Application of Multinomial and Ordinal Regressions to the Data of Japanese Female Labor Market

by

Kayo Fujimoto

BS, Kyoritsu Women's College, 1993

MA, University of Chicago, 1998

Submitted to the Graduate Faculty of

Department of Statistics in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2003

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This thesis will be presented

by

Kayo Fujimoto

It will be defended on

December, 12, 2003

Professor Leon Gleser

Professor Roslyn Stone

Professor Patrick Doreian
Thesis Director

Application of Multinomial and Ordinal Regressions to Data of the Japanese Female Labor Market

Kayo Fujimoto

University of Pittsburgh, 2003

This paper describes the application of ordered and unordered multinomial approaches to Japanese Female Labor Market data with the goal of examining how inter-organizational networks linking schools to large corporations supersede labor market processes in the Japanese female labor market. Two sets of response categories were used for a proportional odds model, a non-proportional odds model, and a multinomial logit model. The results from the six combinations of these models were compared in terms of their goodness of model fit. The results showed that the proportional odds assumption was weakly supported, and the Wald test indicates that the violation of proportional odds assumption seems to be limited to a single variable. My study implies that partially proportional odds model would yield a better fit to my female labor market data.

ACKNOWLEDGEMENTS

I greatly appreciate everyone who has helped me, on both a professional and personal level, and ardently acknowledge some of them here. First, I would like to thank my academic advisor, Professor Patrick Doreian, for providing me an opportunity to learn statistical methods formally and supporting me throughout this process as well as through my graduate studies in general. Second, I would thank to other members of my thesis committees: Professor Leon Gleser and Professor Roslyn Stone. Thanks to Professor Gleser for contributing greatly to my learning of statistics in my graduate study, and willingness to consult about problems encountered. Thanks to Professor Stone for improving my comprehension of applied statistical methods in general by leaps and bounds, and for making me think how to apply these methods to social science research.

Finally, a special thanks to my husband, Erik, for his understanding and support in my studies, as well as for some chocolate mousse during the last part of my writing.

TABLE OF CONTENTS

1. Introduction ... 1
2. Background and Literature Review ... 3
3. Theory and Hypotheses ... 6
   - 3.1. School Recommendation System ... 6
     - 3.1.1. School Nomination System ... 6
     - 3.1.2. Reverse Nomination System ... 9
4. Data and Methods ... 12
   - 4.1. Data ... 12
     - 4.1.1. Latent Response Variable ... 13
     - 4.1.2. Independent Variables ... 14
     - 4.1.3. Control Variables ... 14
     - 4.1.4. Descriptive Statistics ... 15
   - 4.2. Statistical Methods ... 16
     - 4.2.1. Cumulative Logit Model ... 18
       - 4.2.1.1. Proportional Odds Model ... 18
       - 4.2.1.2. Non-proportional Odds Model ... 19
     - 4.2.2. Multinomial Logit Model ... 20
     - 4.2.3. Measures of Fit ... 21
       - 4.2.3.1. Deviance ... 22
       - 4.2.3.2. Akaike Information Criteria ... 22
       - 4.2.3.3. Bayesian Information Criterion ... 22
       - 4.2.3.4. McFadden's R^2 ... 23
5. Results ... 24
   - 5.1. Cumulative Logit Model (with 4 ordered response categories) ... 24
     - 5.1.1. Proportional Odds Model ... 25
     - 5.1.2. Non-proportional Odds Model ... 25
   - 5.2. Multinomial Logit Model (with 4 response categories) ... 30
   - 5.3. Cumulative Logit Model (with 5 ordered response categories) ... 35
     - 5.3.1. Proportional Odds Model ... 35
     - 5.3.2. Non-proportional Odds Model ... 36
   - 5.4. Multinomial Logit Model (with 5 response categories) ... 40
   - 5.5. Comparison between Models ... 46
6. Discussion and Conclusion ... 53

BIBLIOGRAPHY ... 55

LIST OF TABLES

Table 1: Descriptive statistics of latent variable of scale of company for students who applied for clerical jobs 14
Table 2: Cross tabulation of scale of company by job-searching method (with 4 response categories) 15
Table 3: Cross tabulation of scale of company by job-searching method (with 5 response categories) 16
Table 4: Brant test of parallel regression assumption (with 4 response categories) 25
Table 5: Estimates for non-proportional odds model (with 4 ordered response categories) 26
Table 6: Estimated equations for non-proportional odds models (with 4 ordered response categories) 27
Table 7: Estimates for multinomial logit model with "small-medium" as baseline category (with 4 response categories) 30
Table 8: Estimated equations for multinomial logit models (with 4 response categories) 32
Table 9: Brant test of parallel regression assumption (with 5 response categories) 35
Table 10: Estimates for non-proportional odds model (with 5 ordered response categories) 36
Table 11: Estimated equations for non-proportional odds models (with 5 ordered response categories) 37
Table 12: Estimates for multinomial logit model with "small-medium" as baseline category (with 5 response categories) 40
Table 13: Estimated equations for multinomial logit models (with 5 response categories) 43
Table 14: Differences between proportional odds results and non-proportional odds results...... 46
Table 15: Differences between non-proportional odds results and multinomial logit results 48
Table 16: Differences between proportional odds results and multinomial logit results 49
Table 17: Differences between 4 response categories and 5 response categories for proportional odds results........ 49
Table 18: Differences between 4 response categories and 5 response categories for non-proportional odds results and multinomial logit model 50

LIST OF FIGURES

Figure 1: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used SN relative to those who used "other" methods (with 4 response categories).. 28
Figure 2: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used RN relative to those who used "other" methods (with 4 response categories).. 29
Figure 3 : The estimated odds of working for a larger company (solid line) rather than a "small-medium" company (dotted line) for those who used SN relative to those who used "other" methods (with 4 response categories).. 33
Figure 4: The estimated odds of working for a larger company (solid line) rather than a "small-medium" company (dotted line) for those who used RN relative to those who used "other" methods (with 4 response categories).. 33
Figure 5: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used SN relative to those who used "other" methods (with 5 ordered response categories).. 38
Figure 6: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used RN relative to those who used "other" methods (with 5 ordered response categories).. 39
Figure 7: The estimated odds of working for a larger company (solid line) rather than a "small-medium" company (dotted line) for those who used SN relative to those who used "other" methods (with 5 response categories).. 44
Figure 8: The estimated odds of working for a larger company (solid line) rather than a "small-medium" company (dotted line) for those who used RN relative to those who used "other" methods (with 5 response categories).. 45

1. Introduction

In sociology, multinomial logit models are often used to analyze occupational attainment where the dependent variable is a set of categories of occupation type (Schmidt & Strauss 1975; Brown et al 1980; Polachek 1981). Here, categorical response variables are assumed not to be ranked. The multinomial logit model does not take into account any ordering of the response categories. If the response categories can be ordered, ordinal models can take this ordering into account. Benefits of the ordinal approach include fewer parameters to estimate as well as the incorporation of the explicit ordering of the response categories. However, the ordered approach is harder to interpret than the multinomial approach and requires a strong assumption of proportional odds (which will be discussed later).

This paper will describe the application of ordinal and multinomial regression approaches to Japanese Female Labor Market data, using theoretically derived variables from the literature. My paper will show that both the ordinal and multinomial regression approaches yield useful substantive information and that combining the two gives the adequate information about the placement mechanisms in the Japanese Female Labor Market. In addition to substantive benefit of using both approaches, I will also present technical issues by considering how many categorizations of a given response variable fit the data better. In so doing, two sets of response categories will be used for each model, and the results from the four combinations will be compared in terms of goodness of model fit.

In the next chapter, I will provide some background information about the Japanese female labor market and a literature review. In the theory section, I will delineate an important mechanism that operates in women's transition from school to work. Then, I generate some hypotheses to be tested by using both the ordered and unordered extensions of logistic regression to multi-category data. These hypotheses will be tested in the chapter of data and methods, and the final chapter describes the results and states some concluding remarks.

2. Background and Literature Review

Within the sociology of work, much research has been focused on how social networks which link members of a society together provide individuals with differential access to such resources as power and information and therefore influence their economic advancement. For instance, network analysts have investigated the embeddedness of network structures in labor markets by examining social context within which information becomes available and trusted not only through market results but also through social networks (Granovetter 1985, 1974).

Research on the Asian labor market has extended network theory by examining institutional networks between school and employer (Rosenbaum & Kariya 1989; Rosenbaum 1990; Lee & Brinton 1996; Brinton & Kariya 1997; Ishida et al 1997; Ishida 1993). These studies have illustrated how institutional linkages between schools and employers can improve labor market processes via trusted information by implying the effect of educational institutions on individual labor market success.

For instance, Lee and Brinton (1996) examined the effect of attendance at an elite university on an individual's success in entering a large firm in South Korea. They found that job seekers in prestigious South Korean universities are provided with social capital, what they call 'school-related job-search methods,' that allows students to get access to the best jobs. Along similar lines, Brinton and Kariya (1997) also examined the institutional embeddedness pattern for university graduates in Japan. They point out the active involvement of university placement offices and companies with the introduction alumni as contact persons for job-seeking seniors, a process that they call 'semi-institutional embeddedness'.

Rosenbaum and Kariya (1989) introduced still other patterns of institutional embeddedness in the Japanese labor market for high-school graduates by examining the institutional linkages between high schools and companies that are based on mutual trust and cooperation in the process of job placement.

They point out that institutional networks provide employers with good indications of students' human capital in terms of grades upon which the selection criterion is based.

However, little research has shown the way institutional linkages operate in *female* labor market. In addition, little is known about how they interact with the forms of cultural capital that is specific to women. Fujimoto (2004) proposed the concept of 'feminine capital' that is defined as "the extent to which individual female students have been absorbed into the dominant culture of patriarchy" (p.6) and therefore have an advantage in the female labor market competition in Japan by supplementing class-based cultural capital models with gender-based cultural capital models. She argues that gender-specific forms of capital -- called feminine capital-- contribute to maintaining social stratification, and consequently class inequality, in Japanese society.

Japan is an ideal site for investigating the way institutional linkages operate and how they interact with gender-based cultural capital because patriarchal values in Japan prevail in both the workplace and in society as a whole. A relatively large number of big Japanese companies use a "predetermined-career track employment system" based mainly on gender. The majority of female workforce in Japan is employed to fill clerical jobs, and these women are known as 'Office Ladies'[1] (OL). In most cases, these full-time female clerical workers are employed *right after* graduating from school until the time of marriage or childbearing. Throughout this paper, I will refer to the labor market consisting of women filling clerical positions as the "OL Labor Market".

In Japan, a corporation's name, which is usually determined by its scale (number of employees), conveys more prestige than does the occupation in which the employee is engaged. This paper defines 'large' Japanese companies as those that have more than one thousand employees and are sufficiently financially stable as to allow them to offer employee benefits such as housing and family allowances. Among large companies, this study specifically defines 'prestigious' Japanese companies as those that have more than ten thousand employees and have a tradition well known not only by Japanese people but

[1] Ogasawara (2001) defines OL as a vernacular expression for mostly young female clerical workers in offices who are recruited immediately from university and two-year colleges. Their main tasks are serving tea to their male colleagues and company visitors, wiping desks, cleaning ashtrays, and answering telephone calls.

also worldwide. This paper assumes that the larger a corporation's scale (number of employees), the greater its prestige.

It must be noted that this study is limited to the "OL labor market" as this is a common career choice for young Japanese women. This market exists primarily to move women from school into suitable marriages, and hence it is unusual (at least from an American perspective). In this sense, the subject of this paper is relatively narrow, as it does not consider employment patterns for women in the other sections of Japanese labor market.

My data comes from job placement reports filled out by both female junior and four-year college graduates from a traditional X Women's College (XWC) in Tokyo. As a service to future students, graduating students at such colleges are required to fill out a form when they find jobs in Japanese companies. These data provides extensive information on students' characteristics, the companies' characteristics, and how the students received access to employers. Data are restricted to only XWC, and acknowledge this raises concerns about generalizability. However, considering XWC's hundred-year-plus tradition, its reputation and its high success rate at placing graduates in large Japanese companies across industries, I believe that these data can still provide a satisfactory picture of the dynamic structure of the "OL Labor Market" in Japan.

3. Theory and Hypotheses

My paper emphasizes the importance of theory because "[t]heory is essential and should be used to the greatest possible extent to define the model to be used" (Raftery 1995a, p.157), and the choice of control variables should be guided by theory as far as possible (Raftery 1995a). This chapter also derives some hypotheses to be tested using statistical methods by introducing the mechanism of a school recommendation system.

3.1. School Recommendation System

This paper defines School Recommendation System (SRS) as all forms of school support that influence an employer's decision in the hiring of a female student. If students use SRS, they have an advantage in getting jobs in large companies that have ties with the school. This paper calls this advantage 'institutional social capital', borrowing Lee and Brinton's (1996) definition of "social resources or ties that can be acquired only through an individual's attendance at a particular university" (p.182). From here on, this paper will introduce two types of SRS to show how inter-organizational networks between a women's college (hereafter, referred to simply as 'school') and Japanese corporations govern labor market processes and facilitate the employment of young Japanese women in the position of OL.

3.1.1. School Nomination System

The first type of SRS is the system of School Nomination (SN) where employers allocate job offers to individual schools that have a long-term relationship with a corporation. This allocation is called "fixed quarters" (*waku*), by which companies designate the number of students they hire from a specific school. The criterion on which *waku* is based is "the function of schools in preparing youth psychologically for work" (Bowles & Gintis 1976). In the case of the "OL Labor Market", since the employer's aim of hiring women is not only their labor force participation but also candidates for marriage partners of their male employees (Fujimoto 2004), schools teaching appropriate gender roles tend to be selected. This is one of the main reasons why women's colleges are preferable for being allocated into "fixed quarters" for Japanese companies.

In the system of SN, a school selects the students who are most likely to be hired by companies at the first stage. In this sense, the task of screening appropriate women who meet the corporation's requirement is delegated to the school, and companies then select students from among those who are put forward by the school. It must be pointed out that there is no formal or written contract between the institutions involved. In addition, the system of SN is in no way equal to obtaining a job offer from a corporation. Therefore, it is quite possible that employers do not hire school nominees even if students have passed the competitive school screening. However, the system of SN is still advantageous for the school because their students can get access to employers and obtain information about job opportunities before students from other schools that do not have institutional ties with the given employers. Since companies hire appropriate students on a first come, first served basis, students using the SN system have a significant advantage over those who do not use (or are not able to use) it.

There are two main benefits large Japanese firms enjoy by relying on the SN system. The first is that employers can secure dependable sources of labor in a way that reduces their recruitment costs. Because of the continuity of a trusting relationship with particular schools, companies do not have to worry about a shortage of dependable labor. The second benefit is that employers can recruit students of similar quality every year from certain schools. This is different from the corporate practice of recruiting male university graduates of similar quality every year by way of a semi-institutional alumni network (Brinton & Kariya 1997). In the 'OL Labor Market', because the position concerned here is relatively less important for companies in terms of work-related human capital, companies are more concerned with securing enough (perhaps mediocre) female graduates who are up to a certain standard than with having the best students possible.

Schools also benefit from the SN system by gaining a stable supply of job placements from specific large Japanese firms. From an institutionalist point of view (Meyer & Rowan 1977), this continuity means schools acquire legitimacy by conforming to the social norm that one of the major roles of a school is to place as many students as possible with large Japanese firms. By successfully placing its students in the "OL Labor Market," this social recognition helps a school in establishing its reputation with large Japanese companies. If graduates have a good chance of getting jobs in large Japanese firms as OL, the probability of their marrying promising Japanese businessmen is also high. Since parents are also stakeholders in their daughter's choice of school (Brinton 1992), a reputation of high placement rates will also be recognized by parents in the form of tuition payments. Consequently, what matters for a school is the high placement rate of its graduates with prestigious Japanese companies, as opposed to how happy the graduates actually will be with their employers. For this reason, this system works only for those who fill jobs requiring relatively less important work-related human capital for companies.

Since the task of screening appropriate women is delegated to individual schools, they select students that are as desirable as possible to compete with the nominees from other schools – thus maximizing the number of their students being placed in prestigious Japanese companies. The job placement office at a school is in charge of deciding which students are most likely to be hired by the companies. In the system of SN within a school, since teachers are not involved in screening process, it is more likely that the criterion of selection of the women does not depend on academic achievement. If employers do not care about academic achievement, as the Bowles-Gintis (1976) hypothesis suggests, a school is more likely to nominate appropriate students who meet the company's "other" requirements.

Fujimoto (2004) reports that 'feminine capital' is composed of three feminine characteristics: dependency on parents (living with parents rather than living alone)[2]; ability to manage household tasks (majoring in home economics)[3]; and youthfulness (junior college graduate rather than four-year college graduate)[4] were all are important for getting clerical jobs at large Japanese companies[5]. Her concept of

[2] Since dependent daughters have little influence in the household, they tend to be seen by men as obedient and thus suited for jobs ancillary to those of men.
[3] Since large Japanese companies hire women not only as workers but also as potential marriage partners for promising male employees, they regard women who are good at managing household tasks as desirable employees.
[4] From the perspective of an employer, two fewer years of education ensure that junior college students have a longer commitment to their company when they quit jobs for marriage.
[5] She defined "large" companies as having more than a thousand employees.

gender-based cultural capital is especially important in the context of institutional linkages for understanding the contribution of educational institutions to women's aspirations to work as OL. Since the number of its graduates who are placed in prestigious Japanese companies determines the reputation of a given school and influences its survival, the larger the companies (which indicate more prestige) in which a school succeeds in placing their students, the greater the reputation that the school gains. For this reason, I assert that schools use the criteria of feminine characteristics to screen desirable students to maximize the number of students hired at prestigious companies, and hence to maintain the institutional ties. In this way, feminine capital seems to be quite important in determining women's employment outcome, and hence my study control variables describe feminine capital.

3.1.2. Reverse Nomination System

The second type of SRS that maintains institutional linkages is the system of Reverse Nomination (RN). Under this system, unlike the SN, a school neither screens nor nominates students initially. Instead, the students collect information about companies by contacting Old Girls (OG), whose names appear in reports of job placement and in this way get access to employers without an initial nomination from the school. If they pass the final stage of interviews, such students are then asked by the company to get their school's nomination. After turning in their nomination to that corporation, students then obtain their actual job offer. In this sense, the task of screening desirable women is delegated to the companies, and the school then simply issues the formal letter of nomination to those who have been pre-screened by companies, and hence is called 'reverse' nomination. As in the case of SN, there is no formal or written contract between the institutions. However, unlike the system of SN, any student who has been pre-screened by the company can obtain a RN.

Under the system of RN, students can only request the school to issue its nomination once. According to the personnel at the job placement office at XWC, students tend not to ask for a nomination to the first company that makes them an offer. Instead, students tend to carry on looking for even more prestigious companies, and thus the RN system works to avoid student indecision once the nomination is made by trying to force decisions by students. Schools and companies agree that the school will issue a nomination for a given student only once, in exchange for guaranteeing the job to those students who turn in their nomination. From students' point of view, it is a waste of time to look for another job after accepting one, as the school will not issue any additional nominations after the first.

As in the system of SN, both school and companies can enjoy benefits from the system of RN. From the perceptive of a company, the SN system serves to reduce recruitment costs by preventing students' indecisive behavior and twisting their arms to make an immediate decision. From the perspective of the school, the RN system serves to establish its reputation by placing as many students as possible at prestigious Japanese companies without any effort by the school in screening nominees. It must be observed that the RN system actually deprives students of their opportunities of looking for other companies that may better suit them. By forcing their students to decide immediately after being screened by a company through the system of RN, the school maintains institutional linkages with employers. This system is more beneficial to the school because it requires less effort on the college's part than the SN system, but the final result to its reputation is equivalent to that of SN.

To summarize, the difference between the SN and the RN systems is thus the timing of the student's being nominated by the school. Under the SN system, students are nominated by the school *before* getting access to employers, whereas under the RN system students are nominated *after* getting access to employers and securing an informal job offer. More importantly, another difference is the degree of active involvement by the school in the screening process. In the SN system, the school is involved more actively with selecting nominees, whereas in the RN system the school is not involved the selection process (though it does introduce the 'Old Girls' to the new graduates in order to get access to employers). For the former, companies trust schools to screen candidates based on feminine characteristics, and to provide desirable candidates from whom they hire. For the latter, companies select desirable women who are affiliated with a specific school that have ties to that company. This suggests that the companies would not choose these women if they did not have institutional social capital in the first place.

In this way, the placement offices at the schools play an important role in shaping the hiring process since students actually cannot get a formal job offer without the school's involvement. A given school has the goal of increasing its reputation by actively catering to the accepted social norms. However, many schools can only prosper if they possess ties to Japanese companies with high prestige. Considering that schools can establish their reputations by institutional linkages with prestigious companies, they are willing to help Japanese companies in reducing recruitment costs either by being in charge of screening desirable nominees (SN) or collaborating with them in forcing their students to decide immediately (RN). Therefore, the first two hypotheses are:

1. *Students who are recommended by their school by the SN system are more likely to receive an offer from the higher prestige Japanese companies as an OL than students who used other job-searching methods such as public advertisement.*
2. *Students who are recommended by their school by the RN system are more likely to receive an offer from the higher prestige Japanese companies as an OL than students who used other job-searching methods such as public advertisement.*

The next two hypotheses are concerned with a school's specific strategies to achieve its goals efficiently by institutional linkages. Since the system of SN requires a given school to screen suitable students, it is plausible that schools would focus only on the "high-return", "prestigious" Japanese companies – paying little attention to the "smaller" (less prestigious, but still large in size) suitable companies simply because of limited resources. Since the system of RN requires far fewer resources on the college's part, it is plausible that schools would try to catch these "smaller" fish using this method instead of the more costly SN. Therefore, my third hypothesis is:

3. *Schools will focus their placement efforts primarily on SN for the greater prestigious companies, and rely on RN to bring them placements with the lesser prestigious companies.*

In the next section, I will test these three hypotheses using statistical models.

4. Data and Methods

In this section, I will show data description, specifications of variables with basic statistics, and statistical methods that I used in this study.

4.1. Data

I used the data collected from X Women's junior and four-year college that has a tradition of institutional ties to large Japanese companies. X Women's College (XWC) is located in Tokyo, Japan, and was founded with the aim of providing women with specialized knowledge and skills to foster their independence and social self-reliance. XWC is appropriate for my research because it has both a junior and a four-year college. Since the level of tradition and the degree of social reputation are identical for both the junior and the four-year college, my research can compare the two with respect to the probability of job placement at large Japanese companies. When both junior and four-year college students obtain jobs, they are required to fill out a form that reports their job placement outcome and turn it in to the job placement office. The job placement report includes information about students (name, home economic major, staying with parents, qualification, and hobbies), as well as information about companies from which students obtain informal job offers (location, business funds, annual sales, number of employees, type of industry). The job placement report also includes information about how students applied to corporation, how students commute, and what kind of occupation the students applied for.

This research uses job placement reports of individual applicants at XWC (one report per student) as a unit of analysis. There are 787 valid cases of students who applied for clerical work. This study is limited to students who applied for clerical jobs because the scope of research is "OL labor market"[6]. The availability of data restricted my study to 1993, a year of economic recession in Japan. The significance

[6] In addition to this, from technical point of view, there are few cases that used SRS to apply for such positions, which affects the estimation.

of 1993, as the year of my study, is that it well reflects the selection mechanisms in "OL labor market," as opposed to years of the Japanese bubble economy that would confound selection mechanisms with a surplus of available positions and a dearth of applicants.

Although there might be some limitations for generalizing results from data based on job placement reports from XWC and four-year colleges, these data cover almost all the large Japanese companies across industries in Tokyo area. Similar firms may be likely to engage in similar hiring practices for students from other junior or four-year colleges with similar linkages, although this could depend on linkages to the other schools.

4.1.1. Latent Response Variable

Employment outcome: This response variable is the latent variable of <u>scale of company</u> where individual students received their job (for clerical jobs), which is determined by the number of employees. My study uses two ways of ordering. The first way is to order the company scale into 4 categories of "small-medium" (<1000), "large" (1000-4999), "very large" (5000-9999), and "huge" (≥10000). The second way is to order it into 5 categories of "small-medium" (<1000), "large" (1000-4999), "very large" (5000-9999), "huge" (10000-49999), and "giant"(≥50000). Essentially, the last category of "huge" was split into two categories of "huge" and "giant," for a total of 5 response categories. The descriptive statistics for these ranges for students who applied for clerical jobs[7] are shown in Table 1.

[7] For non-clerical jobs, there are no counts in the response category of "very large" with SN, and "small-medium" with RN. In addition, there is only one count in the categories of "small-medium" with SN, "large" with SN, and "huge" with SN, and "large" with RN. The other count is 2 for the category of "huge" with RN and 3 counts for the category "very large" with RN. These few counts for non-clerical jobs with SRS indicate that the SRS is rarely applied to non-clerical jobs.

Table 1: Descriptive statistics of latent variable of scale of company for students who applied for clerical jobs

4 response categories				5 response categories			
Scale of company	counts	percent	Cumulative percent	Scale of company	counts	percent	Cumulative percent
"small-medium" (<1000)	160	20.3	20.3	"small-medium" (<1000)	160	20.3	20.3
"large" (1000-4999)	177	22.5	42.8	"large" (1000-4999)	177	22.5	42.8
"very large" (5000-9999)	150	19.1	61.9	"very large" (5000-9999)	150	19.1	61.9
"huge" (≥10000)	300	38.1	100.0	"huge" (10000-4999)	241	30.6	92.5
				"giant" (≥50000)	59	7.50	100.00
total	787	100.0		total	787	100.0	

4.1.2. Independent Variables

Use of SRS: Job placement report includes the information of how students received the information about their corporation and how they applied for the jobs. I created the variable (SRS) that is composed of three categories of the use of SN, the use of RN, and other job-searching methods such as advertisements, family connections, or the public employment security office. In the analysis, two dummy variables of SN and RN will be used by treating other methods as a baseline category.

4.1.3. Control Variables

Feminine capital: The question of how students commute was used as an indicator of the first form of feminine capital, expressed by students' dependency on their families. I used a dummy variable (parents), indicating commuting from parents' or other family member's home (coded as 1) versus commuting from an apartment where only students live (coded as 0). The second form of feminine capital, expressed by women's ability to manage household tasks, was measured by using the department where the students belong. Since XWC was composed of only two departments in 1993 (home economics and literature and arts), the dummy variable (home) that indicates majoring in home economics (coded as 1) versus majoring in literature and arts (coded as 0) is used. In addition, I used a dummy variable (junior) indicating junior college (coded as 1) versus four-year college (coded as 0) in order to measure the third form of feminine capital.

4.1.4. Descriptive Statistics

As some basic information, Table 2 and Table 3 provide information about the distribution of dependent and independent variables for 4 response categories and 5 response categories respectively.

Table 2: Cross tabulation of scale of company by job-searching method (with 4 response categories)

Company size / Method	"small-med" (<1000)	"large" (1000-4999)	"very large" (5000-9999)	"huge" (≥10000)	Total
School nomination	2	4	10	53	69
	(2.90)	(5.80)	(14.49)	(76.81)	(100.0)
Reverse nomination	10	42	46	96	194
	(5.15)	(21.65)	(23.71)	(49.48)	(100.0)
Other methods	144	129	93	146	512
	(28.13)	(25.20)	(18.16)	(28.52)	(100.0)
Total	156	175	149	295	775
	(20.13)	(22.58)	(19.23)	(38.06)	(100.00)

( ) represents row-wise percentage

The result of a chi-square test shows that there is a significant association between these two variables (chi-square=107.543 with df=6; p-value<0.0001). We can also observe that the count of students who got an offer from a "small-medium" company by school nomination (2) and the count of students who got an offer from a "large" company by school nomination (4) are quite small.

Table 3: Cross tabulation of scale of company by job-searching method (with 5 response categories)

Company size Method	"small-med" (<1000)	"large" (1000-4999)	"very large" (5000-9999)	"huge" (10000-49999)	"giant" ≥50000	Total
School nomination	2	4	10	32	21	69
	(2.90)	(5.80)	(14.49)	(46.38)	(30.43)	(100.0)
Reverse nomination	10	42	46	74	22	194
	(5.15)	(21.65)	(23.71)	(38.14)	(11.34)	(100.0)
Other methods	144	129	93	132	14	512
	(28.13)	(25.20)	(18.16)	(25.78)	(2.73)	(100.0)
Total	156	175	149	238	57	775
	(20.13)	(22.58)	(19.23)	(30.71)	(7.35)	(100.00)

( ) represents row-wise percentage

The result of a chi-square test shows that there is a significant association between these two variables (chi-square= 143.499 with df=6; p-value< 0.0001). We can observe that the count of students who got an offer from a "giant" company by the other methods is small relative to the count of who got an offer in other categories of company scale.

4.2. Statistical Methods

My study applies both ordinal regression and multinomial regression approaches to model the association between the scale of company and the characteristics of OLs that it hires. In the study of occupational attainment, Miller and Volker (1985) estimated both a multinomial logit model and an ordered probit model.[8] They used the multinomial logit model to model occupational outcome by relating subsequent occupations to those of first jobs. They then derived estimates from an ordered probit model[9] by taking ordering of occupations into consideration, which enabled them to interpret vertical (as opposed to horizontal) mobility, and therefore confirm prior notions of job hierarchies. In order to address the

[8] They used data from a random sample of Australian males collected for the Social Mobility in Australian Project in 1973.

[9] In probit model, link function is the standard normal cumulative density function. My study uses logit link. Except for possibly in the tails of the distribution, there are few differences between probit and logit models.

issue “whether the multinomial logit or ordered probit model is more appropriate for the study of occupational attainment” (p.200), they compared the performance of these models. They found that the estimated predictions from the multinomial logit model and those from the ordered probit models vary significantly, and they caution that policy recommendations based on these models will be sensitive to the modeling strategy employed.

In terms of comparing both multinomial logit and ordered probit models, Miller and Volker argue that the choice of modeling strategies should be based on the ability of the model to predict occupational structures. There are mainly three findings from their comparisons. First, they report that the multinomial logit model has more predictive power for their data than that of the ordered probit model, using maximum-probability technique[10]. However, one of the possible reasons for their result would be that their ordinal probit model would not satisfy proportional odds assumption. Second, when they predict categories with relatively small numbers of observations, multinomial logit model also gives the better performance, which is consistent with the result obtained from Powers, Marsh, Huchfeldt, and Johnson (1978). As a final finding, when they use the ordered probit model, they used two separate orderings that involve status-attainment scores and income comparison, and this reveals substantial differences in estimated probabilities in the category where ranking varied little (administration and professional and technical occupations). They attributed this result to the relatively small size of these particular groups. Based on this result, they recommend experimenting with a number of ordering variables to deal with the issue of the effect of the various ordering schemes on outcomes. However, if the ordering is not unambiguous, one probably should not be using ordinal models.

My study also employs both ordered logit model and multinomial logit model to compare the performance of these models. In addition to comparing multinomial logit models to ordered logit models, my study also examines the performance of a different number of response categories for each model to extend the element of model selection. The aim of my study is to show a useful application of ordinal

[10] By using maximum-probability technique, they assigned each individual to the occupational category for which he or she has the highest estimated probability of membership.

regression and multinomial regression approaches, including modeling decisions with respect to how many categories (or the number of thresholds) of a given continuous response variable fit data better when we treat it as categorical, and apply it to multivariate discrete models in social science research. The reason for treating a continuous response variable as categorical is to create a substantively meaningful categorical variable so that the results can be interpreted easily.

4.2.1. Cumulative Logit Model

Ordinal Regression Model (ORM) assumes that the categories of an ordinal response can be ranked, but the distances between the categories are unknown. My study employs cumulative logit model to model cumulative probabilities of employment outcome. Since cumulative logits are correlated asymptotically, simultaneous estimation is required. The motivation to use ordinal regression model is to avoid the assumption that the distances between response categories are equal as opposed to ordinary regression[11]. In ORM, observed response categories (y_i) are tied to the latent variable (y_i^*) by a measurement model that divides y_i^* into J ordinal categories so that J-1 cutpoints[12] are estimated.

My discussion limits to two versions of a cumulative logit model. The first is a proportional odds model (McCullagh 1980), and the second is non-proportional odds model (also called a generalized ordered model) that was proposed by Fu (1998).

4.2.1.1. Proportional Odds Model

Proportional odds model means that covariates have the same effect on the odds as the response variable has at any dividing point by regarding different values of covariates as shifting the response distribution to the right (or left) without changing its spread or shape. In the proportional odds model, the cumulative logits model the effect of covariates on odds of response below or equal to the cutpoint *m* in the latent variable. First, the odds that an outcome is less than or equal to *m*, versus being greater than *m,* given **x** is defined as follow:

$$\Omega_{\leq m|>m}(\mathbf{x}) \equiv \frac{P(y \leq m \mid \mathbf{x})}{P(y > m \mid \mathbf{x})} \qquad for\ m = 1, \cdots, J-1$$

[11] Detailed discussion about ordinal regression model is found in Long & Freese (2001).
[12] This is also called thresholds.

The log of the odds is defined as follows:

$$\log \Omega_{\leq m|>m}(\mathbf{x}) = \tau_m - \mathbf{x\beta} \qquad \text{where } \tau_m \text{ is a cutpoint of m}$$

So,

$$P(y = m \mid \mathbf{x}) = F(\tau_m - \mathbf{x\beta}) - F(\tau_{m-1} - \mathbf{x\beta})$$

And

$$P(y \leq m \mid \mathbf{x}) = F(\tau_m - \mathbf{x\beta})$$

In this model, we cannot estimate J-1 cutpoints as well as an intercept; therefore the intercept is set to zero. Also, it is important to note that $\mathbf{x\beta}$ is subtracted rather than added in $F(\tau_m - \mathbf{x\beta})$, which complicates the interpretation of the log odds ratio.

Proportional odds model employs a strict assumption that odds ratio does not depend on the cutpoint, and therefore we need to test the proportional odds assumption, which is also called "parallel regression assumption." My study uses two methods of (1) approximate likelihood ratio test (Wolfe & Gould 1998), and (2) Wald test (Brant 1996). Approximate likelihood ratio test of proportionality of odds across response categories ominously tests that the coefficients for all variables are simultaneously equal. It compares the log likelihood from proportional odds assumption to the log likelihood obtained from pooling J-1 binary models, adjusting for the correlation between the binary outcomes defined by $y \leq m$. On the other hand, Wald test tests the "parallel regression assumption" for each variable individually. In both tests, if p-value is small, the parallel regression assumption can be rejected. I will start with proportional odds model and test proportional odds assumption, and then I will fit non-proportional odds model.

4.2.1.2. Non-proportional Odds Model

The Non-proportional odds model relaxes the assumption of proportional odds by allowing the effect of the covariates to vary with the point where the categories of the response variable are dichotomized. Fu's (1998) parameterization of cumulative logits models the effect of covariates on odds of response above or equal to a specific cutpoint in the latent variable[13]. In addition, he parameterizes his model in terms of regression parameters plus an intercept (rather than cutpoints) for outcomes 0, 1, 2, · · ·, m:

$$P(Y < k) = F(-XB_k) \qquad for\ k = 1, \cdots, m$$

So the model is:

$$\log \frac{P(Y \geq k)}{P(Y < k)} = XB_k \qquad for\ k = 1, \cdots, m$$

In this model,

$$\begin{aligned} &P(Y = 0) = F(-XB_1) \\ &P(Y = j) = F(-XB_{j+1}) - F(-XB_j) \qquad for\ j = 1, \cdots, m-1 \\ &P(Y = m) = 1 - F(-XB_m) \end{aligned}$$

Fu's parameterization allows me to easily interpret the results and remove the assumption of proportional odds from the equations. My study assumes that distribution of the latent variable is logistic and assumes totally non-proportional odds. Totally non-proportional odds model is basically equal to multinomial logit model, but the parameterization is different, which will be discussed in the next section.

4.2.2. Multinomial Logit Model

In addition to fitting ORM whose assumption is that the values of a response variable can be ordered, this study fits multinomial regression model that models qualitative categorical response variables having more than 2 categories. Multinomial logit models are generalization of logit models for binary responses and fitting the generalized logit model requires simultaneously satisfying the J-1 equations that specify the model. Multinomial logit model is defined:

$$\begin{aligned} &\log\left(\frac{\pi_{ij}}{\pi_{i1}}\right) = x_i \beta_j \qquad for\ j = 2, \cdots, J \\ &\qquad\qquad\qquad\qquad\qquad\quad i = 1, \cdots, N \\ &where\ \pi_{ij}\ is\ P(Y = j \mid \mathbf{x}) \end{aligned}$$

So, probability is obtained by:

---

[13] A macro of his model is available in the statistics software STATA.

$$\pi_{ij} = \frac{\exp\{x_i\beta_j\}}{\sum_{j=1}^{J}\exp\{x_i\beta_j\}}$$

In this model,

$$P_i(Y=0) = \frac{1}{1+\sum_{j=1}^{J}\exp\{x_i\beta_j\}}$$

$$P_i(Y=j) = \frac{\exp\{x_i\beta_j\}}{1+\sum_{j=1}^{J}\exp\{x_i\beta_j\}}$$

$$P_i(Y=J) = \frac{\exp\{x_i\beta_j\}}{1+\sum_{j=1}^{J}\exp\{x_i\beta_j\}}$$

$$\begin{aligned} &where\ j = 2,\cdots,J-1 \\ &\qquad\quad i = 1,\cdots,N \end{aligned}$$

Multinomial logit models assume response counts at each level of covariate combination as multinomial and multinomial counts at different covariate combinations are independent. The benefit of using multinomial logit model is that it models the odds of each category relative to a baseline category as a function of covariates, and it can test the equality of coefficients even if confounders are different unlike in the case of pair-wise logistics where testing equality of coefficients requires assumptions about confounder effects.

4.2.3. Measures of Fit

My study compares the performance of the proportional odds model, the non-proportional odds model, and the multinomial logit models using some measures of model fit. In addition, my study also compares performance due to changing the number of cutpoints (or changing the number of response categories) for each model. As the number of cutpoints (or equations) increases, the number of parameters also increases, and hence I expect that the increased-cutpoint models to fit better. My study uses four scalar measures of model fit: (1) Deviance, (2) Akaike information criterion, (3) The Bayesian information criterion, and (4) McFadden's R^2. There is no convincing evidence that selection of a model that maximizes the value of a given measure necessarily results in a model that is optimal in any sense other than the model having a larger (or smaller value) of that measure (Long & Freese 2001, p.80). However, it is still helpful to see any differences in their level of goodness of fit, and hence provide us some guidelines in choosing an appropriate model.

4.2.3.1. Deviance

As a first measure of model fit, I use residual deviance (D) for model M_k, which is defined as follows:

$$D(M_k) = 2\sum_{i=1}^{n}\sum_{j=1}^{J} I(y_i = j)\log\left(\frac{I(y_i = j)}{\hat{p}_{ij}}\right)$$

where $\hat{p}_{ij}$ is a predicted value and y_i is an observed value for $i=1,\cdots, N$

4.2.3.2. Akaike Information Criteria

As a second measure of fit, I use Akaike (1973) information criterion that is defined as follows:

$$AIC = \frac{-2\ln \hat{L}(M_K) + 2P}{N}$$

where $\hat{L}(M_k)$ is the maximum likelihood of the model M_k and P is the number of parameters in the model. A model having smaller AIC is considered the better fitting model.

4.2.3.3. Bayesian Information Criterion

As a second measure of model fit, I use Bayesian Information Criterion (BIC) (Raftery 1995a) as a simple and accurate large-sample approximation, especially if there are as few as about 40 observations (Raftery 1993). For my study, BIC is defined as follows:

$$BIC_k = D(M_k) - df_k \ln N$$

where $D(M_k)$ is deviance for model M_k and df_k is the degrees of freedom associated with the deviance. The more negative BIC_k is, the better the fit. Raftery (1995a) also provides guidelines for the strength of evidence favoring M_2 against M_1 based on a difference in BIC as follows:

$\lvert BIC_1\text{-}BIC_2 \rvert$	Evidence
0-2	Weak
2-6	Positive
6-10	Strong
>10	Very Strong

4.2.3.4. McFadden's R^2

As a last measure of fit, my study uses McFadden's adjusted R^2. This measure is also known as the "likelihood-ratio index," which compares a full model with all parameters (M_{Full}) with the model just with intercept ($M_{Intercept}$) and defined as:

$$\overline{R}^2_{McF} = 1 - \frac{\ln \hat{L}(M_{Full}) - K^*}{\ln \hat{L}(M_{Intercept})}$$

where K^* is the number of parameters.

5. Results

There are two parts in my analyses. The first part uses 4 categories of a response variable, and fits both ordered logit and multinomial logit models. In ordered logit model, my observed response categories of number of employees (y_i) are tied to the latent variable (y_i^*) by a measurement model that divides y_i^* into 4 ordinal categories so that 3 cutpoints are estimated as follows:

$$y_i = \begin{cases} 1 \Rightarrow \text{"small} - \text{medium"} & \text{if} \quad 0 \le y_i^* < 1000 \\ 2 \Rightarrow \text{"large"} & \text{if } 1000 \le y_i^* < 5000 \\ 3 \Rightarrow \text{"very large"} & \text{if } 5000 \le y_i^* < 10000 \\ 4 \Rightarrow \text{"huge"} & \text{if} \quad y_i^* \ge 10000 \end{cases}$$

The second is to create 5 categories of a response variable, and fit both ordered logit and multinomial logit models. In ordered logit model, measurement model divides y_i^* into 5 ordinal categories so that 4 cutpoints are estimated as follows:

$$y_i = \begin{cases} 1 \Rightarrow \text{"small} - \text{medium"} & \text{if} \quad 0 \le y_i^* < 1000 \\ 2 \Rightarrow \text{"large"} & \text{if } 1000 \le y_i^* < 5000 \\ 3 \Rightarrow \text{"very large"} & \text{if } 5000 \le y_i^* < 10000 \\ 4 \Rightarrow \text{"huge"} & \text{if } 10000 \le y_i^* < 50000 \\ 5 \Rightarrow \text{"giant"} & \text{if} \quad y_i^* \ge 50000 \end{cases}$$

After fitting 4 models (2 ordinal logit models and 2 multinomial logit models), my study calculates measures of fit to see which model is more adequate.

5.1. Cumulative Logit Model (with 4 ordered response categories)

My study uses cumulative logit model in order to test the first hypothesis (*Students who are recommended by their school by the SN system are more likely to receive an offer from the higher prestige Japanese companies as an OL*) and the second hypothesis (*Students who are recommended by their school by the RN system are more likely to receive an offer from the higher prestige Japanese* *companies as an OL*). I performed a regression analysis on SRS as an explanatory variable, using the dependent variable scale of company and controlling for feminine capital of parents, home, and junior. This was done for both the proportional and the non-proportional odds model.

5.1.1. Proportional Odds Model

First, I fit the proportional odds model and test the "parallel regression assumption." The result of approximate likelihood ratio test shows that the parallel regression assumption is on the verge of being rejected at the 0.05 level (chi-square= 7.98 with df=10, p-value= 0.055). Brant's Wald test allows to us to specify which variable violate parallel regression assumption and the result show as follows in Table 2:

Table 4: Brant test of parallel regression assumption (with 4 response categories)

Variables	Chi-square	P < chi-square	df
All	15.10	0.128	10
School Nomination	0.81	0.665	2
Reverse Nomination	9.70	0.008	2
Parents	0.27	0.872	2
Home	2.40	0.301	2
Junior	2.86	0.240	2

The Brant test shows that the largest violation is "reverse nomination," which suggests that a possible violation of proportional odds assumption may be almost entirely due to this variable. Since one of my explanatory variables of interest violates proportional odds assumption, my study also tried a non-proportional odds assumption.

5.1.2. Non-proportional Odds Model

The results of fitting non-proportional odds model[14] and the estimated equations are shown in Table 5 and Table 6, respectively.

[14] I also fit the same model using robust variance estimate. Since the results do not differ by much, I select model-based variance estimates.

Table 5: Estimates for non-proportional odds model (with 4 ordered response categories)

Scale of company		b	S.E.	exp{b}	95% C.I. of odds ratio	
Equation 1 ("small-med")					Lower	Upper
SRS	SN	2.11**	0.73	8.25	1.95	34.81
	RN	1.78***	0.39	5.91	2.73	12.77
Feminine capital	Parents	1.00**	0.37	2.71	1.32	5.54
	Home	0.58**	0.21	1.78	1.17	2.71
	Junior	0.48*	0.21	1.62	1.07	2.47
	Constant	-0.29	0.37			
Equation 2 ("large")						
SRS	SN	2.15***	0.45	8.60	3.59	20.59
	RN	0.90***	0.21	2.45	1.63	3.70
Feminine capital	Parents	0.78*	0.35	2.18	1.09	4.34
	Home	0.35*	0.16	1.42	1.03	1.96
	Junior	0.40*	0.18	1.49	1.05	2.12
	Constant	-1.12**	0.36			
Equation 3 ("very large")						
SRS	SN	1.86***	0.32	6.43	3.41	12.14
	RN	0.59**	0.20	1.80	1.22	2.65
Feminine capital	Parents	0.66	0.40	1.94	0.88	4.28
	Home	0.54**	0.17	1.71	1.23	2.37
	Junior	0.61**	0.20	1.84	1.26	2.70
	Constant	-2.05***	0.42			

* $p<0.05$; ** $p<0.01$; *** $p<0.001$; ( ) represents cutpoint

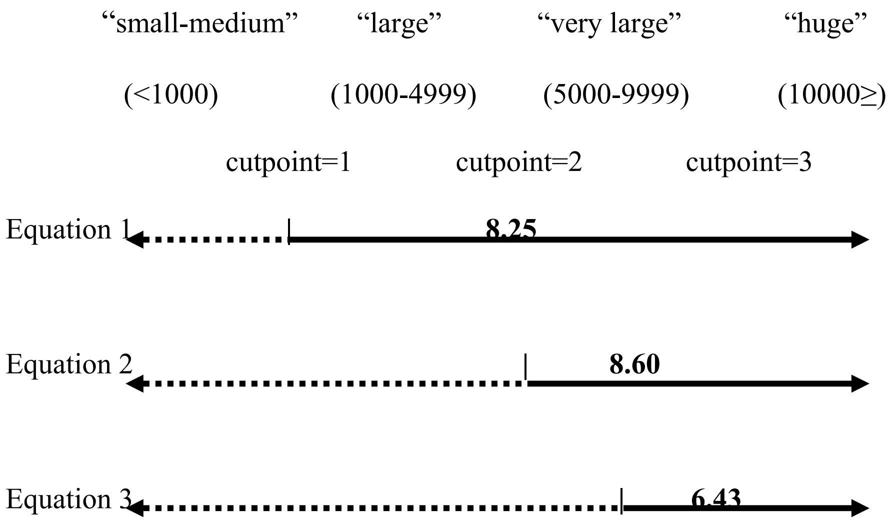
Table 6: Estimated equations for non-proportional odds models (with 4 ordered response categories)

Estimated non-proportional odds model
Equation 1: cutpoint = 1 (<1000)
$\log\{\Pr(Y \geq 1 \mid x) / \Pr(Y < 1 \mid x)\} = -0.29 + 2.11(SN) + 1.78(RN) + 1.00(parents) + 0.58(home) + 0.48(junior)$
Equation 2: cutpoint = 2 (1000-4999)
$\log\{\Pr(Y \geq 2 \mid x) / \Pr(Y < 2 \mid x)\} = -1.12 + 2.15(SN) + 0.90(RN) + 0.78(parents) + 0.35(home) + 0.40(junior)$
Equation 3: cutpoint = 3 (5000-9999)
$\log\{\Pr(Y \geq 3 \mid x) / \Pr(Y < 3 \mid x)\} = -2.05 + 1.86(SN) + 0.59(RN) + 0.66(parents) + 0.54(home) + 0.61(junior)$

The main effects of SN on receiving an offer from a larger company for clerical jobs for given feminine characteristics are significant across all equations. The sign of all coefficients of the SN are positive, which indicate that the odds of receiving an offer from a larger company rather than a smaller company is higher for students who received nomination by the school relative to those who used other methods. This result supports my first hypothesis.

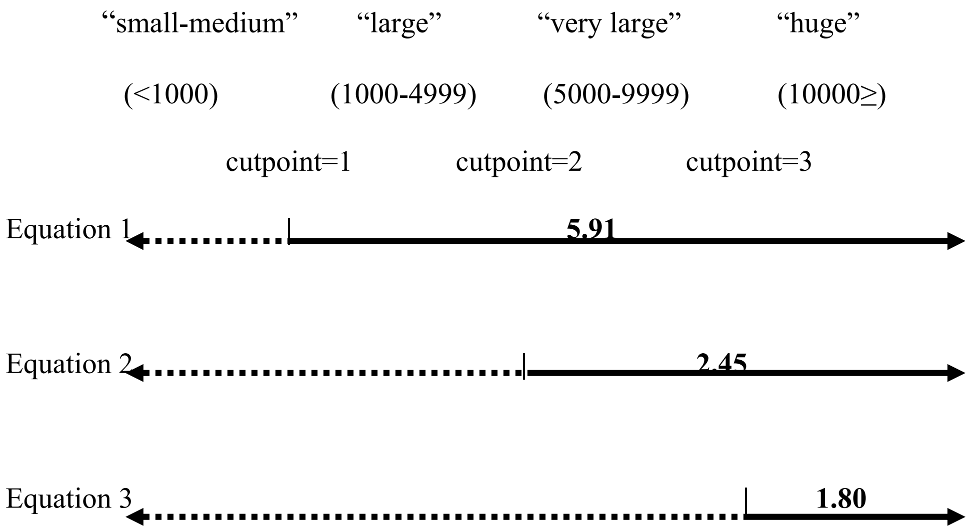
Similarly, the main effects of RN on receiving an offer from a larger company for clerical jobs for given feminine characteristics are also significant across all equations. Again, the sign of all coefficients of the SN are positive, which indicate that the odds of working for a larger company rather than a smaller company is higher for students who received a RN by the school relative to those who used other methods. This result supports my second hypothesis. The graphical presentation of these odds for each equation is represented in Figure 1 (SN) and Figure 2 (RN).

Figure 1: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used SN relative to those who used "other" methods (with 4 response categories)



The odds of receiving an offer from a company of "large" or greater (cutpoint = 1) is 8.25 for students who used the SN system relative to those who used other methods for given feminine characteristics, which is about the same as the odds of receiving an offer from a "very large" or greater company (cutpoint = 2). In addition, the odds of receiving an offer from a "huge" company (cutpoint = 3) is 6.43 for students who used the SN system relative to those who used other methods for given feminine characteristics. I used Wald test to test the null hypothesis that the coefficients are the same in the two equations. The results of testing the proportionality of coefficient of the SN show that slopes for SN term are proportional across outcome categories.

Figure 2: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used RN relative to those who used “other” methods (with 4 response categories)



The odds of receiving an offer from a “large” company or greater (cutpoint = 1) is 5.91 for students who used RN relative to those who used other methods. In addition, it must be noted that for students who used the RN system relative to those who used other methods, the odds of receiving an offer from a “very large” or greater company (2.45), and the odds of receiving an offer from a “huge” company (1.80) become relatively smaller (but still significant).

I used Wald test to test equality of coefficients of RN in the two equations. The results show that the estimated coefficients between equation 1 and equation 2 (chi-square=5.65, p=0.02), and between equation 1 and equation 3 (chi-square=8.98, p=0.003) are significant. It is important to point out that if “large” is included with “small-medium” as comparison group, the odds dropped from 5.91 to 2.45, while the additional inclusion of “very large” as comparison group does not change the odds much (from 2.45 to 1.80). These results indicate that, with respect to the effect of RN on employment outcome, the cutpoint between “small-medium” and for each of larger categories of company scale may be worth investigating using a multinomial logit model that treats the response categories as unordered.

With regard to measures of fit, the deviance is 1756.466, AIC was 2.525, BIC was -2786.697, and McFadden’s R^2 was 0.056. These will be compared with other models in the later section.

5.2. Multinomial Logit Model (with 4 response categories)

In multinomial logit models, the ordering of response categories is no longer assumed, and a specific baseline response category is used in estimating all equations simultaneously. Table 7 represents the results of fitting the multinomial logit model using a given smaller category of scale of company as baseline. In addition, Table 8 shows the estimated equations, indicating those main effects of SN or RN that are significantly associated with the outcome, controlling for feminine characteristics. There are three equations obtained by using the "small-medium" category as a baseline in addition to comparisons of "huge" vs. "large", as well as "huge" vs. "very large".

Table 7: Estimates for multinomial logit model with "small-medium" as baseline category (with 4 response categories)

Scale of company		B	S.E.	exp{b}	95% C.I. of odds ratio	
Equation 1					Lower	Upper
("large" vs. "small-medium")						
SRS	SN	0.57	0.89	1.77	0.31	10.04
	RN	1.43**	0.43	4.18	1.81	9.65
Feminine capital	Parents	0.61	0.44	1.84	0.77	4.37
	Home	0.42	0.25	1.53	0.93	2.52
	Junior	0.32	0.26	1.38	0.83	2.28
	Constant	-0.84	0.45			
Equation 2						
("very large" vs. "small-medium")						
SRS	SN	1.73*	0.81	5.66	1.16	27.58
	RN	1.90***	0.43	6.72	2.89	15.60
Feminine capital	Parents	1.02*	0.52	2.78	1.01	7.69
	Home	0.28	0.27	1.33	0.79	2.24
	Junior	0.16	0.27	1.17	0.69	1.99
	Constant	-1.39**	0.52			

Equation 3						
(“huge” vs. “small-medium”)						
SRS	SN	2.82***	0.75	16.74	3.88	72.29
	RN	1.93***	0.41	6.92	3.11	15.40
Feminine capital	Parents	1.26**	0.48	3.52	1.36	9.10
	Home	0.83***	0.24	2.28	1.42	3.68
	Junior	0.80**	0.25	2.22	1.36	3.65
	Constant	-1.77**	0.50			
Comparison of						
(“huge” vs. “large”)						
SRS	SN	2.25***	0.55	9.48	3.26	27.63
	RN	0.51*	0.25	1.66	1.02	2.70
Feminine capital	Parents	0.65	0.49	1.91	0.73	5.02
	Home	0.40	0.21	1.49	0.99	2.25
	Junior	0.48*	0.24	1.61	1.01	2.58
	Constant	-0.93	0.51			
Comparison of						
(“huge” vs. “very large”)						
SRS	SN	1.09**	0.41	2.96	1.34	6.56
	RN	0.03	0.26	1.03	0.63	1.70
Feminine capital	Parents	0.24	0.55	1.27	0.43	3.72
	Home	0.54*	0.22	1.72	1.13	2.63
	Junior	0.64*	0.25	1.90	1.16	3.13
	Constant	-0.38	0.56			

p<0.05; **p<0.01; ***p<0.001

Table 8: Estimated equations for multinomial logit models (with 4 response categories)

Estimated multinomial odds model
Equation 1 ("large" vs. "small-medium")
$\log\{Pr(Y_{i2}=1 \mid x) / (Y_{i1}=1 \mid x)\} = -0.84+0.57(SN) + 1.43(RN) + 0.61(parents) +0.42(home) + 0.32\ (junior)$
Equation 2 ("very large" vs. "small-medium")
$\log\{Pr(Y_{i3}=1 \mid x) / (Y_{i1}=1 \mid x)\} = -1.39+1.73(SN) + 1.90(RN) + 1.02(parents) +0.28(home) + 0.16\ (junior)$
Equation 3 ("huge" vs. "small-medium")
$\log\{Pr(Y_{i4}=1 \mid x) / (Y_{i1}= 1 \mid x)\} = -1.77+2.82(SN) + 1.93(RN) + 1.26(parents) +0.83(home) + 0.80(junior)$
Comparison of "huge" vs. "large"
$\log\{Pr(Y_{i4}=1 \mid x) / (Y_{i2}= 1 \mid x)\} = -0.93+2.25(SN) + 0.51\ (RN) + 0.65(parents) +0.40(home) +0.48(junior)$
Comparison of "huge" vs. " very large"
$\log\{Pr(Y_{i4}=1 \mid x) / (Y_{i3}= 1 \mid x)\} = -0.38+1.09(SN) + 0.03(RN) + 0.24(parents) +0.54(home) + 0.64(junior)$

Log odds of SN in equation 3 ("huge" vs. "small-medium") is largest (2.82), with those of comparison of "huge" vs. "large" (2.25). In addition, log odds of RN in equation 3 ("huge" vs. "small-medium") is largest (1.93), followed with those in equation 2 ("very large" vs. "small-medium") following (1.73). Figure 3 and Figure 4 show graphical presentation of the significant odds of receiving an offer from a specific category of larger company rather than a certain baseline category of smaller companies for students who used SN relative to those who used other methods, as well as students who used RN relative to those who used other methods.

Figure 3 : The estimated odds of working for a larger company (solid line) rather than a “small-medium” company (dotted line) for those who used SN relative to those who used “other” methods (with 4 response categories)

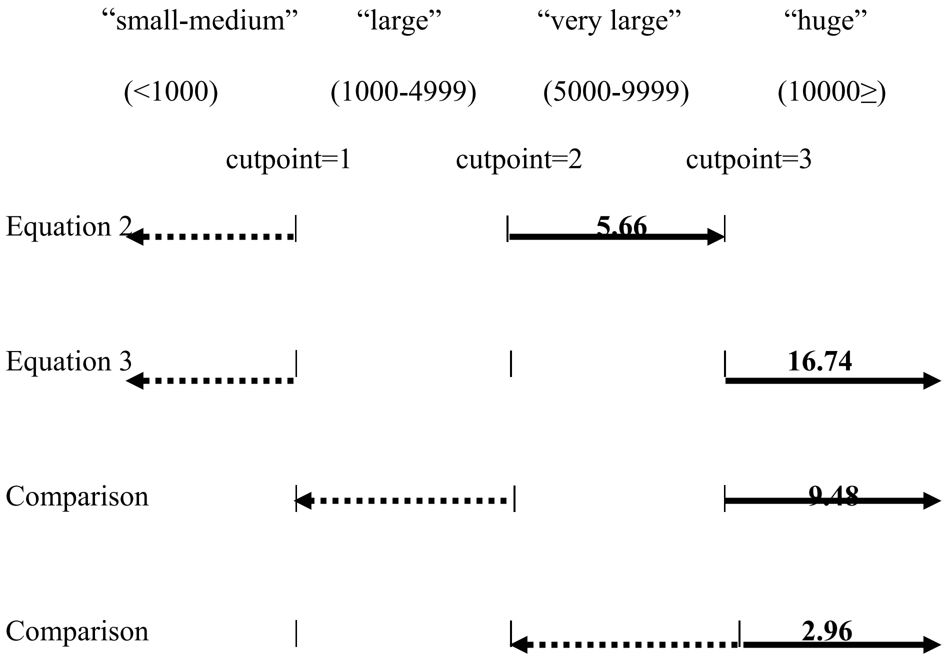
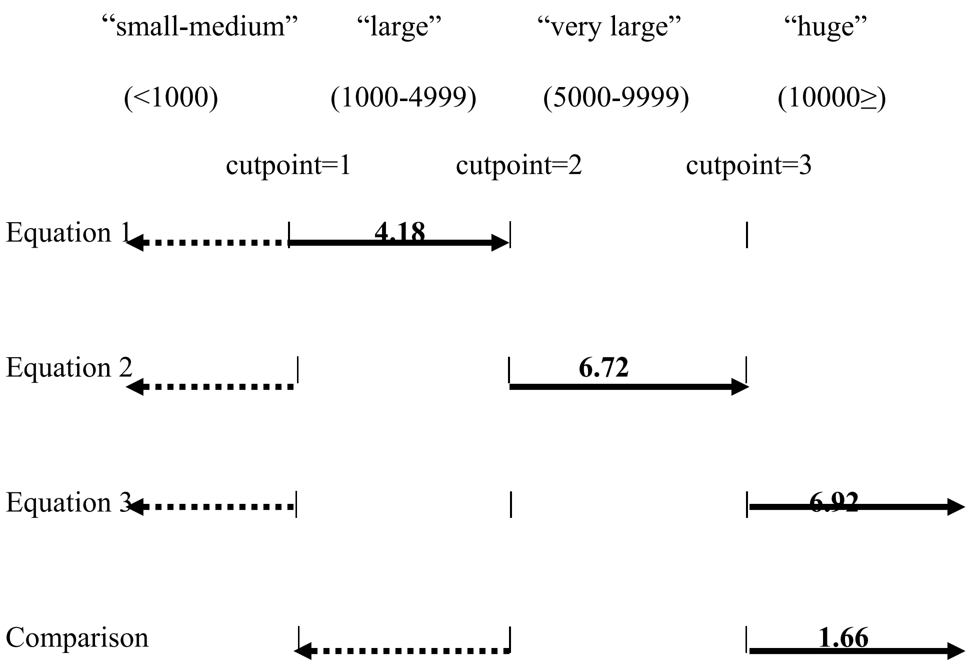


Figure 4: The estimated odds of working for a larger company (solid line) rather than a “small-medium” company (dotted line) for those who used RN relative to those who used “other” methods (with 4 response categories)



The odds of receiving an offer from a "huge" company rather than a "small-medium" company for students who used the SN system relative to those who used other methods (16.74) is almost three times higher than the odds of receiving an offer from a "very large" company rather than a "small-medium" company for students who sued the SN system relative to those who used other methods (5.66). On the other hand, the odds of receiving an offer from a "huge" company rather than a "small-medium" company for students who used the RN system relative to those who used other methods (6.92) is about the same as the odds of receiving an offer from a "very large" company rather than a "small-medium" company for students who used the RN system relative to those who used other methods (6.72).

More importantly, although the effect of receiving an offer from a "huge" company rather than a "very large" company is significant for those who used the SN system relative to those who used other methods (b=1.09), this effect is not significant for those who used the RN system relative to those who used other methods (b=0.03). Contrary to this, although the effect of receiving an offer from a "large" company rather than a "small-medium" company is significant for those who used the RN system relative to those who used other methods (b=1.43), this effect is not significant for those who used the SN system relative to those who used other methods (b=0.57). This result indicates that for students who used the SN system relative to other methods, the dichotomous division between "huge" and other lower-level categories of company's scale is important, while for those who used the RN system relative to those who used other methods, the division between "small-medium" and other higher-level categories of company's scale is important. However, the 95% confidence intervals of these estimated odds ratio are too wide, especially for the effect of SN, and hence they are not precise. Therefore, we need to be cautious about the interpretation of these odds. In addition, the magnitude of the effect is of more interest than significance per se.

To summarize, in addition to the evidence of my first hypothesis provided by cumulative logit model, the result of the multinomial logit model indicates that this holds true especially for students who received an offer from a "huge" company. Similarly, the second hypothesis holds especially for students who received an offer from a company in the category of "large" or greater. These results support my third hypothesis (*Schools will focus their placement efforts primarily on SN for the greater prestigious companies, and rely on RN to bring them placements with the lesser prestigious companies*).

With regard to measures of fit, deviance is 1756.062, AIC was 2.524, BIC was -2787.102, and McFadden's R^2 was 0.056.

5.3. Cumulative Logit Model (with 5 ordered response categories)

My study also uses five ordered response categories to see if an additional cutpoint at the extreme range of the response categories can contribute to model fit, and thus cause changes in parameter estimates. For this I performed a regression analysis on SRS as an explanatory variable, using the dependent variable scale of company and controlling for feminine capital of parents, home, and junior. This was done for both the proportional and the non-proportional odds model.

5.3.1. Proportional Odds Model

The result of approximate likelihood ratio test shows that the parallel regression assumption is on the verge of being rejected at the 0.05 level (chi-square= 24.42 with df=15, p-value= 0.058). Brant's Wald test allows to us to specify which variable violate parallel regression assumption and the result show as follows in Table 9:

Table 9: Brant test of parallel regression assumption (with 5 response categories)

Variables	Chi-square	P < chi-square	df
All	21.12	0.133	15
School Nomination	2.00	0.572	3
Reverse Nomination	10.99	0.012	3
Parents	0.34	0.953	3
Home	3.16	0.368	3
Junior	5.02	0.171	3

The Brant test shows that the largest violation is "reverse nomination" (but less so compared with the case in 4 response categories), which suggests that the violation of the proportional odds assumption may be due to this variable.

5.3.2. Non-proportional Odds Model

The results of fitting non-proportional odds model and estimated equations are shown in Table 10 and Table 11, respectively.

Table 10: Estimates for non-proportional odds model (with 5 ordered response categories)

Scale of company		b	S.E.	exp{b}	95% C.I. of odds ratio	
Equation 1 ("small-med")					Lower	Upper
SRS	SN	2.11**	0.73	8.33	1.97	35.18
	RN	1.78***	0.39	5.90	2.73	12.75
Feminine capital	parents	1.00**	0.37	2.74	1.34	5.61
	home	0.58**	0.21	1.78	1.17	2.72
	junior	0.48*	0.21	1.62	1.07	2.47
	constant	-0.29	0.37			
Equation 2 ("large")						
SRS	SN	2.16***	0.45	8.69	3.63	20.81
	RN	0.90***	0.21	2.45	1.63	3.69
Feminine capital	parents	0.80*	0.35	2.22	1.11	4.42
	home	0.35*	0.16	1.42	1.03	1.96
	junior	0.40*	0.18	1.50	1.05	2.12
	constant	-1.13**	0.36			

Equation 3 ("very large")						
SRS	SN	1.87***	0.32	6.51	3.45	12.28
	RN	0.58**	0.20	1.79	1.22	2.63
Feminine capital	parents	0.69	0.40	2.00	0.91	4.42
	home	0.54**	0.17	1.72	1.25	2.38
	junior	0.61**	0.20	1.84	1.26	2.70
	constant	-2.09***	0.42			
Equation 4 ("huge")						
SRS	SN	2.34***	0.43	10.45	4.52	24.14
	RN	1.07**	0.40	2.91	1.32	6.43
Feminine capital	parents	0.84	1.05	2.30	0.29	18.23
	home	0.28	0.30	1.33	0.74	2.38
	junior	1.40*	0.66	4.04	1.10	14.80
	constant	-5.42***	1.163			

* p<0.05; **p<0.01; ***p<0.001; ( ) represents cutpoint.

Table 11: Estimated equations for non-proportional odds models (with 5 ordered response categories)

Estimated non-proportional odds model

Equation 1: cutpoint = 1 (<1000)

$\log\{Pr(Y \geq 1 \mid x) / Pr(Y < 1 \mid x)\} = -0.29 + 2.12(SN) + 1.78(RN) + 1.00(parents) + 0.58(home) + 0.48(junior)$

Equation 2: cutpoint = 2 (1000-4999)

$\log\{Pr(Y \geq 2 \mid x) / Pr(Y < 2 \mid x)\} = -1.13 + 2.16(SN) + 0.90(RN) + 0.80(parents) + 0.35(home) + 0.40(junior)$

Equation 3: cutpoint = 3 (5000-9999)

$\log\{Pr(Y \geq 3 \mid x) / Pr(Y < 3 \mid x)\} = -2.08 + 1.87(SN) + 0.58(RN) + 0.69(parents) + 0.54(home) + 0.61(junior)$

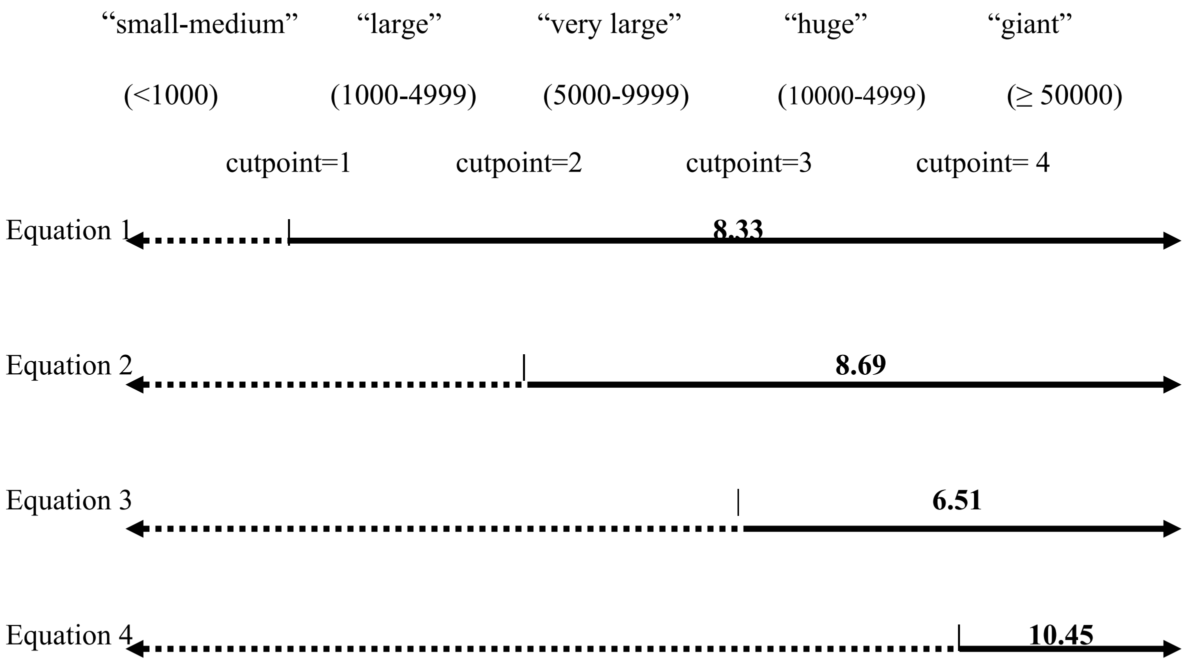
Equation 4: cutpoint = 4 (10000-4999)

$\log\{Pr(Y \geq 4 \mid x) / Pr(Y < 4 \mid x)\} = -5.42 + 2.35(SN) + 1.07(RN) + 0.84(parents) + 0.28(home) + 1.40(junior)$

The main effects of SN are significant across all equations, and SN estimates of equation 1, equation 2 and equation 3 are almost identical with those of ordinal logit model with 4 response categories. It is interesting to note that SN estimate of the additional category (equation 4) is the biggest (b=2.35). This implies that cutpoint 4 seems to be important in telling the effect of SN on employment outcome.

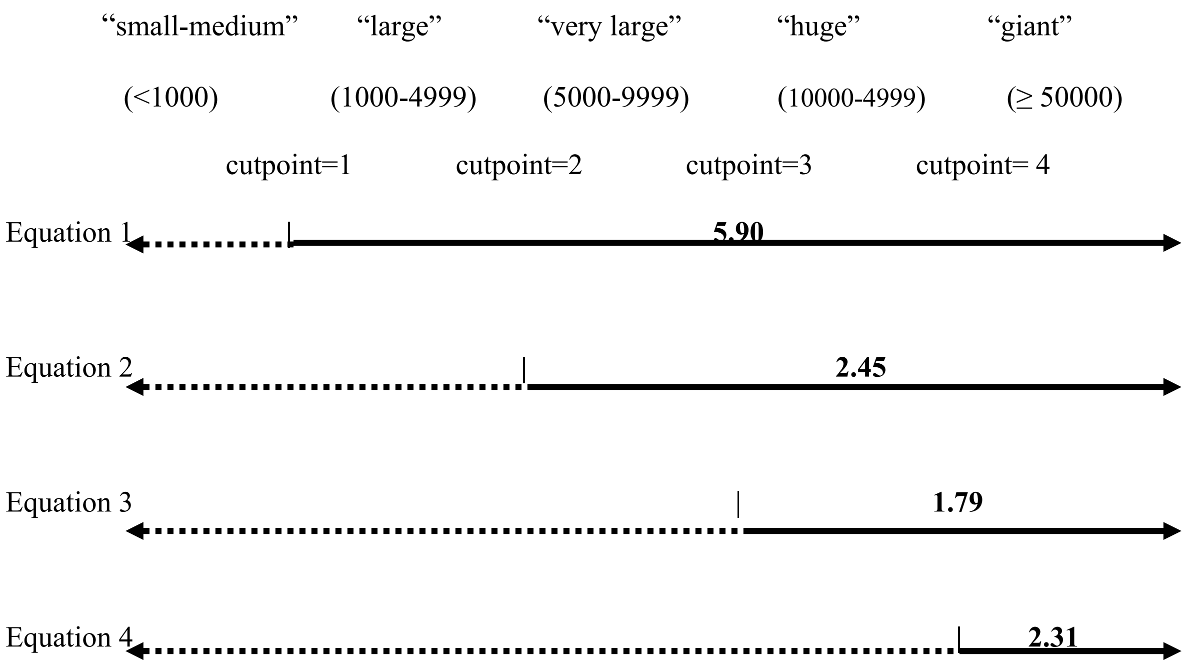
Similarly, the main effects of RN are significant across all equations, and RN estimates of equation 1, equation 2 and equation 3 are almost identical with those of ordinal logit model with 4 response categories. The magnitude of RN estimate of cutpoint 4 (b=1.07) is the next highest to that of cutpoint 1(b=1.78). The graphical presentations of these odds for each equation are shown in Figure 5 and Figure 6 for both SN and RN.

Figure 5: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used SN relative to those who used "other" methods (with 5 ordered response categories)



The odds of receiving an offer from a company of "giant" or greater (cutpoint = 4) is 10.45 for students who used SN compared to those who used other methods for given feminine characteristics, which is the highest odds compared to those of other cutpoints. This result indicates that the effect of SN on employment outcome is strongest when I divide equations into the highest end of response variable and lower than that, which confirms my first hypothesis. However, the 95% C.I of this parameter is too wide (4.52, 24.14), and estimate is not precise. The result of testing the proportionality of coefficient of the SN using a Wald test shows that slopes for SN term are proportional across outcome categories.

Figure 6: The estimated odds of working for a larger company (solid line) rather than a smaller company (dotted line) for those who used RN relative to those who used "other" methods (with 5 ordered response categories)



The odds of being receiving an offer from a company of "giant" or greater (cutpoint = 4) is 2.31 for students who used RN relative to those who used other methods. It is interesting to note that the magnitude of the odds decrease from equation 1 to equation 3, but in equation 4, it increases. However, these estimated odds are not precise, considering the wide 95% confidence intervals. The results of the Wald test to test the null hypothesis that the coefficients are the same in the two equations show that the estimated coefficients of the RN between equation 1 and equation 2 (chi-square=5.65, p=0.02), and between equation 1 and equation 3 (chi-square=9.02, p=0.003) are significant. This result implies that it is safe to assume that the odds of equation 4 is proportional to odds with other equations.

With regard to measures of fit, deviance is 2000.948, AIC was 2.886, BIC was -2502.824, and McFadden's R^2 was 0.056.

5.4. Multinomial Logit Model (with 5 response categories)

Table 12 represents the results of fitting the multinomial logit model using a given smaller category of scale of company as baseline. In addition, Table 13 shows the estimated equations, indicating those main effects of SN or RN that are significantly, associated with the outcome, controlling for feminine characteristics.

Table 12: Estimates for multinomial logit model with "small-medium" as baseline category (with 5 response categories)

Scale of company		B	S.E.	exp{b}	95% C.I. of odds ratio	
Equation 1					Lower	Upper
("large" vs. "small-medium)						
SRS	SN	0.57	0.89	1.78	0.31	10.13
	RN	1.43**	0.43	4.18	1.81	9.66
Feminine capital	parents	0.61	0.44	1.84	0.78	4.40
	home	0.42	0.25	1.53	0.93	2.52
	junior	0.32	0.26	1.38	0.84	2.29
	constant	-0.85	0.45			
Equation 2						
("very large" vs. "small-medium")						
SRS	SN	1.74*	0.81	5.70	1.17	27.80
	RN	1.90***	0.43	6.71	2.89	15.57
Feminine capital	parents	1.03*	0.52	2.81	1.02	7.78
	home	0.28	0.27	1.33	0.79	2.24
	junior	0.16	0.27	1.18	0.69	2.01
	constant	-1.40**	0.52			

Equation 3						
("huge" vs. "small-medium")						
SRS	SN	2.44**	0.76	11.42	2.59	50.33
	RN	1.80***	0.41	6.07	2.70	13.65
Feminine capital	parents	1.18*	0.49	3.25	1.25	8.49
	home	0.82**	0.25	2.28	1.40	3.70
	junior	0.70**	0.26	2.02	1.22	3.33
	constant	-1.73**	0.50			
Equation 4						
("giant" vs. "small-medium")						
SRS	SN	4.10***	0.82	60.27	12.05	301.36
	RN	2.63***	0.54	13.90	4.85	39.81
Feminine capital	parents	1.90	1.12	6.70	0.75	59.64
	home	0.82*	0.37	2.27	1.11	4.66
	junior	1.85**	0.66	6.39	1.75	23.30
	constant	-5.57***	1.25			
Comparison of						
"huge" vs. "large"						
SRS	SN	1.86***	0.56	6.41	2.14	19.17
	RN	0.37	0.26	1.45	0.87	2.41
Feminine capital	parents	0.56	0.50	1.76	0.66	4.67
	home	0.40	0.21	1.49	0.98	2.27
	junior	0.38	0.24	1.46	0.90	2.35
	constant	-0.88	0.51			

Comparison of						
“giant” vs. “large”						
SRS	SN	3.52***	0.64	33.81	9.60	119.04
	RN	1.20**	0.43	3.32	1.43	7.70
Feminine capital	parents	1.29	1.11	3.62	0.41	32.15
	home	0.40	0.34	1.49	0.76	2.89
	junior	1.53*	0.66	4.62	1.28	16.68
	constant	-4.72***	1.25			
Comparison of						
“giant” vs. “very large”						
SRS	SN	2.34***	0.53	10.57	3.75	29.79
	RN	0.73	0.43	2.07	0.89	4.83
Feminine capital	parents	0.87	1.13	2.38	0.26	21.99
	home	0.54	0.34	1.71	0.87	3.34
	junior	1.69**	0.66	5.42	1.49	19.74
	constant	-4.17***	1.27			
Comparison of						
“giant” vs. “huge”						
SRS	SN	1.66***	0.44	5.28	2.25	12.40
	RN	0.83*	0.41	2.30	1.02	5.13
Feminine capital	parents	0.72	1.10	2.06	0.24	17.71
	home	-0.002	0.32	1.00	0.54	1.85
	junior	1.15	0.65	3.17	0.88	11.37
	constant	-3.84**	1.23			

$p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 13: Estimated equations for multinomial logit models (with 5 response categories)

Estimated multinomial odds model

Equation 1 (“large” vs. “small-medium”)

$\log\{\Pr(Y_{i2}=1 \mid \mathbf{x}) / (Y_{i1}=1 \mid \mathbf{x})\} = -0.85+0.57(SN) + 1.43(RN) + 0.61(parents) +0.42(home) + 0.32 (junior)$

Equation 2 (“very large” vs. “small-medium”)

$\log\{\Pr(Y_{i3}=1 \mid \mathbf{x}) / (Y_{i1}=1 \mid \mathbf{x})\} = -1.40+1.74(SN) + 1.90(RN) + 1.03(parents) +0.28(home) + 0.16 (junior)$

Equation 3 (“huge” vs. “small-medium”)

$\log\{\Pr(Y_{i4}=1 \mid \mathbf{x}) / (Y_{i1}= 1 \mid \mathbf{x})\} = -1.73+2.44(SN) + 1.80(RN) + 1.18(parents) +0.82(home) + 0.70(junior)$

Equation 4 (“giant” vs. “small-medium”)

$\log\{\Pr(Y_{i5}=1 \mid \mathbf{x}) / (Y_{i1}= 1 \mid \mathbf{x})\} = -5.57+4.10(SN) + 2.63 (RN) + 1.90(parents) +0.82(home) +1.85(junior)$

Comparison of “huge” vs. “large”

$\log\{\Pr(Y_{i4}=1 \mid \mathbf{x}) / (Y_{i2}= 1 \mid \mathbf{x})\} = -0.88+1.86(SN) +0.37(RN) + 0.56(parents) +0.40(home) + 0.38(junior)$

Comparison of “giant” vs. “large”

$\log\{\Pr(Y_{i5}=1 \mid \mathbf{x}) / (Y_{i2}= 1 \mid \mathbf{x})\} =-4.72+3.52 (SN) + 1.20(RN) + 1.29 (parents) +0.40(home) +1.53(junior)$

Comparison of “giant” vs. “ very large”

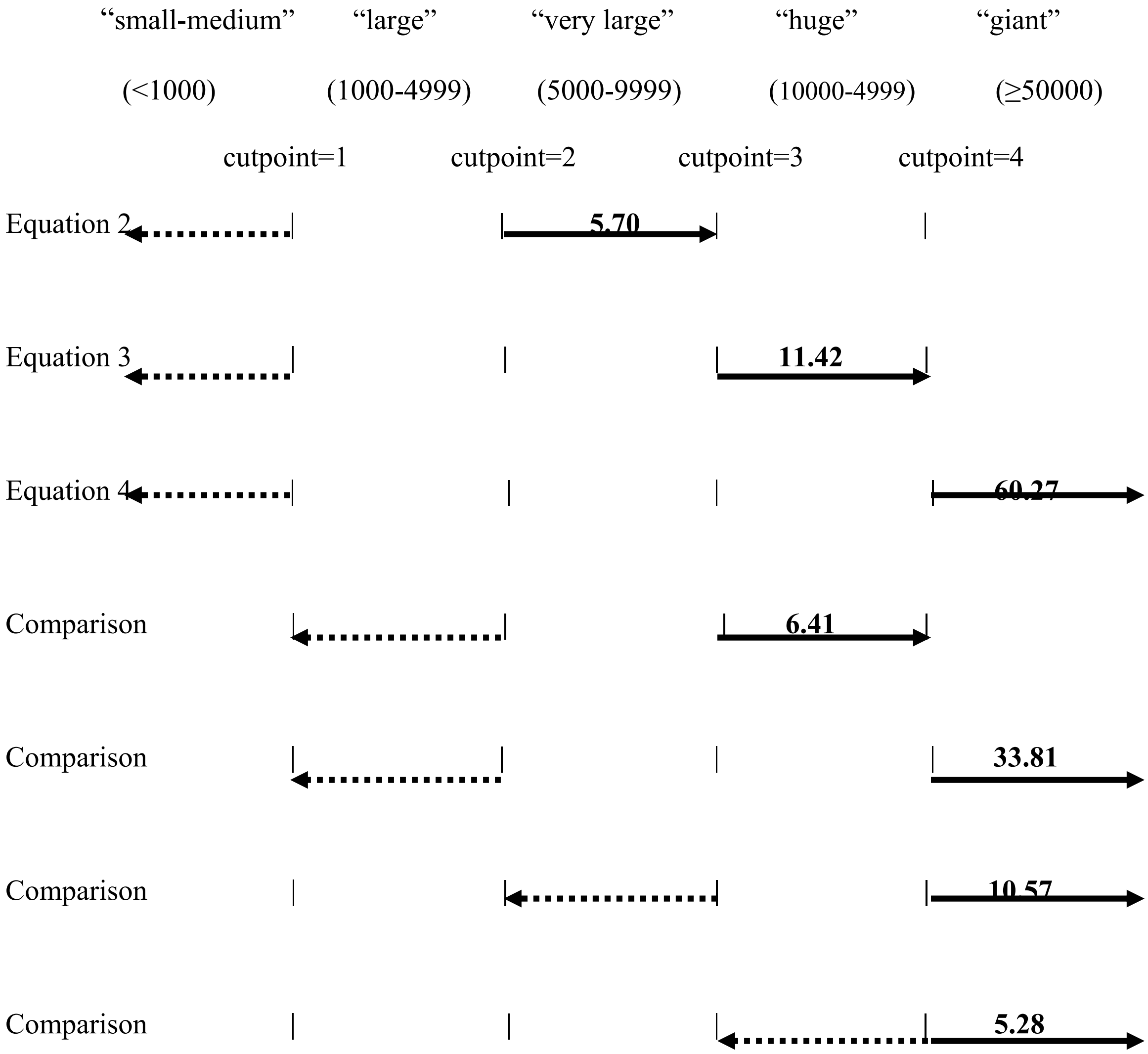
$\log\{\Pr(Y_{i5}=1 \mid \mathbf{x}) / (Y_{i3}= 1 \mid \mathbf{x})\} = -4.17+2.34 (SN) + 0.73 (RN) + 0.87(parents) +0.54 (home) +1.69(junior)$

Comparison of “giant” vs. “huge”

$\log\{\Pr(Y_{i5}=1 \mid \mathbf{x}) / (Y_{i4}= 1 \mid \mathbf{x})\} = -3.84+1.66(SN) +0.83(RN) + 0.72(parents) -0.002 (home) +1.15(junior)$

Figure 7 and Figure 8 show graphical presentation of the odds of receiving jobs at a specific category of a larger company rather than a certain baseline category of smaller company for students who used SN relative to those who used other methods, as well as students who used RN relative to those who used other methods.

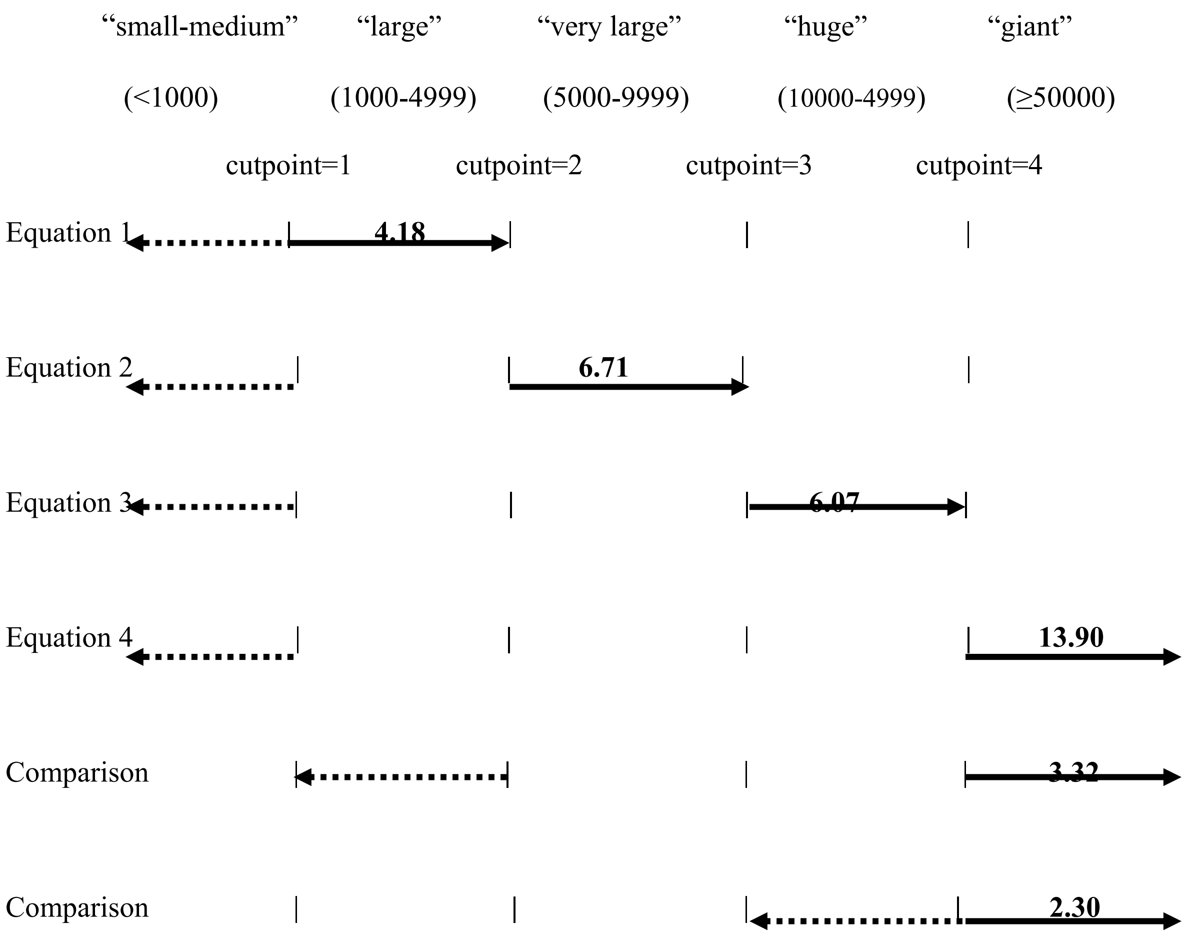
Figure 7: The estimated odds of working for a larger company (solid line) rather than a "small-medium" company (dotted line) for those who used SN relative to those who used "other" methods (with 5 response categories)



It is noticeable that the odds of receiving an offer from a "giant" or a larger company rather than a certain baseline category of smaller companies for students who used SN relative to those who used other methods are all significant (equation 4, 6, 7, and 8). This implies that the category of "giant" is quite important to understand the effect of SN on employment outcome. This is especially evident in the case of the odds of receiving an offer from a "giant" company relative to a "small-medium" company (60.27) that is almost 6 times higher than the odds of receiving an offer from a "huge" company relative to a "small-medium" company (11.42). In addition, the odds of receiving an offer from a "giant" company rather than a “huge” company is 5.28 for those who used SN relative to those who used other methods. However, since 95% C.I. is quite large, one must be cautious of these large values in estimated point odds ratio.

Recall that in previous analysis (multinomial logit model with 4 response categories), the odds of receiving an offer from a “huge” company rather than a “very large” company is significant for those used SN relative to those who used other methods. However, after adding the category of “giant” in response variable, it is no longer significant. This result implies that the different between “very large” and “huge” companies in terms of their effect of SN on employment outcome does not matter so much.

Figure 8: The estimated odds of working for a larger company (solid line) rather than a “small-medium” company (dotted line) for those who used RN relative to those who used “other” methods (with 5 response categories)



The odds of receiving an offer from a larger company rather than a "small-medium" company for students who used RN relative to those who used other methods are all significant (equation 1, 2, 3, and 4), and the odds of receiving an offer from a "giant" company rather than a "small-medium" company (13.90) is about 2 times as much as those with other categories of larger companies. However, again, we have to be cautious that the 95% confidence intervals for these estimated odds are quite large.

It is also interesting to note that the odds of receiving an offer from a "giant" company relative to a "huge" company is significant (2.30), while in previous analysis (with multinomial logit with 4 response categories), the odds of receiving an offer from a "huge" company rather than a "very large" company was not significant. Regarding the odds of receiving an offer from a "huge" company rather than a "large" company, what was significant in the 4-category model is no longer significant in the 5-category model. This result shows that the reverse recommendation system seems also be important in getting jobs at higher end of scale of companies ("giant" vs. "huge").

With regard to measures of fit, Deviance is 1999.529, AIC was 2.884, BIC was -2504.242, and McFadden's R^2 was 0.057.

5.5. Comparison between Models

I calculated the differences between proportional odds model and non-proportional odds model for each response categories in terms of their measures of model fit. The results are shown in Table 14.

Table 14: Differences between proportional odds results and non-proportional odds results

	4 response categories			5 response categories		
Measures of fit	Proportional 1	Non-proportional 1	Difference	Proportional 2	Non-proportional 2	Difference
Deviance	1774.387	*1756.466*	17.921	2025.422	*2000.948*	24.474
AIC	*2.522*	2.525	-0.003	*2.878*	2.886	-0.008
BIC	*-2834.429*	-2786.697	-47.732	*-2576.828*	-2502.824	-74.005
McFadden's R^2	*0.057*	0.056	0.001	*0.059*	0.056	0.003

Italicized value indicates better fit.

According to three measures of goodness of fit (AIC, BIC, McFadden's R^2, proportional odds model fits better than non-proportional odds model. The absolute difference between $BIC_{proportional}$ and $BIC_{non\text{-}proportional}$ for 4 response categories is 47.7, which is a very strong evidence favoring proportional odds model against non-proportional odds model. The similar result was obtained for 5 response categories as well (absolute difference between 2 models in terms of their BICs is 74). However, deviance measure indicates the opposite results.

Recall that previously the result of "omnibus" approximate likelihood ratio test showed that the parallel regression assumption was on the verge of being rejected at the 0.05 level (p-value= 0.055). I also conducted a likelihood-ratio test by comparing the likelihood statistics of a proportional odds model and a non-proportional odds model. The likelihood statistics computed by proportional odds model using 4 response categories is -887.1935 and for non-proportional odds model is -878.2329. Twice the difference of these two statistics follows a chi-square distribution with 10 degrees of freedom since non-proportional odds model has 10 more parameters than the proportional odds model. The result shows that chi-square value we observe is not unlikely (p=.056) under the null hypothesis that the proportional odds model fits as well as the non-proportional odds model. The similar result was obtained for 5 response categories (p=0.060). It seems to be difficult to determine which model fit better than the other, considering these mixed results. In my study, I chose the non-proportional odds model because Brant test shows that one of variables of interest violates non-proportional assumptions.

I also calculated the differences between multinomial logit model and non-proportional odds model for each response categories in terms of their measures of model fit. The results are shown in Table 15.

Table 15: Differences between non-proportional odds results and multinomial logit results

	4 response categories			5 response categories		
Measures of fit	Non-proportional 1	Multinomial 1	Difference	Non-proportional 2	Multinomial 2	Difference
Deviance	1756.466	*1756.062*	0.404	2000.948	*1999.529*	1.419
AIC	2.525	*2.524*	0.001	2.886	*2.884*	0.002
BIC	-2786.697	*-2787.102*	0.404	-2502.824	*-2504.242*	1.419
McFadden's R^2	0.056	0.056	0.000	0.056	*0.057*	-0.001

Italicized value indicates better fit.

When 4 response categories are used, the difference in deviance is 0.404 (multinomial logit model fits better), difference in AIC is 0.001 (multinomial logit model fits slightly better), difference in BIC is 0.404 (multinomial logit model fits better), and difference in McFadden's R^2 is zero.

When 5 response categories were used, the differences in deviance is 1.419 (multinomial logit model fits better), difference in AIC is 0.002 (multinomial logit model fits slightly better), difference in BIC is 1.419 (multinomial logit model fits better), and difference in McFadden's R^2 is -0.001(multinomial logit model fits better). According to these results, all measures agree that multinomial logit models fit better than non-proportional odds model. One of the possible reasons of this result may be that as the number of response categories increases, the number of parameters also increases, which leads to better fit and possible over-fitting. Also, it is possible that the very small difference in measure of fit would be due to the difference in estimation procedures.

Table 16 represents the differences between proportional odds model and multinomial logit model for each response categories in terms of their measures of model fit.

Table 16: Differences between proportional odds results and multinomial logit results

	4 response categories			5 response categories		
Measures of fit	Proportional 1	Multinomial 1	Difference	Proportional 2	Multinomial 2	Difference
Deviance	1774.387	*1756.062*	18.326	2025.422	*1999.529*	25.893
AIC	*2.522*	2.524	-0.002	*2.878*	2.884	-0.006
BIC	*-2834.429*	-2787.102	-47.327	*-2576.828*	-2504.242	-72.586
McFadden's R^2	*0.059*	0.056	0.001	*0.059*	0.057	0.002

Italicized value indicates better fit.

Three measures of fit (AIC, BIC, and McFadden's R^2) indicate that proportional odds model fit better than multinomial logit model for both 4 and 5 response categories although deviance measure indicates the opposite results. One of the possible resons of this result may be due to possible over-fitting.

The following Table 17 shows the results of the differences between the 4 response and the 5 response categories for proportional odds model in terms of measures of fit.

Table 17: Differences between 4 response categories and 5 response categories for proportional odds results

	Proportional odds model		
	4 categories	5 categories	Difference
Deviance	*1774.387*	2025.422	-251.035
AIC	*2.522*	2.878	-0.356
BIC	*-2834.429*	-2576.828	-257.601
McFadden's R^2	0.057	*0.059*	-0.002

Italicized value indicates better fit.

In the proportional odds models, the use of 4 response categories fit better than the use of 5 response categories according to the three measures of deviance, AIC, and BIC, although McFadden's R^2 indicates slightly better fit for the use of 5 response categories. The absolute difference in BICs for two models indicates strong evidence in favor of the model with 4 response categories against 5 response categories.

The following Table 18 shows the results of the differences between 4 response categories and 5 response categories for each ordinal and multinomial logit model in terms of measures of fit.

Table 18: Differences between 4 response categories and 5 response categories for non-proportional odds results and multinomial logit model

	Non-proportional odds model			Multinomial logit model		
	4 categories	5 categories	Difference	4 categories	5 categories	Difference
Deviance	*1756.466*	2000.948	-244.482	*1756.062*	1999.529	-243.468
AIC	*2.525*	2.886	-0.361	*2.524*	2.884	-0.360
BIC	*-2786.697*	-2502.824	-283.874	*-2787.102*	-2504.242	-282.860
McFadden's R^2	0.056	0.056	-0.001	0.056	*0.057*	-0.001

Italicized value indicates better fit.

Among non-proportional odds models, the use of 4 response categories fit better than the use of 5 response categories according to the three measures of deviance, AIC, and BIC. Likewise, among multinomial logit models, the use of 4 response categories also fit better than the use of 5 response categories according to three measures of deviance, AIC, and BIC, although McFadden's R^2 indicates slightly better fit for the use of 5 response categories. These results are slightly more evident in the non-proportional odds model than in the multinomial logit model. In other words, the magnitude of difference between the use of 4 response categories and 5 response categories is greater for non-proportional odds model than multinomial logit model. One of the possible reasons of the model with 4 response categories fitting better than the one with 5 response categories would be that as we increase the number of response categories we have fewer counts in each category. These small counts may not estimate well. Additionally, I checked how stable the estimates of both SN and RN with both 4 and 5 response categories by using bootstrap measure. The results of 95% confidence intervals using distributions based on normal, percentile and bias-corrected with 1000 repetitions produced similar results. They did not provide any clear difference between the model with 4 response categories and the one with 5 response categories in terms of their stabilities of the estimates of both SN and RN to determine which one perform better than the other.

To summarize, my study indicates the evidence of supporting both the first and second hypotheses. Using multinomial logit models with 4 response categories, the results indicate that the dichotomous division between “huge” and other lower-level categories of company’s scale is important for those who used the SN system relative to those who used other methods, while the division between “small-medium” and other higher-level category of company’s scale is important for those who used the RN system relative to those who used other methods. In addition, multinomial logit model shows that the effect of receiving an offer from a “huge” company rather than a “very large” company is significant for those who used the SN system relative to those who used other methods, while this effect is not significant for those who used the RN system relative to those who used other methods. However, we have to be cautious about over interpreting these results considering the wide confidence intervals of these estimates.

By using 5 response categories for the non-proportional odds model, my study found that the effect of SN on employment outcome is strongest for the equation that divides the highest end of the response variable from the rest of the variables. Although this result further confirms my first hypothesis, these estimates are not as precise as I would like. By using 5 response categories in multinomial logit model, my study found additional information in that the odds of receiving an offer from a “huge” company rather than a “large” company is also significant for those who used RN relative to those who used other methods. I also found that by changing the response categories, what was significant in the 4-category model was no longer significant in the 5-category model. However, this result would possibly be a sample size problem considering the small magnitude of estimates. Nonetheless, these results imply that the way model selection (in terms of the number of response categories used) is done can have an effect on the conclusions reached.

Finally, with regard to the comparison of models I used, all measures of model fit agree that the multinomial logit model fit better than the non-proportional odds model, but the difference is quite small, which may due to the difference in the estimation procedure. With regard to the comparison between proportional odds model and non-proportional odds model, it is hard to determine which model should be chosen considering mixed results, depending on which test we use. One way to tackle this problem is to check which variable follows proportional odds assumption and use partially proportional odds model. Tangentially, I must admit the limitation in my labor market data, which contains small cell counts for specific categories, has made estimation imprecise. My study also found that the use of 4 response categories also fit better than the use of 5 response categories.

6. Discussion and Conclusion

Substantively, this paper has presented some evidence of institutional linkages between women's colleges and large Japanese companies that govern female labor market processes by examining the mechanisms of school recommendation (SN) and reverse recommendation (RN) relative to other mechanisms of gaining employment. My study shows that multinomial regression approach allows a more detailed analysis of the outcomes generated by the ordinal approach allowing specification of reference response categories, and hence a more thorough picture of the data. With regard to ordinal regression approach, the issue is whether or not proportional odds assumption is satisfied in the data. My study found that proportional odds assumption was weakly supported and the individual test indicated that only the coefficient of reverse nomination system violates the proportional odds assumption. This result implies that partially proportional odds model would be useful to fit my data[15], and hence my future study would be to fit this model to compare with proportional odds model and non-proportional odds model in terms of their model fit. My study also found that, some measures of model fit indicate that fewer number of response categories fit data better, but a bootstrap measurement to check how stable the estimates with different response categories does not provide an answer concerning which one perform better than the other.

There are basically two major limitations in my study. First, my data contains some small cell counts in response categories, which cause estimation problems and hence generates imprecise coefficients. Thus, we should be cautious in interpreting the results generated by these advanced statistical models. Second, my study examined the association between the company scale and the students' placement characteristics, and hence it is limited to indirect inference about organizational behavior concerning school allocation of students and resources (spent screening, interviewing, etc.) to

[15] In the current version of STATA 8, Fu's non-proportional odds model does not support partially proportional odds model, but it will be available in the future version of STATA.

huge Japanese companies. Since my data lacks the number of unique companies who engage in the School Recommendation System (SRS), my third hypothesis about school's effort to nominate their students to huge companies has limited support.

It should be noted that there is a controversy with regard to the importance of model selection in social science research. For instance, Gelman and Rubin (1995) argue that model selection is "relative unimportant compared to the task of constructing realistic models that agree with both theory and data" (p.171). On the other hand, Raftery (1995b) views model selection "an essential part of the task of building a realistic model in social research" (p.185) because the way model selection is done can have an effect on the conclusions reached. Therefore, according to him, information measures the extent to which data provide evidence in favor of or against particular hypotheses, and therefore provides a good general purpose reporting tool.

My study used four measures of goodness of model fit to see how much each model fit data. Concerning the comparison between multinomial logit model and non-proportional odds model, there was little difference in the value of these measures because the models were essentially the same. However, using both models provides us substantive information to better understand the female labor market data. Concerning the comparison between different response categories for each model, these measures seem to be helpful to indicate which categorization of response variable fits the data better.

BIBLIOGRAPHY

Agresti, Allan. (2002). *Categorical data analysis, Second edition*, Wiley & Sons, Inc, New Jersey.

Bowles, Samuel and Herbert Gintis. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic life*. New York: Basic Books.

Brant, Rollin. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics, 46*, pp.1171-1178.

Brinton, Mary. C., Kariya, Takehiko. (1997). "Institutional Embeddedness in Japanese Labor Markets," in *The New Institutionalism in Economic Sociology*, edited by Brinton, Mary. C., and Nee Victor.

Brinton, Mary. C. (1992). "Christmas Cakes and Wedding Cakes: The social Organization of Japanese Women's Life Course." Pp.79-108, in *Japanese Social Organization*, edited by Sugiyama, Takie L. Honolulu: University of Hawaii Press.

Fu, Vincent Kang. (1998). "Estimating Generalized Ordered Logit Models. *Stata Technical Bulletin*, *44*, 27-30. Reprinted in *Stata Technical Bulletin Reprints*, *8*, pp.160-164.

Fujimoto, Kayo. (2004). "Feminine Capital: Forms of Capital in the Japanese Labor Market," *The Sociological Quarterly*, *45(1)*.

Granovetter, Mark. S. (1985). *Getting a Job*. Cambridge, MA: Harvard University Press.

--------------------------. (1974). "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology, 91(3)*, pp.481-510.

Ishida, Hiroshi. (1993). *Social Mobility in Contemporary Japan*. California: Stanford University Press.

Ishida, Hiroshi, Seymour Spilerman, and Kuo-Hsien Su. (1997). "Educational Credentials and Promotion Changes in Japanese and American Organizations," *American Sociological Review, 62*, pp.866-882.

Lee, Sunhwa and Mary C. Brinton. (1996). Elite Education and Social Capital: The Case of South Korea.

Sociology of Education, 69, pp.177-192.

Long, Scott J., and Freese, Jeremy. (2001). *Regression Models for Categorical Dependent Variables Using Stata*. Texas: Stata Press.

McCullagh, Peter. (1980). Regression Models for Ordinal Data (with Discussion). *Journal of the Royal Statistical Society, 42*, pp.109-142.

Meyer, John W. and Brian Rowan. (1977). "Institutionalized Organizations: Formal Structure as Myth and Ceremony," *American Journal of Sociology, 83(2),* pp.340-63.

Miller, Paul. W. and Volker, Paul. A. (1985). "On the Determination of Occupational Attainment and Mobility." *The Journal of Human Resources, 20(2)*, pp.197-213.

Ogasawara, Yuko. (2001). "Women's Solidarity: Company Policies and Japanese Office Ladies." Pp.151-179, in *Women's Working Lives in East Asia,* edited by Brinton, Mary. C. California: Stanford University Press.

Raftery, Adrian. E. (1993). "Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models." *Technical Report, 255*, Department of Statistics, University of Washington.

--------------------. (1995a). "Bayesian Model Selection in Social Research." *Sociological Methodology, 25*, pp.111-163.

--------------------. (1995b). "Rejoinder: Model Selection in Unavoidable in Social Research." *Sociological Methodology, 25*, pp.185-195.

Rosenbaum, James E., and Takehiko Kariya. (1989). "From High School to Work: Market and Institutional Mechanisms in Japan." *American Journal of Sociology, 94*, pp.1334-65.

Rosenbaum, James E., Takehiko Kariya, Settersten, Rick, and Maier, Tony. (1990). "Market and Network Theories of the Transition from High School to Work: Their Application to Industrialized Societies." *Annual Review of Sociology, 16*, pp.263-99.

Rosenbaum, James E., Deluca, Stefanie, Miller, Shazla R., and Roy, Kevin. (1999). "Pathway into Work: Short- and Long-term Effects of Personal and Institutional Ties," *Sociology of Education, 72*, pp.179-196.

Schmidt, Peter, and Strauss, Robsert. P. "The Prediction of Occupation Using Multiple Logit Models." *International Economic Review, 15*, pp.471-485.

Polachek, Solomon. W. (1981). "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics, 63*, pp.60-69.

Powers, John. A., and Marsh, Lawrence C., Huckfeldt, Robert R., and Johnson, Lincoln C. (1978). "A Comparison of Logit, Probit and Discriminant Analysis in Predicting Family Size." *Proceedings of the Social Statistics Section, American Statistical Association*, pp.693-697.

Wolfe, Rory, and Gould, W. (1998). sg76: An Approximate Likelihood-ratio Test for Ordinal Response Models. *Stata Technical Bulletin, 42*, pp.24-27.