REVEALED PREFERENCES FOR COLLEGE RANKINGS UNDER CENTRAL MECHANISM: EVIDENCE FROM TURKEY

by
HAYRİ ALPER ARSLAN

Submitted to the Social Sciences Institute
in partial fulfillment of the requirements for the degree of
Master of Arts

Sabancı University

Spring, 2012

REVEALED PREFERENCES FOR COLLEGE RANKINGS UNDER CENTRAL MECHANISM: EVIDENCE FROM TURKEY

APPROVED BY:

Ahmet Alkan ……………………….
(Thesis Supervisor)

Fırat İnceoğlu ……………………….
(Thesis Co-Supervisor)

Alpay Filiztekin ………………………..

Abdurrahman Aydemir ………………………..

Özge Kemahlıoğlu …………………………

DATE OF APPROVAL: 25.06.2012

Acknowledgements

First of all, I would like to thank to my thesis supervisors, Ahmet Alkan and Fırat İnceoğlu for their invaluable guidance their throughout this thesis. My work would not have been possible without their motivation and brilliant ideas. I would like to also express my gratitude to them for our enjoyable off-class conversations. I am also appreciative to my thesis jury members, Abdurrahman Aydemir, Alpay Filiztekin and Özge Kemahlıoğlu for their helpful comments about my thesis.

All of my other Sabancı University Professors are also deserving of my gratitude for everything they have ever taught me.

I am also thankful to my classmates and officemates for their friendship and complimentary assistances in any topic.

Finally, my family deserves infinite thanks for their encouragement and endless support throughout my education.

MERKEZİ YERLESTİRME SİSTEMLERİNDE ÜNİVERSİTE SIRALAMALARININ AÇIKLANMIŞ TERCİHLERİ: TÜRKİYE ÖRNEĞİ

Hayri Alper Arslan

Ekonomi, Yüksek Lisans Tezi, 2012

Tez Danışmanları: Ahmet Alkan, Fırat İnceoğlu

Anahtar Kelimeler: üniversite öğrenci kabulu, üniversite tercihi , yerlestirme sonuçları

Özet

Bu çalışma 2005 yılı Ögrenci Seçme Sınavının (ÖSS) veriseti ve Türkiye'yi temsil eden lise son sınıf öğrencileri arasında yapılan bir anket çalışmasından elde edilen verilerle, üniversite ögrenci eşleşme oluşumunu incelemektedir. ÖSYM sistemindeki öğrencilerin ve üniversite bölümlerinin karakteristik özellikleri ve üniversite tercih davranışları kullanılarak öğrencilerin açıklanmış tercihleri incelenecektir. Verisetleri öğrencilerin egitim geçmişleri, sınav performansları, detaylı sosyoekonomik ve demografik bilgilerini içermekte ve böylece öğrencilerin bu bilgileri kullanılarak onların açıklanmış tercihleri tahmin edilmektedir. Son olarak tahmin edilen tercihler yerleştirmelerde ve onların işgücü piyasası üzerindeki etkilerini incelemek için kullanılmaktadır.

REVEALED PREFERENCES FOR COLLEGE RANKINGS UNDER CENTRAL MECHANISM: EVIDENCE FROM TURKEY

Hayri Alper Arslan

Economics, M.A Thesis, 2012

Thesis Supervisors: Ahmet Alkan, Fırat İnceoğlu

Keywords: college admissions, college choice, placement outcomes

Abstract

This paper explores the economics of match formation in the context of university entrance exam using a dataset obtained from Turkish university entrance system and a survey of senior high school students from the representative sample of Turkey. I provide a description of the student and department characteristics in the CSSP (Center for Students Selection and Placement) mechanism and utilize detailed information regarding university selection behavior to infer students' revealed preferences. Data allows me to estimate a very rich preference specification that takes into account a large number of educational background attributes, detailed demographic and socioeconomic information, along with exam performances. I develop consideration sets for students to eliminate strategic behavior from our estimation and compare them with some benchmark choice sets. Finally, placement outcomes and their effects on labor market are investigated by using estimated preferences.

Contents

List of Tables

1 Introduction

Each year millions of high school graduates make decisions about whether to continue their education, and if so, where to apply and enroll. Some students take this decision relatively simple because the existence of a particular academic program in a specific institution, the proximity of an institution of higher education, or a host of other factors make a certain school more favorable than others. For others, the choice process is difficult as they attempt to find an institution that will match their educational goals, interests and financial constraints.

College choice decision making is important for the student because it directly affects students' future career path. A students college choice strongly influences his or her professional career, and there is evidence to indicate that the type of postsecondary education a student completes yields differential outcomes (Hossler et al. 1989 [8]). Even if these differential outcomes may be less pronounced when one controls for confounding factors such as choice of academic major and academic ability, there are considerable differences in choice preferences that cause change in the outcomes.

There are numerous studies that provide explanations for the differences in college going decision and academic major choices. These studies also give insight to understand the outcomes of these differences in students' future lives and labor markets outcomes. Kane (2001) [9]gathered college going literature and showed that differences in students attributes have significant effects on decision. Especially family income and their attitudes towards education, race, and gender change college going decision significantly. Polachek (1978) [12] showed females prefer education and fine arts majors more and they avoid generally high earning majors such as engineering and business. Ma (2009)[10] worked on effects of family socioeconomic status and parental involvement in major choices. Balsamo et al. (2012) [2] analyzed personality effects on the selection of majors. Saygin (2011) [14] showed gender differences in major selection by using Turkish data. On the whole, we can see from the literature that the preferences of students for various social and income groups, ethnic groups and gender are considerably different and some groups are particularly disadvantaged with these preferences (e.g. being female and belonging to a lower socioeconomic group).

On the other hand, college choice preferences and their impacts are not sufficiently explored in the literature. Punj et al. (1978) [13] investigated graduate business college selections and found that distance, cost, and quality were the main criteria. DesJardins et al. (1999) [4] looked at college preferences to assess the marketing strategies of land-grant universities.

In this paper, I attempt to show the differences in college choice preferences, the effects of students' attributes on preferences (such as gender, income, location, etc.) and the impacts of these choice preferences on the outcome of these placements, by estimating revealed preferences of students for college rankings. In order to focus particularly on college choice, major choices and constraints from the university entrance system are controlled. Hence the results obtained allow me to examine purely college choice preferences. I use detailed administrative data from the Turkish university entrance test in 2005 and a representative survey conducted by Alkan et al. including students' socioeconomic, demographic and educational background as well as their preparation for the exam. Data includes applicants' choices over all university programs so that I can directly investigate the potential differences in choices made by different groups.

University placement system is centralized in Turkey and administrated by CSSP (Center for Student Selection and Placement). A standardized exam at national level is conducted every year and the applicants make their college choices after they learn their scores. Because college departments' ranking over students depend on exam scores, students' choices are affected from this knowledge. Also information on the previous years' placement patterns give an insight to students' evaluation of their scores and impacts students' choices. Another crucial information about choices is of students' own ranking of various departments. Students do not just report their choices as a set; they rank them up to 24. Hence these criteria require special attention in analyzing students' choices in the CSSP mechanism.

There is also a benchmark study for this paper in the matching literature. Hitsch et al.(2009) (2010) [6] [7] analyzed mate matching by using data on user attributes and interactions from an online dating site and estimated preferences in mate selection. I use a similar setting for the estimation with their study; however, there are considerable differences between these studies. First, their data provide users' choice sets. On the dating site, users initially browse people in the database and email people whom they would like to date. Thus, people that have been browsed form users' choice sets and people that have been emailed are selected mates from these sets. This is the main advantage of their study because they could estimate mate preferences easily according to these sets. A second advantage is the possibility for users to send unlimited numbers of emails. Hence users send emails without considering they will be refused. In my data, choice sets are not known. Certainty of department side rankings in the CSSP mechanism makes departments inactive in the market. This certainty requires students to calculate their position in the department side when they are making choices. On the other hand, I have an advantage of using students' rankings of departments in the estimation.

The empirical challenge of this thesis is to appoint appropriate choice sets to the students and estimate revealed preferences within these appointed sets. There are almost 7,700 number of departments in the CSSP and all of them are a potential candidate for each student. I develop consideration sets for students to reduce the number of departments and make estimation significant, thus eliminating potential strategic behavior of students born from the CSSP mechanism. Then I compare results obtained from using consideration sets with those obtained from using benchmark random choice sets. Comparisons show that results from estimations using consideration sets are more suitable for my analysis because they explain variation better.

Factors affecting students' university exam performances such as socioeconomic and educational background could be also considered in the college and major choices in the centralized mechanisms. Because of score-dependent choices, some departments are unreachable to some students. Therefore, impacts of students' educational background, exam preparation and their abilities on scores affects college choices indirectly. However, I don't use these factors as choice criteria in the analysis. My setting investigates choices after students learn their scores and we accept scores as fixed attributes of students. Alkan et al. (2008) [1] analyzed these factors' effects on exam performance.

Estimation results show that there are significant differences in college choice preferences. Different attributes of students such as gender, income, parents' education and location change students choices and make clear impacts on placement outcomes. Some of the preferences are common such as avoidance of departments whose entrance scores differ from students' scores, long distance between college and hometown, foundation colleges (i.e. private universities established by foundations in Turkey), and two-year vocational programs. However, the differences in attributes change preferences' magnitudes and these directly change influence placements and future outcomes.

The thesis is organized as follows: in Section 2, I provide details about the institutional setting in Turkey; I describe the data and show some descriptive statistics to motivate the rest of the paper. In Section 3 , I explain the research design and report the main results. In Section 4, I discuss estimated results and their effects. In section 5, I conclude.

2 Institutional Setting and Data Description

In this section, I outline the institutional setting of high school and college education in Turkey, and then describe the summary statistics of the data set of high school graduates in the empirical analysis. This is helpful to understand college and major selection process in Turkey.

In the empirical analysis, I use two data sets on high school students who have graduated and entered University Entrance Exam (OSS) in the period 2005. The first data set is from the survey which is conducted by Alkan et al. The survey asked questions to senior high school students who would be entering OSS exam in that year about socioeconomic, politic, demographic and educational background as well as information on students' preparation for the exam. The second data set is obtained from CSSP and provides students' OSS performance as well as their university and department choices. A rich descriptive analysis of OSS can be found in Alkan et al. (2008)[1].

I restrict the analysis to the representative student survey. By making this restriction, I can add more attributes to the estimation of students' utility function apart from exam performance information. The data sets are linked by using students' identification numbers, thus eliminating loss of observations in the merger. The advantage of using the linked data set is that it enables one to study determinants of college choice by linking rich data sets on students' achievement at high school, socioeconomic background and geo-coded information to track students' mobility choices over time.

2.1 High School

In 2005, compulsory education in Turkey amounted to eight years. Primary school starts at the age of 7 and compulsory school leaving age is at age 15. Students who want go to high school enter an exam, Lycee Entrance Examination (LEE) to be accepted to some specific high schools. After the examination and students' choices, each student enrolls in a High school. If a student could not enroll in a school which takes students through central placement, she could enroll in a public high school or other type of schools which don't require exam scores. Table 1 summarizes high school entrance requirements. High schools last 3 or 4 years depending on their type. High schools, which have foreign language education in addition to their standard curriculum, are one year longer.

Table 1: High School Entrance Requirements

School Type	**Entrance Requirements**
Cok Prog. Lise	N/A
Resmi Lise	N/A
Imam Hatip	N/A
Y.D.A. Lise	GPA
Meslek Lisesi	N/A
Ozel Fen Lisesi	LEE
Anad. Lisesi	LEE
Anad. Meslek Lis	LEE
Fen Lisesi	LEE
Ozel Lise	LEE or N/A
Anad. Ogr. Lises	LEE

There are two important criteria in high school education that directly affect university choices. The first one is high school General Point Average (OBP). Each student obtains a OBP according to his/her high school performance. The second effect is field selection. In Turkey, high school education is differentiated in 4 fields in the second year. These are Math and Science(SAY), Math and Social(EA), Social Sciences(SOZ) and Language(DIL). Students choose one of these fields and their education continues in this direction. Field selection is an important decision because university departments also have the same differentiation and they select students with field scores in the university entrance exam. Moreover, students' OBP change with university field selection. Choices made outside one's field are not favored by the system. Such choices are punished by reducing the OBP and students will be asked to answer questions that are outside their field in the university exam.

2.2 University Exam and Degree Choice

There is a centralized university placement system in Turkey. Students are ranked according to the results they attain in an centralized exam conducted by CSSP. CSSP matches students and departments by using Gale-Shapley Stable Algorithm.(Gale-Shapley, 1958 [5])

University exam is held once a year and students are asked multiple choice questions in different areas. The main areas of the exam are mathematics, natural sciences, social sciences, Turkish and other foreign languages. Each area is differently weighted in each field. Hence a student takes at least 3 different scores in the exam (Language field has a separate exam for students who want to make choices based on Language scores.) Students declare their choices after they learn their university exam performance and OBP. OBP is also weighted accordingly each high schools performance in the exam. The schools whose students have higher scores on average in the exam get more weight in calculating AOBP (called weighted OBP) for their students.

Students' total scores is the sum of university exam scores and AOBP scores. Each student has a field score in the exam. For each field's total score AOBP is added to exam scores by multiplying different coefficients. If students make choice within field, AOBP is multiplied with coefficient 0.8 and with 0.3, if outside field. Also some school types provide additional coefficients to special departments, i.e. Anadolu Ogretmen Lisesi provides coefficients to teacher education departments. At the end of the exam period CSSP calculates all these scores and students know their scores and ranking before making their choice of departments.

University departments' hypothetical ranking over students is crystal clear. Departments are differentiated according to fields similar to students' field options. A department ranks students according to their total scores in the system. Therefore, there are four different students rankings and each department prefers students with higher scores in their fields.

Students have the right to choose up to 24 departments. They rank and list their preferred departments in descending order and report this choice list to the CSSP. After completion of submission of choice lists, CSSP matches departments and students by using Gale-Shapley Stable Algorithm. Placements occur according to results of this mechanism and these placements are binding.

2.3 Data Description

The analysis in this paper is based on a sample of 12,828 students from different geographical areas in Turkey. Sample selection techniques are explained in Alkan et al. [1] Table 2 shows sample and population summary statistic about exam performances.

Table 2: Sample and Population Exam Performances

Field	Database	WE	<Level1	Level1	Level2	Total
SAY	Senior Student	44107 (%18)	69,369 (%29)	39,517 (%17)	86,408 (%36)	239,401 (%35)
	Sample	469 (%9)	933 (%17)	767 (%14)	3,181 (%60)	5,350 (%42)
SOZ	Senior Student	14,916 (%10)	38,232 (%27)	32,421 (%22)	59,158 (%41)	144,777 (%21)
	Sample	260 (%9)	573 (%20)	472 (%17)	1,549 (%54)	2,854 (%22)
EA	Senior Student	29,334 (%10)	65,963 (%23)	52,738 (%19)	136,369 (%48)	284,404 (%41)
	Sample	175 (%4)	507 (%13)	464 (%12)	2,755 (%71)	3,901 (%30)
DIL	Senior Student	541 (%2)	1,089 (%25)	6,617 (%9)	14,882 (%64)	23,129 (%3)
	Sample	8 (%2)	5 (%1)	82 (%11)	638 (%87)	733 (%6)

Continued on next page

Table 2 – *Continued from previous page*

Field	**Database**	**WE**	**<Level1**	**Level1**	**Level2**	**Total**
TOTAL	Senior Student	88,898 (%13)	174,703 (%25)	131,293 (%19)	296,817 (%43)	691,711 (%100)
	Sample	912 (%7)	2,018 (%16)	1,785 (%14)	8,123 (%63)	12,838 (%100)

I refer the reader to Alkan et al.[1] for detailed descriptive analysis of 2005 OSS exam database. In this study, OSS exam database is presented along with various descriptive accounts across gender, region, school type and graduation status.

Approximately 55 percent of students made at least one choice. This means that I have to drop nearly half of the observations for the estimation. Table 3 shows the choice ratios of students from different high school type. Anadolu Ogretmen Lisesi, Ozel Lise and Fen Lisesi have higher application ratio. On the other hand, Cokprogramli Lise and Resmi Lise have the lowest apply ratio. The average scores of each school type in different fields are in Table 4. Fen Lisesi and Ozel Lise are significantly higher in average in Math and Science fields; at the other extreme Meslek Lisesi is the lowest in all average scores.

Table 3: School Type and Choice Numbers

School Type	**Apply Ratio**	**Average Number of Choices**
Cok Prog. Lise	0.20	1.23
Resmi Lise	0.21	2.34
Imam Hatip	0.44	3.72
Y.D.A. Lise	0.56	8.48
Meslek Lisesi	0.57	5.46
Ozel Fen Lisesi	0.66	7.81
Anad. Lisesi	0.70	10.17
Anad. Meslek Lis	0.70	5.99
Fen Lisesi	0.82	8.97
Ozel Lise	0.86	11.11
Anad. Ogr. Lises	0.91	11.76
Average	0.55	6.57

Table 4: School Type and Average Scores

School Type	SAY	SOZ	EA	DIL
Meslek Lisesi	94.60	104.22	99.31	0.00
Cok Prog. Lise	128.53	148.40	139.46	0.00
Anad. Meslek Lis	128.91	143.99	138.05	1.76
Imam Hatip	131.35	169.35	151.15	0.95
Resmi Lise	149.87	174.17	164.33	2.83
Ozel Fen Lisesi	182.43	206.77	201.21	13.56
Y.D.A. Lise	185.03	212.98	204.80	36.76
Anad. Ogr. Lises	212.00	233.26	228.84	57.15
Anad. Lisesi	215.87	224.55	227.70	16.09
Ozel Lise	243.98	222.08	239.73	0.00
Fen Lisesi	271.36	249.80	264.67	0.00
Average	165.85	184.85	179.15	13.41

3 Student Revealed Preference Estimation

My estimation approach is based on a sequence of ranking decisions. For each student I appoint choice sets. Because we know students' ranking from their choice lists, appointed sets provide possibly browsed but unselected departments. If the cost of adding department to the choice list is 0 in the choice set, a student s truly ranks department d up to 24 choices where the utility of being accepted from this department $U_s(s,d)$ is greater or equal to reservation utility level of entering a department $v_s(s)$. The reservation utility is the utility level where a student is indifferent to being accepted from a particular department and being unmatched. It can be the last choice of a student in his/her choice list or lower than this level for students who fill their choice list. Hence a student ranks departments in the choice list if the utility level from that department is at least as high as reservation utility level up to 24. If the cost of adding is not 0, students rank departments strategically and the ranking does not represent students' true preferences.

There is one more thing to be considered regarding the students who fill their choice lists. Because of the 24-department limit, the cost of adding departments to the choice list can change for these students' choices. They withdraw some departments because the probability of being accepted to any department drops by adding additional higher scored departments to the list. For example, a student could exchange a department which has higher entrance score with one that has lower entrance score even if it is a less preferred one. Hence this student increases the probability of being accepted to a department by forgoing more preferred departments. This situation makes it harder to assume that the cost of adding department is 0 for students who have completely filled their choice lists. However with the help of the ranking of departments in choice lists, the ordering of choices in the list does not change significantly. Even if the complete ordering of departments changes when a student drops a more preferred department due to the 24-department limitation, the partial ordering of departments within consideration sets does not change. Moreover, the relatively low ratio of applicants (around 15%) who fill their lists reduces this problem.

Why some students make choice and why some don't

Before going into the estimation of students' departmental choices, I briefly look at the motives of students in making their choices. Because our data set also includes students who enter the exam, earn sufficient scores in the CSSP but don't make any departmental choice; we can also make an analysis based on the same data. In this estimation, I use students as unit of analysis. Students are just represented as making the choice or not. The number of choices are insignificant if it is bigger than 0. Logistic regression is a tool for this analysis. As I expected, total scores are the most significant factors in choice motivation (Table 5). However, there are also important factors in making choices such as income level, parents' education levels, and the size of the place where the student lives. Odds ratios are represented in Table 6.

Table 5: Apply Ratio

Variables	Apply a College Department
CSSP score	0.00729***
	(0.000318)
Income	0.0274***
	(0.0101)

Continued on next page

Table 5 – *Continued from previous page*

Variables	Apply a College Department
Parents education	0.134***
	(0.0248)
Location type	0.116***
	(0.0236)
Extra private education	-0.0750***
	(0.0205)
Gender	0.0989**
	(0.045)
Study possibilities	-0.00637
	(0.0104)
Siblings number	-0.0633***
	(0.0115)
Anxious level for exam	-0.0503***
	(0.00853)
Constant	-1.812***
	(0.19)
Observations	10,518

Table 6: Apply Ratio Odd Ratios

Variables	Odd Ratios
CSSP Score	1.007
Income	1.028
Parents Education	1.143
Location type	1.123
Extra private education	0.928
Gender	1.104
Study possibilities	0.994
Siblings number	0.939
Anxious level for exam	0.951

Reporting choices to the CSSP mechanism is costless. If a student makes choices and s/he is not placed in any department, there will be no additional cost according to non-reporting choice conditions for the next years. Therefore, after earning a sufficient score in the CSSP, submitting college choices is important to understand the eagerness of students to move on to tertiary education. Because students who make a choice have a probability to enter a department, these students show their eagerness to go to college. On the other hand, students who don't make any choices directly exit from the university market or postpone entering the market without trying their chances. Estimation results show that students who come from small towns; who have lower income levels; who have more siblings; and who have relatively less educated parents give up their chance to enter college more easily than others. Female students are more eager than male students to continue tertiary education. Interestingly, students who are more concerned about being unsuccessful in the exam tend to avoid making choices. These results are the initial impacts of differences in attributes on students' choice preferences.

3.1 Discrete Choice Model

For discrete choice model we use very similar techniques based on the paper by Hirsth et al. [6] [7] They estimate mate preferences by using interactions of users of an online dating site. In this paper, I use students' choices and their revealed preferences in order to estimate college preferences.

Given the ranking decision rule, student preferences can be estimated using discrete choice methods. We assume that student preferences depend on their own observed attributes as well as the department's observed attributes, and on an idiosyncratic preference shock: $U_s(s,d) = U_s(X_s, X_d; \theta_s) + \varepsilon_{sd}$. We split the attribute vector and the parameter vector into separate components: $X_d = (x_d, d_d)$, $\theta_S = (\beta_s, \gamma_s^+, \gamma_s^-, \upsilon_s)$. The latent utility of student s from a match with university department d is parameterized as (1)

$$U_s(s,d) = x_d'\beta + |(x_s - x_d)|_+\gamma^+ + |(x_s - x_d)|_-\gamma^- + \sum_{k,l=1}^{N}(d_{sk} = 1 \text{ and } d_{dl} = 1)\upsilon^{kl} + \varepsilon_{sd} \quad (1)$$

The first component of utility is a simple linear valuation of the department's attributes. The other components relate the student's preferences to his own characteristics. The difference between the department's and student's attributes is $|x_d - x_s|_+$ if this difference is positive, and $|x_d - x_s|_-$ denotes the absolute value of this difference if the difference is negative. For example, consider the difference in points between the total score of student s and the minimum entry score of department d. If the coefficients corresponding to the score difference in both γ_+ and γ_- are negative, it means that students prefer departments closer to their own scores. The fourth component in the utility function relates preferences to categorical attributes of both sides. Dummy variables indicating that student s and department d possess a certain trait are represented by d_{sk} and d_{dl}. For example, if $d_{sk} = 1$ and $d_{dl} = 1$ indicate that s is from a high income group and that d is a foundation college, then the parameter υ_s^{kl} expresses the relative preference of rich family children for foundation college departments.

I use a rank-ordered conditional logit model to estimate how applicants value college characteristics and how the weights placed on these characteristics vary across various attributes. Rank-ordered logistic model is also known as exploded logit model. Exploded refers to a logit model that incorporates multiple-ranked choices for each person but not only the first choice that gives the highest utility. (McFadden and Train 2000 [11], Train 2003 [16])

The setting of rank-ordered conditional logit model is very similar to a conditional logit model where a coefficient is obtained for each attribute of the alternatives. In this rank-ordered model, each applicant is assumed to have an individual choice set and the individual choice set is assumed to include the university programs that are chosen by the applicant and coefficients are mapped from the ranking of these alternatives. Using this method, I obtain the coefficients for university program attributes such as scholarship status, distance from high school city, instruction language, whether university is a public or foundation college, cost of living index of the college city, etc.

Assuming that ranking a department in the choice list is costless, the choice probabilities then take the rank ordered logit form:

$$
\begin{aligned}
&Prob(\text{ranking of departments}|\text{considering departments}) \qquad (2)\\
&= \int \left(\frac{e^{\beta x_{n1}}}{\sum_{j=1,2,3\ldots,24} e^{\beta x_{nj}}} \frac{e^{\beta x_{n2}}}{\sum_{j=2,3\ldots,24} e^{\beta x_{nj}}} \cdots \frac{e^{\beta x_{n23}}}{\sum_{j=23,24} e^{\beta x_{nj}}} \right)\\
&\times g(\beta|\theta)\mathrm{d}\beta
\end{aligned}
$$

I also estimate the model with standard conditional logistic regression in order to understand effects of the ranking. In the estimation results, the second column in Table 11 represents this estimation.

3.2 A Modeling Framework for Analyzing Students Behavior

My data are in the form of students' choice lists. Students choose and rank departments from the CSSP department set. In order to interpret the data using a revealed preference framework, I make the following assumption:

Assumption: Suppose a student considers two departments, d and d', and ranks d in her/his choice list. Then the student prefers a potential match with d over a potential match with d'.

However, I don't have any data or information about students' choice sets. There are approximately 7,700 departments from 4 fields and each student has no restriction to choose any of them. Even within the field analysis, potential choice sets are not less than 284 (Table 7). 284 does not seem to be a huge number, but most of the students make choice from other fields and the number of departments reaches an enormous number. It is not realistic to assume all departments are browsed by students in order to complete choice lists. Also, working with these huge numbered choice sets is not possible for an estimation methodology. For these reasons, appointing appropriate choice sets is the crucial part of this analysis.

Table 7: Field Department Number

Field	Number of Departments
DIL	284
EA	2509
SAY	3876
SOZ	1008
Total	7677

Strategic behavior in choices is another important point that cannot be neglected when creating choice sets. If all of the choices represent students' true preferences over university departments, it is easy to reach a conclusion with this data. However, dependence of students' decisions on their CSSP scores could divert their choices from their true preferences. Because CSSP mechanism allows everyone to reach past years' placement scores, these data generate a baseline for students to form beliefs on the probability of being accepted by a department. Hence students can calculate the probability of being matched with their choices and they do not add unrealistic departments in their choice lists even these departments are more preferred. This makes it difficult to make the assumption that if a student adds a department in his/her choice list, this department is more preferred compared to departments which are not included in the student's choice list.

Table 8: Reverse Rankings

Score Differences	With Voc. Programs	Without Voc. Programs
>2	3121	2921
>5	2425	2238
>10	1782	1600
>20	1093	937
>30	666	531
>40	440	314
>50	360	235

A final crucial information in the choice lists is the rank of departments. Each student submits choice list in a descending ordered rank. This format is useful to make inferences about students' preference on selected department choice. Since CSSP mechanism uses only scores for matching, it is expected that students rank departments in the descending order of their minimum entry scores. If a student ranking is not in a descending order, acceptance probability to departments with higher minimum entry scores and to departments with lower ranks is lower, compared to students whose rankings were made in descending order. To understand students' order of rankings according to minimum entry scores, I check the number of students who have at least one reverse ranking in their choice lists. Reverse ranking means ranking departments with lower minimum entry scores higher than departments with higher minimum entry scores. Table 8 shows the numbers and the magnitude of reversals. Vocational programs have lower entrance scores and so they may cause the differences to be enhanced. Second column in Table 6 is for the number of students in the exclusion of the vocational programs. Even for big differences, there is a considerable number of reversals in students' rankings. Thus, it is inferred that rankings of departments do not only depend on minimum entry scores. By this way, I can exploit the ranking information in estimation without suspecting too much whether student ranks departments strategically or not.

I develop a unique consideration set for each student based on students' attributes revealed from their choice lists, in order to overcome huge amount of department sets and eliminate strategic behavior of students in their choice processes.

3.3 Consideration Sets

I construct consideration sets for each student in order to increase the explanatory power of estimation and to be able to manage the huge number of department set. Students' choices and their revealed attributes from their reported lists, give me clues about set of departments that students possibly consider for their choices. By using this information, I can reduce the whole set into individual and smaller sets that I can work without worrying big sets and without imposing any selection process to students' choice sets. To construct these sets, I benefited from the choice function $C()$ of students. This function selects some departments from a whole department set $\mathscr{D}$. $C_i(\mathscr{D})$ is the choice list of student i. It gives some revealed attributes (A_i) student i's department selection. I use a subset of these revealed attributes $(A_i^* \subset A_i)$ to create consideration set $D_i \subset \mathscr{D}$.

Table 9: Students Fixed Attributes

No	Attributes	Description
1.	***Exam Scores***	*Students scores in OSS Exam in all fields*
2.	***Demographic attributes***	
	Age	
	Sex	*1 for male, 2 for female*
	Location (which city student live)	*Center of the city used in the calculations*
	Family income	*income index from 1 to 12.*
3.	***Educational History Attributes***	
	High school type	
	OBP(High school GPA)	
	Students field type	
4.	***Students perception about the exam***	
	Concern	
	Importance for the success in future life	
5.	***Family attributes***	
	Siblings number	
	Siblings education levels	
	Parents occupations	
	Parents education levels	

Table 10: College Fixed Attributes

No	Attributes	Description
1.	***Previous year entrance Scores***	
	Minimum entry scores	
	Maximum entry scores	
2.	***University department quota***	
3.	***Field of department***	
4.	***Location***	*Center of the city used in the calculations*
5.	***City attributes***	
	Average income levels	
	Average cultural facilities	
	Cost of living index	
6.	***College type***	*State or Foundation*
7.	***College Country***	*University place in Turkey or not*
8.	***University specific attributes***	
	Name of university	
	Popularity of university	

Table 9 and 10 give descriptions of students and college attributes from the survey and various other databases([17], [3]). I call these attributes fixed attributes because they are independent from any interaction with students and colleges and same for all observations.

In Table 11, interactions between students and college attributes can be seen. The interactions are formed by multiplying attributes of students and colleges or by generating new variables by using these attributes. The reason to use interaction variables is the fixed effects regression. Because students' fixed attributes drop in the estimation, using them together with college fixed attributes provides an environment so as to insert students' attributes in the estimations. Most of our interaction variables in the form of multiplication of attributes; however, distance variables require special attention in our analysis. First one of these variables is the standard kilometer distance of city centers. Each student's high school city and colleges cities are taken as the basis for the distances. Second distance variables are score variables. A students' CSSP score can be in the three regions with respect to previous year entrance scores of departments. I split score variables in three parts to investigate the differences in these regions. One of the score variables is a dummy variable which represents whether students' scores from the previous year are in the departmental acceptance range or not. The remaining two variables are continuous and they show the distance of students' scores from departments' minimum and maximum entry scores from the previous year. Hence, a student's score variable is the only one of these three variables. For example, when a student has a score lower than the minimum entry score of selected department from the previous year, the only score variable is the distance of student score and departmental minimum entry score among the student variables. The other two score variables get 0. Other interactions are just multiplication of attributes.

Revealed attributes are applied score range, applied diploma programs and distance variables. Applied score range is obtained by students' choice lists and departmental minimum entry scores from the previous year. Range between student's maximum scored choice and minimum scored choice give us the applied score range.

Revealed attributes are applied score range, applied diploma programs and distance variables. Applied score range is obtained by students' choice lists and departmental mini-

Table 11: Attributes Interaction Table

			College Attributes								
			Scores		Location			Other			
			Min Score	Max Score	*Location (university)*	*Metropolitan*	*Cost of Living*	*Foundation*	*Scholarship*	*Evening Education*	*Two-Year Programs*
Students Attributes	**Education**	*Exam scores and GPA*	Min Score Distance	Max Score Distance		Score*Mc		Scores*F	Scores*sch	Score*eve	
		High school type (HST)						HST*F	HST*sch	HST*eve	
	Location	*Location (School)*			distance(100km) City diff(inc)						
		Location Type				ltype*Mc		ltype*F			
		Gender				Gen*Mc	Gen*col	Gen*F	Gen*sch	Gen*eve	Gen*two
	Family	*Income*				inc*Mc	inc*col	inc*F	inc*sch	inc*eve	inc*two
		Parents education				p_educ*Mc		p_educ*F	p_educ*sch	p_educ*eve	
		Number of siblings				sib*Mc		sib*F	sib*sch	sib*eve	
	Revealed	*Diploma Programs Revealed*									
		Score Range Revealed									

			Distance				
			Scores			Location	
			Min Score Diff	*Interval*	*Max Score Diff*	*Distance*	*City Differences*
Students Attributes	**Education**	*Exam scores and GPA*				Scores*dis	
		High school type (HST)					
	Location	*Location (School)*					
		Location Type				ltype*dis	ltype*cd
		Gender	Gen*minsd		Gen*minsd	Gen*dis	Gen*cd
	Family	*Income*	inc*minsd		inc*maxsd	inc*dis	
		Parents education	p_educ*minsd		p_educ*minsd	p_educ*dis	
		Number of siblings			sib*dis		
	Revealed	*Diploma Programs Revealed*					
		Score Range Revealed					

mum entry scores from the previous year. Range between student's maximum scored choice and minimum scored choice give us the applied score range. Diploma programs are generated in order to work easier with too much variety of departments. We call diploma programs a set of departments which are generally the same majors with different names in different colleges or they are so similar to be considered within the same major. Hence we reduce the number of selected programs to 138. On average, 3.5 degree programs were selected by students in our sample.

The subset of revealed attributes used in the formation of consideration set contains revealed score range and revealed applied departments. Reduction of choice sets with these revealed attributes relieve our problems as much as possible. Because we don't have data or a method to learn students' belief about the probability of being accepted to departments, it is difficult to appoint sets excluded from score effects. By using applied score range, it is possible to reduce sets without imposing any score range on students' decisions about their beliefs. Also, students could apply to all departments in this range independent from score effects. Limiting interest to only applied programs makes consideration sets independent from diploma program preferences. Hence these sets provide an environment for a study independent from scores' effects and major preferences.

3.4 Comparison of Choice Sets

In order to control consideration set suitability for the analysis, I compare them with some random choice sets. The two random choice sets are generated from the same observation numbers with consideration sets to reach results independent from choice set numbers. The first choice sets are random sets which the unselected departments in the choice sets are randomly selected from the whole CSSP department set. In this randomization there is no restriction such a field or score. The second random set is formed by randomization of departments from the selected diploma programs. Students reveal their preferred programs in their choice lists when they report their list to the CSSP. As explained before I reduce the number of programs to 138. By using this information, random sets within selected diploma programs are constructed. I looked at the explanatory power of whole estimations and some variables separately in these choice sets. Pseudo R^2 measures of regressions from the different sets provide a method for comparison. Additionally, I control log likelihood ratios of regression to check whether my method is correct.

Table 12: Estimation Results from Different Choice Sets

Variables	**Random Set**	**Random within dip.**	**Consideration Set**
Distance (dis)	-0.521***	-0.111***	-0.166***
	(0.0112)	(0.000101)	(0.0103)
Foundation (F)	-6.658***	-6.212***	-4.855***
	(0.195)	(0.148)	(0.173)
Min Score Distance	-0.0443***	-0.0223***	-0.0233***
	(0.00027)	(0.000243)	(0.000253)
Interval	0.0104	0.0959***	0.169***
	(0.0111)	(0.0105)	(0.0109)
Max Score Distance	-0.00846***	-0.00971***	-0.00891***
	(0.000156)	(0.000158)	(0.000158)
Evening Education (eve)	-0.852***	-0.430***	-0.477***
	(0.0815)	(0.0716)	(0.0728)
Scholarship (sch)	-0.0625	-1.956***	-3.293***
	(0.322)	(0.314)	(0.319)
Cost of living	-0.487***	-0.105***	-0.163***
	(0.0201)	(0.0201)	(0.0202)
Two-year program	-1.309***	-1.265***	-1.387***
	(0.0665)	(0.0722)	(0.0744)
Parents_educ*dis	-0.0184***	-0.0061***	-0.0065***
	(0.00151)	(0.00143)	(0.00147)
Parents_educ*F	0.252***	0.389***	0.204***
	(0.0236)	(0.0201)	(0.0241)
Parents_educ*MetropolCity	0.0374***	0.164***	0.179***
	(0.0101)	(0.0095)	(0.00976)
Parents_educ*sch	-0.168***	-0.302***	-0.141***
	(0.0299)	(0.0313)	(0.032)
Parents_educ*eve	-0.0690***	-0.00349	0.00299
	(0.0118)	(0.0114)	(0.0115)
Income*F	0.401***	0.308***	0.399***
	(0.0104)	(0.00864)	(0.0104)
Income*dis	-0.0104***	-0.0091***	-0.0118***
	(0.000682)	(0.000639)	(0.000646)
Income*MetropolCity	-0.0201***	0.0199***	0.0363***
	(0.00507)	(0.00497)	(0.00499)
Income*sch	-0.180***	-0.160***	-0.203***

Continued on next page

Table 12 – *Continued from previous page*

Variables	Random Set	Random within dip.	Consideration Set
	(0.0127)	(0.0132)	(0.014)
Income*eve	0.0746***	0.0796***	0.0845***
	(0.00532)	(0.00506)	(0.00515)
Distance*MetropolCity	-0.0376***	0.0508***	0.0385***
	(0.00275)	(0.00272)	(0.00274)
Score*F	-0.000619	0.000901***	-0.00631***
	(0.000487)	(0.000336)	(0.000371)
Score*dis	0.002***	0.0006***	0.0006***
	(0.0000275)	(0.0000245)	(0.0000246)
Score*MetropolCity	0.00548***	-0.00144***	-0.00118***
	(0.000147)	(0.000131)	(0.000135)
Score*sch	0.0113***	0.0141***	0.0200***
	(0.000869)	(0.000821)	(0.000829)
Score*eve	0.00121***	-0.00212***	-0.00219***
	(0.000221)	(0.000183)	(0.000186)
Gender*eve	0.0322	0.0908***	0.0965***
	(0.0215)	(0.0206)	(0.0208)
Gender*F	-0.000373	-0.106***	0.170***
	(0.039)	(0.0325)	(0.0388)
Gender*mc	-0.395***	-0.0972***	-0.115***
	(0.02)	(0.0195)	(0.0197)
Gender*dis	-0.0378***	-0.0464***	-0.0459***
	(0.00278)	(0.00262)	(0.00265)
Gender*sch	-0.115**	-0.0833	-0.376***
	(0.0506)	(0.0514)	(0.0536)
Locationtype*F	0.357***	0.210***	0.241***
	(0.0359)	(0.02)	(0.0356)
Locationtype*dis	-0.0159***	-0.0467***	-0.0240***
	(0.00217)	(0.00155)	(0.00205)
Locationtype*MetropolCity	-0.149***	0.0546***	0.00108
	(0.0142)	(0.011)	(0.0139)
Siblingsnumber*dis	0.0036***	0.0061***	0.0074**
	(0.000656)	(0.000629)	(0.000646)
Siblingsnumber*eve	-0.0348***	-0.0368***	-0.0394***
	(0.00639)	(0.00598)	(0.0062)
Siblingsnumber*MetropolCity	-0.0528***	-0.0180***	-0.00215
	(0.00494)	(0.00461)	(0.00464)
Siblingsnumber*F	-0.0756***	-0.0200**	-0.0934***
	(0.0137)	(0.00885)	(0.0143)
Siblingsnumber*sch	-0.0176	-0.0939***	-0.0304
	(0.0218)	(0.0212)	(0.0223)
Income*costofliving	0.0295***	0.0219***	0.0149***

Continued on next page

Table 12 – *Continued from previous page*

Variables	Random Set	Random within dip.	Consideration Set
	(0.00191)	(0.00192)	(0.00193)
Gender*costofliving	0.0488***	-0.0238***	-0.0197**
	(0.00869)	(0.00875)	(0.00883)
Gender*two-yearprog	0.134***	0.427***	0.419***
	(0.0296)	(0.0332)	(0.0336)
Income*two-yearprog	-0.0693***	-0.0607***	-0.0589***
	(0.00605)	(0.00667)	(0.00681)
Locationtype*costofliving	0.0509***	0.0107**	0.0458***
	(0.00452)	(0.00501)	(0.00457)
Pseudo R^2	0.2428	0.1209	0.1209
Observations	611,402	637,581	630,633
Number of groups	5,040	5,092	5,039

Table 12 shows regression results from three different choice sets. Pseudo R^2 ratios of regressions are significantly different. Random sets have highest pseudo R^2 and they explain nearly as twice as the other two sets. Consideration sets and random set within diploma programs have almost the same pseudo R^2. However, when we look at Table 13, the explanatory power of the variables show differences. Even if pseudo R^2 ratio is highest in the random set, most of the explanatory power of the estimation comes from score variables. This is the clear proof of the strategic behavior of students and our aim is to avoid this. Random set within diploma programs are eliminated from major preferences, but score variables have a powerful explanatory power in these sets. On the other hand, when we reduce our set within applied score range we almost eliminate all of the score variables variation in the data. This elimination provides us to construct our estimation approach without further struggling with scores' effects on choices. For the other variables, again consideration sets are better. Explanatory power of each variable is increased. Hence we can explain better the impact of attributes on choice behavior of students with excluded score effects.

Table 13: Comparison of Choice Sets

Variables	Random	Random within Diploma	Consideration Set
Distance	0.0285	0.0418	0.0449
Score	0.1453	0.0387	0.001
Foundation	0.0007	0.0001	0.0007
Evening education	0.0001	0.0038	0.0018
Scholarship	0.00006	0.0033	0.0085
Distance crosses	0.031	0.0954	0.0509
Cost of living	0.0003	0.0063	0.0048
Two-year program	0.0003	0.0128	0.0011

When I compare coefficients of the variables, it is seen that there are considerable differences. In particular, distance, scholarship, night education, cost of living, and score variables' coefficients are differentiated in the random set. Variety of departments in the random sets causes formation of choice sets with departments which have many different attributes. These change estimation coefficients for random choice set from the other. For instance, distance avoidance is almost 4 times more than other choice sets. Because of the random assignment of departments to the choice set, possibly many far away colleges' departments do not represent selected departments and distance avoidance is enhanced.

Furthermore, in order to check consideration sets I change the size of the set by playing with the scores range of our sets. I looked at regression outcomes by expanding the score range by adding extra ranges to revealed ranges. The results of these choice sets show that the explanatory power of variables decreased apart from score variables with the expansion of the ranges. Since the strategic effects of choices show their power in the larger score range sets, using consideration sets within revealed applied ranges is a better way to estimate preferences.

3.5 Estimation Results

Table 14 shows the 2 different estimation results. In the first column all students who report choices are included in the estimation.The second column is for students who did not live in metropolitan areas. Because the demographic, socioeconomic and educational background characteristics of students from metropolises are fairly different from students from other cities, I want to check the effects of these differences on choices.

Table 14: Estimation Results

Variables	1	2
Distance(dis)	-0.168***	-0.00288***
	(0.0115)	(0.0135)
Foundation(F)	-7.016***	-6.274***
	(0.18)	(0.25)
Min Score Distance	-0.00336***	-0.00253***
	(0.000317)	(0.000396)
Interval	0.0720***	0.135***
	(0.0118)	(0.0143)
Max Score Distance	-0.00278***	-0.00113***
	(0.000247)	(0.000315)
Evening Education(eve)	-1.056***	-1.107***
	(0.0788)	(0.0998)
Scholarship(sch)	1.365***	0.616
	(0.357)	(0.52)
Cost of living	-0.150***	-0.196***
	(0.022)	(0.027)
Two-year program	-1.086***	-1.113***
	(0.149)	(0.183)
Parents_educ*dis	-0.0049***	-0.006***
	(0.00162)	(0.00188)
Parents_educ*F	0.438***	0.416***
	(0.0251)	(0.0357)
Parents_educ*MetropolitanCity	0.115***	0.101***
	(0.0108)	(0.013)
Parents_educ*sch	-0.335***	-0.354***
	(0.038)	(0.0517)
Parents_educ*eve	0.00647	-0.0388***
	(0.0125)	(0.0149)
Income*F	0.300***	0.220***
	(0.0105)	(0.0149)
Income*dis	-0.0131***	-0.008***
	(0.000723)	(0.000856)
Income*MetropolitanCity	0.0278***	0.0314***
	(0.00557)	(0.00679)
Income*sch	-0.147***	-0.0441*
	(0.0166)	(0.0241)
Income*eve	0.0878***	0.119***

Continued on next page

Table 14 – *Continued from previous page*

Variables	1	2
	(0.00554)	(0.00693)
Distance*MetropolitanCity	0.0359***	0.104***
	(0.00306)	-0.00358
Score*F	0.00623***	0.00556***
	(0.000411)	(0.000595)
Score*dis	0.0008***	0.0008***
	(0.0000281)	(0.0000331)
Score*MetropolitanCity	-0.00118***	-0.00122***
	(0.000146)	(0.000185)
Score*sch	0.0008	0.00187
	(0.000923)	(0.00139)
Score*eve	-0.000146	-0.000101
	(0.000202)	(0.000253)
Gender*eve	0.0993***	0.146***
	(0.0224)	(0.0267)
Gender*F	-0.180***	-0.0989*
	(0.0402)	(0.059)
Gender*MetropolitanCity	-0.0561***	-0.113***
	(0.0215)	(0.0256)
Gender*dis	-0.0442***	-0.0337***
	(0.00296)	(0.00341)
Gender*sch	-0.0254	-0.257***
	(0.0624)	(0.0913)
Locationtype*F	0.122***	0.127***
	(0.0232)	(0.0428)
Locationtype*dis	-0.0483***	-0.0139***
	(0.00175)	(0.00251)
Locationtype*MetropolitanCity	0.0642***	-0.00933
	(0.0122)	(0.0193)
Siblingsnumber*dis	0.0076***	0.0061***
	(0.000722)	(0.00085)
Siblingsnumber*eve	-0.0351***	-0.0645***
	(0.00645)	(0.00834)
Siblingsnumber*MetropolitanCity	-0.0225***	-0.0548***
	(0.00521)	(0.0067)
Siblingsnumber*F	-0.0262**	0.0359***
	(0.0105)	(0.0121)
Siblingsnumber*sch	-0.0989***	-0.147***
	(0.0234)	(0.0303)
Income*costofliving	0.0226***	0.0158***
	(0.00209)	(0.00254)
Gender*costofliving	-0.0158*	-0.000517

Continued on next page

Table 14 – *Continued from previous page*

Variables	**1**	**2**
	(0.00951)	(0.0112)
Gender*two-yearprog	0.487***	0.671***
	(0.0708)	(0.0952)
Income*two-yearprog	-0.0583***	-0.111***
	(0.0131)	(0.0183)
Locationtype*costofliving	0.0194***	0.0494***
	(0.00545)	(0.00825)
Observations	753518	466537
Number of groups	5047	3374

As expected, I find that the applicants of the university entrance exam prefer departments whose last year entrance scores are similar to students' own exam scores. Students do not prefer departments whose last year minimum entry score is higher than their exam scores or when the last year maximum entry scores are lower than students' scores. Even though we constrain our consideration sets with students' revealed score ranges, the preference over departments is more for departments whose last year acceptance score range contains their exam scores. Students avoid departments far away from their hometown. However, the distance preferences show differences according to certain attributes. Increase in average parents' education level and income level, living in a higher populated area, and being female strengthen this avoidance. On the other hand, higher exam scores and having more siblings reduce the avoidance of distance in preferences. Colleges administered by foundations and which require paying higher tuition fees are not preferred by applicants. Conversely, students from higher income levels, from educated families and who live in bigger cities show less avoidance to foundation colleges as expected. Moreover, departments which provide scholarships to students are the more preferred ones among all others. Again expectedly, night education and two-year vocational programs are not preferred from the general part of the sample.

Student preferences are not only related with colleges, but students also make their selection according to the city. In order to understand city preferences of students we add some city attributes to our estimation. The most important city attribute is the cost of living index. Students' choices show that studying in an expensive city is not preferred. Especially for female students and students from lower income groups, this avoidance is enhanced. Metropolitan cities (Istanbul, Ankara and Izmir) are the most expensive places to live and study. Despite the high number of colleges in these cities, higher living costs make them less preferable.

In order to understand risk preferences of students, I investigate students' score distance from the previous year's minimum entry scores. It is a suitable parameter to analyze risk because the amount of distance directly gives the risk level. The more the distance between scores, the lower the probability of students to enter the department. Therefore, it is reasonable to say that students whose department choices have, on average, more distance from the minimum entry scores are more risk lover than others. Results show that female students and interestingly students who come from higher income groups avoid more riskier choices.

Table 15: Estimation Results

Variables	**Rank-Logit**	**Conditional-Logit**
Distance	-0.214***	-0.248***
	(0.0013)	(0.00152)
Foundation	-0.0786***	-0.149***
	(0.0151)	(0.0175)
Min Score Distance	-0.00461***	-0.00536***
	(0.000269)	(0.000293)
Interval	0.106***	0.179***
	(0.0109)	(0.0126)
Max Score Distance	-0.00177***	-0.00148***
	(0.000232)	(0.000242)
Evening Education	-0.429***	-0.470***
	(0.01)	(0.0111)
Scholarship	-1.558***	-1.677***
	(0.028)	(0.0306)
Two-year program	-0.726***	-0.828***
	(0.0319)	(0.0346)

Finally, estimation results show differences with the changes in the methods. In order to increase understanding and compare results in different estimations, I reduce number of independent variables in the estimations. Therefore, I eliminate various effects of students attributes on college preferences. Table 15 shows two basic estimation results without analyzing students attributes impacts on college attributes. The regressions are made with the consideration sets and in both regression same data is used. In the first column the results come from rank ordered regression. In the second column the estimation results are from the conditional logistic regression. These results shows that canceling ranking weights in estimation increases coefficients magnitudes. One possible meaning of these increase is the departments which are ranked higher in the choice list are not same preferences with the total choice list. Results show that students avoid less distance, foundation college, and departments which support scholarship. Interestingly, evening education and two-year vocational programs are also less avoided from student for higher ranked choices. Avoidance from departments with scholarship in both regressions needs further investigation. Decreasing in minimum score distance coefficient is expected, because students ranks departments which have higher minimum entry score in the first places in the their choice lists. Similarly, increase in the maximum score distance coefficient shows the reverse of this change.

4 A Discussion on Effects of Attribute Differences in College Choice Preferences

As it was previously noted above, having different attributes change student preferences. Gender, income level, location type and size, parents' educational levels and exam scores have significant effects on student choices. In this section, I seek to analyze whether these estimated preferences create a stratification in labor market outcomes.

Colleges have impact an impact on their students when they are studying and after they graduate. There are clear-cut differences in many directions between students from more prestigious or provincial universities. The situation is the same for Turkey. There are some high-quality colleges, which are perceived better in the labor market. Therefore, we can say that probabilities of finding better job opportunities of students who have graduated from these colleges are higher. Another criterion for the job market is the possibility of finding a job in the area where the college is located. Students who are studying in colleges located where job possibilities are concentrated and which require more high educated labor have more chances for finding better jobs by networking. These are the main two factors in order to evaluate labor market outcomes of colleges.

TEPAV (2007) [15] conducted a study about higher education and labor market in Turkey.

The report shows differences in college graduates' earnings from selected universities and departments. The study is based on data generated by interviewing and making questionnaires with firms, colleges and students. The report finds that ”for university graduates, private sector thinks that school of graduation is more important than the match between the job and the department graduated from.” Moreover, ”the average high school, vocational high school, vocational college, open-university and university graduates earn 734, 772, 818, 1109 and 1450 YTL respectively, there are significant differences between universities and departments.” According to the findings, it is shown that regional business and industrial needs change the wages of graduates significantly.

In Turkey, metropolitan districts (Istanbul, Ankara, Izmir) host both almost all of the prestigious colleges and the main part of various industries. Therefore, these areas are the best places in Turkey to study and find jobs. This clustering of colleges provides a useful comparison criterion. Estimation results show that students who are from lower income groups; who live smaller locations; who have less educated parents; and who are female have a significant negative preference for metropolitan areas. An interesting result is that students with higher scores may also show an avoidance for metropolitan cities. Students with higher scores have higher probabilities to be accepted to departments which always have higher minimum entry scores compared to previous years. Although most of these departments are located in the metropolitan cities of Turkey, students with higher scores show more avoidance to these cities than students with lower scores. One of the reasons for this is the higher cost of living in the metropolitan areas. When I look at the cost of living index of cities, it can be shown that students from lower income levels and female students in particular have an extra avoidance for expensive cities.

In order to eliminate the effects of students living in metropolitan areas, I look at the estimation results of students from other cities of the country (Column 2 in Table 14). In this estimation, again female students and low-income students don't prefer to choose departments in the metropolitan cities of Turkey. Interestingly, for students who are not from metropolitan cities, the increase in the size of residential areas increases the avoidance of metropolitan colleges.

Distance preferences is also a constraint for some groups. Estimation results show a strong avoidance from the long distance choices. This preference makes students more immobile, especially female students and forces them to find colleges in a restricted area. Hence students living in the periphery of the country do not prefer to go to colleges far away even if they are better.

Two-year vocational programs and foundation college preferences give a clue about the educational quality and future job finding of students. Students who have graduated from vocational programs are less likely to find high-earning jobs. It is shown from the estimation results that female students show less avoidance to these programs. This creates a gap between male and female student outcomes in the job market. Foundation colleges are avoided more by women and lower income group students. Being a student in these colleges requires significant amount of funds and it is expected that students from poorer families will have a negative preference to them. However, surprisingly, female students tend to avoid them more than male students.

In the end, it is seen that preferences of students who are female; who are living away from metropolitan areas; who are from low-income groups cause them to be in a disadvantaged position for the labor market. Conversely, students who live in metropolitan cities and who have higher educated parents have more chances to find higher earned jobs.

5 Conclusion

Using a data set from the CSSP system and a representative survey of students, I first estimate students' college preferences. By appointing consideration sets to each student to help dealing with the huge number of departments alternative in the CSSP system, I am able to reach more significant results from the estimations. Comparisons of consideration sets with the random sets indicate the strength of consideration sets in estimation.

I document that differences in attributes (such as gender, income, type of location, etc.) among students affect their college choices under the control of their scores and major preferences. According to estimation results, a variety of attributes among students causes them to make their choices differently. Especially, differences in income, gender and location create a significant preference variety. On the other hand, avoidance from out of range departments, long distance from students' location, foundation colleges and two-year vocational programs are common preferences.

Finally, I analyze the outcomes of these preferences by using possible placements. Quality, ranking of colleges, job possibilities of cities, and perception of colleges in the labor market are used in the assessment of these outcomes. The findings show that preferences cause preservation of gaps between females and males, rich and poor, and those living in disadvantaged areas and metropolitan areas. Hence highly selective colleges consequently result in high-wage occupations and industries continue to be dominated by males, from higher income groups and those who have been brought up in bigger cities.

Based on the results obtained in this study, I conclude that CSSP data provide valuable insights toward understanding the heterogeneity in college choices. Reported evidence on differences in college choices do not only provide an explanation for the persistent attribute gaps in highly selective college enrollments, high-wage occupations and industries, but it also offers a new perspective on heterogeneity in school choice.

References

[1] A. Alkan, A. Çarkoglu, A. Filiztekin, and F. Inceoglu. Turkiye ortaögretim sektöründe katma deger olusumu; Üniversiteye giris yarisi. Technical report, Sabancı University, 2008.

[2] M. Balsamo, M. Lauriola, and A. Saggino. Personality and college major choice: Which come first? *Psychology*, 3:399–405, 2012.

[3] CSSP. *OSS User Guidance Booklet*.

[4] S. L. DesJardins, H. Dundar, and D. Hendel. Modeling the college application decision process in a land grant university. *Economics of Education Review*, 18(1):117–132, 1999.

[5] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):915, 1962.

[6] G. J. Hitsch, A. Hortaçsu, and D. Ariely. What makes you click?-Mate preferences in online dating. 1, 2009.

[7] G. J. Hitsch, A. Hortaçsu, and D. Ariely. Matching and sorting in online dating. *American Economic Review*, 100(1):130–63, March 2010.

[8] D. Hossler, J. Braxton, and G. Coppersmith. Understanding student college choice. *Handbook of Theory and Research*, 5:231 288, 1987.

[9] T. J. Kane. College-going and inequality. A literature review, UCLA, 2001.

[10] Y. Ma. Family socioeconomic status, parental involvement, and college major choicesgender, race/ethnic, and nativity patterns. *Sociological Perspectives*, 52(2):211–234, Summer 2009.

[11] D. McFadden and K. E. Train. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15:447–470, 2000.

[12] S. W. Polachek. Sex differences in college major. *Industrial and Labor Relations Review*, 31(4):498–508, 1978.

[13] G. N. Punj and R. Staelin. The choice process for graduate business schools. *Journal of Marketing Research*, 15:88–98, 1978.

[14] P. O. Saygin. Gender differences in college applications: Evidence from the centralized system in Turkey, 2011. Job market paper.

[15] Tepav. Higher education and the labor market in Turkey. Technical report, Tepav, 2007.

[16] K. E. Train. *Discrete choice methods with simulation*. Cambridge University Press, Cambridge, United Kingdom, 2003.

[17] S. Tüylüoglu and A. S. Albayrak. Cost of living and examining with ridge regression analysis the most important factors that determine cost of living rankings of provinces in Turkey. *The Journal of Faculty of Economics and Administarive Sciences*, 2010.