroeder, “Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks,” American Journal of Agricultural Economics, Vol.86, pp.467–482 (2004).
(37) Marley AAJ, JJ Louviere, “Some Probabilistic Models of Best, Worst, and Best-Worst Choices,” Journal of Mathematical Psychology, Vol.49, pp.464–480 (2005).
(38) Marley AAJ, D Pihlens, “Models of Best-Worst Choice and Ranking Among Multiattribute Options (Profiles),” Journal of Mathematical Psychology, Vol.56, pp.24–34 (2012).
(39) Marti J, “A Best-Worst Scaling Survey of Adolescents’ Level of Concern for Health and Non-Health Consequences of Smoking,” Social Science and Medicine, Vol.75, pp.87–97 (2012).
(40) Mazanov J, T Huybers, J Connor, “Prioritising Health in Anti-Doping: What Australians Think,” Journal of Science and Medicine in Sport, Vol.15, pp.381–385 (2012).
(41) Mielby LH, M Edelenbos, AK Thybo, “Comparison of Rating, Best-Worst Scaling, and Adolescents’ Real Choices of Snacks,” Food Quality and Preference, Vol.25, pp.140–147 (2012).
(42) Ministry of Health, Labour and Welfare, “Equal Employment and Child Welfare,” In “Annual Health, Labour and Welfare Report 2011–2012,” pp.168–190, URL: http://www.mhlw.go.jp/english/wp/wp-hw6/dl/07e.pdf, retrieved on Sep 12th 2013 (2012).
(43) Mueller S, C Rungie, “Is There More Information in Best-Worst Choice Data?: Using the Attitude Heterogeneity Structure to Identify Consumer Segments,” International Journal of Wine Business Research, Vol.21, No.1, pp.24–40 (2009).
(44) Poe GL, KL Giraud, JB Loomis, “Computational Methods for Measuring the Difference of Empirical Distributions,” American Journal of Agricultural Economics, Vol.87, No.2, pp.353–365 (2005).
(45) Potoglou D, P Burge, TN Flynn, A Netten, J Malley, J Forder, JE Brazier, “Best–Worst Scaling vs. Discrete Choice Experiments: An Empirical Comparison Using Social Care Data,” Social Science and Medicine, Vol.72, pp.1717–1727 (2011).
(46) Rasch D, Pilz J, Verdooren LR, Gebhardt A, “Optimal Experimental Design with R,” Chapman and Hall/CRC (2011).
(47) Revelt D, KE Train, “Mixed Logit with Repeated Choice: Households’ Choices of Appliance Efficiency Level,” Review of Economics and Statistics, Vol.80, No.4, pp.647–657 (1998).
(48) Sawada M, H Aizaki, K Sato, “Consumers’ Evaluation on the Advantages of Improving Feed Self-Sufficiency in Beef Production,” Research Bulletin of Obihoro University, Vol.31, pp.18–24, URL: http://ir.obihiro.ac.jp/dspace/handle/10322/2721, retrieved on Sep 12th 2013 (2010) [Japanese Only].
(49) Scott A, S Vick, “Patients, Doctors and Contracts: An Application of Principal-Agent Theory to the Doctor-Patient Relationship,” Scottish Journal of Political Economy, Vol.46, No.2, pp.111–134 (1999).

(50) Sirieix L, H Remaud, L Lockshin, L Thach, T Lease, "Determinants of Restaurant's Owners/Managers Selection of Wines to be Offered on the Wine List," Journal of Retailing and Consumer Services, Vol.18, pp.500–508 (2011).
(51) Swait JD, JJ Louviere, "The Role of the Scale Parameter in the Estimation and Use of Multinomial Logit Models," Journal of Marketing Research, Vol.30, pp.305–314 (1993).
(52) Thomson DMH, C Crocker, CG Marketo, "Linking Sensory Characteristics to Emotions: An Example Using Dark Chocolate," Food Quality and Preference, Vol.21, pp.1117–1125 (2010).
(53) Train KE, "Discrete Choice Methods With Simulation," 2nd edition, Cambridge University Press, New York (2009).
(54) Van der Wulp I, WB van der Hout, M de Vries, AM Stiggelbout, EM van den Akker-van Marle, "Societal Preferences for Standard Health Insurance Coverage in the Netherlands: A Cross-Sectional Study," BMJ Open 2, e001021, doi:10.1136/bmjopen-2012-001021 (2012).
(55) Van Hulst LTC, W Kievit, R van Bommel, PLCM van Riel, L Fraenkel, "Rheumatoid Arthritis Patients and Rheumatologists Approach the Decision to Escalate Care Differently: Results of a Maximum Difference Scaling Experiment," Arthritis Care and Research, Vol.63, No.10, pp.1407–1414 (2011).

Appendix: Best–Worst Scaling Question

The subsequent questions were provided to our respondents as the best–worst exercises.

"We will provide seven choice sets consisting of three out of seven items on work-time-related corporate support for employees' child care and upbringing. Please choose you think the most and the least important item in accordance with the example below. Even if you don't have any children, if you finished your own child upbringing, or your workplace don't have any such supports, please answer the questions with thorough consideration, which will enable us to obtain more meaningful result of this survey. Example when you think advancing or delaying start or finish hours is the most and exemption of overtime work is the least important item out of three."

Most Important		Least Important
☑	Advancing or delaying start or finish hours	□
□	Working from home	□
□	Exemption of overtime work	☑

Q.1 Please choose you think the most and the least important item out of three.

Most Important		Least Important
□	Child care leave	□
□	Sick and/or injured child care leave	□
□	Reduced working hours	□

Q.2 How about the three items below?

Most Important		Least Important
□	Child care leave	□
□	Flexible working hours	□
□	Advance or delay starting or finishing times	□

Q.3 How about the three items below?

Most Important		Least Important
□	Child care leave	□
□	Working from home	□
□	Exemption of overtime work	□

Q.4 How about the three items below?

Most Important		Least Important
□	Sick and/or injured child care leave	□
□	Flexible working hours	□
□	Working from home	□

Q.5 How about the three items below?

Most Important		Least Important
□	Sick and/or injured child care leave	□
□	Advance or delay starting or finishing times	□
□	Exemption of overtime work	□

Q.6 How about the three items below?

Most Important		Least Important
□	Reduced working hours	□
□	Flexible working hours	□
□	Exemption of overtime work	□

Q.7 How about the three items below?

Most Important		Least Important
□	Reduced working hours	□
□	Advance or delay starting or finishing times	□
□	Working from home	□

(2013 年 9 月 30 日受付)
(2013 年 12 月 18 日採録)