ISGUC.ORG

"İŞ, GÜÇ" ENDÜSTRİ İLİŞKİLERİ VE İNSAN KAYNAKLARI DERGİSİ
"IS, GUC" INDUSTRIAL RELATIONS AND HUMAN RESOURCES JOURNAL

Standardized Test-Based Student Selection and Gender Differences in Academic Achievement

F. Kemal KIZILCA
Ph.D., Ankara University, Department of Economics

Ekim/October 2013, Cilt/Vol: 15, Sayı/Num: 4, Page: 102-115
ISSN: 1303-2860, DOI: 10.4026/1303-2860.2013.0240.x

Makalenin on-line kopyasına erişmek için / To reach the on-line copy of article:
http://www.isguc.org/index.php?p=article&id=538&cilt=15&sayi=4&yil=2013

Makale İçin İletişim/Correspondence to:
Dr.F.Kemal KIZILCA/ Email: kkizilca@law.ankara.edu.tr

"İş,Güç" Endüstri İlişkileri ve İnsan Kaynakları Dergisi
"İş,Güç" Industrial Relations and Human Resources Journal

Ekim/October 2013, Cilt/Vol: 15, Sayı/Num: 4,
Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

"İş,Güç" Endüstri İlişkileri ve İnsan Kaynakları Dergisi, yılda dört kez yayınlanan hakemli, bilimsel elektronik dergidir. Çalışma hayatına ilişkin makalelere yer verilen derginin temel amacı; belirlenen alanda akademik gelişime ve paylaşıma katkıda bulunmaktır.
"İş,Güç" Endüstri İlişkileri ve İnsan Kaynakları Dergisi'nde,'Türkçe' ve 'İngilizce' olarak iki dilde makale yayınlanmaktadır.
"İş,Güç" Endüstri İlişkileri ve İnsan Kaynakları Dergisi, ulusal ve uluslararası birçok indekste taranmaktadır. (Cabell's Directories, Ebsco Socindex, Index Islamicus, Index Copernicus International, Worldwide Political Science Abstracts, Sociological Abstract, Ulakbim Sosyal Bilimler Veritabanı, ASOS Index)

Tarandığı Indeksler / Indexes
ASOS INDEX
CABELL'S DIRECTORIES
EBSCO SOCINDEX
Index ISLAMICUS
Index COPERNICUS Int.
Sociological Abstract
ULAKBİM Sosyal Bilimler Veritanı
Worldwide Political Science Abstracts

"İŞ, GÜÇ" Endüstri İlişkileri ve İnsan Kaynakları Dergisi Ekim/October 2013 - Cilt/Vol: 15 - Sayı/Num: 04
"IS, GUC" Industrial Relations and Human Resources Journal Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

Standardized Test-Based Student Selection and Gender Differences in Academic Achievement[1]

Standartlaşmış Sınavlarla Öğrenci Seçimi ve Öğrenci Başarısında Cinsiyet Farkları

F. Kemal KIZILCA
Ph.D., Ankara University, Department of Economics

Abstract:

Previous studies suggest that standardized test-based student selection creates a bias against female applicants. The underlying reason is the standardized tests' inability to measure some student aptitudes on which female students have advantages. In this study I propose a methodology for measuring the gender bias that standardized tests cause. The methodology is based on comparing gender differentials in student grades between two types of subjects, namely, subjects that standardized tests can predict student success, and the others for which non-standardized measuring gain importance. I apply this methodology to the data I collected from the Law School of Ankara University. The results show that the standardized test scores do not have any predictive power for the student grades in Law subjects where a gender bias in favor of female students exist. The test scores have predictive power for the Economics grades only, where the gender gap disappears. I conclude that, the use of standardized tests results in admission of relatively higher share of low-skilled male students to the Department.

Keywords: *Standardized tests, student selection, gender differentials, SAT*

Özet:

Standartlaşmış sınavlarla öğrenci seçimi üzerine daha önce yapılmış çalışmalar, bu sınavlarla öğrenci seçmenin bir cinsiyet ayrımcılığına yol açtığını ileri sürmüşlerdir. Ayrımcılığın kaynağı, bu sınavların, kadın adayların üstün olduğu ve öğrenci başarısı için önemli bazı yetenekleri ölçmede yetersiz kalmasıdır. Bu çalışmada, ayrımcılığın tespiti ve ölçümü için bir yöntem önerilmektedir. Önerilen yöntem, standartlaşmış sınavların öğrenci başarısını tahmin edebildiği derslerdeki cinsiyet farklarıyla, tahmin edemediği derslerdeki cinsiyet farkının kıyaslanmasına dayanmaktadır. Ankara Üniversitesi Hukuk Fakültesi öğrencilerinin notları üzerine yaptığımız tahlil, üniversite giriş sınavı sonuçlarıyla, hukuk derslerinden alınan notlar arasında bir ilişki olmadığını ve bütün hukuk derslerinde kadın öğrencilerin daha başarılı olduğuna işaret etmektedir. Giriş sınavı puanı ile alınan not arasında anlamlı bir ilişkinin bulunduğu tek ders İktisat'tır. Bu dersten alınan puanlarda anlamlı bir cinsiyet farkı ise mevcut değildir. Bulgular, standartlaşmış sınavla öğrenci seçmenin, görece fazla sayıda yetersiz donanıma sahip erkek öğrencinin Fakülte'ye girmesiyle sonuçlandığını doğrular niteliktedir.

Anahtar kelimeler: *Standartlaşmış sınavlar, öğrenci seçimi, cinsiyet farkları, Öğrenci Seçme ve Yerleştirme Sistemi (ÖSYS)*

[1] The author is thankful to an anonymous referee for helpful suggestions.

Ekim/October 2013 - Cilt/Vol: 15 - Sayı/Num: 04
Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

1. Introduction

Various studies suggest that the heavy reliance on Scholastic Aptitude Test (SAT) scores in university admissions and scholarship grants in the USA creates a bias against female applicants (Moss, 1988-1990; Young, 1991; Stricker *et al*., 1993; Connor & Vargyas, 1992; Silverstein, 2000-2001). These studies rely on the observations that females systematically had higher high-school GPAs and lower SAT scores compared to their male counterparts, and SAT underpredicted the success of first year female students in the U.S. This gender bias is caused by SAT's inability to measure many student qualifications that female students excel. Based on these criticisms, the authors favored substitution of standardized SAT scores with non-standardized exam results or GPAs as admission criteria.

While the above-mentioned studies concern with the legal steps to be taken to tackle the issue of gender bias, the current study proposes an empirical strategy to test and measure the bias caused by standardized tests in student selection. I can summarize the strategy as follows: There are certain student characteristics which affect academic achievement positively that standardized tests cannot measure. These include – but are not limited to – written and oral expression, self-discipline for long-term study, creative thinking, and ability to teamwork. If female students have advantages in – at least some of – these characteristics as the above-mentioned studies suggest, once being admitted to the program, they should get higher grades in those courses for which these qualifications are important, and be in par with male students in courses where standardized tests can predict achievement. Therefore, the methodology applied here is based on comparison of gender differentials in academic achievement for the same cohort of students between these two types of courses. For example, in most sociology departments, students take statistics courses in which standardized tests are useful to measure academic achievement, and history courses in which narrative is of much importance. If the student selection process is based on multiple-choice exams, we shouldn't observe any systematic differences in achievement among any sub-groups of students in statistics course. In history, however, students with higher ability for written expression would get higher grades. There are various studies showing that females in general, indeed, perform better in those fields where language skills and narrative are important (Guiso et al., 2008; and Holmlund & Sund, 2008). Thus, a standardized test-based student selection may result in a gender differentiation in academic performance in history courses in favor of females, whereas a similar grade distribution by gender is expected in statistics.

The case of Turkish Student Selection and Placement System (*ÖSYS*) is a perfect laboratory for testing this hypothesis. In Turkey, students are selected to higher education institutions by a multiple-choice centralized exam. Unless the department requires a specific talent exam (fine arts and sports departments only), all departments accept their students based on the applicants' *ÖSYS* grade. *ÖSYS* basically covers five fields, namely, mathematics, Turkish language and literature, science, social sciences, and foreign language. There are different weights attached to each field, depending on the department's required grade type. The foreign language field grade is used only for foreign language and literature departments, and the contribution of high school GPA on *ÖSYS* grade is very limited.

The data is used in this study comes from a questionnaire I administered to the first-year students of Faculty of Law of Ankara University. Like most of the other ones, the law departments in Turkey select their students based on their *ÖSYS* scores. However, the standard law curriculum comprises many courses in which success is difficult to measure by standardized tests. In their first year, students take four mandatory law courses, namely, Introduction to Law, Civil Law, Roman Law, and Constitutional Law besides some other additional courses. Among the additional courses is a two-semester Introduction to Economics, which

differs from all the others with its mathematical, statistical and graphical content as well as its methodology. Because of its nature, this paper considers economics, among the first-year courses, as the only course for which standardized tests are expected predict success.

Thus, the empirical strategy of this paper is based on estimating the relationship between students' standardized university entrance exam scores and their grades from the mid-term exams; and, as a second stage, compare the gender differentials in Economics and all other mandatory Law courses. The results show that the only course which shows a statistically significant relationship with *ÖSYS* score is Economics. This finding is in line with my expectation that it is difficult to predict students' achievements in Law courses by means of standardized tests. For the second step, the estimations show that in Law courses, without any exception, there is a gender differentiation in favor of females. In Constitutional Law and Introduction to Law, this difference is statistically significant and in Roman Law the level of significance is on the margin. The only course that females have a lower grade in average is Economics, while the difference is not statistically significant.

Based on these findings, I conclude that standardized tests are not good predictors of student achievement for the Law departments, and their inadequacy in measuring the student aptitude creates a gender bias in selection.

2. A Brief History of the Use of Standardized Tests in Student Selection

In the United States, the first use of standardized exams was the IQ tests administered to the army draftees. It was in 1926 that College Entrance Examination Board, the institution in charge of conducting centralized exams for student selection for the higher-education programs, adopted this test. Today, Educational Testing Services, a tax exempt company, conducts SAT.[2]

Turkey followed the US model in the use of standardized tests both in the Armed Forces exams and student selections[3]. In 1950, a multiple-choice exam was administered at the School of Commissioned Officers of the Army[4]. The first use of multiple choice exams at a higher education institution was in 1951 at Istanbul University.[5] Following the establishment of the Inter-University Commission for the Entrance Test, all universities in Turkey –except for two– began to select students based on a centralized multiple-choice test. Since 1974, all universities and departments in Turkey –with few exceptions – are obliged to accept their students based on one centralized multiple-choice exam. *ÖSYS* administers this exam once a year in all regions of Turkey simultaneously.

Although the application of standardized exams for student selection in Turkey followed the American SAT model, there remain some fundamental differences between the two systems. The score obtained from the *ÖSYS* test is the sole determinant of a student's admission to a program in Turkey, with the exception of some sports and fine arts departments which require additional talent exams. In the USA, however, it is the university (or the department) who decides how much weight to put on SAT scores of different fields. Moreover, it's in the university's discretion to decide whether to conduct an additional exam, take into account of high-school GPAs, reference letters, or ask for any additional documentation.

Another important difference between the two admission systems is that, throughout the years, SAT became much richer in its variety of subjects, while the *ÖSYS* remained relatively unchanged. The exam conducted by *ÖSYS* has five main sections, namely, math (including geometry), Turkish language and literature, science subjects (physics, chemistry

[2] See Silverstein (2000-2001) and sources cited therein for the history of the use of standardized tests in the USA.
[3] Milliyet, 18.10.1951, p.2 and Milliyet, 17.03.1953, p.2.
[4] Milliyet, 23.11.1950, p. 5.
[5] Milliyet, 23.08.1951, p. 2.

and biology), social sciences (mostly history, geography, psychology questions), and foreign language. The SAT, however, includes 20 optional sections in addition to its reading, math and writing fields.

Because of the graduate-level nature of law education in the USA, Law School Admission Test (LSAT) is governed by another institution. This exam consists of multiple choice tests on reading, logical thinking and analytical thinking fields. There is also an unscored section in which the applicant writes an essay.[6]

In order to be admitted to a Law department in Turkey, an applicant must take the Turkish Language, Social Sciences and the Math sections of the *ÖSYS* test. The score that a student obtains from the math section of this test is more important from the other sections for two reasons. Firstly, because the other two sections are relatively easier[7], it is the math section which determines the differences in total score among the applicants. Secondly, the weight attached to the math section in calculation of the score for the law schools is the highest among all sections. The weight is 36 percent for the Turkish Language and Literature sections, 8 percent for geography, 7 percent for social sciences, 5 percent for sciences, while it is 44 percent for the math and geometry. It is the joint result of these two reasons that the differences among the applicants' *ÖSYS* scores mostly reflect their success in mathematics.

3. Standardized Tests and Gender Differences in Achievement

The majority of the studies on the gender differences at schools concentrate on the mathematics subjects. While these studies tend to point out a gender gap in favor of males in average, this conclusion is not universal. In their meta-analysis, Hyde et al. (1990: 151) conclude that "females are superior in computation, there are no gender differences in understanding of mathematical concepts, and gender differences favoring males in problem solving do not emerge until the high school years". It is the males' advantage in problem solving which creates the observed gap in average. The gap, however, is decreasing in time (Hyde *et al.*, 1990; Cole, 1997).

Gender differences in the mathematics also vary across countries. Based on the panel results from PISA test which is administered in all OECD countries, Guiso *et al.* (2008) show that there is a correlation between a country's level gender equality (measured as a lower value of Gender Gap Index, GGI) and gender differences in mathematics. The gender gap in mathematics approaches to zero in countries such as Norway, Sweden and Netherlands where the GGI is the highest.

Contrary to the findings regarding the gap in mathematics, results obtained from written expression, reading and foreign language subjects systematically point out the advantage of females in these areas. In the reading section of the PISA test, female students from all 40 countries obtained higher scores. This gap widens even further in countries with higher GGI. While, the gap in favor of male students in math and science subjects in the USA narrowed from 1960 to 1990, the gap in favor female students in writing, remained persistent (Cole, 1997, p.12). Holdmlund and Sund (2008) also show that, in Swedish high schools, female students are more successful compared to their male counterparts both in mathematics and language courses, while the gap is even bigger in the latter.

Because of female students' advantage in written expression, their multiple-choice test results tend to be relatively lower than their scores in essay-writing or constructed response exams (Williams et al., 1992; Gamer and Engelhard, 1999). Thus, conducting a multiple-choice exam in an area where constructed

[6] <www.lsac.org> , reached: 26.04.2011.

[7] In the 2010 Undergraduate Placement Test, which is the second and the last step of *ÖSYS*, the average success rate was 27.8 correct answers out of 56 questions (50 %) for the Turkish Language and Literature section, 41.5 out of 114 questions (36 %), in the Social Sciences section while the same rate was only 14.7 out of 50 questions (29 %) for the Math section.
<http://osym.gov.tr/dosya/1-55969/h/lyssonuc2010sunum.pps>, reached: 30.04.2012.

answers are essential will create a gender bias. One method of eliminating this bias would be use of non-standardized tests results (i.e. essays, interviews, etc.) or high school GPAs in student selection[8].

In 1989, Khadijah Sharif *et al.* sued New York State Education Department on the grounds that, the Department's policy to grant scholarships solely based on SAT scores created a discrimination against female applicants. The plaintiffs asserted that the gender gap in the SAT scores were systematically higher than the gap in high school GPAs, thus, SAT did not reflect female students' true aptitudes. The claim that SAT created a gender bias was actually a matter of debate even before the start of the lawsuit. The legislation made a decision in 1987 to run an experiment to check the validity of these claims. As an experiment in 1988, the State took applicants' high-school GPAs into account in addition to their SAT scores in scholarship grants. As a result, the number of female students who were awarded a scholarship increased from *71* to *86*, while the same number for the males decreased from *129* to *113* in one year[9]. It was after this outcome that the court found the plaintiffs right and the experimental change in the legislation became permanent.

Another important step to eliminate the gender gap in SAT scores was in 1999, when Preliminary SAT was extended to include a "writing skills" section with the efforts of FairTest, a non-governmental organization dedicated to fight discrimination in student selection. After inclusion of this section there was a *40* percent reduction in the gender gap in the next year, with a further *26* percent in the following. Despite all these efforts, the gender gap in the SAT scores remains higher compared to the gap in high-school GPAs.[10]

As stated previously, standardized test results are the sole admission criteria for university admissions in Turkey for the majority of the programs. Moreover, the score obtained from the math section of the *ÖSYS* exam remains as the most important among all other sections –even for social sciences, law and humanities departments– as long as the applicant targets a top-rated department. The standardized multiple-choice tests, however, are not able to measure many valuable student characteristics such as long term study discipline, analytical thinking, logical thinking, written or oral expression, all of which may affect the achievement at university-level positively. The literature summarized above shows that female students had better qualifications in these aptitudes. Thus, standardized tests which are unable to catch these aptitudes will create a bias against females. In the next section I provide the empirical evidence on the gender bias as a result of student selection based on standardized testing.

4. Data and the Results

4.1 Gender Gap in Student Achievement by Subjects

In 2010-2011 academic year, the Faculty of Law of Ankara University admitted 800 new students. Students were sorted based on their *ÖSYS* scores; and starting from the first, each of them were placed in three sections (A, B and C) one by one. This placement method prevented any systemic differences across sections with regard to the mean *ÖSYS*. The data used in this study comes from the section C only; but thanks to the section placement method, I assume that the

[8] Kuncell and Hezlett (2007) propose that standardized tests which are tailored for graduate programs –such as Graduate Record Examination (GRE-T), Graduate Record Examination Subject tests (GRE-S), the Law School Admissions Test (LSAT), the Pharmacy College Admissions Test (PCAT), the Miller Analogies Test (MAT), the Graduate Management Admissions Test (GMAT) and the Medical College Admissions Test (MCAT)- predict graduate school GPAs successfully. However, the authors stress the importance of the use of multiple selection criteria such as recommendation letters, statements of purposes and interviews. Moreover, they also point out that gender bias remains in standardized tests.

[9] 709 F. Supp. 345 (S.D.N.Y 1989).

[10] Fairtest (1999), "PSAT Revisions Further Narrow Gender Gap", source: <http://fairtest.org/psat-revisions-further-narrow-gender-gap>, reached: 29.04.2011.

results reached here have representative power for the rest of the first year students.

Economics is a mandatory course in the first year of the Faculty of Law. There were *352* registered students in total at the C section of the first-year students in 2010-2011 academic year. When I exclude the students coming from foreign countries, who were not subject to the *ÖSYS* system, and students who were repeating the course, this figure falls to *272*. I conducted the survey at the first week of the year, when the attendance rate was at its peak. The response rate to the survey from the Turkish-national students was *249*, among whom only *4* were repeating the course[11]. The seating capacity of the classroom was 256, which was below the total student number. This means that there were only *7* students who refused to participate in the survey, with a resulting response rate of *97* percent[12].

Limiting the survey with the first-year students eliminates any sample-selection problems resulting from class repetitions or school drop-outs. After the first year, further gender gaps may arise because of course selections, drop-outs and class repetitions, etc. (see Dayıoğlu and Türüt-Aşık, 2007). I further restrict the data with the mid-term exams only, considering the make-up option for the finals, which may raise further methodological complications. By doing this, I assume that if *ÖSYS* scores predict student achievement successfully, the resulting correlation should be observed in the first semester grades. This remains as a realistic assumption as long as problems arising from students' adaptation period in the first year do not mask the "real" predictive success of standardized tests.

Since this study does not aim to make any investigations on the differences in field and course selections by gender, the data is limited with only mandatory law courses and Economics, which is also mandatory. Because of its mathematical and graphical content, it is relatively easier to measure achievement in Economics by standardized tests. So, I expect the Economics grades of the students be correlated with their *ÖSYS* scores. The other course scores investigated in the study are of Civil Law, Roman Law, Introduction to Law and Constitutional Law, which constitute all of the mandatory law courses in the curriculum.

Figure 1 plots the distribution of grades for each course by gender. It is visible in the graphs that the females are concentrated at higher grades in every law course compared to the males. The gap is at the highest in Introduction to Law. In Economics, however, the graph shows that the female grades are distributed in a narrower range while the graph it is not much informative in comparing the means.

Table 1 reports the T-test results for each course to check the significance of differences in means by gender. It is clear that the female students have an advantage in law courses without an exception. The difference is statistically significant at 1 percent in Introduction to Law (*introduction*) and 10 percent in Constitutional Law (*constitutional*). In Roman Law (*roman*) the significance level is at the margin (with a *p* value of *0.107*), while the difference is not significant but still in favor of the females in Civil Law (*civil*). The only course that male students have an advantage is Economics (*econ*). The next section provides evidence on the link between the standardized-test based student selection and the female advantage in Law courses.

[11] The foreign origin students were absent because the registration procedures were not complete in the first week of the semester.

[12] I exclude one student from the sample who was an outlier in age and income. The resulting sample size varies between 240 and 244 depending on the responses to the specific questions.

Ekim/October 2013 - Cilt/Vol: 15 - Sayı/Num: 04
Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

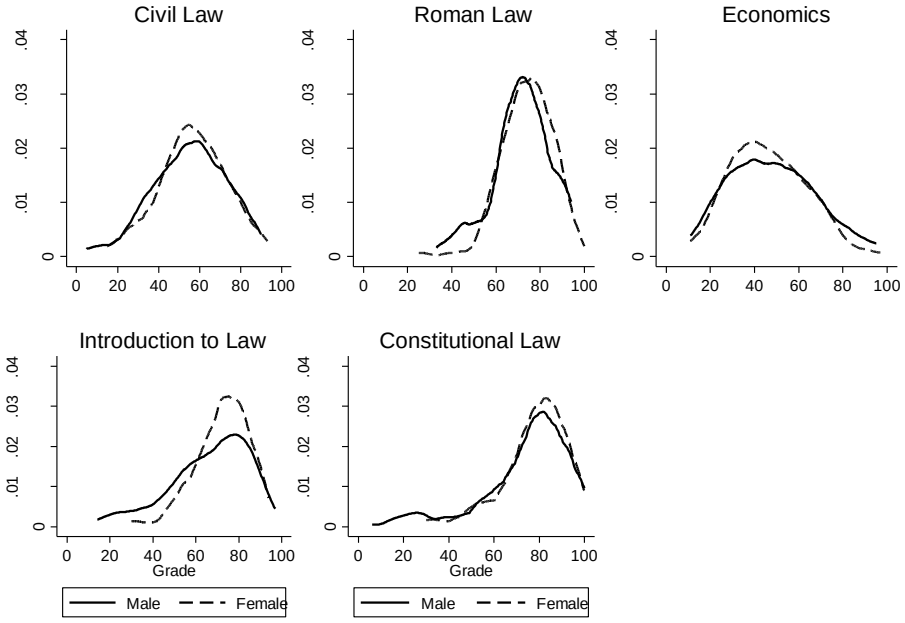

Table 1: Average Grades (T-Test Results)

	Male	Female	Male - Female	st. error
civil	54.96	56.77	-1.81	(2.19)
constitutional	74.62	78.39	-3.77*	(2.18)
introduction	66.30	72.51	-6.21***	(2.00)
roman	71.66	74.23	-2.57	(1.59)
econ	47.99	46.18	1.81	(2.36)
N	244			

Notes: The top grade is 100. * $p < .1$, *** $p < .01$

4.2 Multivariate Estimations

This section investigates the factors affecting the gender gap documented in the previous section establishes its causality with the student selection process. To do so, I estimate student grades using the control variables that I obtained from the survey and their *ÖSYS* scores. The set of controls include sex, age, accommodation type, distance from the faculty (in minutes), and their monthly expenditure plan (in Turkish Liras). The data also included a dummy variable taking the value one if the student is working, but because only three students declared that they had jobs, this variable was later excluded from the analysis. Table 2 and 3 report the descriptive data on the remaining variables.

Ekim/October 2013 - Cilt/Vol: 15 - Sayı/Num: 04
Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

Table 2: Descriptive Statistics (T-Test Results)

	Male	Female	Male – Female	st. error
Age	18.74	18.36	0.38***	(0.11)
OSYS	512.50	513.30	-0.81	(1.28)
Distance	23.67	23.67	-0.00	(2.31)
Expenditure	513.68	551.05	-37.36	(69.51)
N	243			

*** $p < .01$

Table 3: Accommodation Preferences (Column Percentages)

	Male	Female	Total
With family	23.28	32.54	28.10
Dormitory	36.21	53.97	45.45
Lease	40.52	13.49	26.45
Total	100.00	100.00	100.00
N	242		

There are no statistically significant differences between the two sexes in their *ÖSYS* scores (*OSYS*), distances from the faculty (*distance*), and their month expenditure plans (*expenditure*), as Table 2 reports. However, male students are *4.5* months older than the females; and this difference is highly significant. The difference in mean ages might be reflecting the gender bias within the families of the applicants with regard to the university applications. Knowing that the *55* percent of the total applicants are male[13], we can infer that, a greater percentage of males take the exam multiple times while a higher fraction of families allow girls to take the exam only once.

As reported in Table 3, there are also important differences between the male and female students with regard to their accommodation preferences. A much smaller share of females (*13.5* percent) prefer to live in a leased apartment, probably because of gender-specific disadvantages of using that option, such as required heavy-lifting while moving, concerns about the security of the apartment and the neighborhood, or the need to deal with some "male-intensive" local small-scale in the moving process. It is also probable that female students are more likely to apply for a university in their home cities because of the extra non-pecuniary cost of moving for them, or as a result of social or within-household discrimination. These might explain the reasons why, a higher percentage of females live with their families or relatives (the second raw of Table 3). As the differences in accommodation may have an effect on the students' time use, I will take them into account in the estimations of course grades.

Table 4 and 5 report the estimation results for the student grades with control variables. In the first column of the each panel are our main coefficients of interest, namely, female student dummy (*female*) and *ÖSYS* score (*OSYS*). In the second column, student's age (*age*) is added. Lastly, the third column additionally controls for the distance from the faculty (*distance*) and dummies of accommodation choice. Since the *expenditure* variable had inconsistent signs, high standard

[13] < http://www.osym.gov.tr/dosya/1-555969/h/lyssonuc2010sunum.pps> ; reached: 30.04.2012

errors and it has caused lower adjusted R^2 statistics, I excluded this variable from the estimations.

Adding the *ÖSYS* score and student's age in estimations did not create much of a difference from the estimated gender gap in student grades reported in Table 2. The females have a systematic advantage in Law courses, while only in Economics the gap is in favor of the males. In Introduction to Law and Constitutional Law, the gap is statistically significant.

Another important finding of these estimations is that the coefficient of the *ÖSYS* score for Law courses is close to zero and not statistically significant. In other words, there is no correlation between a student's *ÖSYS* score, which is the only admission criterion, and his/ her success in first-year Law courses. This is a result I relate to both to the inadequacy of standardized tests in student selection to a Law school *and* the excessive weight attached to mathematics in those tests[14]. However, only in Economics course student grades show a strong and significant relationship with the ÖSYS scores (Table 5). In the sample, the ÖSYS grades of the students have a range between 495 and 548. This implies an effect on the Economics exam result up to *27* points.

The negative sign of *age* in each course is consistent with the findings with Dayıoğlu ve Türüt-Aşık (2007), who estimated the student grades in Middle East Technical University. It is probable that this result reflects the negative correlation between the number of *ÖSYS* test a student takes before the admission and his/ her success at the university.

The effect of the type of accommodation on grades is smaller than expected. Students who live with their families or relatives (*With family*) and those who live in dormitories (*Dormitory*) have higher grades in average, but the coefficients are not statistically significant. The coefficient of the distance, however, is negative, except for the Economics and significant at Civil Law, Introduction to Law and Constitutional Law estimations. This implies that the time spent on the way between the school and the place of accommodation does have a negative effect on the success. Yet, I do not expect this conclusion have an influence on the gender gap, considering the previous finding that there was not a significant difference between the male and female students in their distances from the Faculty (Table 2).

A joint evaluation of the above findings implies that, *ÖSYS* scores are not good predictors of success for the first year law courses. The only course that ÖSYS predict success is Economics. The systematic gender gap in favor of female students in law courses show that the reliance solely on *ÖSYS* scores in admission for Law departments results in overrating of some male applicants in the selection process.

[14] Kendir and Tuncer (1969) conducted a similar study for the 1965-1966 academic year, when the weight attached to social sciences for entrance to the Law faculties was higher. The authors found a weak, yet a positive correlation between the *ÖSYS* score and student achievement.

Ekim/October 2013 - Cilt/Vol: 15 - Sayı/Num: 04
Sayfa/Page: 102-115, DOI: 10.4026/1303-2860.2013.0240.x

Table 4: Estimation of Grades for the Law Courses

	Civil Law			Introduction to Law		
	(1)	(2)	(3)	(1)	(2)	(3)
female	1.872	0.786	-0.069	6.277***	5.634***	5.473**
	(2.201)	(2.243)	(2.355)	(2.003)	(2.054)	(2.179)
OSYS	0.097	0.106	0.127	0.095	0.101	0.115
	(0.111)	(0.110)	(0.110)	(0.101)	(0.101)	(0.102)
age		-2.848**	-2.989**		-1.685	-1.780
		(1.335)	(1.332)		(1.224)	(1.235)
distance			-0.127*			-0.048
			(0.074)			(0.067)
accommodation==With family			2.354			-0.466
			(3.338)			(3.093)
accommodation==Dormitory			3.667			1.339
			(2.793)			(2.588)
Constant	5.307	54.045	46.758	17.326	46.137	41.429
	(56.799)	(60.832)	(60.712)	(51.734)	(55.720)	(56.196)
Observations	242	242	240	243	243	241
R^2	0.006	0.025	0.051	0.044	0.051	0.060

Table 4 (cont'd)

	Roman Law			Constitutional Law		
	(1)	(2)	(3)	(1)	(2)	(3)
female	2.614	2.289	1.528	3.840*	3.761*	3.071
	(1.591)	(1.635)	(1.697)	(2.187)	(2.251)	(2.347)
OSYS	0.092	0.095	0.102	0.076	0.077	0.097
	(0.080)	(0.080)	(0.079)	(0.110)	(0.110)	(0.110)
age		-0.847	-0.954		-0.208	-0.287
		(0.971)	(0.961)		(1.342)	(1.330)
distance			-0.116**			-0.164**
			(0.052)			(0.072)
accommodation==With family			2.732			3.500
			(2.414)			(3.333)
accommodation==Dormitory			3.250			3.549
			(2.023)			(2.789)
Constant	24.425	38.970	38.164	35.464	39.022	31.848
	(41.068)	(44.344)	(43.792)	(56.481)	(61.070)	(60.553)
Observations	241	241	240	243	243	241
R^2	0.017	0.020	0.059	0.015	0.015	0.052

Standard errors in parentheses
The base accommodation category is 'Lease'
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 5: Estimation of Economics Grades

	(1)	(2)	(3)
female	-2.238	-3.334	-3.595
	(2.292)	(2.333)	(2.457)
OSYS	0.488***	0.495***	0.507***
	(0.115)	(0.115)	(0.115)
age		-2.999**	-2.902**
		(1.404)	(1.408)
distance			0.091
			(0.076)
accommodation==With family			0.501
			(3.513)
accommodation==Dormitory			2.947
			(2.930)
Constant	-201.978***	-149.722**	-160.857**
	(59.136)	(63.593)	(63.759)
Observations	242	242	240
R^2	0.072	0.089	0.102

Standard errors in parentheses
The base accommodation category is 'Lease'
* $p < .1$, ** $p < .05$, *** $p < .01$

5. Conclusion

The capacity of standardized tests to measure student aptitude and achievement is limited. Previous studies on the subject maintained that, the inadequacy of these tests in student selection and scholarship grants worked against the female applicants through disregarding some of their aptitudes which are valuable at higher-education level. Two empirical findings based on the US selection system pioneered that literature: Firstly, high school GPAs of female applicants remained systematically higher than those of males, compared to the difference in the SAT scores. Secondly, the statistical power of the SAT score in predicting the first-year university grades of the females was lower than its power for male grades.

The main contribution of this study is to suggest an empirical strategy to investigate the gender-bias that standardized-based student selections cause. The investigations based on the data from the Law Department of Ankara University data support that standardized test-based student selection cause a gender bias indeed, and this bias is observable through comparison of gender differentials in student achievement among different courses. Female students tend to receive higher grades in courses where standardized tests have no predictive power in first-year student achievement, whereas the grade gap disappears in economics, the only field where *ÖSYS* scores can predict the success.

This study also provides some limited information on the gender differences in some student characteristics and their external conditions, such as age, accommodation and their expenditure plans. We observed major differences between two sexes in their accommodation choices, though the effect on the grades remained limited. A limitation of this study is that, we were not able to collect data on the time use and class attendance of the students to measure contribution to the gender gap in achievement. Extending this study to cover these issues would be a contribution of a future study.

References

Cole, N. S. (1997). *The ETS Gender Study: How Females and Males Perform in Educational Settings*, Educational Testing Service.

Connor, K., & E. J. Vargyas (1992). The Legal Implications of Gender Bias in Standardized Testing, *Berkeley Women's Law Journal*, 7, 13-89.

Dayıoğlu, M. ve S. Türüt-Aşık (2007). Gender differences in academic performance in a large public university in Turkey, *Higher Education*, 53, 255-277.

Gamer, M., & G. Engelhard Jr. (1999). Gender Differences in Performance on Multiple-Choice and Constructed Response Mathematics Items, *Applied Measurement in Education*, 12:1, 29-51.

Guiso, L., F. Monte, P. Sapienza, & L. Zingales (2008). Culture, Gender and Math, *Science*, 320, 1164-1165.

Holmlund, H., & K. Sund (2008). Is the gender gap in school performance affected by the sex of the teacher?, *Labour Economics*, 15, 37-53.

Hyde, J. S., E. Fennema, & S. J. Lamon (1990). Gender Differences in Mathematics Performance: A Meta-Analysis, *Psychological Bulletin*, 107(2), 139-155.

Kendir, S. ve C. Tuncer (1969). A.Ü. Hukuk Fakültesi Öğrencilerinin Üniversite Giriş Sınavlarında Aldıkları Puvanlarla Birinci Sınıftaki Başarılarının Karşılaştırılması (A Comparison of the University Entrance Test Grades and the Achievements of the First Year Student of the Law Faculty of Ankara University), *Ankara Üniversitesi Hukuk Fakültesi Dergisi*, XXVI (1-2), 365-376.

Kuncel, N.R. & S.A. Hezlett (2007). Standardized Tests Predict Graduate Students' Success, *Science*, vol.315, 1080-81.

Moss, K., L. (1988-1990). Standardized Tests as a Tool of Exclusion: Improper Use of the SAT in New York, *Berkeley Women's Law Journal*, 4, 230-244.

Poe, K. (1998). Blinded by Results: Is Looking to GPA in Addition to Standardized Test Scores Truly a Less Discriminatory Solution to Merit Scholarship Selection?, *Women's Rights Law Reporter*, 19(2), 181-191.

Silverstein, A. L. (2000-2001). Standardized Tests: The Continuation of Gender Bias in Higher Education, *Hofstra Law Review.*, 29, 669-700.

Stricker, L.J., D.A. Rock, & N.W. Burton (1993). Sex differences in predictions of college grades from scholastic aptitude test scores, *Journal of Education Psychology*, 85(4), 710-18.

Williams, M. L., C. Waldauer, & V. G. Duggal (1992). Gender Differences in Economic Knowledge: An Extension of the Analysis, *Journal of Economic Education*, Summer, 219-231.

Young, J. W. (1991). Gender Bias in Predicting College Academic Performance: A New Approach Using Item Response Theory, *Journal of Educational Measurement*, 28(1), 37-47.