PRIVATE SCHOOL EFFECTS ON HUMAN CAPITAL PRODUCTION EVIDENCE FROM TÜRKİYE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BİLGEHAN ÖZBAYLANLI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF ECONOMICS

APRIL 2023

Approval of the thesis:

PRIVATE SCHOOL EFFECTS ON HUMAN CAPITAL PRODUCTION EVIDENCE FROM TÜRKİYE

submitted by **BİLGEHAN ÖZBAYLANLI** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Economics, the Graduate School of Social Sciences of Middle East Technical University** by,

Prof. Dr. Sadettin KİRAZCI
Dean
Graduate School of Social Sciences _____

Prof. Dr. Şirin SARAÇOĞLU
Head of Department
Department of Economics _____

Assoc. Prof. Dr. Hakan ERCAN
Supervisor
Department of Economics _____

Examining Committee Members:

Prof. Dr. Meltem DAYIOĞLU TAYFUR (Head of the Examining Committee)
Middle East Technical University
Department of Economics _____

Assoc. Prof. Dr. Hakan ERCAN (Supervisor)
Middle East Technical University
Department of Economics _____

Prof. Dr. Erol TAYMAZ
Middle East Technical University
Department of Economics _____

Prof. Dr. Semih TÜMEN
TED University
Department of Economics _____

Assoc. Prof. Dr. Ali Murat BERKER
Hacettepe University
Department of Economics _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Bilgehan ÖZBAYLANLI

Signature:

ABSTRACT

PRIVATE SCHOOL EFFECTS ON HUMAN CAPITAL PRODUCTION
EVIDENCE FROM TÜRKİYE

ÖZBAYLANLI, Bilgehan
Ph.D., The Department of Economics
Supervisor: Assoc. Prof. Dr. Hakan ERCAN

April 2023, 219 pages

The human capital theory establishes the vital impacts of the human capital produced by education on productive skills valued in the labor market and long-term economic growth. This study focuses on private school effects on human capital production. The studied period 2013-2017 covers an exam (TEOG) taken by all private and public school students, a national-scale voucher program, and a rapid expansion of private sector enrolment. The study utilizes population-level data covering 4.8 million 15 years olds. The education production function approach frames the conceptual model for analysis. The human capital measure is the standardized TEOG score. I first explored whether private schools produce more human capital than public schools *ceteris paribus*. I find a combined effect of 0.62 standard deviations (sd) due to both school inputs and peer inputs of private schools and a lower bound of 0.25 sd for school inputs alone. Second, I estimate the net competition effect of the private school share on the mean public school achievement in each province. I find an OLS estimate of -0.045 sd and an IV estimate of -0.036 sd. Third, I explored whether the gender achievement gap differs by school sector through a joint test scores and grades model.

I estimate a gap of 0.30 sd for TEOG scores and 0.44 sd for GPAs in public schools. These reduce to 0.13 and 0.12, respectively, in private schools. When differenced, the remaining gender grading gap is 0.14 sd in public schools and 0.01 sd in private schools.

Keywords: Human Capital, Private School Effect, Cognitive and Non-Cognitive Skills, Education Production Function, Gender Achievement Gap

ÖZ

ÖZEL OKULLARIN İNSAN SERMAYESİ ÜRETİMİNE ETKİSİ
TÜRKİYE ÖRNEĞİ

ÖZBAYLANLI, Bilgehan
Doktora, İktisat Bölümü
Danışman: Doç. Dr. Hakan ERCAN

Nisan 2023, 219 sayfa

İnsan sermayesi teorisi, eğitim sayesinde üretilen insan sermayesi ile işgücü piyasasında değer verilen üretken beceriler ve uzun vadeli ekonomik büyüme arasındaki güçlü bağlantıyı ortaya koymaktadır. Bu çalışma, özel okulların insan sermayesi üretimi üzerindeki etkilerine odaklanmaktadır. Çalışmada kapsanan 2013-2017 dönemi, özel okullardakiler dahil tüm öğrencilerin ortak bir sınava (TEOG) girdiği, özel okullara yönelik ulusal ölçekte bir teşvik programının yürütüldüğü ve özel okul öğrenci sayısının çok hızla arttığı bir döneme karşılık gelmektedir. Çalışma verisi 15 yaş grubu 4,8 milyon öğrenciden oluşmakta ve tüm ortaokul öğrenci popülasyonunu kapsamaktadır. Eğitim üretim fonksiyonu yaklaşımı, analizler için kavramsal modeli sağlamaktadır. İnsan sermayesi ölçüsü standartlaştırılmış TEOG puanıdır. İlk olarak, diğer tüm unsurlar sabitken özel okulların devlet okullarına kıyasla daha fazla insan sermayesi üretip üretmediğini araştırılmıştır. Özel okulların sağladığı okul girdileri ve akran grubu girdileri bir arada değerlendirildiğinde, özel okulların devlet okullarına göre 0,62 standard sapma (ss) daha fazla insan sermayesi ürettiği tespit edilmiştir. Sadece daha iyi okul girdilerine dayalı fark içinse 0,25 sd'lik

alt sınır saptanmıştır. İkinci olarak, iller bazında özel okul öğrenci oranının, ortalama devlet okulu kazanımı üzerindeki net rekabet etkisi tahmin edilmiştir. Sıradan en küçük kareler yöntemi ile -0.045 ss ve araç değişken yöntemi ile -0.036 ss'lik fark bulunmuştur. Üçüncü olarak, cinsiyet kazanım farkının okul sektörüne göre farklılık gösterip göstermediği, test skorları ve ağırlıklı not ortalamalarını birlikte ele alan ortak bir modelle araştırılmıştır. Devlet okullarında TEOG skorları açısından 0,30 ss ve ağırlıklı not ortalamaları açısında 0.44 ss'lik bir fark tespit edilmştir. Bu fark özel okullar için sırasıyla 0,13 ss ve 0,12 ss'e düşmektedir. Her iki eşitlik birbirinden çıkarıldığında, kalan cinsiyet notlandırma farkının devlet okullarında 0.14 ss, özel okullarda 0.01 ss olduğu saptanmıştır.

Anahtar Kelimeler: İnsan Sermayesi, Özel Okul Etkisi, Bilişsel ve Bilişsel Olmayan Beceriler, Eğitim Üretim Fonksiyonu, Cinsiyet Kazanım Farkı

DEDICATION

To My Beloved Wife Nurgül and Lovely Daughter Beyza,

ACKNOWLEDGMENTS

I want to express my most sincere thanks and gratitude to my supervisor Assoc. Prof. Dr. Hakan Ercan for his guidance, support, and patience throughout the long process of conducting this study.

I would like to thank my examining committee members Prof. Dr. Meltem Dayıoğlu Tayfur, Prof. Dr. Semih Tümen, Prof. Dr. Erol Taymaz, and Assoc. Prof. Dr. Ali Murat Berker heartily for their guidance and inspiring contributions, which have significantly shaped the final form of this dissertation.

I am also grateful to my family for their continuous support and faith in me. I owe a lot to my wife Rabia Nurgül and my daughter Ayşe Beyza for their patience and constant support. My parents İsmail and Ayşe Özbaylanlı, and my parents-in-law, Hanifi and Nuran Karaveli, were always in check with my progress. Their enthusiasm was always inspiring.

I want to thank my previous co-authors, Prof. Dr. Mustafa Hilmi Çolakoğlu and Doç. Dr. Mehmet Cansız, for their encouragement as well as their assistance in accessing resources for this research.

I am also grateful to Funda İnce for her continuous assistance and to Şamil Demircan and Onur Karabacakoğlu for their technological support. I am thankful for Nihal Ercan and Dr. Cumhur Çiçekçi, who have always provided the motivational impetus for me throughout the challenging writing process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CC	Complete Cases
FIML	Full Information Maximum Likelihood
GPA	Grand Point Average
LGS	Liselere Geçiş Sistemi
MAR	Missing at Random
MCAR	Missing at Completely Random
MNAR	Missing at Not Random
MEB	Milli Eğitim Bakanlığı
MoNE	Ministry of National Education
NAEP	The National Assessment of Educational Progress
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares
PIAAC	Programme of the International Assessment of Adult Competencies
PISA	The Programme for International Student Assessment
SEM	Structural Equation Modeling
SUREG	Seemingly Unrelated Regression
TEOG	Transition from Basic Education to Secondary Education
TIMMS	The Trends in International Mathematics and Science Study
TL	Turkish Liras

CHAPTER 1

INTRODUCTION

For a long time, education was primarily valued for its social role. Pioneers of human capital theory, Schultz (1960) and Becker (1964), pointed out the crucial links between the education of students and the productive skills valued in their future professional careers. They proposed that the role of education should be envisaged as an income-generating form of human capital investment (Hansen, 1970).

Education since then has been called upon to increase earnings, accelerate economic growth and equalize income distribution. Two important lines of inquiry have evolved from human capital theory to form together the research area known as the economics of education. The first pillar provided a practical analytical framework to understand the relationship between education and income at the individual level (Mincer, 1974) and the aggregate level (Lucas, 1988; Mankiw, Romer, & Weil, 1992). The second pillar focused on how education is produced (Bowles, 1970; Hanushek, 1979). The two pillars are complementary in that human capital produced as an output of education (in schools and households) later enters as an input into the production of wages and the gross domestic product per capita. Focusing on private school effects on human capital production, this study is located in the second pillar of the economics of education literature.

Human capital corresponds to the stock of cognitive and non-cognitive skills, innate ability, personality traits, and knowledge an individual possesses. At the micro level, the distribution of human capital determines the distribution of earnings. At the macro level, a country's aggregate stock of skills is crucial for economic growth and socioeconomic development (Hanushek & Woessmann, 2015). Note that human capital production is not only restricted to schools, as it is also produced in households.

However, human capital produced in schools inherently has more policy relevance and is better positioned to generate social returns.

The economic perspective and its analysis techniques best inform household and government decisions on how much and how to invest in human capital. One such significant decision is the choice of the school sector. Private schools are generally presumed to provide a better education quality compared to public schools. They are under more vital competitive forces to do so. Moreover, private schools select students, which allows them to form potentially a better peer composition. In return, they demand tuition.

Households weigh benefits against costs in their school sector decisions. While they can directly deduce the costs from the observed tuition rates on the market, there is no such clear indicator regarding benefits. When provided with the estimates of expected benefits from learning achievements and other outcome measures, families can better assess if the academic and other benefits of a private school decision are worth it.

Governments, on the other hand, invest in voucher programs to expand school choice. The intended direct benefit is higher human capital levels for students participating in the program, known as the *participation effect*. The anticipated indirect benefit is the increased quality of education in public schools due to the increased numbers of and higher enrolment in private schools, which is known as the *competition effect*. However, unintended negative consequences may arise due to the sorting of better students and teachers from public schools into private schools, known as the *cream-skimming effect*. The estimates of the net impacts of school choice programs of governments shed light on whether the participation and positive competition effects overcome sorting and adverse competition effects. Estimates of the impact of voucher programs coupled with the information about their design also have international value as they will help governments to design better interventions (Egalite & Mills, 2021).

This study explores the following three research questions. The first two questions are the core questions of economic research on private school effects. Urquiloa (2016)

concludes in his survey prepared for the fifth handbook of the economics of education that answers to both questions are still not settled. Moreover, while the interventions in this policy area depend heavily on context, evidence remains primarily restricted to a few countries like the United States, has attracted recent attention, and more evidence of gender gaps in cognitive and non-cognitive achievement is called for.

Research Question 1: Are private schools more effective than public schools in producing human capital?

Research Question 2: Does competition from private schools increase public school effectiveness in producing human capital?

Research Question 3: Does the gender achievement gap differ by school sector?

For all three questions, human capital will be measured by achievement on standardized test scores from the TEOG exam held between the 2013-14 and 2016-2017 academic years in Türkiye. For the third question, GPA obtained in the last three years will be used as an additional measure of achievement.

Burgess (2016) and Hanushek (2020) both underline that the most remarkable innovation recently is the utilization of national, large-scale administrative databases. A critical contribution of this study is introducing student population level data covering over 4.8 million students and 18 thousand individual schools spread over four successive cohorts.

The period covered in the study is also unique and ideally suited for studying private school effects on human capital production. First, the TEOG exam was mandatory both for private and public school students. As it replaced one of the course exams in both semesters of the last year of lower-secondary school (grade 8), the whole student population participated in the same nationwide standardized test examination. Attendance of all students, especially private school students, in the nationally-held standardized test was not mandatory for the previous and successive periods, which reflected a substantial restriction against studying private school effects.

Second, there was a large-scale, nationwide voucher program, which aided expanding the private sector share. In just a few years, the number of students graduating from private lower-secondary schools increased by 63.2 percent, and the proportion of private school students in the total student population increased by 78.9 percent. Similar increments happen in terms of the number of schools, too. The second research question requires large-scale expansion in voucher programs and cross-market variation in the utilization of those vouchers within the country. The design of the voucher program had aspects highly suited to this requirement.

The next chapter will cover a literature review and will include four subsections. The first subsection focuses on quantitative measures of human capital. Here I introduce the human capital theory and briefly summarize its profound influence on microeconomic and macroeconomic theory-building and empirical research. I will also discuss how it is criticized. Then I report on the empirical findings that relate quantitative measures of human capital (attainment: years of schooling or highest degree achieved) to individual earnings and gross domestic product per capita. The second subsection describes the qualitative measures of human capital, where the primary measure is the standardized test scores as employed in this study. Once again, I will report on the empirical findings that relate qualitative measures of human capital (achievement: scores from standardized tests) to individual earnings and gross domestic product per capita. The third subsection introduces the education production function framework, which is the workhorse model of the economics of education. It guides how to specify a theoretical and an empirical model to analyze human capital production. It guides which input factors shall enter the production function and which proxy variables shall be included in the analysis if we cannot observe these direct inputs. The last subsection will summarize the different effects of private schooling on human capital production and the econometric challenges associated with estimating these effects. This subsection will also elaborate on how students and schools select for the market. I will also report on the empirical evidence of these effects, which will be used as a benchmark when interpreting my estimates.

I will present a chapter focusing on the institutional framework next. The institutional and political backdrop was dynamic right before and within the few years covered in this study. Understanding the institutional setup is essential to interpret the study's findings.

Next, I will roll out a separate data analysis chapter before the methodology chapter. Since this is primarily an applied study and as this data is first introduced in this scale, I presented data in detail and aimed to link this presentation with the literature survey and institutional setup sections.

The methodology chapter first summarizes the overall aspects of the methodological strategy. It is expressed that this is a quantitative study based on observational data, which is founded on the education production function framework. The econometric fixes to solve the problems related to data limitations are summarized. This chapter puts a particular emphasis on addressing missing data in student variables. I discuss in detail the alternatives, specify my strategy, and implement it. For the first question, I define two types of private school effects. First is the Type-1 private school effect, which is the combined effect of school inputs and peer inputs. I use a Two-Step estimator (Castellano, Rabe-Hesketh, & Skrondal, 2014). The Type-2 private school effect aims to isolate the effect only due to school inputs via a correlated random effects model (Wooldridge, 2010; Snijders & Bosker, 2012). For the second question, I estimate the effect of private school student share of the province on the average public school achievement in that province to learn about the net competition effects of the voucher program. I utilize the (conditionally exogenous) quota allocation rule at the province level of the voucher program to supplement the model specified based on the education production function, which also includes additional controls for the factors that determine the expansion of private school share. I jointly model TEOG test scores and GPAs in a within-school-cohorts fixed effects setup for the final research question. I simultaneously use structural equation modeling (SEM), and seemingly unrelated regression (SUREG) approaches to alternatively handle missing data while enabling correlated errors between the two outcome measures.

In terms of findings, using a standardized test score measure (to zero mean and one standard deviation), I find a positive effect of around 0.62 standard deviations for the Type-1 private school effect, which corresponds to the combined effect of school inputs and peer inputs. This type of effect is generally not estimated in the literature, although it is directly relevant to the school-sector decisions of parents. I find a lower bound of 0.25 standard deviations for the Type-2 effect, which corresponds only to the private school effect due to school inputs. The typical estimate for this type of private school effect is 0.20 standard deviations. Type-2 effects are relevant for the assessment of school accountability and evaluation of school performance.

I find a competition effect of -0.045 standard deviations on the average TEOG score of a typical school due to one standard deviation increase in the share of private school students in the same province. The instrumental variable estimate of the same effect is close to this value at -0.036. The effect size aligns with the modest impact sizes seen in the literature. A large part of the literature finds modest positive competition effects. However, some of the most influential contributions, like Dee (1998) found -0.023 sd, Husted & Kenny (2000) document -0.060 sd, and Geller, Sjoquist & Walker (2001) detect between -0.018 to -0.137 sd competition effects. Both the effect size and the sign of the competition effect depend heavily on the country's context at the time of voucher program implementation and on the specific design of the voucher program. Hence, the effect estimates need to be assessed in conjunction with these circumstances described in detail in the chapter for institutional background.

Regarding gender achievement gaps, when the human capital measure is the TEOG achievement score, I find a 0.30 standard deviation of the gender achievement gap in public schools. In comparison, this reduces to 0.12 standard deviations in private schools. When the measure is GPA, the gender gap is 0.44 standard deviations in public schools, decreasing to 0.13 standard deviations in private schools. The errors of these two measures have a correlation coefficient of 0.91. No school factor has been shown to make a significant difference in the gender achievement gap in the literature before (Burgess et al., 2004; Bharadwaj et al., 2016). The findings here provide

evidence for such a school factor in the Turkish context: the school sector. On the other hand, when I use the take-the-difference strategy, the unobserved effects of innate ability and the historical human capital inputs cancel. In the remaining variation, which is argued to reflect non-cognitive skills based on previous literature, the gap is 0.14 for public schools and 0.01 for private schools. These differences in gender achievement gaps are called gender grading gaps (Terrier, 2020) and are regarded to be based on gender-average non-cognitive skill differences among girls and boys. These types of gaps are mainly discussed in terms of equity concerns.

I finally conclude in the last chapter.

CHAPTER 2

LITERATURE SURVEY

The literature survey chapter will present theoretical constructs and empirical evidence related to the measures of human capital, how human capital is produced, what private school effects are, and how we learn about these effects. There are four subsections. In the first subsection, I describe quantitative measures of human capital. This part also summarizes human capital theory's emergence, development, and strong influence on microeconomic and macroeconomic research. The second part focuses on the qualitative measures of human capital, which generally refer to cognitive skills measured mainly by standardized test scores. I will also touch upon the non-cognitive skills, which are argued to be very influential for future education and labor market outcomes. Similar to before, I will summarize how qualitative human capital measures influence microeconomic and macroeconomic analysis. The third subsection describes the education production function, the workhorse model used to analyze production relations between inputs and outputs of education. The final subsection will summarize the private school effects and the consequences of policies that promote private school choice. All four subsections are closely interrelated and, as a whole, aim to provide the understanding and the means for the analysis to be carried out in later chapters.

2.1 Educational Quantity as the Measure of Human Capital

The quantitative measures of human capital include the years spent in education within schools or the highest degree or diploma achieved. These are known as *attainment* measures. From the micro-perspective, this corresponds to the attainment of individuals. On the other hand, the macro-perspective focuses on the aggregate amounts of the stock of human capital, which are measured by the country-average

years of schooling or share of citizens with a particular degree/diploma. In the following two subsections, I will summarize the major theoretical underpinnings and empirical findings from both perspectives.

2.1.1 Micro-perspective

Historically, education was a peripheral issue in analyzing economic phenomena or the labor markets. The complementary works of Shultz (1960), Becker (1964) and Mincer (1974) were consequential in transforming this perspective.

In his seminal work on capital formation by education, Schultz (1960, p.571) argues that education should be seen as an investment in individuals. Skills and knowledge, which are the products of educational investments, should then be regarded as "*a form of capital*". As this type of capital becomes an integral part of its investors, it is called human capital. Schultz (1961, p.1) further argues that human capital investment is the principal trigger of the increases in earnings per worker.

Becker (1964) puts the fundamental ideas of human capital into a framework. He suggests that education and on-the-job training, the leading producers of human capital, enhance an individual's cognitive skills to meet future job requirements and increase worker productivity. Increased productivity then leads to increased earnings.

The workhorse empirical model for analyzing investments in human capital at the individual level is the Mincer equation (Mincer, 1958, 1975):

$$\text{Ln}Wage_i = \beta_0 + \beta_1\, Schooling_i + \beta_2\, Experience_i + \beta_3\ \ Experience_i^2 + e_i$$

The outcome variable $\text{Ln}Wage_i$ stands for the natural logarithm of earnings of individual *i*. The coefficient β_0 stands for the logarithm of earnings of an individual with no education and no experience. Schooling indicates the number of years of schooling the individual *i* received, referred to as "educational attainment". It is the measure of human capital produced from formal schooling. The coefficient $\beta_1$is the rate of return to each year of schooling, and it is assumed to be the constant along all

schooling levels. Experience, on the other hand, is years of work experience. Note that adding the squared term of experience reflects that wage is assumed to increase with more experience, albeit with a decreasing rate. Experience also reflects human capital formation due to on-the-job training and learning by doing. The Mincer model, hence, is a model that principally relates wage outcomes to human capital inputs. Additional variables like gender, race, etc., can be added to the model. The estimate of β_1, the percentage increment in wages associated with each additional year of schooling, lies in the range of 5 to 15 in a large pool of studies (for a list of references, see Hanushek & Woessman, 2008, p.615; Chattopadhyay, 2012). This estimate, on average, is found to be higher for girls, primary school students, and less developed countries (Psarcharopoulos & Patrinos, 2018).

Human capital theory suggests that the coefficient β_1 reflects the improvement in the productive capacity of the workers due to education. However, Krueger and Lindahl (2001, p. 1104) argue that it can also reflect, at least partly, the unobserved innate ability or personality traits of the worker, which is correlated both with the education level and the wage level. It could be the case that those who are better educated might be earning more because they have higher levels of education; however, there is another possibility that they would have done better in the labor market anyway due to their higher inherent ability and motivation. In other words, schooling might be endogenous to innate ability or personality traits. The human capital theory further implies that education, being an investment good, should be received until the gain in the marginal productivity of schooling equals the marginal opportunity cost of the additional schooling. Card (1999) theorizes that the ideal schooling level varies between individuals for two main reasons: (i) ability varies among individuals, and more-able individuals gain more from additional schooling (as they learn at lower non-pecuniary costs, i.e., it is easier for them to study and learn; and thus they more efficiently convert schooling into human capital), (ii) marginal rates of substitution between current and future earnings may differ across individuals, which leads variation in discount rates of foregone earnings during education among individuals. An implication for this is that the bias in the ordinary least squares estimates of the

returns to an extra year of schooling becomes unclear (Harmon, Oosterbekk & Walker, 2003, p.119).

As a rival theory, the *sorting hypothesis* suggests that education is essential not because it augments human capital but signals intrinsic productivity (Spence, 1973; Stiglitz, 1975). Education aids in screening and sorting workers with higher ability/motivation from workers with lower ability/motivation. Employers face information asymmetry about the intrinsic productivity of potential employees. It could be the case that educators might better assess inherent productivity. Hence, educational outcomes serve as credentials that sort workers according to their unobserved attributes.

Indeed, regardless of whether education only sorts out or, in fact, augments productivity, education represents a good investment for workers as it increases individual or sector-level earnings (Lang, 1994; Psacharopoulos, 1994). However, if education augments productivity, then longer years of schooling might be desirable; otherwise, if education only sorts productivity, then the shortest years of education sufficient to sort the workers are desirable. On the other hand, both mechanisms are expected to cooperate in most real-world cases. Spence (2002), the pioneer of signaling theory, allows for the concurrent existence of the human capital augmenting and the signaling roles. Page (2010: 36) also argues that much is still unknown about the relative weight put by firms on each component.

The consensus suggests that education is not merely a proxy for unobserved ability. Card (1999), Oosterbeek (1992), and Woessmann (2016), among several other studies, provide evidence that even controlling for ability, it will be profitable to spend longer schooling time. These results contradict the screening hypothesis, which implies that the shorter study duration will enhance earnings more. Harmon, Oosterbekk & Walker (2003) reports that the returns to the signaling component indicate minor effects. From the Marxist perspective, Bowles and Gintis (1975) also agree that schooling augments the levels of human capital. However, they criticize human capital theory by arguing that the education system leads to the segmentation

of the workforce, acts as a barrier against the development of working-class consciousness, and forms the basis for the fragmentation of economic classes.

Moreover, human capital investments influence the distribution of earnings in society. Becker (1993) argues that although endowments for ability are likely to be distributed evenly in the population, earnings tend to be unevenly distributed because individuals with higher ability receive higher returns on their human capital investments. Accordingly, they make these investments more intensively. If individuals with higher ability are also provided with more opportunities, human capital investments would aggravate the inequality in earnings in society (Weiss, 2015).

As another concern, it might be the case that the increased supply of educated might exceed the skills demanded in the job market. In this case, the return to schooling for the excess stock of skills would be lower than expected (Harmon, Oosterbekk & Walker (2003) and lead to a waste of human capital due to over-qualified employment in jobs that in fact fewer skills (Leuven and Osterbeek, 2011).

The main takeaways for this dissertation from a microeconomic perspective are as follows: Human capital theory suggests that what students learn in schools and homes has a productive value that can be utilized to earn better wages in the future. Indeed, Schultz (1960, p.573) views students as self-employed producers of capital who work when they study and create human capital.

Moreover, during the analysis to be carried out in the following chapters, we will use as control the mother's educational attainment and father's educational attainment. Here we understand the relevance and importance of utilizing these variables in an education production function framework.

Economic returns to attainment measures also underline the importance of preventing drop-outs, which is mainly a concern for girls. Finally, we have seen that human capital theory has received several criticisms but remains the most influential theoretical and empirical construct.

2.1.2 Macro-perspective

As summarized above, labor economists provide ample evidence for substantive pecuniary returns to individual investments in education. Further, the benefits of human capital investments by individuals might provide positive externalities to other individuals within the same economy. If there are such social returns on top of private returns, then this would justify public investments in education. These social returns may be in the form of learning externalities that provide direct economic benefits (Sianesi & Van Reenen, 2003, p.160) or societal impacts (like less crime, a better environment, higher social cohesion, better parenting, etc.) that have indirect economic benefits (OECD, 1998). A macroeconomic perspective is needed to assess such social returns.

Schultz (1961) argued that the part of the output that cannot be explained by the traditional inputs (as documented by Denison, 1962) observed in the macro-level empirical analysis was also related to the increases in the stock of human capital. Despite his early emphasis on the macro perspective, the first ten years of human capital theory development mainly focused on microeconomic investigations. On the other hand, the next two decades until the 90s were quite prolific regarding the theoretical and empirical macroeconomic studies under the influence of human capital theory.

In this part, I will summarize the perspectives of three main macroeconomic models in describing the role of human capital in economic growth. The common aspect of the three different perspectives is the measure of human capital. All models utilize quantitative measures such as average years of schooling or degree attained (i.e., secondary school diploma, etc.) to represent the human capital formed by education. I will use a unified notation so that the innovation of each model can be discerned.

The reference model in the macroeconomic analysis is the Solow model (Solow, 1956), where there are two inputs, labor, L, and capital, K, to produce one aggregate output, Y. A represents total factor productivity and is the critical determinant of long-

run growth. Technological progress in *A* is assumed to be determined exogenously, i.e., not specified by the agents within the model:

$$Y = K^{\alpha}(AL)^{1-\alpha}$$

The first model that explicitly utilizes human capital (Mankiw, Romer & Weil, 1992) extends the above model by considering two types of labor: educated and uneducated. Education here has become a factor of production. The human capital component of the growth is driven by the accumulation of more education (Hanushek and Woessman, 2021):

$$Y = K^{\alpha}H^{\alpha}(AL)^{1-\alpha-\beta}$$

where *H* is the stock of human capital, measured by the proportion of the labor force who have completed secondary education. The returns to the human and the physical capital are both assumed to be lower than one. Thus, even if both types of capital inputs increase without limit, the aggregate output growth is determined ultimately by the growth rate of the exogenous term, A. Changes in education move the economy from one steady-state level to another but once reached that level, education does not lead to further growth. In this framework, changes in GDP per capita are regressed on changes in the level of education (Hanushek, 2015).

Another model that emphasizes human capital is the endogenous growth model by Lucas (1988):

$$Y = AK^{\alpha}(h\ell L)^{1-\alpha}$$

Where *h* is the stock of human capital. ℓ is the fraction of the population working in conventional jobs, while $1-\ell$ is the fraction of the population who produces knowledge through research and development (R&D). In effect, there are two sectors in Lucas's model where one sector uses human capital to produce output, while the other sector, human capital, is used to create new human capital. In contrast to Mankiw's model, this model implies that output can grow without limit, provided that

the stock of human capital increases. A similar endogenous growth model developed by Romer (1990) incorporates human capital as the input of the R&D sector. The core function of human capital is boosting the economy's innovative capacity. These models foresee that the economies in which the ratio of human capital to physical capital is higher would enjoy faster growth (Barro & Salai-Martin, 2004).

A third explanation suggested by Benhabib and Spiegel (1994) incorporates human capital as an input to the growth rate of total factor productivity, which depends on the level of a nation's stock of human capital. Human capital influences the catch-up rate of a country to the level of total factor productivity of the technology leaders.

$$Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha} \quad \text{where} \quad A_t = A_0 e^{g(H)t}$$

The above model focuses on the diffusion of technologies (Nelson and Phelps, 1966). Like new technologies increase firm productivity, countries that adopt new technologies at scale are bound to grow rapidly. Education facilitates the diffusion of knowledge to exploit force new technologies invented by others (Benhabib and Spiegel, 2005). The human-capital measure is the average adult population with at least a secondary school diploma. Economies that depend on high-tech have a more significant demand for skilled workers (Hanushek, 2009, p.41).

In terms of the measure of human capital, studies focusing on the impact of stock (i.e., level) of human capital generally use average educational attainment (i.e., years of schooling) in the population over 25 or in the labor force (for instance Barro, 2001; Bassanini & Scarpetta, 2001). Studies on the flow (i.e., investments into) of human capital typically rely on enrollment rates (for instance, Barro, 1991; Englander & Gurney, 1994). As an essential input into these studies, Barro and Lee (1993) identified and compiled the internationally comparable attainment data, refining it every ten years (see Barro and Lee, 2001, 2013).

Macro impacts of human capital are either assessed via growth regressions or internal social rates of return. Among thirteen studies surveyed in Sianesi and Van Reenen

(2003, pp.177-180), a one-year increase in average educational attainment is found to increase the level of output per capita by 3 to 6 percent under augmented neo-classical specifications. On the other hand, one percentage point faster growth is estimated by models based on the new-growth theories.

Arguing that the macro evidence on the augmented neo-classical and endogenous growth models remained inconclusive and did not allow to precisely distinguish between models, Sala-i-Martin, Doppelhofer, and Miller (2003) applied Bayesian model averaging technique to incorporate the fact that there is uncertainty in which model is the true model. Out of 67 different explanatory factors, they found the primary school enrollment rate to be the second most crucial factor for growth after the East-Asia dummy (see Barro & Sala-i Martin, 2004, p. 544).

A set of studies have explored whether the effects of human capital on growth differ by gender. Knowles, Lorgelly, and Owen (2002) derive a neoclassical growth model where education enters as a gender gap. They find that female education has a statistically significant positive effect on labor productivity, which is less clear for boys. They also conclude that the gender attainment gap impedes economic growth. Klasen (2002), with an analysis based on the Augmented Solow model, finds that girls' education leads to higher growth effects than boys'. Moreover, the gender attainment gap accounts for between 0.4 to 0.9 percentage points of differential in the growth rate between East Asia and less developed regions.

2.2 Educational Quality as the Measure of Human Capital

The previous subsection measured human capital in the form of the years of schooling a student received, known as the attainment measure. Human capital theory suggests that what we learn at school provides productive skills valued highly in the labor market. Hence, more attainment is expected to create higher economic value for the individual and society. In its extreme form, signaling/screening theory argues that schools do not help students acquire productive skills relevant to the labor market. Hence, the least amount of years spent in schooling that is just enough to signal high

innate ability is preferable to any longer terms. Empirical evidence summarized in the previous subsection indicates that a combination of both perspectives is in action in actual circumstances. Thus, a substantial amount of additional human capital over innate ability can potentially be accumulated in each pupil during a period spent in school.

Whether the mentioned potential for learning turns into actually learned skills in school is crucial. Attainment only accounts for the years utilized in schools without considering what has happened during that time. It presumes each schooling year produces an equivalent human capital over every country's schooling period. Brown and Saks (1987, p.319) and Hanushek (2012) suggest that the emphasis should be on the quality over the quantity of schooling for measuring human capital production through education. As an early contribution confirming this view, Taubman and Wales (1973) detected considerable earning variation based on the quality of the college while controlling for the level of education.

The achievement is measured by standardized test scores (Hanushek, 2017)[1]. There are two major international testing programs: The Programme for International Student Assessment (PISA) evaluates 15-year-olds in their graduation year from lower-secondary school (eight-graders) numeracy, science, and literacy performance every three starting from 2000. The Trends in International Mathematics and Science Study (TIMSS) assesses the level of science and numeracy skills of grade eight students every four years since 1995.

In addition to international standardized tests, national-level standardized tests assess skills within individual countries. The National Assessment of Educational Progress (NAEP) of the United States evaluates numeracy, reading, science, writing, civics, and arts. The transition from Basic Education to Secondary Education (TEOG)

[1] As mentioned in the previous subsection, Becker (1964) states that education and on-the-job training enhance an individual's cognitive skills to meet (current or future) job requirements, which in turn increases worker's productivity and earnings.

examination of Türkiye (which provides data for this study) is a national-level assessment of mathematics, science, literacy, English proficiency, history and religious knowledge. It is similar to NAEP regarding national-level coverage and comparable to the TIMMS in evaluating basic skills combined with content knowledge.

Programme of the International Assessment of Adult Competencies (PIAAC) focuses on the skills of *adults* aged 15-65 and evaluates how those skills help them process information in complex work environments. Three cycles of PIAAC studies have been carried out; first in 2011-12, second in 2014-15 (Türkiye participated in this cycle), and third in 2017. It assesses three broad skills: numeracy, literacy and problem-solving (in technology-intensive workplaces for the first cycle and in an adaptive way in the later two cycles) (OECD, 2021).

When assessing PISA or TIMMS outcomes, we assume educational qualifications as a proxy for individuals' future productive potential. In contrast, the PIAAC directly measures skills. Hence, it can relate an individual's current level of cognitive skills to the current probability of being employed and to the actual wages (Quintini, 2014). PIAAC data readily shows how skills utilized in the workplace are associated with labor productivity or wage gaps: While only half of the adults having up to Level 1 proficiency in literacy is employed, eighty percent of adults with more than Level 4 proficiency are employed. Numeracy and literacy score averages of employed adults are also higher than those of nonemployed adults. Differences in mean reading skills account for thirty percent of the discrepancies in labor productivity across countries (OECD, 2013). One standard deviation improvement in numeracy skills is associated with an 18 percent rise in wages. Compared to problem-solving, returns to literacy and numeracy (PIAAC skills overlapping with PISA or TEOG) are distinctly more pronounced (Hanushek et al., 2015).

Jencks and Philips (1998) and Mayer and Paterson (1999, p.4) argue that the score on a cognitive test depends on the combination of the innate ability of the pupil to grasp productive skills and the extent of opportunities provided by the human capital

investment of the parents (or the government) into the pupil. High test scores are relevant for indicating that the pupil has learned productive skills in school applicable in the future workplace or that the pupil has a high innate ability for learning new skills during on-the-job training in the future workplace[2]. Hence, compared to the conflicting views from the human capital theory and the signaling theory on the desired level of attainment, both perspectives favor strictly higher standardized test scores at any specific attainment level.

2.2.1 Micro-perspective

An established finding in the microeconomics literature is that cognitive test scores strongly predict employment status and wages in the future. The critical question is whether the association between the scores of students on standardized tests and their future performance in the labor market is causal (Hanushek, 2009, p.42).

The Mincer equation (Mincer,1975), which is the empirical workhorse model, is extended to incorporate the effect of the actual stock of individual skills by adding the score from the standardized achievement test:

$$\text{Ln}Wage_i = \beta_0 + \delta\text{Score} + \beta_1 Schooling_i + \beta_2 Experience_i + \beta_3 Experience_i^2 + e_i$$

Behrman (2010) argues that the above formulation has been adopted slowly in the literature because of the lack of datasets that include achievement, attainment and wages at the same time. Furthermore, the coefficient of the achievement measure is generally interpreted as a measure of ability and is incorporated to address concerns related to school selection bias of high-ability individuals. High-ability students are expected to have extended education due to lower costs and higher benefits for them to do so. Behrman & Birdsall (1983) was the first to show that excluding school quality (measured by achievement scores) in the Mincer wage production function will lead to an upward bias for the parameter estimate of the years of schooling. Such

[2] A methodological implication of this is that if we cannot control for innate ability for learning in an analysis of school effects on test score outcomes, then we cannot separate out the effect of a school effect readily.

investigations include Lazear (2001), Mulligan (1999), and Murnane et al. (2000). However, the consensus in these studies is that school attainment is still the primary human capital measure.

There is a vast literature linking test scores and earnings. Chetty, Friedman & Rockoff (2014), Dougherty (2003), Hanushek (2009), Heckman, Stixrud & Urzua (2006), Lin, Lutter & Ruhm (2019), Rose (2006), Ozawa et al. (2022), and Watts (2020) provide reviews of the literature on the relationship between test scores and labor market outcomes. In overall, these reviews conclude that the evidence essentially confirms that standardized test scores are systematically related to the likelihood of employment and level of wages both in developed and developing countries.

The gender achievement gap, i.e., the average test score differences between girls and boys, is also explored by economists since it predicts future gender gaps in employment status or wages and reflects inequality in the education provision (Lai, 2010). Recent evidence shows that girls have made impressive progress in many dimensions of education and have even outperformed boys in most educational achievement measures (DiPrete & Buchmann, 2013; Rose, 2006).

Furthermore, Gintis (1971) and Kautz et al. (2014) underline the importance of non-cognitive skills and argue that they are related to future educational and labor market outcomes as much as cognitive skills. Importantly, non-cognitive skills are malleable by families, schools, and social environments throughout K-12. Most cited non-cognitive skills are the group known as The Big Five, which include conscientiousness, agreeableness, extraversion, neuroticism, and experience (Alderoti, Rapallini & Traverso, 2023). Conscientiousness, which corresponds to being hardworking, responsible, and organized is regarded as the prominent non-cognitive skill that is shown to be directly associated with job performance and wages (Almlund, Duckworth, Heckman & Kautz, 2011).

On the other hand, teachers generally react to non-cognitive skills, which are reflected in pupil behaviors (like showing up in school, paying attention during lectures,

handing in homework timely, and behaving appropriately in class). Furthermore, they reflect that valuation in student grades (Jackson, 2018, Willingham, Pollack & Lewis, 2002). Grades given by individual teachers versus standardized test scores obtained from nationally-held tests are theorized to be indicators of different achievement dimensions. For instance, Westphal, Vock and Kretschmann (2021) argue that conscientiousness partly accounts for why some students obtain higher grades than their scores from standardized achievement tests. Borghansett al. (2016) underline that the very reason for the existence of achievement tests is that while the grades depend on the students' personalities, achievement tests are personality-neutral.

The gap between test scores and grades, known as the grading gap, has been explored to understand the achievement gap between black and white students (Jencks & Phillips, 1998), the gender wage gap (Bertrand, Goldin, & Katz, 2010), and socioeconomic status related gaps (Kautz et al. 2014). Findings of Heckman, Pinto, & Savelyev (2013) also indicate that verdicts of teachers reflected in GPA are strong predictors of adult outcomes.

2.2.2 Macro-perspective

In this subsection, I will summarize the literature that relates standardized achievement test scores to long-term economic growth.

There is a concern that aggregate attainment alone might be a lacking measure, and the quality dimension of education might significantly impact the aggregate level of human capital. Combined with the fact that education quality differs considerably across countries, mentioned concern led growth researchers to question incorporating aggregate attainment as the ubiquitous human capital measure.

Pritchett (2001), in his influential study on the growth in developing countries, shows that human capital investments have negligible impacts on economic growth due to low educational quality; i.e., the time spent in school was not raising cognitive skills or productivity.

Based on the above concern, growth research has incorporated indicators of educational quality in the last two decades. Hanushek and Woessmann (2021) suggest a canonical growth model relating the GDP per capita g to labor skills H (human capital stock) and remaining factors X as such:

$$g = \alpha H + \beta \boldsymbol{X} + \boldsymbol{\varepsilon}$$

In their seminal study, Hanushek and Kimko (2000) proposed an approach to aggregate the individual test scores at the country level to allow international comparisons. As attainment and human capital were used interchangeably, they called aggregate achievement (or aggregate cognitive skills) *knowledge capital* to differentiate from the aggregate attainment measure. They find that adding the countries' test score averages into the model boosts the variance that the model accounts for from 33 to 73 percent. Moreover, the coefficient value of years of schooling (attainment) almost became zero. Jamison, Jamison & Hanushek (2007) extended this analysis with more control variables and prolonged periods. They concluded that cognitive skills affect income growth by accelerating technological innovations. Hanushek and Woessmann (2008, 2015) find that a one standard deviation rise in standardized test scores leads to a two percentage point increase in economic growth.

Hanushek and Woessmann (2015) draw attention to whether to focus on the students with lower or higher achievement scores. High achievers are valuable as they have the potential to play the roles of business leaders, innovators, or star scientists. Endogenous growth models (Lucas, 1998; Romer, 1990) highlight the importance of the research and development workforce fed by high achievers. Similarly, technological diffusion models (Nelson & Phelps, 1996) underline that imitation or catch-up strategy is best implied with high-skilled human force. On the other hand, concentrating on much broader groups is also valuable as they can put established technologies into use across different industries. Thus, the Augmented Neoclassical Model (Mankiw et al., 1992) emphasizes the importance of basic skills, especially

literacy skills. Hanushek & Woessmann (2015) find that focusing on both groups will contribute to economic growth. Moreover, they are complementary strategies that fortify the effectiveness of each other[3]. In this line of thought, PISA scores, for instance, encourage countries to have a large base of at least level 2 proficiency students *and* a solid amount of level 5 or 6 proficiency students.

2.3 Education Production Function Framework

The extensive evidence for the economic importance of skills accumulated by the education process was reported in the previous two subsections. The first subsection has outlined how education measures are incorporated as a right-hand side (i.e., as an input) variable into the production of (i) earnings in microeconomic analyses and (ii) gross domestic product per capita in macroeconomic studies. Quantity measures of education, mainly the years of schooling or the highest graduation level, were discussed in this subsection. The second subsection elaborated on the quality measures such as standardized test scores or course grades. The literature surveyed within both of those subsections provided ample evidence for the unconfounded association between both types of education measures and later life outcomes. Specific to this study, previous sections show that standardized test scores and school grades are relevant and essential concepts for economic analysis.

This subsection focuses on the production of education itself, which implies that the education measure is now the left-hand side variable. Now it is the outcome. The Education production function model is the core theoretical economic framework for studying human capital production through education. An educational production function is the productivity relationship that links schooling inputs to outputs. Often explored output is a standardized test score (Bowles, 1970, p.12; Todd & Wolpin, 2003, p.F3). Thus, the education production function can alternatively be called as *achievement production function* or *knowledge and skills production function*

[3] Heller-Sahlgren & Jordahl (2022) find that a ten percentage point rise in the share of students achieve basic literacy increases the yearly growth rate by 0.18 percentage points, while the same improvement for high-achievers is 0.87 percentage points. Note that the increase in the latter is harder to achieve.

(Behrman, 2010, p.4891). The following and final subsection of the literature review will present the economic perspective for understanding the effects of private schooling on educational outcomes based on the production function framework to be laid out in this subsection. These closely linked subsections aim to provide the basis of empirical analysis to be covered in the later chapters.

Education economists draw an analogy between students' human capital accumulation process in schools/households and a firm's production process. By the production function analogy, they aim to understand the production technology that combines student and school inputs to develop new cognitive and non-cognitive skills (Todd & Wolpin, 2003). Schools are regarded as core production units in the education production framework (Hanushek, 1979). The objective function of a school is a means for maximizing student outcomes given the resource inputs on hand. Achievement outcome differences measured by test scores or grades are considered here as more immediate measures of human capital differences. Still, at the same time, they act as proxies for skills valued in the labor market (Hanushek, 2020, p.163). Everything that determines the achievement outcome is regarded as an *input* into the production process. These inputs involve student inputs, family inputs, peer inputs, school inputs, and neighborhood inputs. The mentioned input-output approach is a distinct aspect of economic thinking for skill formation in schools (Becker, 1999). The above perspective also makes explicit that the human capital production of students in a specific school is not entirely within the control of that school. In that sense, the education production framework admits that complementary production processes run within households or neighborhoods in addition to the school production processes.

Specification and estimation of an education production function are built on the theoretical underpinnings that present pupil skills correspond to the cumulative of all historical household and school inputs on top of children's inherited endowments (Todd & Wolpin, 2003, 2007). Existing human capital stock and the current input flow determine the current level of learning. Households maximize their utility over

the life cycle by making optimal human capital investments in their children (Ben-Porath, 1967), The government's objective is to maximize the country's human capital stock (Brown & Saks, 1987, p.487).

Based on the notation of Harris (2010), the education production model is:

$$A_{it} = g(F_{it}, F_{it-1}, \dots, F_{i0}, S_{it}, S_{it-1}, \dots, S_{i0}, P_{it}, P_{it-1}, \dots, P_{i0}, I_i, \varepsilon_{it})$$

A_{it} is the achievement for student *i* at time *t*, which is a function $g(\cdot)$ of the family inputs *F*, school inputs *S*, peer inputs *P* in all periods, a fixed student contribution I_i, and an error term, ε_{it}. The fixed student contribution, I_i, corresponds to any fixed differences that affect learning. This term primarily covers innate cognitive ability and personality traits but also involves any aspect of the student that affects achievement formation but remains intact during the lower-secondary period. Finally, ε_{it} is the random error term provided that all inputs from all periods are being observed.

Human capital accumulation in each period equals to the cumulative sum of all student skills obtained from conception to the period *t*[4]. If we are willing to assume there is no decay (i.e., some of the previous learning becomes absolute as time passes) in the last stock of achievement, we can rewrite the production function as such:

$$A_{it} = g(A_{it-1}, F_{it}, S_{it}, , P_{it}, , I_i, \varepsilon_{it})$$

But if there is a geometric decay by some constant λ:

$$A_{it} = g(\lambda A_{it-1}, F_{it}, S_{it}, , P_{it}, , I_i, \varepsilon_{it}\text{-}\lambda\varepsilon_{it-1})$$

Harris & Sass (2005) and McCaffrey (2004) conclude that the assumptions about geometric decay are not strong, and, hence, it is reasonable to assume λ=1.[5]

[4] Due to lack of data for past periods, past inputs are generally assumed to be irrelevant to current outputs (Harris, 2010).

[5] There is some evidence that it is around 0.80 (Harris, 2010).

I will now summarize the literature on the proxy measures of these inputs. Ideally, what shall be utilized is direct input measures of education like parents' verbal interaction on education matters at home, teachers' instructional effectiveness in the classroom, and so on. However, proxy measures must be utilized as reliable and sufficiently large data is generally unavailable for direct input measures. In the first two subsections, we have seen that inputs into production are labor and capital investments, where capital investments involve *human* and *physical* capital. I will follow the same course summarizing the educational input proxies for human capital production.

Home inputs correspond to the capital and labor inputs provided by the students and their parents in their homes. *Home capital inputs* are provided only by parents and involve physical and human capital dimensions. Both the parents and students provide *home labor inputs*, which include the quantity and quality of time parents allocate to the educational tasks of children as well as the quantity of time children dedicate to studying on their own[6] (Ritzen and Winkler, 1976, p.430).

Physical capital inputs in the household are proxied by the number of books (Bowles, 1970)[7], home possessions like a computer or a student's study desk (Broer, Bai, & Fonseca, 2019; Yang, 2003), parental occupation (Bowles, 1970), family income (Todd and Wolpin, 2003), and home ownership of the family (Ritzen & Winkler, 1976).

Human capital inputs in the household are proxied by the attainment (years of schooling) of the parents (Goldhaber, 1996, p.96) or their final degree of graduation (Bowles, 1970; Ritzen and Winkler, 1976).

---

[6] Note also Schultz (1960, p.573), while introducing the notion of human capital, underlined that students, as self-employed producers of capital, actually work when they study and create human capital.

[7] Broer, Bai, and Fonseca (2019) report (based on TIMMS experience) that this measure is strongly associated with the parental attainment.

Time inputs of parents are proxied by the family size (Bowles, 1980) or the number of siblings (Ritzen & Winkler, 1976, p.431)[8], and the absence of one or more parents (Bowles, 1970). Becker (1993)'s theory of time allocation within the family implies a quality-quantity trade-off regarding family size. Thus, working mothers may have less time to spare for their children. Still, Feinstein & Symons (1999) argue the quality time (i.e., productive time devoted to helping with the educational tasks of their children) they spend with their children seems to be comparable with non-working mothers. Overall, a mother's labor market participation status is a candidate control variable for parental time inputs.

The labor input of the student at home is measured by the time allocated for the homework (*input quantity*) and the grade obtained from it (*input quality*) (Coleman, Hoffer, & Kilgore, 1982). Betts (1995) finds a significant positive association between the amount of homework and the level of achievement, which also underlines the importance of student time input at home and the means of facilitating this[9]. The existence of a child's own room at home, which provides an opportunity for the child to make focused time inputs, can be a proxy for both the quantity and the quality of time inputs. Noell (1982), on the other hand, argues that the gender of the student might be another factor that proxies time inputs from the students at home.

A school's contribution to its students' human capital development is also a function of the capital and labor inputs provided within the school. Capital inputs include physical and human capital dimensions. School labor inputs include the quantity and quality of time a teacher allocates to a given child and the student's participation and attempt to learn within the classroom (Ritzen and Winkler, 1976).

Physical capital inputs in the school involve school facilities, number of classrooms, labs, libraries, books per student, computers, and amenities for sports and arts. (Hanushek, 1986, p.1155). Human capital embodied in teachers and administrators

[8] The rationale is that there remains less time the parents can allocate to any one child as the number of children in the home increases, *ceteris paribus*.

[9] Coleman, Hoffer & Kilgore (1982) notes that private (catholic) schools demands more homework.

involve education level (master, Ph.D.), years of experience, certificates, teacher test scores, project or research work, teacher salary, teacher motivation, and expenditure per pupil (Behrman, 2010; Bowles, 1970; Hanushek 2020). The school's region may also predict both physical and human capital investments (Noell, 1982).

Labor input from the teacher is proxied by the teacher-pupil ratio or class size (Hanushek, 1986, p.1160), which are highly correlated. The teacher-pupil ratio might be a better proxy since different teachers instruct different lower-secondary school courses, and a teacher's total course load may affect the quality of their inputs. Lazear (2001) also suggests that the higher the number of students in a classroom, the higher the probability that a lecture is interrupted (due to frequent questions or unruly behavior), which will reduce the time input for learning.

Regarding student labor inputs, obtaining data for direct measures of the student's involvement and effort in the school is hard. The absence rate is suggested as one potential measure for the lack of these inputs (Cain & Goldberger, 1983, p.216). However, the lower scores associated with absences might be due to the direct effect of these absences or may reflect unobserved factors that lead to high levels of absenteeism (Coleman, Hoffer, & Kilgore, 1982).

Ritzen & Winkler (1976) argue that students routinely make work-leisure decisions within the classroom, where they value the praise from their peers and adapt their focus and effort accordingly. If the peer group changes, the reward structure or the academic norms within the group may also change. In reaction to the new peer group, the child may change the proportion of time spent on studies instead of leisure (Robertson and Symons, 2003). As another mechanism, teachers may set the pace of teaching according to the classroom's average learning capacity, knowledge, and behavior of the students.

The standard strategy of controlling for peer effects in the literature entails including school and neighborhood compositions (such as school averages of input measures) as additional inputs into the production function (Castellano, Rabe-Hesketh, &

Skrondal, 2014, p.348; Crain & Ferrer, 1982, p.3-4; Hanushek, 1986, p.1155). Paloyo (2020, p.298) argues that if ending up in a particular peer group is (approximately) random within a school, we can overlook the reflection problem. The evidence in terms of peer effects is mixed (Sacerdote, 2014, p.269), mainly due to the context-specific nature of peer effect existence and the inability to differentiate peer effects in the case of endogenous selection into schools[10]. Evans et al. (1992) conclude that ordinary lest squares estimation of peer group effects is positively biased due to the selection of peer groups by parents and the selection of peer groups by schools. Among the studies with robust research designs utilizing education production function modeling, Ding & Lehrer (2007), Feinstein and Symons (1999), and Jackson (2013) find that 7 to14 percent of the contribution of school factors to the test score outcomes can be directly linked with the peer inputs.

Adequate measures of innate ability are hard to obtain. Hanushek (1986, p.1156) and Harris (2010) argue that they have never been available. Hence, innate ability remains part of the error term in the econometric models, which leads to an upward bias in the parameter estimates of student variables (Levin, 1976, p.152), as innate abilities are positively correlated with family background through both the genes and the environment. Early test scores are sometimes suggested as proxy measures for innate ability. However, early test scores might not be an appropriate proxy for innate ability because of their profound interplay with socioeconomic status (Currie and Thomas, 1999). On the other hand, Hanushek (1986, p.1156) argues that the correlations between innate abilities and school attributes will likely be small after allowing for family background variables. Hence, he further claims that the size of the bias on the parameters of school inputs due to unobserved innate ability would be much more limited.

[10] Jackson (2013) and Tonello (2016) exploits random variation in peer inputs between cohorts. Jackson (2013) further utilized IV strategy to account for the case that this variation may not be random.

Finally, the literature indicates that the level of student inputs may be moderated by the student demographic characteristics such as gender or ethnicity (Bradley and Taylor, 2004). Feinstein & Symons (1999) find that girls are more responsive to parental inputs than boys. They also show that girls from private schools lag behind girls from public schools, *ceteris paribus*.

Monk (1989, p.31) states that two research traditions mark the implementation of education production functions. The first is to obtain parameter estimates of the education production function, which was heavily pursued in the first two decades of research in this area. There is administrative value in knowing fully about the education production function, which potentially enables school administrators to make least-cost input allocation decisions on an operative basis by being able to compare marginal products with prices. However, the inconsistency across several studies which estimate these functions frustrated these normative investigations aiming to determine optimal educational resource allocation. Moreover, (Levin, 1976, p.153) argues that the fundamental assumption behind the exact application of the production function analogy is that schools are *technically efficient*, i.e., they maximize output subject to their input combination alternatives. Market competition forces the individual firms and industries they form to achieve technical and allocative efficiency.

Competition between schools, however, is generally not as extensive as the competition between industrial firms. The school managers do not readily know input/output prices, and schools do not operate at the production possibilities frontier while it is assumed so. Hanushek (1986) argues that these aspects do not preclude the effective utilization of the education production function framework for input-output analysis. Indeed, he underlines that schools do not operate on the production possibilities frontier itself provides the variation, which enables us to learn from statistical analysis. Hence, the second approach to education production functions is

to utilize it as the conceptual basis of policy research[11] and focus on estimating the coefficient of the policy variable consistently using econometric fixes as required.

A second digression from the generic industrial firm production framework is the multi-product nature of educational production, which differs from the production lines of firms optimized for a single type of product. The majority of studies incorporate only a single measure of output, the cognitive skills measured by the scores on standardized achievement tests, even though the schools also produce non-cognitive skills (Levin, 1980, p.152). A system of equations can address the multi-product nature of school production. However, data limitations often prevent the use of this strategy.

A third digression is that the process is not much of an issue in industrial firms as it is written in engineering blueprints and applied straightly. However, the processes that transform inputs into outputs in educational production are not so straightforward. There is no engineering blueprint for teaching excellence. Engineers deal with machines and professional adults who are experts in their craftmanship; teachers deal with students and parents. Hence, the education production function should not be regarded as an exact replication of an industrial production function. It shall be considered as a framework for input-output analysis, where processes are essential but, in general, unobserved elements of production that go into the school random error term. These shall be kept in mind while conducting research in this framework.

2.4 Private Schools and Human Capital Formation

2.4.1 Private school effects on individual student outcomes

Hsieh and Urquiola (2006) report that private schools are generally expected to be more effective in increasing achievement than public schools. As private schools face

[11] To sum up, education production function analysis won't replicate production manager's routine monthly optimized production plan in an industrial firm. But it provides a framework to learn about average input-output relations in educational production.

competitive pressure to a higher degree than public schools, they are incentivized to provide higher quality education. They are better positioned to improve the human capital of their students (DeAngelis & Wolf, 2020). In addition to the achievement criteria, private schools offer religious education for the families who demand it, offer better peers, more school safety and discipline, and more varied extracurricular activities (Figlio & Stone, 1999). Based on this expectation (or presumption), parents send their children to those schools based on the trade-off between assumed benefits and income/credit constraints. Learning about actual private school effects will inform families and let them optimize human capital investment decisions for their children. This knowledge will also help achieve allocative efficiency.

There may be students and parents who are at the margin of making private school decision. The private school choice may lead to higher achievement for some of these students. Then aggregate achievement can be increased by providing more choice to families through vouchers (Hoxby, 1994). Education vouchers are a demand-side intervention aiming to expand parental school choice by supplying governmental subsidies for the private schooling decisions of families (Patrinos, 2007). Different voucher designs have diverging goals and varying *modus operandi* (Arenas, 2004). If the voucher cost is on par with the expenses of educating those students in public schools, then the government may opt for this policy[12]. This rationale focuses on the *participation effect* of the voucher. School choice benefits participating students in two ways; first, they can study in a school that better matches their particular needs (thus improving their motivation). Secondly, they can enroll in a higher quality school than their reservation public school (Cullen, Jacob, & Levitt, 2005). As long as these benefits lead to achievement gains, more human capital will be invested in the participating children. Hence, informing educational planners on the estimates of participation effects of voucher programs will help them gauge their program's allocation rules and operations. Aggregate participation effects will inform strategies focusing on increasing the country's level of human capital stock.

[12] Provided that private schools are also at least as good at improving non-cognitive outcomes of students as public schools.

Understanding the factors that lead families to make private school choices is essential to make valid inferences about the private school effect. Parental education is among the most cited factors (Coleman, Hoffer, & Kilgore, 1982). Well-educated parents, on average, make better decisions about getting their children into a school that will add the most skills. At the same time, they demand private schools more on average.

Private schools require tuition; hence, factors like family income, the number of siblings already enrolled in private school, ownership of the home where the family lives, and credit constraints determine whether the private school decision is feasible for a specific child. On the other hand, adding tuition as a control variable in the analysis model is found to be problematic. Because tuition is strongly correlated with household income and parental schooling, the collinearity problem likely arises, which leads to a flip in the coefficient signs of family income and parental education when the tuition is included in the model. Moreover, tuition also proxies the quality or reputation of the school, and its coefficient partly reflect those factors . (Goldhaber, 1996, p.101).

The region's socioeconomic status and the home's specific location are suggested as other factors influencing parents' decisions on private schooling Whether the parent is employed in the public sector is important since public sector officials are more inclined to support and opt for public schooling (Toma & Long, 1988). Lankford and Wyckoff (1992) report that the gender of the student seems to have no significant effect on enrolment into a public or religious school[13].

Parents also weigh the observed quality of the school against the tuition cost. The propensity of private schools in the locality and the differences in quality between the two sectors are also suggested as decisive factors in the literature. For instance, West and Palsson (1988) incorporate the public school teacher-to-student ratio as an input factor and find that when this ratio gets lower for public schools, more students attend

[13] Note that the majority of religious schools have private school status in the United States.

private schools in that particular locality. Frenette & Chan (2015) reports that the province of school accounts for the average standardized test score differences between the two school sectors.

Parents are inclined to choose schools with high socioeconomic status families but also prefer families with similar racial or ethnic characteristics (Goldhaber, 1996, Hamilton & Guin, 2006). Parents are reported to believe that small class size is essential (Kleitz et al., 2000).

Whether or not the school is religiously oriented may also be crucial for parents if they emphasize education that integrates religious values. The percentage of religious-oriented private schools in the household's residential state and whether public alternatives to those schools are also suggested as a factor (Martinez-Vazquez and Seaman, 1985). Chiswick (1988) reports a significant influence of religion on achievement.

After presenting the factors determining private school choice, I will summarize the findings of the empirical literature on private school effects. In their seminal and controversial study, Coleman, Hoffer, & Kilgore (1982) found that students attending Catholic schools tended to obtain higher achievement than those with similar backgrounds attending public schools. Goldberber & Cain (1982), Noell (1982), and Murnane (1985) heavily criticized the methodological approach of this seminal paper. Discussions and revisits continued longer (McEvan, 2004, p.110). Reanalysis of the original data by Noell (1982) produced an estimate of around 0.13 standard deviations of increase in standardized test scores due to private school status. Gamoran (1996) also found positive effects around 0.1 standard deviations, while Rouse (1998), Jepsen (2000), Angrist et al. (2002), Howell et al. (2002), and Anand, Mizala, & Repetto (2009) find effect sizes around 0.2 standard deviations. Somers, McEwan, & Willms (2004) find an average private school effect of 0.3 standard deviations for ten Latin American countries holding constant the level of student inputs like

socioeconomic status.[14] The meta-analysis of voucher participation effects by Anderson, Guzman, & Ringuist (2013) reports a range of effect sizes between 0.05-0.07 standard deviations. Shakeel, Anderson, & Wolf (2021)'s review, on the other hand, documents modest positive effect sizes between 0.11 to 0.33 standard deviations[15]. In general, estimates are lower for programs from the United States and are higher for developing countries. Morgan, Petrosino, & Fronius (2015) highlight the need for more rigorous research on voucher design and its impacts in developing country contexts.

Among the influential studies that do not report significant positive results, Goldhaber (1996), Figlio & Stone (1999), and Elder & Jepsen (2014) find no statistically significant effects, while Sander (1996), Lubienski et al. (2008), and Reardon (2009) find negative but insignificant estimates. The OECD (2011) documents that with similar levels of autonomy and student composition between the two school types, the private school advantage disappears in 13 of the 16 OECD countries. On the other hand, Nguyen et al. (2003) find that Catholic schools improve math test scores (over public schools) only for boys but not for girls. Abdulkadiroglu, Pathak, & Walters (2018) show that the expansion of school choice can reduce student achievement drastically (0.40 sd) if private schools recruiting students were failing/declining private schools right before the start of the voucher program.

Regarding the private school effects on non-cognitive outcome measures, Elder & Jepsen (2014) find a slightly positive impact on absences and suspension and a small negative effect on tardiness. As future research, they call for more investigation into the effects of private schooling on non-cognitive outcomes and whether these effects differ among demographic groups. DeAngelis & Wolf (2020) use as the measure the effort spent on student questionnaires that require patience (within the PISA exam) to assess private schools' effect on non-cognitive skills, which are mainly

[14] However, this effect reduces to 0.04 standard deviations when the school's average socioeconomic status is controlled for.

[15] These are predominantly intention-to-treat effects.

conscientiousness and effort. They find that survey nonresponse reduces achievement by 0.25 standard deviations. Miller (2022), on the other hand, focuses on non-cognitive skills such as self-esteem and self-efficacy to find an improving effect of 0.16 and 0.22 standard deviations between ages twelve and fifteen.

2.4.2 Private schools' competition and sorting effects on public schools

States get involved in education sector by (i) designing curriculum and conducting standardized examinations, (ii) influencing attainment levels through the compulsory schooling laws, (iii) direct provision of education services by building schools or employing teachers, and (iv) funding by providing non-tuition public schooling or education vouchers to (partly or totally) cover the tuition of private schools (Tooley, 1996). When a government reduces its involvement in any of these, then it would be a move toward the market.

A market-based system is criticized due to the concerns that it will lead more segregated education system, less social cohesion, and uneven income distribution (Gradstein and Justman, 2000). However, a more centralized education provision is criticized, too, due to the claims that educational opportunities are unevenly distributed even under a centralized education system. Furthermore, public provision is argued to be associated with lower allocative and productive efficiency. Freedman (1962) made the resounding call for a more significant move towards the market, more private school choice and the provision of education vouchers. His main argument was that increased competition created by increased private school involvement would improve the quality of public education, called the competition effect. Public schools act as a system of local monopolies. Increased parental choice is expected to make them more responsive to the needs and demands of students and parents. (Cullen, Jacob, & Levitt, 2005, p.729). Their efficiency is also assumed to increase. Yet the effectiveness of private choice programs remains debated (Pianta & Ansari, 2018; Urquilo, 2016).

Sander (1999) raised the critical concern that if private schools attract better public school students, mean achievement in public schools will decrease due to the reduced

number of high-ability students (composition effect) and the adverse peer effects (contextual effect). This is known as the cream-skimming effect and acts contrary to the competition effect (Epple & Romano, 2008). The cream-skimming effect can lead to a state where disadvantaged students are left behind in the most lagging schools and deprived of peer contribution (Henriquez, Lara, Mizala, & Repetto, 2012). Cream-skimming effect may dominate the participation and competition effects of voucher programs. Hence, the net effect of a voucher program is an empirical question.

Moreover, in addition to adverse peer effects, the effectiveness of public schools may decrease even due to the competition effect alone. McMillan (2001) theorizes a set-up where if a public school makes a high effort, both low and high-socioeconomic-status households would choose this school. However, if they exert low effort, only low socioeconomic status households would choose it. Higher effort leads to higher enrollment (bringing more resources or reputation), but it is costly as it requires higher effort. When a voucher program takes away high socioeconomic status students at the margin in terms of their quality-cost tradeoff, public school teachers and administrators may opt to supply only low effort based on their cost-effort tradeoff. The author underlines that there would be no such risk if the voucher was assigned to only low socioeconomic status households.

Borland and Howsen (1992) is a seminal study that suggests including measures of the degree of concentration in educational markets in econometric models. They find weak evidence for the competition effect. Blair and Staley (1995) explore evidence of competition from neighboring districts, reflected by the average test score of borderline school districts. They detect competition effects in school districts surrounded by high-achieving districts. Zanzig (1997) finds that just a few school districts are sufficient to observe a competitive market. Note that an education market generally corresponds to a school district in the United States. Since schools are funded by local property tax, students cannot cross the boundaries of their school district. Therefore, school districts compete for students via their local public schools. Housing and school choice go in hand, leading to Tiebout type of sorting (Tiebout,

1956; Martinez-Vazquez and Seaman, 1985), while private school choice might allow a lower-cost housing solution. In contrast, in the UK or Türkiye, students are not restricted by the education district boundaries. In that case, schools compete for the students. Even in the case of no school boundaries, time and monetary costs of transportation (of student plus the parent) generally set a natural distance limit of 25-40 miles at most. Hence, schools within a certain distance compete with each other in that case. Bradley & Taylor (2002) show that a three percentage points increase in the average exam score of its competing schools leads to a one percentage point improvement in a specific school.

The above studies were focused on more generic competition between districts or schools. Estimating private schools' competition effects on public schools is more involved. An important challenge for the estimation of competition effect is that private school share in a particular locality might be endogenous with achievement in the public sector in that locality for the reason that quality composition of the public sector affects private school demand (Downes and Greenstein, 1996; Sander & Cohen-Zada, 2020, p.513)[16]. This raises the requirement of isolating the supply side of the private school share.

To address the endogeneity of private school involvement with public school quality, it is required to focus on the factors that affect the supply decisions of private school entrepreneurs. Moreover, the voucher programs are based on the assumption that new private schools will open near low-quality public schools, and the empty slots in incumbent private schools will be filled. This will lead to a more competitive education sector and increased effectiveness/efficiency in public schools. Hence, factors that determine the location decisions of private schools shall be understood.

[16] On the other hand, endogeneity of private competition effect with public school quality may or may not be present in any specific context. High quality public schools *ceteris paribus* reduce the demand for private schools. However, religious values, for instance, might trigger certain groups to supply private schools in a highly strong public school district and certain group of parents might choice school based on those values instead of achievement levels. Also, in a fast expanding voucher program circumstances, private sector entrepreneurs might be convinced that new demand will meet them, too.

Downes and Greenstein (1996) is the seminal contribution that examines the location choice of private schools. They find that the properties of the population and the average public school quality in a specific locality influence private school location patterns. Parents' educational attainment levels are found to be especially important for the number of private schools in a locality.

Demographic factors, primarily religious affiliation, may lead individuals to choose a specific type of private school independently of the average public school quality in that locality. They also detect a significant variation in the educational objectives of private schools based on their religious affiliation[17]. Hoxby (1993) has shown that the magnitude of a subsidy from a sponsoring religious organization depends on the relative size of the local population belonging to it[18].

Both Barrow (2006) and Downes and Greenstein (1996) put emphasis on the measures of demographic factors that correlate with the heterogeneity of demand in the community, which lead to a feeling for some groups that the quality of education is inadequate. The family income or the parental attainment distributions are suggested as indicators of such factors[19]. Moreover, as within-community heterogeneity can be larger in the urban areas or specific regions of the country, urban dummies and regional dummies are suggested to control both urban-rural differences and differences between urban areas. Zhan (2018) concludes that school choice does not necessarily make private schools favor low income neighborhoods; instead, they prefer places with higher average income. However, if the voucher program is large enough, the appeal of poorer localities to private entry increases. Zhan (2018)

[17] This difference is generally context dependent. In the United States these differences are between Protestant and Catholic-oriented private schools. In Türkiye, for instance, that corresponds to differences between religious and secular private schools.

[18] It can be argued that a factor separating religious schools from secular ones is that there is an ongoing voucher-like system operating behind the scenes. So, religious orientation of schools reflects to a degree the exogeneity (to the public school quality) that is brought about by voucher programs. Religious communities also fund the establishment of these schools.

[19] Barrow (2006) also finds that mean income in a community is an important factor for the private sector share in that community, while Downes and Greenstein (1996) do not. However, the latter authors find the mean income levels of neighboring communities as a relevant factor.

documents that private school propensity is higher in localities where the teacher-student ratio and the expenditure per student in public schools are lower.

Hoxby (1994) and Dee (1998) were among the first studies to focus on competition precisely due to increased numbers and enrolment shares of private schools, both of whom concluded positive effects of such competition on public school quality. Both Hoxby (1994) and Dee (1998) used religious population densities as the instrument for the rationale that a substantial portion of private schools is sectarian. This class of instruments remains the mainstay for estimating competition effects[20].

In their comprehensive literature surveys, McEwan (2000) and Belfield & Levin (2002) reported that the most of the first group of studies concluded with a modest-sized positive private competition effect. Recent reviews by Egalite (2013), Epple et al. (2017), and Jabbar et al. (2019) confirm previous literature surveys in concluding that modest positive competition effects on average. Goldhaber (1996), Sander (1999), Jepsen (2000), McMillan (2001), Geller, Sjoquist, & Walker (2006), Hsieh & Urquiola (2006), and Bowen & Trivitt (2014) are among the studies with robust research designs that find no effect or a modest negative effect.

The long-term effects of voucher programs are considerably less explored. Carr (2011) and Figlio et al. (2020) find positive long-term effects, while Canbolat (2021) and Gray et al. (2016) conclude that after an initial positive short-term effect, the benefits on the public school sector fade away. The major difference between these voucher programs lies in their design. The first pair has an accountability-tied voucher design, while the second pair has a voucher-shock design. Accountability-tied voucher designs require failing school status, are longer-term by design and improvement in test scores is expected[21]. In contrast, all the students meeting the

[20] However, Cohen-Zada (2009) argues that the common use of the current Catholic share in the local population as an instrument for private school competition will not be valid and instead, historical shares should be used.

[21] It is also methodologically challenging to isolate the long-run effect of an accountability-tied voucher from the stigma effect of accountability (Chakrabarti, 2008). It could be the case that vouchers had no

voucher criteria are eligible for voucher-shock designs. These are generally shorter-term, intensive programs whose positive effects in the initial years might turn negative in the longer term.

Hisieh and Urquiola (2006) underline that the productivity effects of school choice policy on schools and students shall be analyzed at the aggregate market level (Hsieh and Urquiola 2006). Voucher programs aim to increase the demand for private education. Based on the eligibility rules and priorities, demand from different groups of students may vary. The demand also depends on how the public perceives private schooling in a specific locality. Market-level variation in these factors provides identification for the effects of school choice programs.

effect, but the stigma of being a failing school led to an increase in the average achievement level of the school.

CHAPTER 3

INSTITUTIONAL SETTING

TEOG examination provided a unique opportunity to explore private school effects because the design of the examination led all private schools and all of their students to participate in the same exam together with all public schools and their students. The period covering TEOG examination is also of particular interest as it corresponded to a major transformation period in terms of private school markets in Türkiye. In this chapter, I will describe the institutional setting which provided a basis for these developments. I will first present a summary of the historical developments regarding the private school sector to better put the later reform period into context. I will focus on the reform period of 2012-2019 to set an institutional backdrop for the analysis of the TEOG exam. The final two subsections will cover the details of the voucher program covering the 2014-2018 period and the TEOG exams held between 2013-2017.

3.1 Historical Background of Private Schooling in Türkiye

The first appearance of private schools can be traced back to the reform period of 1839-1876, known as the *Tanzimat* period. The Edict of Reforms (*Islahat Fermanı*) of 1856 was the first official document talking about private schools (Uygun, 2003:108), which permitted the establishment of private schools by minorities (i.e., non-Muslim communities) within the Ottoman Empire. The first law for the education sector, which came into power in 1869 (*Maarif-i Umumiye Nizamnamesi*), also had articles that permitted the establishment of private schools both by Turkish educational entrepreneurs as well as ones from minority groups. However, private schools were mainly established by minority groups in that period. These initiatives indicated that private schooling was a significant subject early in the country.

The activities of private schools of minority groups within the period of 1876-1903 (during the reign of Abdulhamit II) were deemed offensive. However, their governmental supervision remained limited (Altınok, 2019:13). These schools were generally conceived to provide high quality education. Akyüz (2012:240) argues that this perception of higher education quality triggered Turkish educational entrepreneurship to establish 28 private schools of Turkish origin, which inhabited around 4500 students by the year of 1903. In 1915, the Private Schools Circular (*Hususi Mektepler Talimatnamesi*) was enacted to promote the activities of Turkish educational entrepreneurs (Özkaya, 1986:37). However, private sector involvement remained limited until the 1950's despite an ongoing positive stance towards such investments.

Interest in private schools began to increase during the 1960s. The "Committee for Private Schools", convened in 1961, reported that the law in action dating back to 1915 had become outdated (MoNE, 1961:3), which triggered the process that produced Private Education Institutions Law coded 625 in 1965.

After the first private college (*özel yüksek okul*) opened in 1962, the period of 1962 to 1969 witnessed the proliferation of private colleges. In 1971, there were 70.000 students in 44 different private colleges. The student movement demanding the closure of these private colleges in 1971 involved both left and right-viewing students. These events triggered the university reform process in 1972-73, which further led to the founding of Higher Education Council, a milestone in the tertiary education history of the Türkiye.

Under all this turmoil, private colleges remained in the system with their titles turned into "universities". The political instabilities of the 1970s, however, reduced the general interest in private schooling during that period. The Fourth Five-Yearly Development Plan emphasized the prevention of education from turning into a traded commodity, and it prioritized equal opportunity (State Planning Organization, 1979:455).

However, the neo-liberal political and economic environment of the 1980s that strongly favored privatization renewed interest in private schools (Gök, 2004:97). Moreover, the process of globalization and the knowledge-intensive labor force demanded by globalization had significantly impacted the worldviews of the governments of this era (Gümüş, 2015:168). The Decree of Jan 24th, 1980 opened up the economy of Turkey to international trade as well as to global capital markets. Contraction of the state, privatization of public economic entities, and promotion of the private sector were the accompanying policies.

A new Private Education Institutions Law coded 5580 was enacted in 1985 and remained operative until 2007. This version of the law made the rules, eligibility, and procedures for investing in the private school sector easier. Even foreign direct investment in the Turkish private education sector was made possible as long as foreign investors had Turkish partners. In 1988, the free determination of private school fees by the school owners was permitted (Küçükçayır & Cemaloğlu, 2017).

The political and economic perspective of the 1980s also emphasized the role of market competition to maximize the quality of goods and services at lower costs. Accordingly, the state should ensure the establishment and uninterrupted operation of free markets. This should then even be more important for an expensive public service like education. So, competition from private schools was regarded as a means to amplify the competitive pressure on the public schools so that the overall quality of education would improve and costs would reduce. Private schools would act as exemplary organizations for public schools and lead them in the process (Altınok, 2019). This line of thinking inherently carries the belief into the superiority of private schools in terms of education quality.

Governments in the 90s continued on the positive and supportive perspective on private sector involvement in the education sector. The government program of 1991 explicitly stated that private and philanthropic involvement in establishing private schools would be encouraged. The government program for the 1993-1995 period reiterated the importance of private sector investments in terms of alleviating the

financial burden of the government in the education sector. The Fifth Five Yearly Development Plan covering the period 1990-1994 and the government programs of 1996-1997 underlined the need for expanding the role of the private sector in developing and improving innovative educational approaches and technologies as well as the domestic production of educational materials and equipment.

The Seventh Five-Yearly Development Program (1996-2000) documented that both the participation and the contribution of the private sector in the education sector up to that date remained rather limited since the students enrolled in private schools corresponded to only 1.5 percent of the total student population and that half of the capacity of private schools could not be tapped. The lack of incentives was suggested as one of the important reasons for these outcomes.

As a milestone in Turkish educational history, a compulsory schooling law (coded 4306) was enacted in 1997, which was binding on both school sectors. One year before, the Welfare Party, assessed as a reactionary party by the secular state organizations, became the highest-voted party in the elections and led the formation of a coalition government. The pressures from these organizations dissolved the coalition government. A schooling law (4306) stipulating eight years of uninterrupted education was shortly followed (Ozturk, 2017). It was widely deemed that the main motivation for the law was to restrict religious education as it also mandated the closing of three-year lower secondary *Imam Hatip* schools[22], which were religious-education focused vocational public schools whose graduates may work as officials in mosques (as *imams* or *muezzins*) or become Koran teachers/instructors (Pak, 2004, p.326)[23]. The structure of this law diverged from typical compulsory schooling laws

[22] There were 609 Imam-Hatip schools just before the law was enacted, and more than 200 was waiting in the queue for approval from the government to open up their services.

[23] Imam Hatip schools were so popular among conservative and religious families that the number of students enrolled in those schools exceeded quite an extent the potential positions implied by the mentioned-above occupations. These occupations are especially not suitable for the employment of girls as, for instance, girls cannot act as imams. However, demand from girls for these schools was almost

because it neither altered the legal dropout age (it remained at 15) nor elongated the duration of compulsory schooling. Rather, it combined primary school with lower-secondary school and introduced a single "eight-year basic education". It required an eight-years of basic education in a single school building. Furthermore, implementation of the law stipulated a major expansion in school buildings to meet the demand from 1.5 million out-of-school lower secondary school students (Ozturk, A, 2017, p.7). This led to a more than 30 percent increase in the number of classrooms in the years between 1998-2002. The contributions from the private school sector were welcomed. The Eight Five-Yearly Development Plan (2001-2005) indicated the continued need for increasing the proportion of private schooling in all levels of education, which was still 1.7 percent and showed almost no improvement since the previous five-yearly plan period. The government program for 1998-2002, on the other hand, emphasized the efforts to establish a "fair fee structure" in private schools, which can be regarded as one of the early concerns about the equitability of the private school system.

The major proliferation of private schools was witnessed after 2002. The government program for 2002-2003 underlined the importance of "competition" within the education sector. The Ninth Five-Yearly Development Plan (2007-2013) documented that the private sector share and the support in investments in the education sector had seen a substantive boost by 2006, which led to wider adoption of information technologies in schools and to improvements in the curriculum. Moreover, it was stated that due to the increased share of private investments, public investments could focus more on students from socio-economically disadvantageous families and the prospects for "equal opportunity in education" would be enhanced.

The Ninth Plan also envisaged incentives for transforming private tutoring institutions (*özel dershaneler*) into private schools. These institutions were mainstays for lower-

equal to the demand from boys. In that sense, Imam-Hatips were regarded as supportive for increasing attainment of girls from conservative families.

secondary and higher-secondary school students in their preparation for the standardized examinations held nationwide.

The accession process to the European Union starting in 2005 also positively affected the enlargement of the private schooling sector as well as the governmental support provided to this sector (Gümüş, 2015:167). The dominant perspective of EU institutions or the international organizations (World Bank, United Nations, etc.) in that period was called for the contraction of the public sector and the reduction in public investments. The private sector should be supported to fill the gap created by the contracting public sector.

In 2007, Private Institutions Law coded 5580, which is the most recent law that organizes and regulates the private school sector, was put into force. This law gave private schools autonomy in terms of finance and administration as long as they acted according to the articles of the mentioned law. The original version of the law did not involve major changes. However, amendments to it in the following years were consequential and will be discussed in depth in the next subsection.

The government program of 2008 emphasized the role of potential contributions from the private sector to meet the "schooling-rate targets" of the country. After September 2012, high school attendance was made compulsory nationwide in Turkey, further increasing the need for private investments into the education sector. Mandatory higher-secondary schooling also increased the importance given to TEOG exam taken at the end of lower-secondary school, which was decisive on the higher-secondary school to be attended.

Moreover, the Syrian refugee concentration sharply increased after 2012, raising refugee students' public education demand[24]. Initially, the refugee population was concentrated in the south-east border regions. However, after 2015, refugees started self-selecting into locations. The increasing rate of immigrant propensity in public

[24] The term refugee here corresponds to the general usage in the literature. Their status in Türkiye is officially specified as *Syrians under temporary protection*.

schools resulted in a *native flight* from public to private schools. Tümen (2019) used provincial enrollment registry data to find that higher refugee presence led to a statistically significant native flight from public to private primary schools: One native student switched to a private school for every 31.6 refugee pupils. The magnitude of the effect was modest compared to previous studies around the world.

As the dataset to be analyzed in this study does not have an indicator for foreign student status (including refugees) for individual students, in Figure 3.1 below, I present the numbers of (major five groups of) foreign students enrolled in the grade-8 (lower-secondary school seniors) at each instance. The colored region corresponds to the period analyzed in this dissertation. The numbers in this period remain limited compared to the numbers seen afterward.

Table 3-1 Foreign Student Population Enrolled at the Grade-8 Level

	Syria	Iraq	Afghanistan	Iran	Russia	Total
Oct 2012	25	110	181	80	82	478
Oct 2013	193	151	299	108	55	806
Oct 2014	295	272	411	79	93	1150
Oct 2015	668	582	418	132	130	1930
Oct 2016	2127	1222	702	210	186	4447
Jun 2017	2210	1224	761	208	194	4597
Nov 2017	5071	1864	1172	351	237	8695
Nov 2018	9666	2957	1889	487	255	15254
Nov 2019	19291	3510	2511	658	276	26246
Feb 2020	30562	4134	2863	688	289	38536
Oct 2021	44544	4744	3015	797	333	53433
Oct 2022	58552	5368	3190	800	459	68369

Source: Ministry of National Education

As mentioned earlier, the original version of the law coded 5580 did not make many changes to the previous law. As seen in Figure 3.1, only slight increases were observed in the share of private schools from 2008 to 2012. However, the amendment to private school law coded 5580 that got into force in 2012 was significant in (i) clearly stating the purpose of private schools and (ii) providing new government support schemes.

The purpose of private schools was stated as such: "*Private schools shall not only operate to generate profits. They are permitted to earn revenues for the purposes of increasing the educational quality and to further invest into their future operations*".

The first of the new government support instruments, on the other hand, was to equate the fee rates of electricity, water and natural gas applied to private schools to the rates applied to public schools. The second instrument provided vouchers for the vocational school students enrolled in private vocational schools (at the upper-secondary level) located within organized industrial regions. The size of the voucher could be up to 1.5 times the actual cost of a student when enrolled in a comparable public vocational school. The scope of this subsidy scheme was first extended to cover all private vocational schools in the same year; then, it extended to all the types of private schools at all levels (even including the private pre-schools) in 2014 and hence turned into a generic private school voucher program (*özel okullara yönelik eğitim ve öğretim desteği*).

In parallel with the Ministry of National Education's amendments to the law coded 5580, the Ministry of Economy also enacted a new set of incentives to increase the rate of private investment in the Turkish education sector. These incentives involved income and corporate tax exemption for five years, investment support (for those who could obtain an investment incentive certificate), value-added tax and custom tax exemptions in machinery and equipment procurement, and general value-added tax exemption.

From the policy perspective, The Strategic Plan for the Ministry of National Education (2015-2019) regarded the lower than desired levels of private sector investment into the education sector as a "threat" and adopted the target of 7 percent private school share (in terms of the number of students) in 2019 from a baseline of 3.51 in 2014 (MEB, 2015). The Minister of National Education also stated that the long-term objective was to reach the respective statistics of the European Union and the OECD, which were around 15 percent. The incentives described in the preceding

paragraphs were all in line with achieving this objective. Indeed, the share was raised to 8.3 percent in the 2017-2018 academic year. The actual number of students enrolled in private schools increased from 662 thousand to 1.3 million. The president of the Private Schools Association of Turkey argued that around 300 thousand of this increase during 2014-2018 could be attributed to the Private School Voucher Program (Dal, 2018, p.6 in Altınok, 2019, p.93). As seen in Figure 3.1 below, the share of private schools displayed even more impressive progress in the mentioned period. It was doubled. This quantitative increase was declared to be "sufficient" by the Minister of National Education in 2018. The voucher program instantaneously closed for new applications, while the existing beneficiaries would be benefiting until they graduate. The new objectives for the following three years for private schools were stated as (i) increasing their own quality and flexibility and (ii) leading public schools and helping improve their performance.

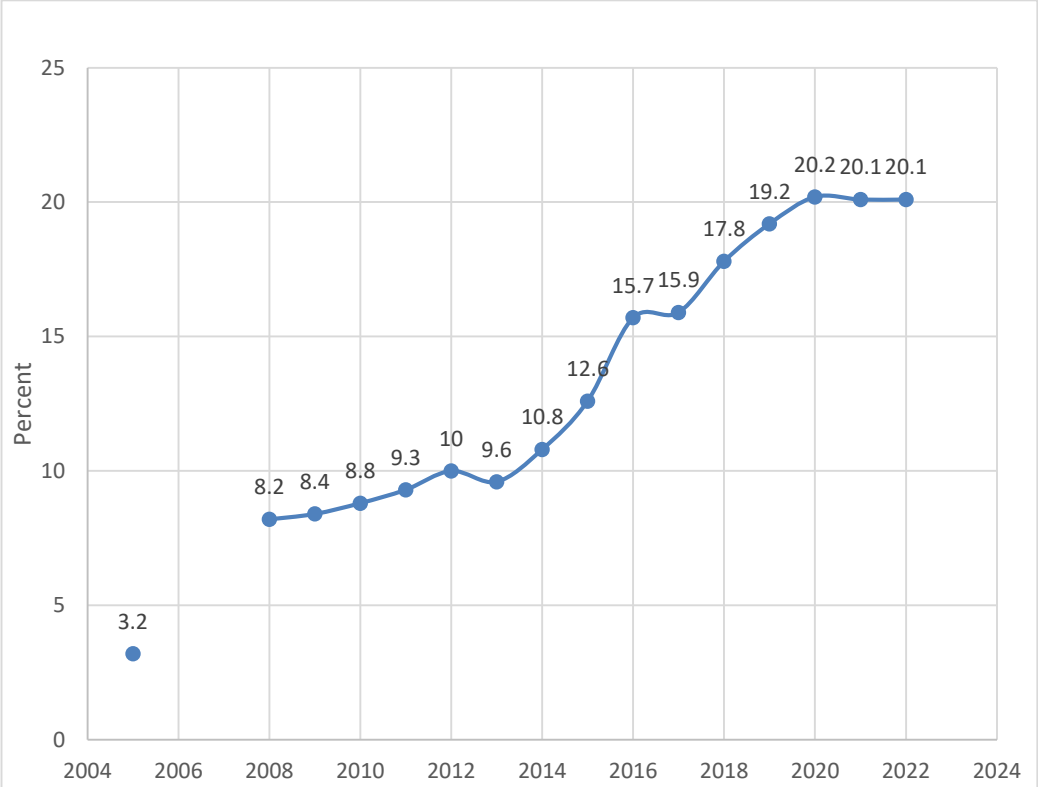

Figure 3.1 Share of private schools in overall number of schools

Another structural and remarkable change between 2012 and 2018 was the dramatic boost in the quantity of lower-secondary Imam Hatip schools. As indicated in the previous subsection, all lower-secondary Imam Hatip schools were closed during the compulsory schooling reform in 1997. As mentioned above, Law coded 6287 in 2012 introduced 4+4+4 (primary school + lower-secondary school + higher secondary school) years of mandatory education. This law also re-opened the lower-secondary Imam Hatip schools.

As observed from Figure 3.2, the period of 2012-2018 witnessed very rapid growth in those schools, both in terms of the number of schools and the number of students. Especially striking was the eight-times increase in terms of the number of students enrolled in these schools. As mentioned earlier, around 680 Imam Hatip lower-secondary schools were present before they were abolished in 1997. This time the number of such schools increased by more than five times the old stock of those schools. The share of Imam Hatip schools within all lower-secondary schools also increased from 6.47 to 18.1 percent in this period.

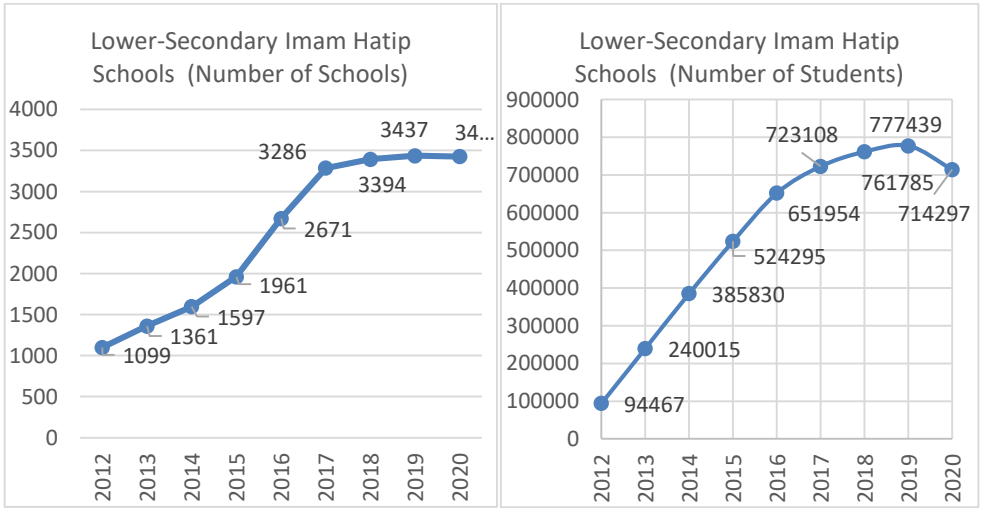

Figure 3.2 Statistics of Imam Hatip Schools for the period of 2012-2020

Another crucial aspect of the period of 2013-2017 is the coup attempt in July 2016 and its impact on the private schooling sector. A large group of private schools were abolished by the Statutory Decree coded 667 right after the unsuccessful coup

attempt. 287 of those schools were lower-secondary schools. They corresponded to 18.2 percent of the total private lower-secondary school population and 22 percent of the total private lower-secondary student population. Buildings, facilities and amenities of those schools were made public and in most of the cases handed over to the lower-secondary Imam Hatip schools in the following years.

On the other hand, the OECD (2020) reports that from 2010 to 2018, Türkiye, together with Peru and Portugal, was among the three countries that saw a major decline in its public share accompanied by a corresponding increase in the private share in education provision. The increase in the share of schools that compete with two or more schools in Türkiye was above 20 percentage points in the mentioned period. This was an increase also observed in only four other countries.

As a third finding, it compares the reading performance, and indicates that public school students on average scored 5 points higher in the unadjusted case and approximately 75 points higher when the scores are adjusted for the socio-economic status of students and schools.[25]

The PISA performance of Turkey is highly relevant for the period of this study. PISA classifies students into proficiency levels, where level 6 is the highest level. A two-tier strategy is implied by how the OECD reports about these proficiency levels: Countries should ensure as many of their students meet at least level 2 proficiency. Countries should also try to increase the number of top performers indicated by at least level 5 proficiency.

From Türkiye, 6,890 students in 186 schools finished the examination, representing 884,971 15-year-old students, which corresponds to 73% of the total population of 15-year-olds)[26]. Among participating Turkish students, 74% (vs. the OECD average

[25] PISA exam results are standardized to have a mean of 500 and standard deviation of 100 across OECD countries.
[26] Approximately 600,000 students participated in PISA 2018, representing about 32 million 15-year olds in the schools of the 79 participating countries.

of 77%[27]) of students attained at least Level 2 proficiency in reading, while 3% were top performers (vs. 9%). In mathematics, the same figures were 63% (vs. 76%) and 5% (vs. %11), while in science, they were 75% (vs. %78) and 2% (7%). Turkey's average performance in PISA 2018, in all three proficiency areas, was not substantially different compared to its performance in 2009 or 2012 and was higher than the performance in 2003, 2006, and an anomalously low outcome in 2015. Socio-economically advantaged students in Türkiye scored 76 points higher on average than the socioeconomically disadvantaged students in reading, which is not significantly different from OECD average of 89.

The ratio of teachers with at least a master's degree is comparable in advantaged and disadvantaged schools. In Türkiye, low- and high-performing students are clustered in the same schools more often than the OECD average (OECD, 2020). In all countries and economies that participated in PISA 2018, girls significantly outperformed boys in reading by 30 score points on average. A gender gap of 25 points in Turkey is in line with OECD trends. In Turkey, girls scored similarly to boys in mathematics, while boys outperformed girls by five score points on average in OECD countries. Turkish girls outperformed boys in science by 7 points (vs. OECD mean of 2).

Before concluding this subsection on the Turkish private education system, it should be noted that private schools in Türkiye have to follow the same curriculum as public schools, especially at the lower-secondary school level. Moreover, the duration of each of the two school terms in each year is identical in both the private and public school sectors. Hence, in contrast to other country contexts (for instance United States), differences in curriculum or in the length of the schooling period cannot play a critical role in the level of achievement differences between the two school sectors. The majority of the potential differences might potentially be due to the differences in the quality of instruction and learning materials/amenities.

[27] All of the values in parentheses will correspond to OECD averages from now on.

3.2 Private School Voucher Program (2014-2015 to 2018-2019)

As the assessment of the Private School Voucher Program is one of the research objectives of this study, I will summarize the institutional background of this support program with more details in this subheading.

The objectives of the voucher program were stated by the Minister of National Education as such: (i) establish a competitive educational system focused on educational quality, (ii) enable all students who want to benefit from the opportunities of private schools, and (iii) to encourage and trigger educational investments from the private sector.

The assignment of the subsidy involved a set of steps. In the first step, the quotas for all provinces were determined according to the development level of the provinces. The provinces were grouped into six groups in terms of their socio-economic development level, where the level one group had the highest development level and the level six group had the lowest. The weights for each group were as follows, respectively, from level one to level six: {0.95, 0.95, 1.00, 1.00, 1.20, 1.30}. Hence, less developed provinces received more weight. Ministry of Industry and Technology classifies provinces into six development levels via principal components analysis of comprehensive, multi-sectoral data (Sanayi ve Teknoloji Bakanlığı, 2019). Based on the weights and the number of students in each province in each year, the Ministry of National Education declared in each year the quotas for each province before the start of the applications by the schools and the students (see MoNE, 2014, p.4-6).

For lower-secondary schools, the total quota in each year was 75.000 for the first two years and 15.000 for the last three years. In each school year, 40 percent of the subsidy should be assigned to the freshman-year students (which was the 5th grade for lower-secondary schools) and 20 percent to each of the grades 6, 7 and 8. Hence, students could benefit from the subsidy starting from any grade. When a subsidy was assigned to a student at any grade, it was provided until the end of the lower-secondary school period, provided they did not fail and repeat during any grade.

The timeline (for instance, for of 2014-2015 academic year) started with simultaneous applications of schools and students between 8-29 of August. The approvals from district branches of the Ministry of National Education were between 19-29 August. On 1st of September, the list of schools and students to be supported schools are announced. On 4th of September, the quotas of each private school (among the ones declared as eligible). Eligible students submit their list of preferred schools (up to 15 schools). On 10th of September, the allocations are announced. Between 1-18 September the registrations are carried out. On 19th of September, the backup replacements are announced and between the 22-30 September, the backup registrations are finalized.

The criteria for schools' eligibility and quotas involved the socio-economic development status of the province (changes from 1 to 6, where 1 is most developed; determined by Ministry of Industry and Technology), transformation (from private tutoring institutions into private schools) program appraisal point, mean student GPA in the school, the number of children of martyr/war-wounded parents, student/classroom ratio of the district, the quantity of private schools and number of students per teacher in the district, number of teachers in the schools, the overall student quota of the school, and the group of the private school (A, B, C or D), where the exact group is determined according to the amenities and physical qualities of the private school (see Annex-11 in MEB, 2014).

The criteria for students involved academic success in the previous year (85-100, 70-84, 55-69), being a competitor in Olympiads or other international competitions, being a competitor in national-level competitions, being a competitor in provincial competitions, monthly family income, number of siblings in the education system, life status of parents (dead/alive), marital status of parents, disciplinary penalties, and being a child of martyr/war-wounded or being a child under protection school (see Annex-11 in MEB, 2014). In the first implementation year, the contribution of academic achievement was 50 percent, but in the second implementation year, it was

reduced to 26 percent. The weight of family income was, on the other hand, 10 percent in the initial implementation year, which was increased to 34 percent. Hence, in the second year, the socioeconomic background became a higher priority over achievement. In the following years, previous academic achievement was discarded from the criteria set. In the first year, the documentation and proof for the accuracy of family income were not so in-depth, but in the following years these requirements (and the sanctions) became gradually more sophisticated as the weight of this criterion increased. The Ministry of National Education collaborated with the Ministry of Finance to enforce the criteria for income levels of families.

In the first year of the subsidy program, only students who were enrolled in public schools were eligible to apply for the subsidy. However, starting from the second year, students who were already enrolled in private schools were also allowed to apply for the subsidy. The main reason for this decision was that the mentioned restriction was non-enforceable. A parent whose student was already enrolled in a private school could first transfer the student to the public school assigned by default to this student and then apply for the subsidy. No parent could be denied from doing this.

The yearly monetary value of the subsidy was {3500, 3750, 4000, 4280, 4610} Turkish Liras (TL) for the five-year period between the academic years of 2014-15 and 2018-19. These values corresponded to {3.92, 3.75, 3.23, 3.04, 2.88} times the net minimum wage effective at the beginning of each respective academic year. These amounts were especially meaningful for private schools that have the 12,000-15,000 TL range. The mentioned range corresponded to the majority of the private schools located in mid-tier provinces, and Altınok (2019) concludes that the subsidy program impacted schools in those provinces the most.

It was prominent that the period of transformation of private tutoring institutions (*özel dershaneler*) into private schools perfectly overlapped with the duration of the voucher program. Indeed, one of the rationales of the subsidy program was to support this transformation process, especially at the high school level. When the

transformation process of private tutoring institutions was announced, the ending time was announced as the end of the 2018-19 schooling year. However, the ending time of the subsidy program was not declared at the beginning of it. The Minister of National Education stated in the budget discussions in the parliament speech that the subsidy program was successful in terms of reducing the number of students in public school classrooms and alleviating the load of double-session school days in the public schools. He also argued that a test-based view of private school performance was undesirable. Finally, he expressed that the new focus of investments would be on improving the conditions of disadvantaged public schools.

3.3 The TEOG Examination (2013-2014 to 2016-2017)

This subsection will provide the details about TEOG examination and its special relevance for exploring private school effects. In the 2013-2014 education year, the examination system held at the end of lower secondary school was changed. The new exam system was titled the "Transition from Basic Education to Secondary Education" (TEOG) examination.

The TEOG examination was a standardized cognitive test. It was distributed into two semesters in the last (fourth) year of lower-secondary school education. In each semester, students took 6 tests in 6 different learning areas over two days, where three tests were taken each day. Six test items included mathematics, science, literacy, recent history, religious knowledge, and English proficiency. Each test involved 20 questions and 40 minutes per test were allocated. Half-hour breaks were provided for students so that they could rest and replenish. There was a compensation exam for students who could not enter the exams for valid reasons, which was the first in the history of centralized examinations in the country[28]. This reduced the stress associated with exam anxiety, which is shown to be an important factor that also has different impacts based on a student's gender. On average, girls are impacted more than boys in a negative way when they are subjected to high-stakes tests (OECD, 2017; Niederle

[28] The LGS examination, which replaced TEOG in 2018 does not have a compensation examination.

& Vesterlund, 2010; Saygın, 2018). Hence, the TEOG exam reduces the potential of bias in achievement estimation due to the difference between boys and girls regarding high-stakes exam stress.

The overall score for achievement was calculated by a weighted average of the TEOG test scores (weight = 0.7) and the grade point average (GPA) obtained from regular classes in the school over the last three years of four-years long lower secondary school education (weight = 0.3). The TEOG questions depended on curriculum of the last academic year of secondary school (grade 8). The incorporation of school scores was regarded as strengthening the school-student relationship.

TEOG test was decisive as it was centrally allocating *all* of the students who will choose public schools. The one-point difference might mean being enrolled in a high-secondary school far away from home. When a student cannot obtain the grade needed for the elite public school nearby, then her/his parents might have to look for a private school alternative, which would bear a significant cost difference compared to an elite public school. This qualifies TEOG as a high-stakes exam. On the other hand, private high schools also require TEOG exam results for their selection process for higher-secondary school enrollments. Hence, entering into a private high school with a higher reputation also required a high score on TEOG exam.

TEOG exam was also regarded as a metric where schools are ranked against each other. The Ministry of National Education determined the success of each lower-secondary school based on the school-average test scores obtained in TEOG (Sen, Yildirim & Karacabey, 2020, p.81). Upper secondary schools were ranked according to the minimum TEOG score (*TEOG taban puanı*) among their new students[29]. So this was also a reason why both public and private schools took it seriously and pushed their enrolled students to perform well in the examination. LGS exam replaced TEOG in 2018, and the policy of "central allocation of students into high schools based on a centralized exam" was abandoned. Instead, about 10 percent of students

[29] An example: ANKARA TEOG 2016-2017 LİSE TABAN PUANLARI (meb.gov.tr)

would be assigned to elite schools; the remaining students would be assigned to the closest school to their address on the e-School system of the Ministry of National Education. If the students did not perform well in LGS and could not be assigned to an elite upper-secondary public school, then parents would have two more options over enrolling their child into the school address-based assigned. They can either send her to a private school (LGS score is optional) or can enroll in an Imam Hatip school. Otherwise, they need to change (for real or for fake) their address.

TEOG was also seen as a "measuring" exam (i.e., measuring achievement), while LGS replacing it was seen as a "selecting" exam. All students had to enter TEOG, while participation was not mandatory for LGS. Only 110,000 out of 1.1 million students could enroll in elite or selective schools during the LGS period, so it would be nonsensical to force every student to enter[30].

It should be noted that TEOG exam was the first ever standardized, objective, nationally held exam for those students in that period. Before the switch to the mandatory eight-year education period in 1997, primary school students were taking a nationwide held, high-stakes, standardized exam titled *Anadolu Liseleri Sınavı* after graduating from primary school in grade five. By the inception of eight-year mandatory education in 1997, this exam was abolished[31].

By design, the TEOG exam measured the same achievements (kazanımlar) that must be learned during courses required by the curriculum. TEOG was also replacing one of the in-class exams for both semesters in the final year of lower-secondary school. Hence, they were testing the same cognitive content.

The highest level cognitive skills are analysis, assessment and creativeness. Questions assessing these aspects are less in number in TEOG compared to PISA. TEOG is more

[30] For more details, Kuzu, Kuzu, & Gelbal (2019) compares TEOG and LGS exams over student, parent, and teacher opinions.

[31] From the methodological perspective, this eliminates the opportunity to utilize value-added approach (controlling for past achievement) in econometric estimation models.

involved in testing mid-tier cognitive skills (Çelik, Kul & Çalık Uzun, 2018; Tunç & Baydar, 2022). LGS, on the other hand, is more similar to PISA. It even has some open-ended questions and tests higher-level cognitive skills compared to TEOG. This is in line with the more international emphasis on PISA exam. Exceptionally strong "national PISA score averages and economic growth linkage", suggested by Hanushek & Woessmann (2015, 2019) have been strongly embraced by the OECD, the World Bank, and the European Commission. This pushes countries to adopt "we should align with PISA" policies. It should have the same effect for Türkiye, too.

CHAPTER 4

THE DATA

4.1 Introduction

The unique relevance of TEOG examination scores for studying private school effects in Turkiye is described in the previous chapter. The data recruited for this study covers the whole student population in grade 8 (i.e., in the graduating year of lower-secondary school) of each successive student cohort from the 2013-2014 academic year to the 2016-2017 academic year, which corresponds to the TEOG exam period.

Burgess (2016) argues that the survey data, especially the data from Labor Force Surveys, were the mainstay of the economics of education. However, important new learning came from the utilization of other types of data that include administrative or register data. He underlines the importance of recruitment and use of such data and calls for research funding to support such initiatives. In line with this call, this chapter aims to contribute to data availability of school and student administrative data on a very large scale for the first time.

In this chapter, I will provide the descriptive aspects of the variables provided by the data set, including their missingness, and will relate them with the outputs and inputs of the education production function. Generally, the data for the study is provided in one or a few summary tables within one of the subheadings of the methodology chapter in dissertations. One reason for this is that the data utilized is already provided and described elsewhere, and excessive re-description of data is aimed to be avoided. Or the data aspects might be secondary, and the main focus might be on the analysis technique. In that case, the important data limitations like missingness might be lightly addressed. However, since this study introduces the population level data and

the main objective of the study is applied analysis, I opt to present the data in detail in a separate chapter right before the methodology section.

I will first present the variables of interest, which are the outcome variables (student TEOG score, student GPA of the last three years) and the input variables whose parameters are of special interest (private school status, subsidy status, gender). Then I will describe the other variables at student and school levels that correspond to the remaining inputs into the education production function or into the decision rules of households and schools.

4.2 The Variables of Interest

To start with, Table 4.1 presents the total number of lower secondary-schools as well as their distribution into the public and private sectors among the lower secondary school population in Türkiye. The year 2014 corresponds to the end of the 2013-2014 academic year, and so on.

The TEOG period covers four successive academic years. Remarkably, in just a few years, the number of private schools increased by 63.4 percent and the proportion of private schools in the total number of schools increased 59.4 percent. There was a fall in the last year which is mainly due to the coup attempt between the third year and the fourth year of implementation, which led to the abolition of a high number of lower-secondary private schools.

The sharp fall from 2014 to 2015 in public schools, on the other hand, is mainly associated with the termination of an eight-year-long primary schooling, and switching to 4+4+4 system. With the separation of primary and lower-secondary schools like in the past, the last cohort of primary schools graduated in 2014, and primary schools were separated from lower-secondary schooling. It is seen that the number of public schools regressed to its previous averages gradually in the following years.

Table 4-1 Number of lower-secondary schools and ratio of private schools

Year	*# Total Sch.*	*# Public Sch.*	*# Private Sch.*	*% Private Sch.*
2014	15,803	14,907	896	5.67
2015	15,050	14,025	1,025	6.81
2016	16,188	14,724	1,464	9.04
2017	16,280	14,947	1,333	8.19

In addition to number of schools, Table 4.2 presents the number of students in their graduation year from the lower secondary school in each cohort. It is once more remarkable that in just a few years, the number of students graduating from private lower-secondary schools increased 63.2 percent and the proportion of private school students in the total number of schools increased by 78.9 percent.

Table 4-2 Number of students at grade-8 (last academic year) in each cohort

Year	*# Total Stu.*	*# Public Stu.*	*# Private Stu.*	*% Private Stu.*
2014	1,272,923	1,227,710	45,213	3.55
2015	1,275,827	1,224,003	51,824	4.06
2016	1,162,796	1,088,995	73,801	6.35
2017	1,182,937	1,115,991	66,946	5.66

Table 4.3 summarizes the empirical content of the dataset in terms of the two outcome variables, which are the TEOG Exam Score and the Grade Point Averages (GPA) at school for the last three years of the lower-secondary schooling period. In the raw data, there are no students declared as having missing values from the TEOG exam. However, some students have obtained score of a zero points. TEOG exam had a compensation exam one week later than the first exam, so I assume the students who obtain a zero score from both chances should not have taken the TEOG exam ever. Also, it should be noted that four wrong answers did not eliminate a wrong answer (which is the case in the most standardized tests held nationally in Türkiye) in the TEOG exam, which makes obtaining a zero score implausible on this exam if the student had ever attended the exam. Hence, I created a new TEOG Score variable, which involves a missing value if the score in the raw data is zero.

Table 4-3 Missing observations in outcome variables

	TEOG=0 (total)		***TEOG=0 (public)***		***TEOG=0 (private)***	
Year	***# Missing***	***% Missing***	***# Missing***	***% Missing***	***# Missing***	***% Missing***
2014	19,504	1.532	19,475	1.586	29	0.064
2015	22,796	1.787	22,760	1.859	36	0.069
2016	20,396	1.754	20,359	1.870	37	0.050
2017	19,292	1.631	19,240	1.724	52	0.078
	GPA (total)		***GPA (public)***		***GPA (private)***	
Year	***# Missing***	***% Missing***	***# Missing***	***% Missing***	***# Missing***	***% Missing***
2014	256	0.020	250	0.020	6	0.013
2015	301	0.024	279	0.023	22	0.042
2016	148	0.013	146	0.013	2	0.003
2017	201	0.017	199	0.018	2	0.003

As seen in Table 4.3 above, missing TEOG score observations belonging to private school students are negligible in number, which indicates that almost all private school students participated in the exam. This is highly important for studying private school effects since the nationwide standardized exams held both before and after TEOG period do not have this specialty. The participation of private school students in the standardized exam was not mandatory in the preceding and succeeding periods. Leaving the participation decision to the students/parents highly compromises the potential comparison between public and private schools and the students enrolled in them.

The proportion of students who did not take the TEOG exam is around 1.7 percent and this ratio has remained stable over the years. Although this ratio of public schools is not totally negligible, as in the case of private school students, it is not inhibitive either. On the other hand, since the primary focus of this study is on private school effects, the non-missingness in private school students is more important. Quite a large pool of public school students can still provide counterfactual information for private school students. The missingness in the GPA outcome measure is negligible for both school sectors. Overall, the data is rich ino reflecting the output factors to be employed in the education production function. Moreover, the data provides a unique opportunity to incorporate private sector students into the data analysis.

In Table 4.4, the overall summary of TEOG test scores is provided. The TEOG test has a range of 0-350 points. The upper section of the table indicates the overall public and private means as well as the gaps between both school sectors over the four years. The gaps are pretty sizeable. In the middle part of the table, I present standardized test scores, where the standardization is done separately each year. The bottom part reflects the standardized test scores when all data from four cohorts are combined.

The unadjusted difference of the overall means of public and private schools in the combined data is 1.31 standard deviations (sd) to the advantage of private schools. According to Cohen's (1969, p.25-26) conventions for social sciences, an effect size of 0.2 sd is a small, 0.5 sd is a medium, and 0.8 is a large magnitude. Hence, the unadjusted difference is substantively large. We will see how this difference will change after adjustments are made by inferential data analysis under the conceptual framework provided by the education production function approach.

Table 4-4 TEOG test scores – overall summary

	Test Scores (Absolute values within range: 0-350)			
Year	***Overall mean***	***Public mean***	***Private mean***	***Gap***
2014	191.55	187.89	289.57	101.68
2015	192.18	187.84	292.64	104.80
2016	198.73	192.72	285.04	92.32
2017	216.99	212.00	298.69	86.69
	Test Scores (Each cohort ind. standardized to mean 0 and sd. 1)			
	Overall mean	***Public mean***	***Private mean***	***Gap***
2014	0.00	-0.05	1.42	1.47
2015	0.00	-0.06	1.37	1.43
2016	0.00	-0.08	1.18	1.26
2017	0.00	-0.07	1.08	1.14
	Test Scores (Combined cohorts standardized to mean 0 and sd. 1)			
	Overall mean	***Public mean***	***Private mean***	***Gap***
2014	-0.11	-0.16	1.22	1.38
2015	-0.10	-0.16	1.26	1.42
2016	-0.01	-0.09	1.16	1.25
2017	0.24	0.17	1.35	1.18

Two issues are noticeable in Table 4.4. The first is that the mean test score values in the first two years are close to each other, the third cohort differs slightly, and the last cohort score average seems to differ remarkably more. Second is that the gap between

private schools and public schools has reduced over the years. To investigate the first one, I present cohort-to-cohort comparisons of the whole score distribution for each cohort combination in Figure 4.1 below. The red-colored region is the intersectional region of the two distributions. The comparisons of 2014-2015, 2014-2016, and 2015-2016 all exhibit a large amount of common region and appear to be replications of a standardized test that remain stable over the three years.

The last cohort's distribution has a bump at the higher levels of the score distribution. Figure 4.1 thus also indicates the need for incorporating time (i.e., cohort) fixed effects into inferential analysis when four cohorts of data are combined, which is suggested by Hanushek (1996) to deal with such changes in scoring structures of standardized tests.

It is expected to some degree that as the teachers get more experienced with a specific standardized exam, their teaching-to-test skills improve. Also, better auxiliary study items in terms of supplementary books or internet resources have become available as more sample exams are in place. This might be the reason for the bumps in TEOG scores in the later years.

Another reason for the bump in TEOG score distribution might be the lack of more challenging questions that otherwise would ensure better differentiation among the upper quantiles of the score distribution. We should not forget that the 2016-2017 academic year coincided with a turbulent period due to law enforcement after the coup attempt.

The other outcome variable is the Grade Point Average (GPA) of each student, which reflects the mean of grades a student obtains over the last three years of the lower-secondary school period. As seen from Table 4.5 below, the standardized means and the standardized gaps of GPA closely resemble the ones for TEOG scores. As seen from the ridgeline plots in Figure 4.2, the overall distributions from year to year and the bumps in the last cohort are comparable between TEOG and GPA distributions.

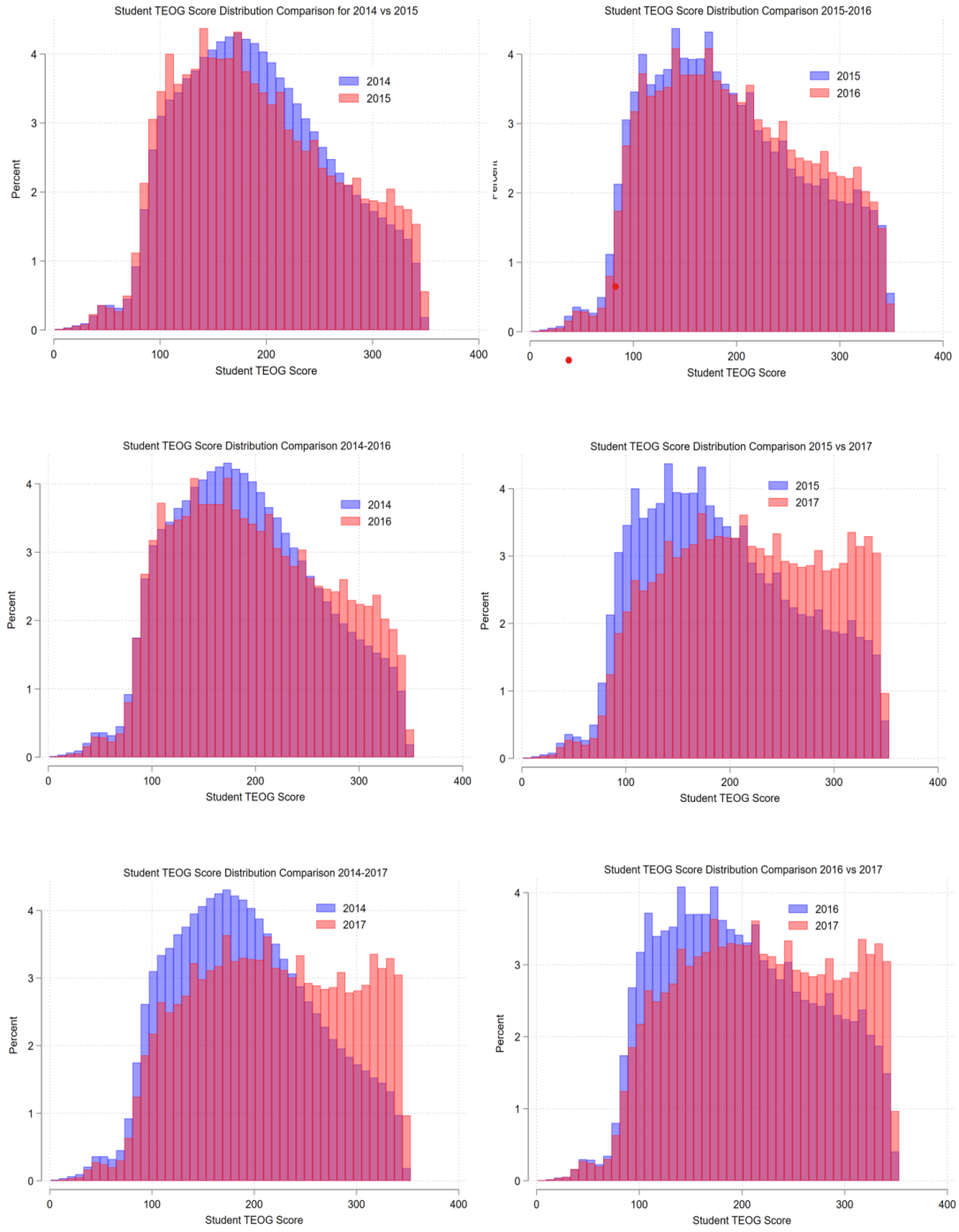

Figure 4.1 TEOG Test Score Distribution Year-to-Year Comparisons

The bump in the GPA distribution is interesting. GPA grades can be subjective, especially across schools. In other words, they might not be objectively comparable among students from two different schools, but the comparison between two students

within a paricular school is more plausible since, at the lower-secondary school level, either the same teacher from each discipline (math, literacy, science, etc.) teaches to all classes or if there is more than one teacher, then they should be similar in teaching effectiveness as teachers are largely sorted into schools according to their effectiveness in Türkiye. Their teaching effectiveness is also expected to converge with each other due to close collaboration and pressure from the parents or the principal of the school.

Table 4-5 Student GPAs- overall summary

	GPA (Absolute values within range: 0-100)			
Year	***Overall mean***	***Public mean***	***Private mean***	***Gap***
2014	68.81	68.02	90.26	22.24
2015	70.40	69.53	91.15	21.62
2016	72.78	71.54	91.00	19.46
2017	75.83	74.79	93.22	18.43
	GPA (Each cohort indiv. standardized to mean 0 and sd. 1)			
	Overall mean	***Public mean***	***Private mean***	***Gap***
2014	0.00	-0.05	1.38	1.43
2015	0.00	-0.06	1.32	1.38
2016	0.00	-0.08	1.16	1.24
2017	0.00	-0.07	1.13	1.20
	GPA (Combined cohorts standardized to mean 0 and sd. 1)			
	Overall mean	***Public mean***	***Private mean***	***Gap***
2014	-0.19	-0.24	1.16	1.40
2015	-0.09	-0.15	1.22	1.37
2016	0.06	-0.02	1.21	1.23
2017	0.25	0.18	1.35	1.17

There are reasons, on the other hand, for the case that students from different schools are not comparable in terms of GPA. A weighted average of the TEOG score and the GPA was taken into attention in the selection of upper-secondary schools of choice. These constitute two groups of schools: the group of private high schools and the group of selective (or elite) public upper-secondary schools in Turkey during the TEOG period. Hence, the bump might also reflect an intended extra boost to scores.

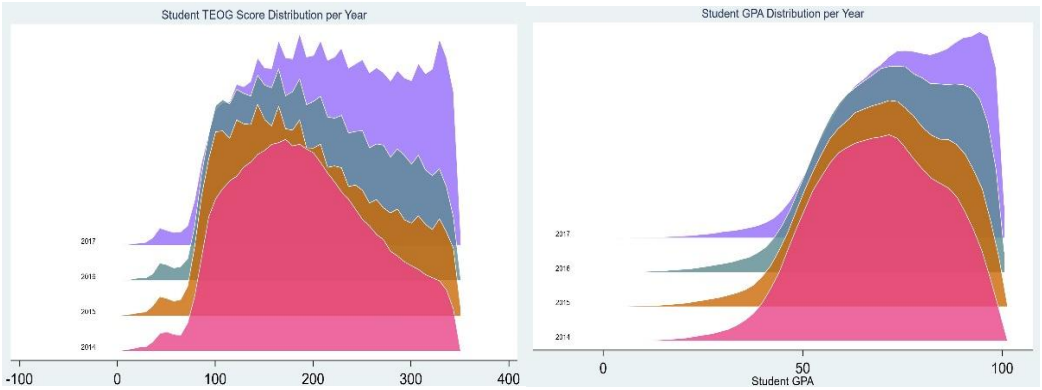

Figure 4.2 Ridgeline Plots for TEOG Score and GPA distributions (2014 to 2017)

Note that when the standardized test system switched to TEOG in 2013-2014, only a single year of boost was possible as only one year was left before the exam. In 2015-2016, two years of boost was possible. In 2015-2016 and in 2016-2017, a full three years could be boosted. This might at least partially explain the increasing bump in the later years. Both private and public school teachers and administrators might opt for such a boost.

What might it imply for the private-public school comparison? Even if that is the case, it is seen from the raw data summaries in Table 4.5 and in Figure 4.3 (left column) that the mean rate of increase (i.e., boost) in GPA was more pronounced in public schools compared to private schools.

After comparing the TEOG score and GPA distributions in different years, let's now compare the distributions of such distributions between public and private schools in each year. Figure 4.3 represents the public-private comparisons of students' TEOG scores on the left and students' GPA values in the right column. Both columns resemble each other to a great degree. Only the direction of skewness of public schools does differ between them. It is remarkable that public and private distributions are vastly different and their region of intersection region limited for both the TEOG-based and the GPA-based comparisons.

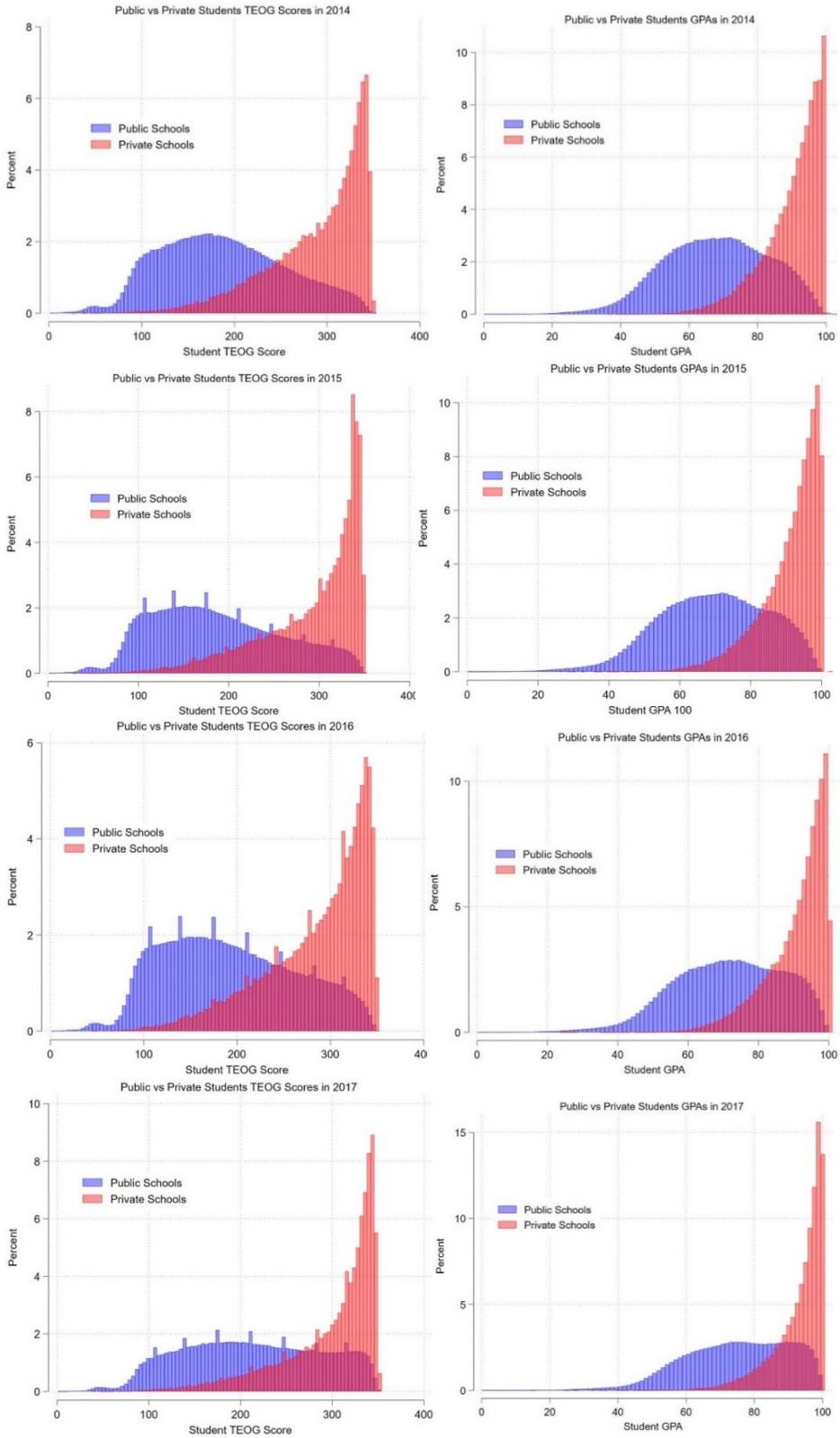

Figure 4.3 Public vs Private Comparisons of TEOG Scores & GPA for 2014 to 2017

One concern related to the TEOG score and GPA distributions is the potential existence of ceiling effects on test or GPA scores. Heyns and Hilton (1982, p.96) state that the ceiling effect shall be assessed by calculating the difference between the maximum score possible and the observed mean, divided by the standard deviation. Ceiling effects might be a problem if this ratio is less than one. Table 4.6 below represents these ratios for each year and for each school sector. There are no ceiling effects on public schools' outcomes. Private school students' TEOG scores and GPAs do not show a ceiling effect since the ceiling effect measure is at the borderline. Only for the GPA of 2017, there is a slight ceiling effect, which I acknowledge as a modest limitation.

Table 4-6 Assesment of potential ceiling effects in outcome measures

Score Type: / Year:	**2014**	**2015**	**2016**	**2017**
Public Schools' TEOG Scores	2.43	2.28	2.22	1.86
Private Schools' TEOG Scores	1.18	1.05	1.19	1.01
Public Schools' GPAs	2.11	1.99	1.86	1.66
Private Schools' GPAs	1.15	1.07	1.08	0.94

After observing that the standardized summaries of TEOG score and GPA resemble each other substantially based on the comparison of Table 4.4 with Table 4.5, let's examine how the whole standardized distributions compare in Figure 4.4 below. The common regions in both the public schools and the private schools constitute a large part of each distribution. However, there are also non-overlapping parts. As discussed in the literature review section, this is an expected observation. Both the standardized TEOG test score and the grade point averages are primarily measures of cognitive skills and domain knowledge. However, those measures also reflect non-cognitive skills to a certain degree. In that line, it is likely that teachers incorporate non-cognitive skills (behavior, attendance, respect, perseverance, etc.) into the final grade. Hence, the empirical content of GPA is expected to reflect the contribution (or detraction) by non-cognitive skills more than the standardized TEOG test score. As discussed earlier, a more appropriate comparison between TEOG score and GPA

distribution can be made within each school, and Figure 4.4 might also reflect differences among schools in the same school sector in terms of their inclination to give some boost to the GPAs of their students. Finally, we also notice that neither TEOG score nor GPA distribution strictly dominates each other throughout the whole range of the standardized x-axis.

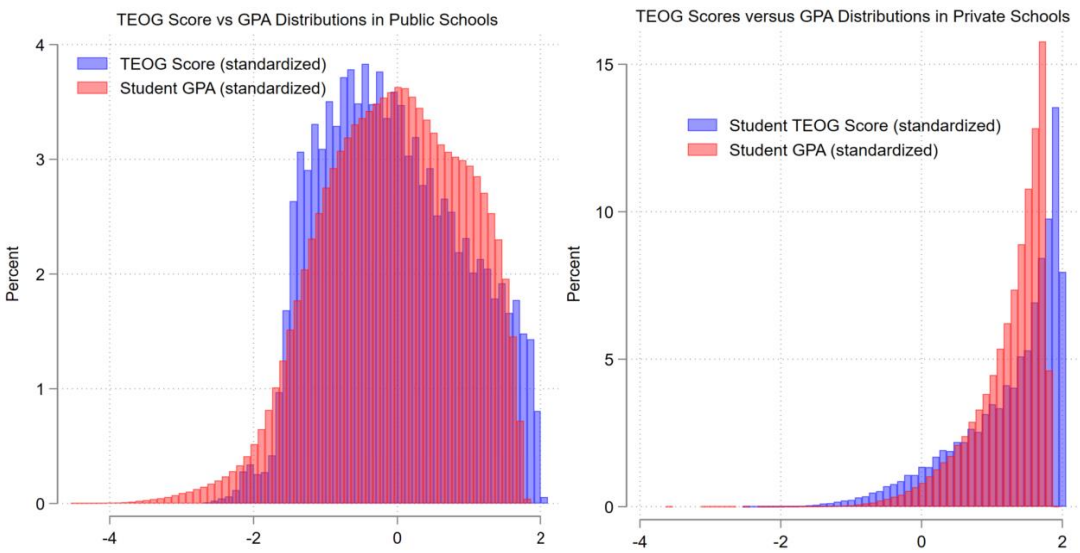

Figure 4.4 TEOG & GPA Distribution Comparison by School Sector

Now let's move to the other important variable of interest, which is the student gender. As indicated in Table 4.7, the first thing to notice is the ratio of the number of girls to the number of boys (girls to boys ratio). Although it should be very close to one when all children are under education (in Türkiye, 49.8 of the population was female and 50.2 percent were male over the TEOG period. Hence female to male ratio was 0.992), there is a substantial loss of girl attendance in lower-secondary schooling. The overall girls-to-boys ratio during the TEOG period in the whole student population was 0.929, in contrast with the expected ratio of 0.992. Moreover, this ratio was even lower in private schools:

Table 4-7 Numbers of boys and girls in public and private schools

	Public Schools			*Private Schools*		
Year	***Boys***	***Girls***	***Girls/Boys***	***Boys***	***Girls***	***Girls/Boys***
2014	634,649	593,061	0.934	24,194	21,019	0.869
2015	633,558	590,445	0.932	27,985	23,839	0.852
2016	566,056	523,940	0.926	39,697	34,104	0.859
2017	575,325	540,666	0.940	36,341	30,605	0.842

Table 4.7 also indicates the need for investigating the potential impact of the differential selection of girls (compared to boys) on the observed lower-secondary school population as well as their differential selection into the population of private school students. Figure 4.5 below presents the TEOG score and GPA distributions of the whole population of boys and girls over the whole TEOG period. The mean (and standard deviation - sd) for TEOG scores of boys is 188.33 (74.25), and for girls it is 211.72 (71.01). The mean (sd) for the GPAs of boys is 68.59 (16.16), and for girls it is 75.38 (14.65). Girls are observed to achieve considerably more, both in terms of standardized test scores and GPAs. In the unadjusted case, both distributions for girls strictly dominate boys over the whole range of the distribution.

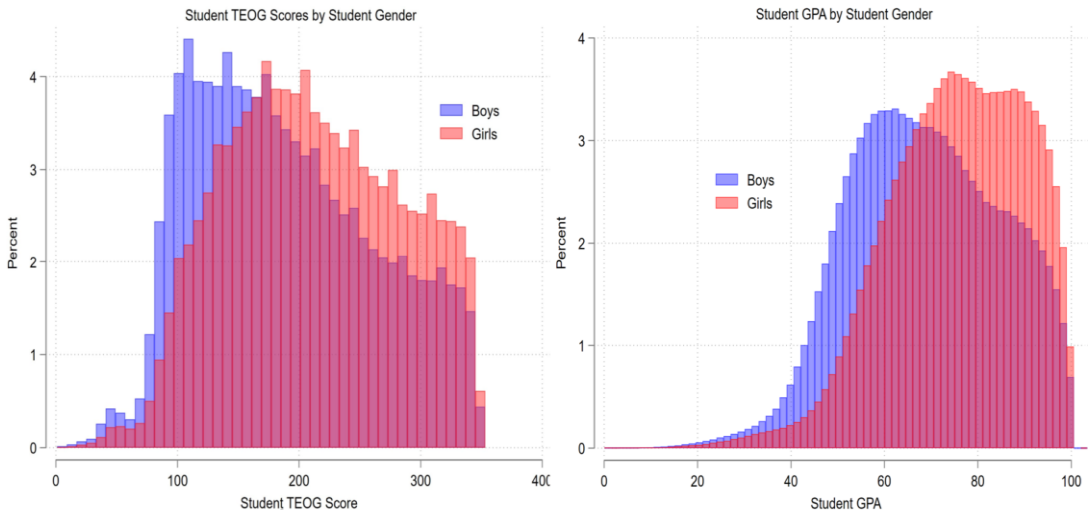

Figure 4.5.TEOG & GPA Distribution Comparison by Student Gender

The mean (and S.D.) for standardized TEOG scores of boys is -0.15 (1.01), and for girls it is 0.16 (0.96). The mean (and S.D.) for GPAs of boys is -0.21 (1.02), and for girls it is 0.22 (0.96). This translates into an unadjusted 0.32 sd gender achievement gap in terms of TEOG scores and a 0.44 sd gender achievement gap in GPA to the advantage of girls.

Table 4.8 below represents how the TEOG scores of girls and boys differentiate across public and private schools for each cohort. We observe that the average gender achievement gap (for TEOG scores) within private schools is less than half of the same gap in public schools. This observation remains stable in each of the cohorts.

Table 4-8 Average TEOG scores of boys and girls in public and private schools

	Public Schools			***Private Schools***		
Year	***Boys***	***Girls***	***Gap***	***Boys***	***Girls***	***Gap***
2014	176.57	199.94	23.37	285.24	294.54	9.30
2015	176.59	199.98	23.39	287.84	298.28	10.44
2016	180.85	205.71	24.86	279.44	291.57	12.13
2017	199.30	225.58	26.28	293.53	304.79	11.26
	***Public Schools** (standardized)*			***Private Schools** (standardized)*		
Year	***Boys***	***Girls***	***Gap***	***Boys***	***Girls***	***Gap***
2014	-0.311	0.005	0.316	1.163	1.290	0.127
2015	-0.312	0.006	0.318	1.199	1.340	0.141
2016	-0.254	0.083	0.337	1.085	1.249	0.164
2017	-0.004	0.353	0.357	1.279	1.429	0.150

In the Figure 4.6 below, we observe, side by side, the how TEOG score distributions of boys and girls compare in both public and private schools (for combined data of four cohorts). In public schools, girls' TEOG scores distribution dominates boys' distribution all over the range (i.e., it is always concentrated more in the higher quantiles of the score distribution). However, in private schools, girls' distribution concentrates more on the higher parts of the range of x-axis. However, for the highest range of the TEOG score distributions, we observe that boys' distribution takes over.

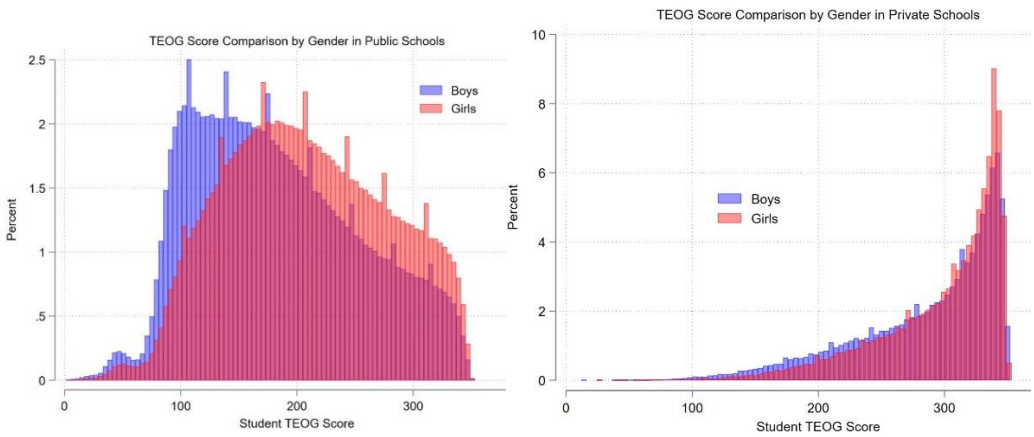

Figure 4.6 TEOG Score Comparison by Gender in Each School Sector

On the other hand, Table 4.9 represents how the GPAs of girls and boys differentiate between public and private schools for each cohort. We observe that the average gender achievement gap (for GPAs) in private schools is less than half of the gap in public schools in each of the cohorts. This observation once more remains stable in each of the cohorts.

Table 4-9 Grade Point Average (GPA) means by school type and gender

	Public Schools			***Private Schools***		
Year	***Boys***	***Girls***	***Gap***	***Boys***	***Girls***	***Gap***
2014	64.47	71.82	7.35	89.24	91.44	2.20
2015	65.99	73.32	7.33	90.09	92.40	2.31
2016	68.16	74.20	6.04	89.79	92.40	2.61
2017	71.57	78.21	6.64	92.26	94.37	2.11
	Public Schools (standardized)			***Private Schools (standardized)***		
Year	***Boys***	***Girls***	***Gap***	***Boys***	***Girls***	***Gap***
2014	-0.468	-0.003	0.465	1.099	1.234	0.135
2015	-0.371	0.092	0.463	1.152	1.298	0.146
2016	-0.234	0.211	0.445	1.133	1.299	0.166
2017	-0.019	0.401	0.420	1.289	1.423	0.134

Figure 4.7 below represents this time how GPA score distributions of boys and girls compare in both public and private schools (for combined data of four cohorts). In both the public schools and the private schools, girls' GPA distribution dominates boys' distribution all over the range.

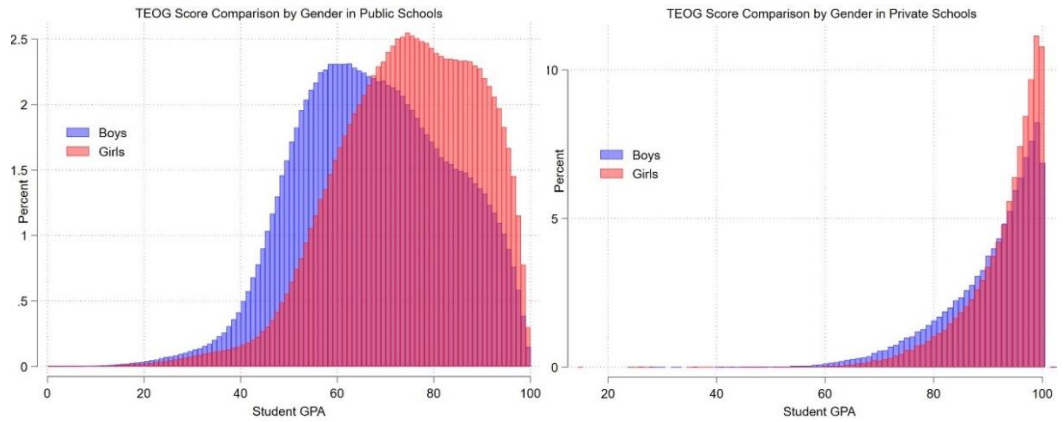

Figure 4.7 GPA Comparison by Gender in Each School Sector

The final variable whose parameter estimates are of interest is the private school voucher (subsidy) assignment indicator for each student. This variable takes the value of one only in private schools, and it is always zero for public school students, as the voucher assignment is only observable for students graduating from private schools.

The data of subsidized students could only be made available for the last two cohorts of the study period. Figure 4.8 represents comparisons of subsidized and unsubsidized students in terms of their TEOG score and GPA distributions for the combination of the 2015-2016 and 2016-2017 cohorts.

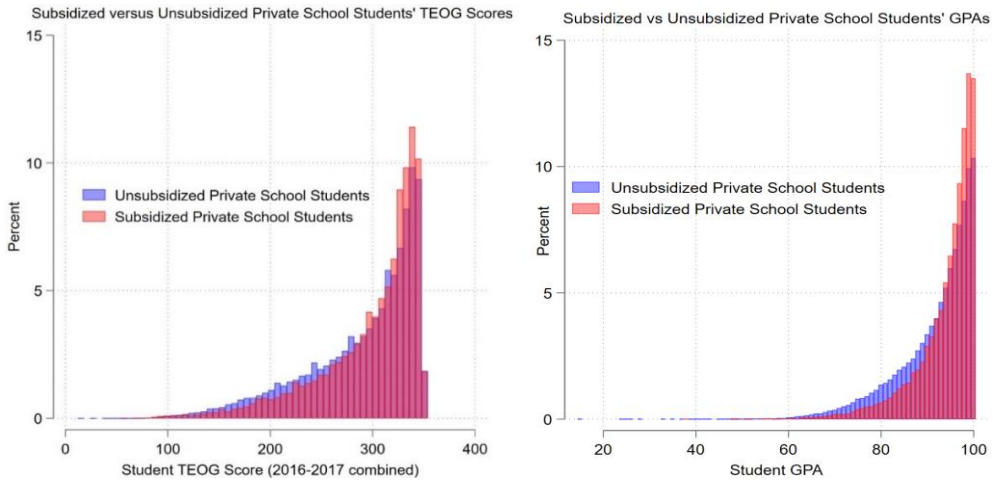

Figure 4.8 TEOG and GPA Comparisons for Subsidized Students

Table 4.10 presents summary statistics for subsidized students. Numbers, TEOG score averages and GPA averages of subsidized students are provided for the total subsidized student population as well as separately for boys and girls. Averages of the student TEOG Scores and GPAs are observed to be higher for the subsidized group compared to the whole private school averages. The gender achievement gap, on the other hand, is smaller both for TEOG scores and GPAs.

Table 4-10 Numbers and score-averages of subsidized students

Number of subsidized students in private schools				
Year	***Total***	***Boys***	***Girls***	***Girls/Boys***
2016	11,797	6,453	5,344	0.828
2017	11,942	6,554	5,388	0.822
TEOG score averages of subsidized students (0-350)				
Year	***Total***	***Boys***	***Girls***	***Gender Gap***
2016	292.17	287.98	297.22	9.24
2017	304.12	300.15	308.95	8.80
Standardized TEOG score averages of subsidized students				
Year	***Total***	***Boys***	***Girls***	***Gender Gap***
2016	1.257	1.200	1.326	0.126
2017	1.420	1.366	1.485	0.119
GPA averages of subsidized students (0-100)				
Year	***Total***	***Boys***	***Girls***	***Gender Gap***
2016	92.95	92.03	94.07	2.04
2017	94.92	94.26	95.72	1.46
Standardized GPA averages of subsidized students				
Year	***Total***	***Boys***	***Girls***	***Gender Gap***
2016	1.333	1.275	1.404	0.129
2017	1.457	1.415	1.508	0.093

4.3 Input Variables for Education Production Function

This subsection will describe control variables, which constitute inputs into the education production function. I will start with family and student inputs, then continue with school inputs.

Parental investment is a key determinant of children's human capital (Becker, 1993). Parents invest time and money in their child's education as well as genetically transfer the endowment (inherent ability) for learning capacity. Data includes several

variables to control for various dimensions of parental investment that we have covered in the literature review chapter (pages 22-24 in this study): parental education, parental occupation, income status of the family, family intactness (i.e., parents are married and living together), parental bereavement, number of siblings, home ownership status and the existence of own room for the student at home.

Table 4.11 below presents the levels of learning in the data, how students are sorted to different levels of mother's education and how the average TEOG test-score for each mother's education level varies within each cohort. We observe that private school distribution is concentrated more at the higher educational levels compared to public schools. The most crowded graduation level for mothers is the "primary school" level for public school students, whereas the same is the "undergraduate level" for private school students' mothers. Average TEOG scores are increasing in higher levels of mother's education except for the Ph.D. level in public schools. There is a gap between private and public school students across the board at all education levels. However, the gap becomes smaller as the level of education gets higher.

Table 4-11 Mother's education level statistics by school sector

	2014				2015				2016				2017			
Mother's Education Level	**Public**		**Private**		**Public**		**Private**		**Public**		**Private**		**Public**		**Private**	
	%	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*13.31*	156.40	*17.75*	284.95	*14.15*	168.24	*18.55*	289.72	*15.14*	174.70	*16.15*	282.97	*15.62*	192.35	*17.25*	296.36
Illiterate	*10.57*	149.18	*0.37*	235.56	*10.02*	146.82	*0.30*	233.60	*9.42*	152.51	*0.40*	237.24	*8.92*	168.99	*0.42*	245.40
Some Primary Sch.	*5.39*	162.01	*0.48*	257.32	*5.24*	160.32	*0.57*	256.85	*5.12*	166.55	*0.69*	247.34	*4.93*	184.56	*0.61*	259.31
Primary School	*52.06*	186.98	*13.49*	263.42	*51.24*	186.14	*13.64*	262.00	*50.45*	191.04	*15.48*	255.44	*49.20*	209.65	*16.08*	268.71
Middle School	*5.28*	204.49	*3.63*	264.35	*5.44*	204.20	*3.66*	266.67	*5.63*	207.55	*4.37*	259.13	*5.74*	226.19	*3.70*	272.99
High School	*10.21*	232.90	*23.59*	283.95	*10.66*	234.80	*24.10*	285.28	*10.98*	236.60	*25.61*	277.98	*11.70*	255.23	*24.66*	290.95
Institute	*1.04*	267.04	*6.09*	303.12	*1.04*	271.48	*6.18*	306.26	*1.04*	270.80	*6.20*	299.33	*1.17*	288.56	*5.95*	311.17
Undergraduate	*1.95*	277.55	*28.02*	308.50	*2.04*	283.02	*28.99*	313.13	*2.05*	281.35	*27.81*	308.28	*2.56*	297.49	*30.06*	318.62
Master	*0.14*	271.44	*3.22*	309.94	*0.13*	277.88	*3.17*	317.97	*0.13*	279.79	*2.62*	312.54	*0.14*	296.51	*3.17*	321.39
PhD	*0.05*	220.19	*0.93*	318.18	*0.05*	220.31	*0.83*	321.99	*0.04*	227.22	*0.69*	320.08	*0.03*	243.29	*0.88*	326.46

It is striking that all the statistics provided in Table 4.9, as well as the ratio of missingness, remain similar and stable throughout the four succeeding cohorts within the group of public schools and within the group of private schools. This observation seems to be in line with the conjecture of a data-generating function from which a random realization comes out each year.

Table 4.12 below represents analogous statistics for the father's education. A highly similar pattern to a mother's education is observed. However, this time the test score average of the Ph.D. group is more similar to the Master's group.

Table 4-12 Father's education level statistics by school sector

Father's Education Level	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score
Missing	*12.84*	169.23	*20.49*	285.38	*13.78*	169.21	*18.64*	290.28	*14.75*	175.20	*16.59*	282.74	*15.28*	193.28	*17.87*	296.76
Illiterate	*1.87*	140.61	*0.03*	233.66	*1.77*	137.42	*0.02*	235.28	*1.62*	137.42	*0.06*	243.29	*1.56*	157.85	*0.04*	248.02
Some Primary Schooling	*3.74*	149.66	*0.20*	255.23	*3.66*	147.54	*0.22*	252.77	*3.57*	147.54	*0.30*	242.60	*3.35*	169.22	*0.25*	259.92
Primary School	*48.31*	174.61	*7.15*	254.11	*47.11*	173.60	*7.59*	252.18	*45.85*	173.40	*9.29*	247.29	*44.25*	196.89	*8.39*	260.69
Middle School	*9.23*	193.31	*3.65*	258.62	*.9.07*	191.80	*3.56*	257.87	*9.10*	191.79	*4.26*	252.88	*8.96*	214.96	*3.90*	266.14
High School	*16.49*	217.89	*19.46*	275.01	*17.07*	218.77	*20.23*	276.63	*17.53*	218.77	*22.84*	269.47	*18.35*	241.22	*22.10*	284.84
Institute	*2.57*	249.88	*5.81*	297.52	*2.55*	253.58	*5.96*	298.38	*2.54*	253.58	*6.26*	292.63	*2.54*	274.00	*5.49*	305.83
Undergraduate	*4.50*	264.54	*35.72*	304.65	*4.56*	269.51	*36.66*	309.08	*4.60*	269.51	*34.10*	304.86	*5.27*	287.22	*35.45*	315.75
Master	*0.37*	268.97	*5.52*	309.52	*0.35*	273.79	*5.26*	315.31	*0.35*	273.79	*4.81*	311.47	*0.36*	292.91	*4.83*	321.01
PhD	*0.09*	256.25	*1.95*	314.87	*0.09*	263.72	*1.85*	320.80	*0.08*	263.72	*1.49*	315.97	*0.09*	280.37	*1.68*	325.40

To better grasp how the distributions vary across different parental education levels in terms of TEOG test scores, the ridgeline plots are provided in Figure 4.9 below. The shape of the distribution takes a drastic change after the upper-secondary school level for the mother's education and at the institute level for the father's education.

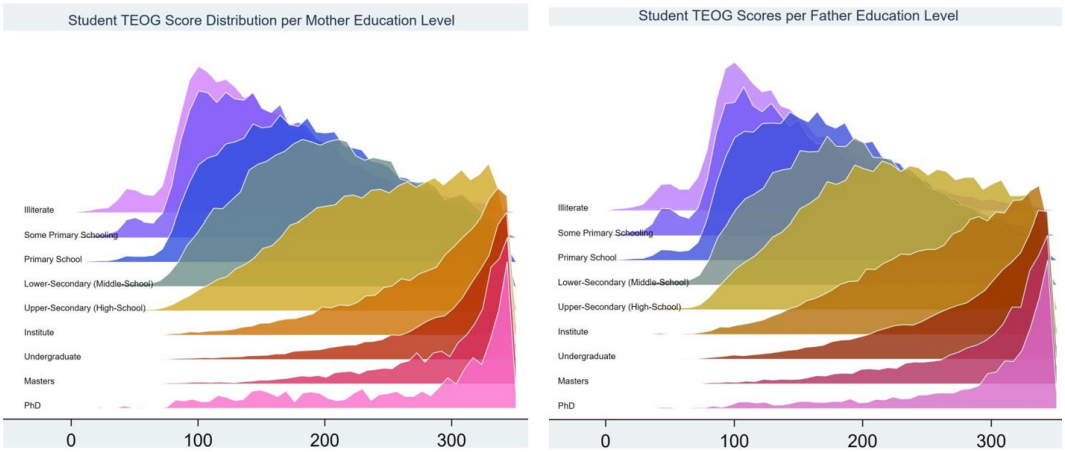

Figure 4.9 TEOG Score Distributions Per Mother and Father Education Levels

One important issue with both the mother and father's education is the percentage of missingness in data. As the information on these is collected via a questionnaire,

missingness and mis-measurement to a certain degree are indispensable. Jerrim and Mickelwrigth (2014, p.770) provide missingness for the father's education in PISA exam, which is 12 percent when rounded. This is very close to the approximately 13 percent missingness in TEOG in 2014, as reported in Table 4.12.

As I will discuss in the methodology section next, the case that "*missingness in a certain predictor variable in a regression setting is not related to the outcome variable*" provides leverage for addressing the limitations emanating from the missingness. In Figure 4.10 below, the test score distributions of observed and missing mother education are compared within each school sector. There is evidence that the test score distribution does not differ between observed and missing mother education in private schools. So it fits the above-mentioned case. For public school students, on the other hand, the two distributions do not overlap as much.

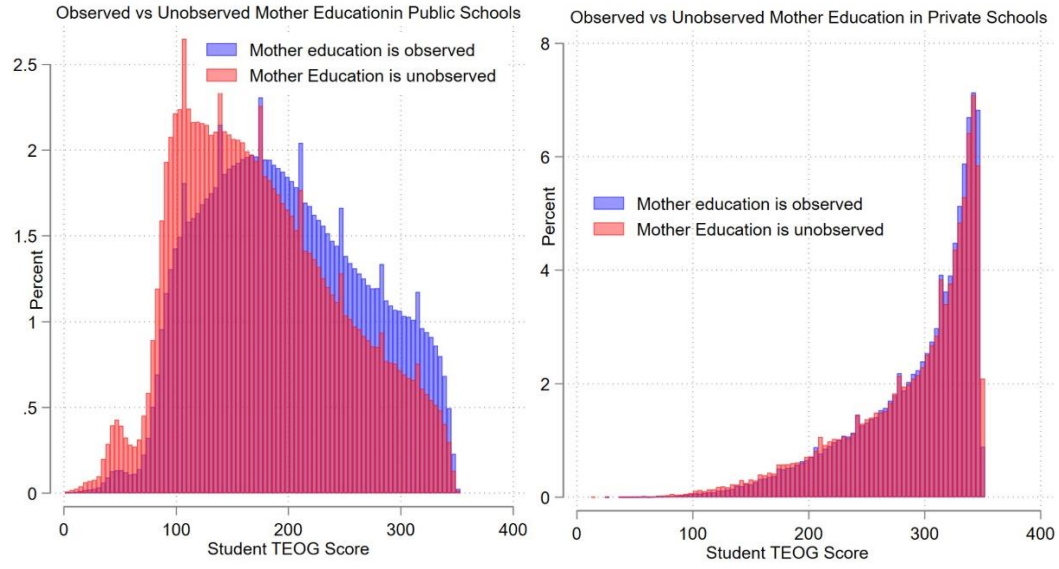

Figure 4.10 Observed vs Missing Mother Education Comparison Per School Sector

However, when we restrict the above-mentioned comparison in public schools to the students whose family income status is reported as "high" and who study in schools residing in the most developed districts, then the overlap of distributions increases substantially, as seen in Figure 4.11. This implies, conditional on reported family

income, missingness in mother education becomes much closer to not related to the outcome variable. Hence, the methodological requirement is better met.

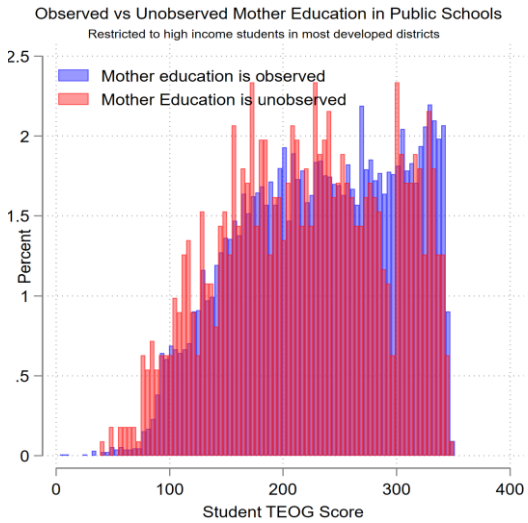

Figure 4.11 Observed vs Missing Mother Education: Restricted Contrast (Public Schools)

On the other hand, the missingness in the mother's education largely overlaps with the missingness in the father's education. Among 4,865,484 students in the data, both mother and father education values are missing for 556,638 students. For 161,584 students, only mother education and for 144,109 students, only father education values are missing.

In addition to the concern about missingness, another concern is the validity of the not-missing data. In PISA exams, for instance, the family background information is collected by a questionnaire given to students during the actual exam. Due to concerns about potential differences between the students' replies and correct answers, a smaller sample in PISA includes the replies from parents in addition to their pupils. Differences between students' responses and parental responses were detected for all countries participating in PISA. However, the replies of Turkish students are found to be the most congruent among the set of all PISA-participator countries. The mentioned difference was only 0.20 for Turkish students, whereas it was 0.71 for

German students, who were the most non-congruent (see Jerrim and Mickelwright, 2014, p.771). The family background data reported for TEOG data in this chapter is not collected directly from the students; hence this type of non-congruence is not such a risk. However, if we assume that the congruence of Turkish students is largely due to a cultural trait (being careful and sincere about expressing the right information in terms of this kind of personal information), then this would support the argument that information declared by Turkish parents tends to be reliable.

The important question on data reliability is how the family background data ends up in the administrative database. The student and family background data is recruited from the e-School (*e-Okul*) module of the Ministry of National Education (MoNE) of Turkiye. Only school administrations are allowed to enter information into this system. Some of the information comes from national databases run by other ministries (like the MERNIS registry system of the Ministry of Internal Affairs). In terms of other information, when a student is first registered in a primary school, a form asking for family background information is filled in and submitted to the secretary of the school. During primary school, the primary school teacher (single teacher for four years) is responsible for sending home the form, getting it filled back and signed by one of the parents, and entering the student's information into the system (or handing it to someone who is responsible for that). In the lower-secondary period, each classroom had many teachers. However, each class of students has an advisory teacher responsible for them. Advisory teachers should send the form home, get it back, and input the data into the mentioned management information system.

Moreover, after the TEOG system was initiated, the checking and correcting of this information by school administrations before each of the TEOG exams was made required by MoNE[32]. Hence, the TEOG system had a better design to reliably collect student and family background information compared to other examples like the PISA exam. It is evident by the existence of missing observations in some cases that the

[32] See, for instance, page 7 of the following official document of MoNE (in Turkish): Ortaöğretim Kurumlarına Geçiş Sistemi-Seviye Belirleme Sınavı e-Başvuru Kılavuzu/2103 (meb.gov.tr)

execution did not perfectly match the design. However, the fact that the information either comes from other reliable databases or is directly reported by the families and the entries into the system made by school personnel should have eliminated the erroneous entries to a certain degree. Another important supplementary factor for reliability is the stability of missingness and other findings all through the four cohorts. As feedback for the future, summaries of missingness in this chapter call for increased awareness of and enforcement by school administration for the quality and completeness of the data entries. This is easily achievable, and the pay-off in terms of the quality of policy analysis in the future is large.

Next to parental education, parental occupation is the other primary proxy variable suggested in the literature for the level of family investments in the child's human capital development process. In Table 4.13, the categories for mother's occupation is presented together with the percentage of mothers in each category and the average TEOG test score of each category separately for public and private schools.

Table 4-13 Mother's occupation category statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Mother's Occupation	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*13.23*	169.66	*20.09*	285.50	*14.03*	169.34	*18.51*	290.23	*15.01*	175.50	*16.34*	283.39	*15.38*	193.11	*17.50*	296.63
Unemployed	*68.99*	187.85	*35.53*	277.66	*67.76*	187.89	*36.26*	278.91	*66.96*	193.04	*39.22*	271.60	*66.19*	212.31	*36.12*	285.44
Other Occupation	*12.66*	188.33	*23.49*	294.91	*12.98*	188.83	*23.31*	298.86	*12.80*	194.05	*22.35*	291.69	*12.89*	213.79	*24.44*	305.10
Private Sector Worker	*2.10*	198.31	*1.65*	293.54	*2.29*	197.43	*2.19*	294.50	*2.46*	201.94	*2.85*	278.73	*2.59*	222.50	*2.85*	293.41
Self-Employed	*0.47*	199.33	*0.78*	291.39	*0.43*	201.85	*0.84*	293.29	*0.38*	208.84	*0.79*	282.71	*0.33*	232.90	*0.71*	301.65
Public Sector Worker	*0.13*	216.15	*0.25*	297.39	*0.14*	214.56	*0.34*	298.75	*0.15*	221.17	*0.32*	288.96	*0.15*	239.98	*0.33*	307.05
Retired	*0.36*	239.16	*1.22*	297.89	*0.34*	239.82	*1.14*	301.23	*0.29*	238.96	*1.09*	295.71	*0.27*	255.85	*1.00*	305.16
Public Sector Official	*0.24*	262.79	*1.74*	311.76	*0.26*	264.93	*1.88*	313.57	*0.27*	268.52	*1.93*	308.83	*0.29*	284.09	*1.81*	318.64
Army/Internal Affairs	*1.08*	265.14	*6.92*	304.81	*1.04*	271.15	*7.15*	309.10	*1.00*	272.29	*7.06*	303.07	*1.07*	289.84	*7.01*	315.11
Teacher/Lawyer/Prosec.	*0.73*	288.71	*8.83*	315.33	*0.72*	295.73	*8.39*	319.33	*0.68*	294.65	*8.05*	314.77	*0.83*	310.46	*8.22*	324.16

The categories within the e-School database include self-employed, private-sector (blue-collar) workers and other occupations as the three broad groups of private-sector occupations. The *other occupations* category seems to cover quite a number of occupations. However, for the mother's education variable, it seems to be a lesser issue. For the public sector, the occupation categories are public sector (blue-collar) worker (*kamu işçisi*), public sector official (*memur*), military officials, internal affairs

officers, lawyers/prosecutors and teachers. In addition to these, it also has non-employed and retired categories. In order to avoid too small group sizes, I grouped army and internal officers in a single group as well as the teacher mothers and lawyer/prosecutor mothers in a single group while reporting the statistics above. These groups represent internal similarities in terms of the outcome and the predictors. It is again noticeable that the proportions and average test scores are highly similar and stable over the four cohorts within each school sector.

The category involving the highest proportion of mothers is non-employed both for public and private school students. However, non-employed public school mothers constitute 69 percent of all public school mothers, whereas the respective value is 35.5 percent for private school mothers. Larger proportions of private school mothers are concentrated in the more prestigious occupations that are associated with higher average TEOG test scores.

Table 4.14 presents a father's occupation information in the data. The most striking difference from the mother's education is the change in the non-employed and other occupation categories. The non-employed category is less than 5 percent for public school students and less than 0.5 percent for private school students. The other occupation category includes nearly half of the fathers, both for public and private school students. As other occupations may involve both farmers at one end and nanotechnology researchers or software developers at the other end of the spectrum. One strategy could be using interaction terms for occupation and education. Otherwise, similar to mothers' case, larger proportions of private school fathers are concentrated in the more prestigious occupations that are associated with higher average TEOG test scores.

Table 4-14 Father's occupation category statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Father's Occupation	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*12.54*	169.83	*20.25*	285.93	*13.44*	169.82	*18.46*	290.53	*14.41*	175.84	*16.55*	283.25	*14.87*	193.75	*17.83*	296.96
Unemployed	*4.90*	159.75	*0.40*	269.92	*4.72*	158.01	*0.38*	266.09	*4.64*	164.23	*0.49*	259.09	*4.48*	180.33	*0.48*	278.77
Other Occupation	*52.55*	182.30	*47.75*	285.64	*51.25*	182.19	*48.32*	288.80	*50.13*	187.23	*46.62*	281.93	*49.89*	207.03	*49.19*	295.89
Private Sector Worker	*14.71*	192.13	*4.98*	287.95	*15.74*	191.41	*5.85*	286.45	*16.44*	196.77	*8.37*	271.48	*16.72*	216.08	*7.50*	287.42
Self-Employed	*3.49*	206.84	*5.00*	277.72	*3.47*	207.22	*5.07*	278.65	*3.46*	212.54	*4.94*	272.68	*3.16*	232.11	*4.34*	288.97
Public Sector Worker	*1.82*	203.57	*0.88*	288.17	*1.82*	204.20	*1.03*	289.66	*1.69*	208.79	*1.11*	275.47	*1.65*	228.87	*0.91*	293.57
Retired	*2.41*	207.99	*1.76*	286.98	*2.26*	207.69	*1.79*	292.44	*2.15*	211.85	*2.01*	281.24	*2.01*	230.94	*1.84*	295.97
Public Sector Official	*1.56*	245.52	*4.11*	308.97	*1.52*	248.86	*4.16*	312.21	*1.47*	251.94	*4.77*	305.72	*1.52*	271.05	*4.44*	315.77
Army/Internal Affairs	*4.79*	237.69	*9.04*	303.58	*4.56*	242.55	*8.91*	305.64	*4.41*	246.90	*9.22*	298.13	*4.36*	266.52	*7.95*	312.18
Teacher/Lawyer/Prosec.	*1.24*	275.00	*5.84*	312.76	*1.21*	281.61	*6.03*	317.24	*1.20*	282.89	*5.91*	312.28	*1.35*	300.28	*5.52*	322.25

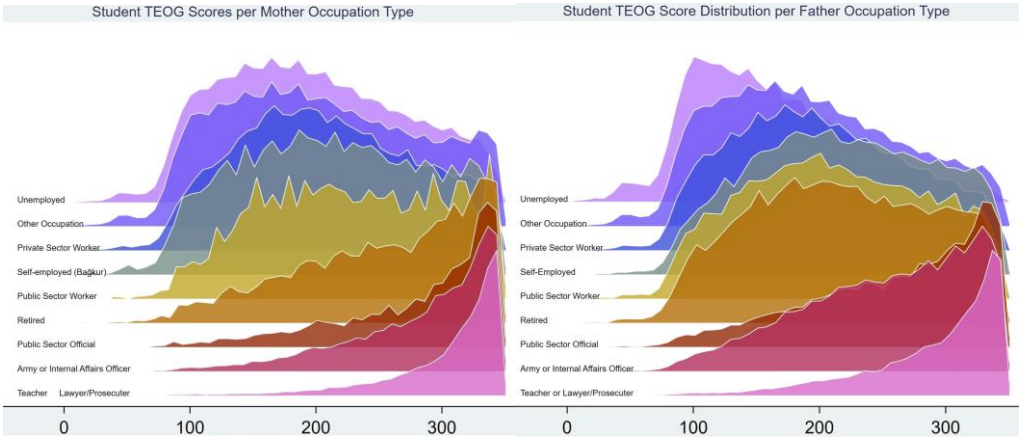

Figure 4.12 TEOG Score Distributions Per Mother And Father Occupation

Figure 4.12 above represents the distribution of TEOG scores specific for each category, both for the mother's occupation and the father's occupation. For mothers, the shape of the distributions substantively changes for the retired and public sector employees. For fathers, it changes only for the group of public sector employees (i.e., for the last three distributions).

Parental occupation data overlap with parental education data to a large extent. Among 4,865,484 students in the data, 510,914 students have missing data in all four variables of {mother's education, father's education, mother's occupation, father's occupation}. The total number of missingness that reflect all possible combinations between these variables is 889,667 (i.e., 18 percent of whole data).

The next proxy for family investments in the human capital of their children is the family income. Table 4.15 represents the categories in the e-School database for family income status (*aile gelir durumu*) indicator. School administrations enter this information based on their interview during the registry period and can update later on based on the interactions with the family. Hence, this variable makes more sense when the family incomes of students are compared within the same school. On the other hand, we observe that reports for private school students' family income are concentrated on the higher levels of family income compared to public schools, as expected. Figure 4.13, on the other hand, represents the test score distributions.

Table 4-15 Family income status by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Family Income	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*8.13*	177.21	*17.63*	288.89	*9.42*	176.47	*16.68*	292.99	*11.17*	182.85	*17.70*	286.86	*13.05*	199.82	*19.49*	299.39
Very Low	*3.76*	152.72	*0.17*	266.77	*3.56*	149.62	*0.17*	267.82	*3.58*	156.18	*0.18*	255.15	*3.26*	172.28	*0.23*	269.02
Low	*25.72*	166.99	*1.77*	276.20	*25.28*	165.67	*1.62*	271.07	*25.47*	171.55	*2.62*	254.30	*24.36*	189.35	*2.29*	272.61
Medium	*40.68*	192.44	*19.79*	289.07	*40.17*	192.30	*20.98*	289.43	*39.39*	197.91	*25.13*	276.96	*38.59*	217.44	*22.62*	291.05
High	*20.36*	213.41	*51.80*	291.76	*20.26*	215.78	*52.13*	295.29	*19.21*	220.35	*47.89*	290.16	*22.44*	241.45	*48.37*	303.14
Very High	*1.36*	214.72	*8.84*	282.27	*1.31*	217.59	*8.23*	288.12	*1.18*	220.66	*6.48*	286.89	*1.44*	242.80	*6.99*	300.14

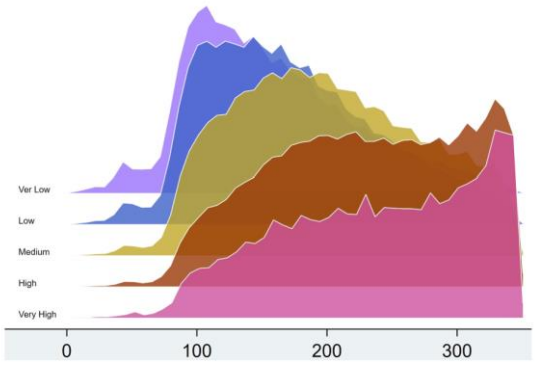

Figure 4.13 TEOG Score Distributions By Family Income Status

We observe from Table 4.15 that the missingness in the family income status variable is lower than the missingness in the previous family input variables. As seen in Figure 4.14, the overlaps in outcome measures between observed and missing cases both for public and private schools (without any restrictions) are also less problematic.

On the other hand, there are 525,538 missing cases for family income, and 307,547 of these are jointly missing with mother education. However, for 410,675 missing cases for mother education, there exists an observation for family income. This would be beneficial in handling missingness in parental education and occupation variables.

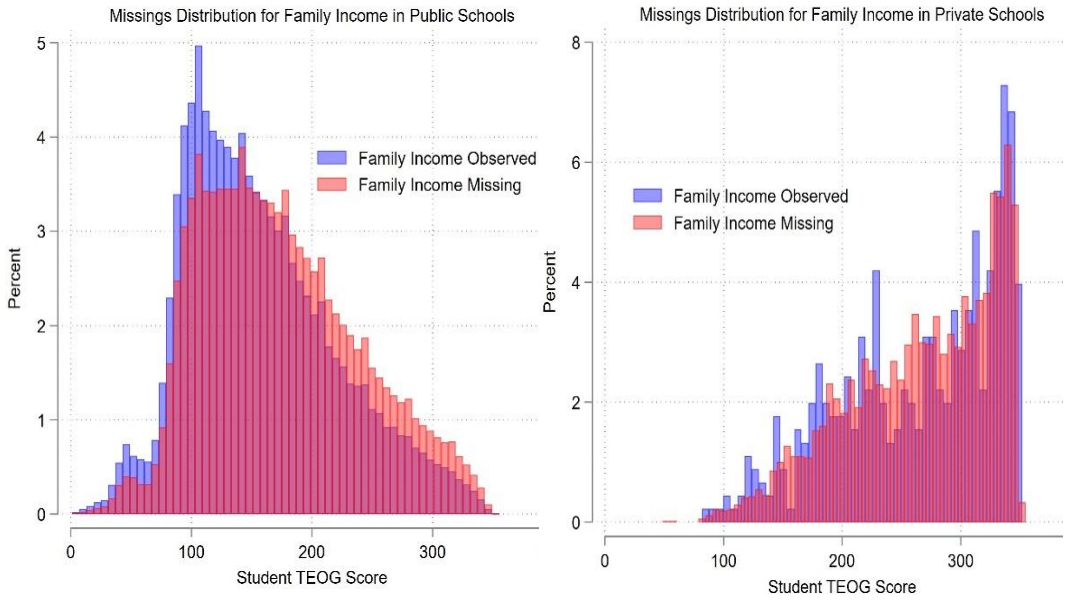

Figure 4.14 Missing vs Observed Cases of Family Income per School Sector

Sibship size, on the other hand, is important for resource allocation to pupils within the household. It can constrain financial as well as time resources. Table 4.16 represents the proportion of different numbers of siblings and associated average TEOG test scores in each school sector. Public school students tend to have higher numbers of siblings.

Unadjusted comparisons indicate that lower numbers of siblings are associated with higher average test scores both for both school types. Figure 4.15 presents how the whole test score distribution takes shape for each sibship size and indicates that after two and more siblings, the test score distribution substantially changes. Finally, there is no missing data regarding the number of siblings.

Table 4-16 Number of siblings statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Number of Siblings	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Single Child	*8.97*	210.54	*26.33*	295.39	*9.97*	211.23	*27.96*	297.74	*9.86*	212.62	*27.83*	289.22	*11.36*	232.51	*32.71*	303.22
One Sibling	*35.57*	204.76	*50.95*	293.60	*36.12*	204.56	*50.93*	296.45	*26.17*	208.84	*51.73*	288.27	*36.13*	228.13	*49.40*	301.32
Two Siblings	*24.12*	187.59	*16.16*	278.32	*23.80*	186.27	*15.15*	281.61	*23.75*	192.10	*14.68*	275.33	*23.21*	210.47	*12.64*	287.36
Three Siblings	*12.42*	172.83	*4.24*	267.13	*12.02*	171.28	*3.85*	270.43	*11.96*	177.25	*3.77*	266.05	*11.64*	195.54	*3.35*	276.35
Four or More Sibligns	*18.93*	154.07	*2.31*	254.03	*18.08*	152.73	*2.10*	252.79	*18.26*	159.64	*1.98*	250.20	*17.66*	177.15	*1.90*	266.37

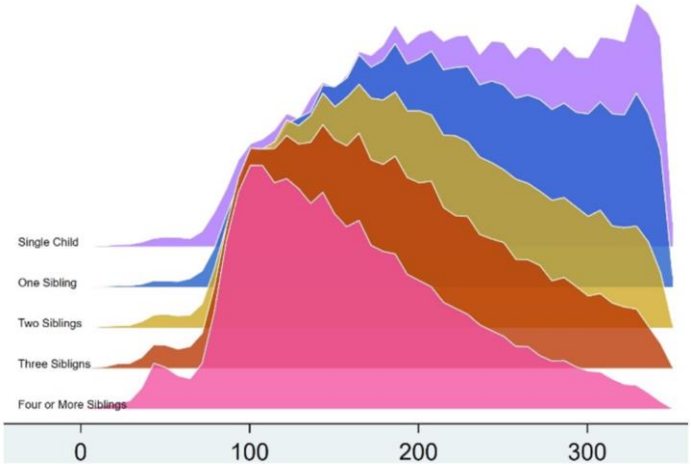

Figure 4.15 Student TEOG Test Score Distribution Per Sibship Size

Table 4.17 represents parental life status and family intactness statistics. Bereavement of a parent or living with a single parent due to divorce can reduce the financial and time resources available to pupils as well as cause psychological challenges for children. Both the proportions and test score effects of the loss of a parent are drastically different for the first two cohorts and the last two cohorts. This was never the case for other control variables described up to now[33].

The missingness is negligible for parent life status variables. For family intactness, on the other hand, the proportions and test score averages are similar and stable over

[33] I discussed this with the related database experts. Their suggested explanation is that first two years reflects the stock variable, so, contain students who lost a parent since the beginning of lower-secondary school, while the last two years reflect a flow from a time very close to the exam. Sometimes higher management demands an update on a random day (for a high level briefing, etc.), which takes at least a day. After each TEOG exam, a snapshot is taken. Most probably an erroneous "select" code is written in 2016 and the same code was used in 2017, too.

the course of all cohorts. Non-intact family structure is associated with lower average test scores for both school sectors. Both the proportion of single-parent families and the score gap within each school sector are similar for both school sectors.

Table 4-17 Parental demographic statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score
Mother Life Status																
Missing	*0.63*	171.43	*0.81*	266.49	*0.62*	170.22	*0.85*	264.22	*0.62*	177.36	*0.68*	263.18	*0.60*	191.47	*0.65*	273.42
Mother is Alive	*98.62*	188.12	*98.81*	289.80	*98.66*	188.09	*98.71*	292.96	*99.31*	192.91	*99.29*	285.20	*99.33*	212.18	*99.26*	298.83
Mother is Dead	*0.75*	171.04	*0.38*	277.79	*0.71*	168.51	*0.43*	275.61	*0.07*	121.75	*0.02*	270.10	*0.07*	132.46	*0.09*	318.04
Father Life Status																
Missing	*0.48*	158.23	*0.45*	269.12	*48.00*	156.55	*0.34*	269.17	*0.51*	162.94	*0.37*	257.93	*0.50*	181.09	*0.37*	277.00
Father is Alive	*97.03*	188.33	*97.85*	289.81	*97.15*	188.30	*98.11*	292.90	*99.29*	193.03	*99.32*	285.15	*99.30*	212.29	*99.28*	298.74
Father is Dead	*2.48*	176.05	*1.69*	280.74	*2.38*	175.56	*1.55*	281.30	*0.19*	121.13	*0.31*	284.72	*0.20*	136.63	*0.35*	303.37
Family Intactness																
Missing	*0.00*	-	*0.00*	-	*0.00*	-	*0.00*	-	*0.00*	-	*0.00*	-	*0.00*	-	*0.00*	-
Live with Both Parents	*91.45*	188.61	*90.02*	290.73	*91.33*	188.71	*89.93*	294.17	*91.05*	193.68	*89.97*	286.51	*90.88*	213.20	*89.20*	300.22
Live with Single Parent	*8.55*	180.19	*9.98*	279.11	*8.67*	178.75	*10.07*	278.92	*8.95*	183.44	*10.03*	271.89	*9.12*	199.98	*10.20*	285.94

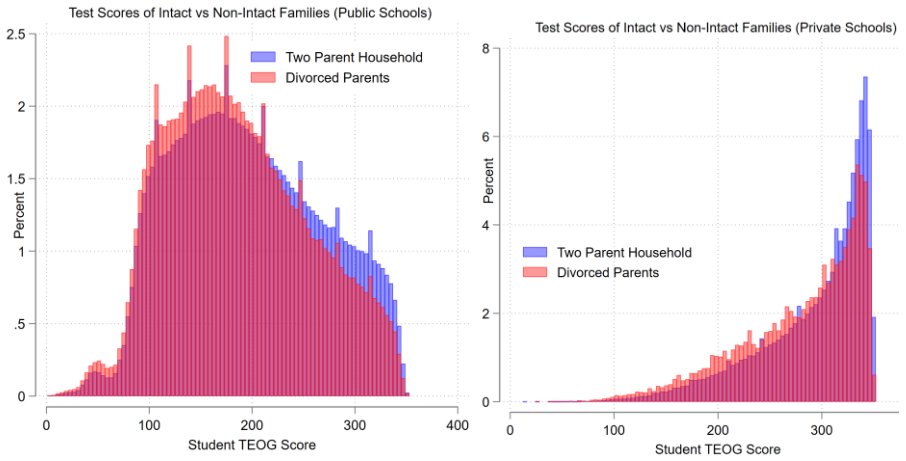

Figure 4.16 Family Intactness and Test Score Distribution per School Sector

Figure 4.16 compares TEOG test score distributions of intact and non-intact families per school sector. While the common region in both graphs is large, the distribution of intact families is more concentrated over higher scores and dominates the distribution of students from non-intact families for both school sectors. On the other hand, there is also no missing data in terms of family intactness, and hence no analysis of missingness is needed.

Home status statistics in Table 4.18 describe the proportion of students whose family is a tenant or an owner of a house or live in a public housing/lodging (*lojman*). Roughly two-thirds of students live in homes owned by their families, a little less than one-third in rented homes and a small portion of students in public lodgments. The proportion of students living in public lodgings is higher for private school students compared to public school students. The average test scores of owners and tenants are close to each other, whereas the average test scores of students living in public lodging are higher than those two groups. Figure 4.17 shows the distribution of each home status type, and we observe how close the distributions of homeowners and tenants are; and how different the distribution of students living in public lodgment is. Finally, Figure 4.18 presents missingness for each school sector. While missingness in the private sector can be seen not to be associated with the outcome variable, it seems so for the public sector.

Table 4-18 Home status statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Home Status	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*4.06*	168.82	*9.63*	289.33	*4.72*	167.93	*8.76*	293.31	*5.58*	176.21	*8.81*	288.25	*6.80*	193.29	*10.34*	298.09
House on Rent	*30.21*	187.18	*19.00*	292.52	*30.94*	186.32	*20.41*	294.07	*31.29*	191.77	*22.03*	283.42	*30.77*	210.86	*22.75*	295.95
Own House	*64.12*	188.14	*66.94*	287.54	*62.76*	188.75	*67.10*	291.09	*61.63*	193.55	*65.52*	284.04	*60.97*	213.44	*65.97*	298.72
Public Housing	*1.61*	235.33	*4.92*	309.54	*1.58*	237.73	*3.73*	311.24	*1.50*	240.24	*3.63*	305.16	*1.46*	259.72	*3.66*	316.51

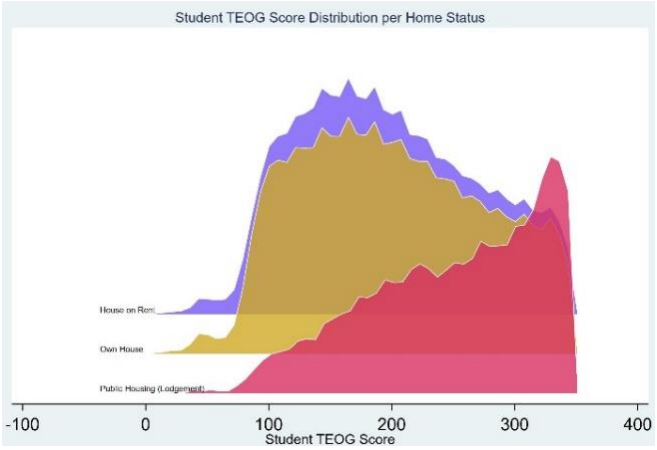

Figure 4.17 Home Status and TEOG Test Score Distribution

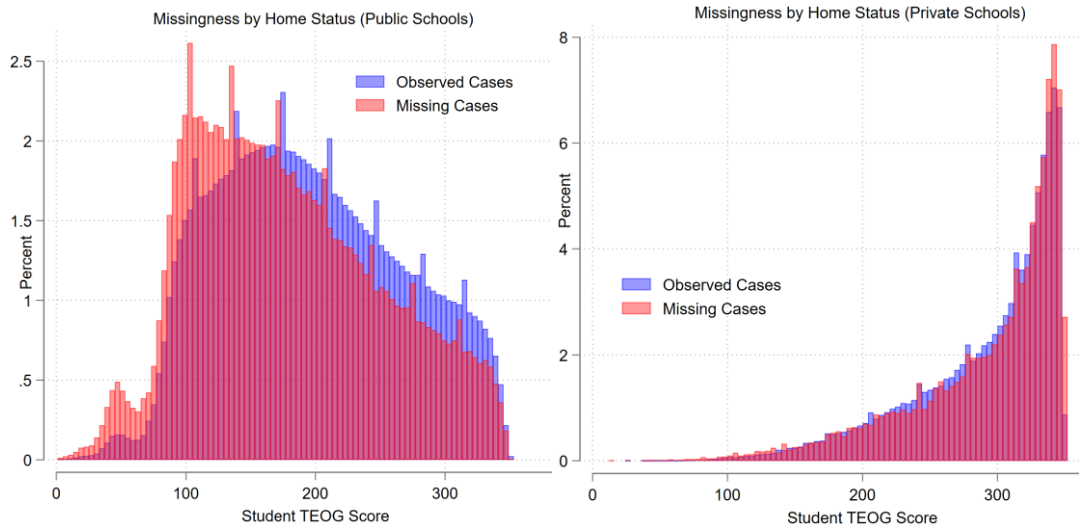

Figure 4.18 Missing vs Nonmissing Home Status Distributions by School Sector

Studies done at home are complementary to the face-to-face education carried out in school. Home environment, which supports rehearsal and learning, is found to have an impact on the human capital development of the student (Collins, 2007). The existence of a separate own student room can increase the focus and motivation of the student to carry out learning tasks at home.

Table 4.19 represents the proportion of students having their own room and associated mean test scores for each school sector over four different cohorts. It is immediately noticeable that the majority of private school students have their own rooms. Their proportion is more than double the proportion of public school students who have separate room at home. Having an own room is associated with higher average test scores for both school sectors. However, the gap is higher for public school children compared to private school counterparts.

Table 4-19 Student's own room status statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Own Room at Home	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Missing	*4.05*	169.42	*9.19*	288.53	*4.68*	168.30	*8.44*	292.86	*5.54*	176.36	*8.23*	286.80	*6.80*	193.51	*9.91*	297.35
No Own Room	*57.42*	171.63	*7.80*	276.87	*56.70*	170.94	*8.34*	278.29	*55.07*	176.15	*9.72*	264.37	*53.12*	194.80	*9.16*	279.71
Has Own Room	*38.53*	213.50	*83.01*	290.87	*38.62*	214.39	*83.22*	294.05	*39.39*	217.73	*82.06*	287.32	*40.09*	237.38	*80.94*	300.99

If we examine not just the averages but whole distributions, Figure 4.19 indicates that the distribution of test scores for students with a separate room stochastically dominates the distribution of test scores for students who do not have a separate room. This tendency is observed both within the group of public schools and within the group of private schools.

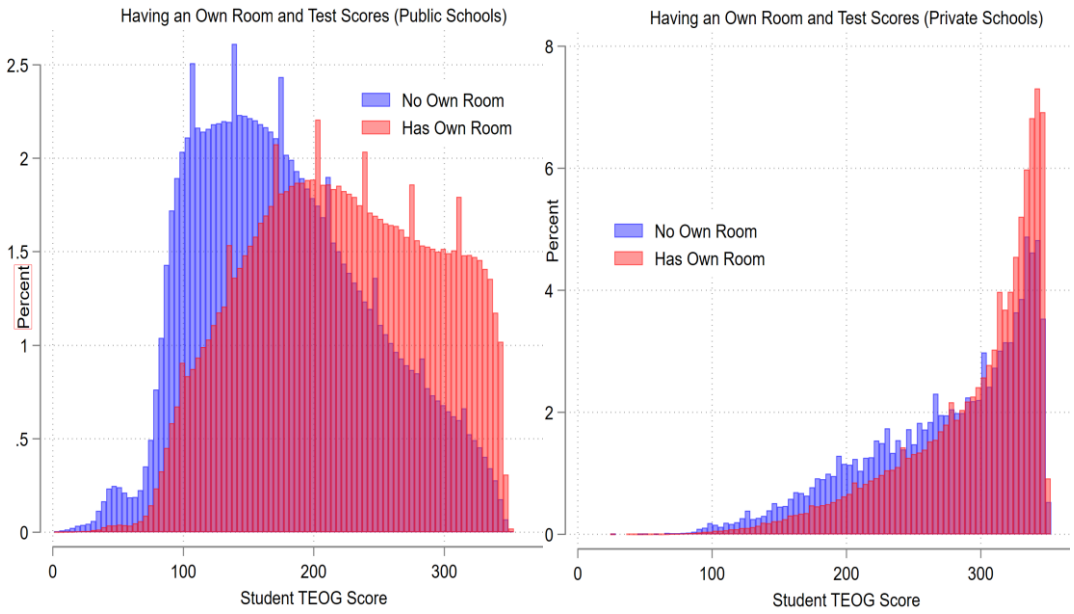

Figure 4.19 Existence of an Own Room and TEOG Test Score Distribution

If we compare the test score distribution of students with missing data for the own room variable with the outcome distribution of students with observed data for the same variable, as in Figure 4.20 below, we see that missingness in terms of the existence of the own room is not associated to a significant degree with test score outcome in private schools, which is evident from strongly overlapped distributions between observed and missing values of the existence of an own room at home.

However, the same cannot be argued for public schools. The test score distribution of the students with observed own room cases stochastically dominates the distribution of the students with missing room cases. This is a strong indication that missing data may not be random.

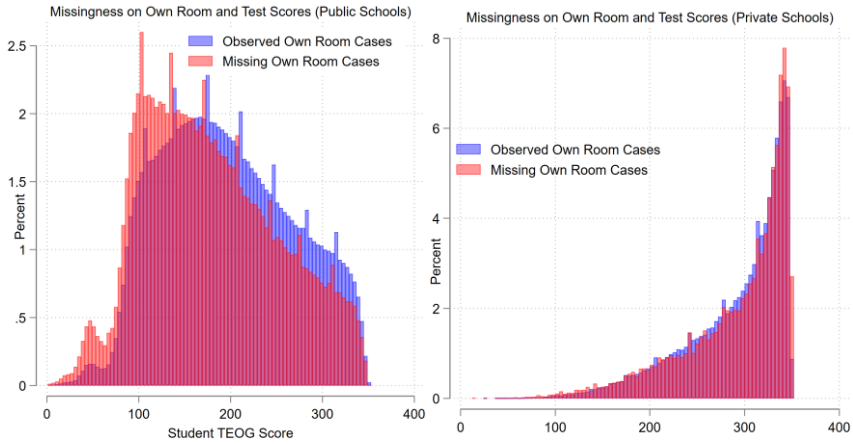

Figure 4.20 Missingness of Own Room and TEOG Test Score Distribution

However, when we restrict the comparison, for instance, between missing and observed cases within the public school sector to families with high incomes and from most developed districts (which are predominantly the distinctive aspects of private schools), the overlap in distribution becomes much better, as presented in Figure 4.21.

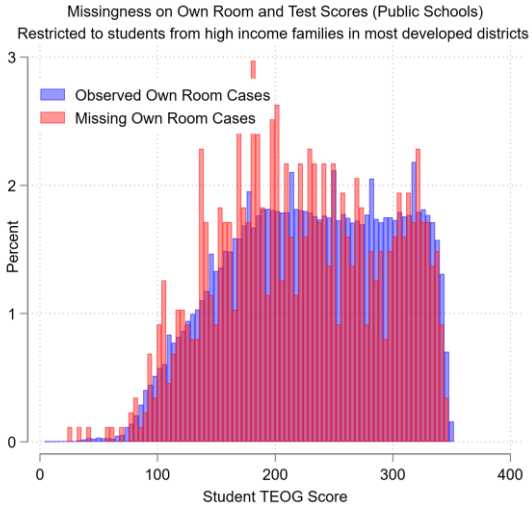

Figure 4.21 Missingness of Own Room and Test Score Distribution (Restricted)

In addition to student-level proxies, the TEOG dataset also offers school-level proxies relevant to education production function analysis. A policy of increased utilization of double school sessions in schools (i.e., a morning group followed by an afternoon group of students in the same school) was in effect during the TEOG period. Table

4.20 represents that this policy is almost entirely present only in public schools. The average test scores, as seen in Table 4.20, and the distribution of test scores, as seen in Figure 4.22, indicates that single-session schools are associated with better achievement levels, but the difference is not so punctuated.

Table 4-20 School session statistics by school sector

	2014				2015				2016				2017			
Number of School Sessisons	**Public**		**Private**		**Public**		**Private**		**Public**		**Private**		**Public**		**Private**	
	%	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**	**%**	**Score**
Single Session	*76.54*	187.58	*99.78*	289.55	*78.22*	187.93	*99.80*	292.32	*80.76*	193.47	*99.86*	285.05	*81.15*	212.81	*100.00*	298.69
Double Sessions	*23.46*	188.30	*0.22*	292.51	*21.78*	187.73	*0.20*	298.75	*19.24*	191.73	*0.14*	282.04	*18.85*	210.00	*0.00*	-

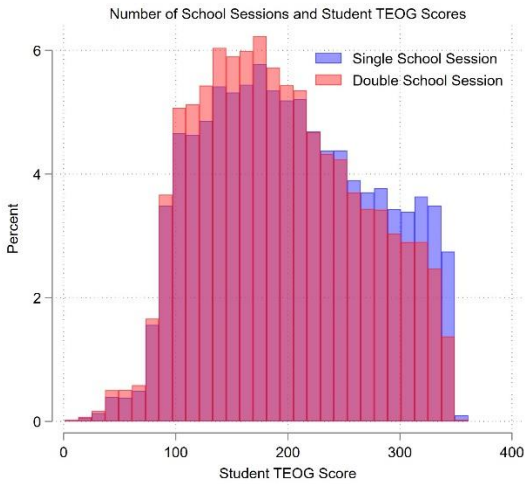

Figure 4.22 Number of School Sessions and TEOG Test Score Distribution

School size, as discussed in the literature review chapter, is one of the structural characteristics of schools. It also receives substantive research interest[34]. The data involves the size of the school in each cohort, which is a good proxy for the school size as often the student capacity of each school is stable (because of being bound by the classroom capacity).

[34] One reason is related to the competition argument which says successful schools attract students and enlarge, failing schools lose students and shrink. Another reason is the scale economies of school resources.

Table 4.21 represents the proportion of schools in each size category, formed by the quarters of the distribution of public and private schools.

Table 4-21 School size category statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
School Size Category	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score
0 < School Size ≤ 23	*22.53*	176.84	*25.78*	270.89	*24.32*	176.16	*23.90*	269.64	*30.76*	180.74	*25.07*	267.73	*30.56*	200.22	*23.03*	281.61
23 < School Size ≤ 43	*22.45*	174.99	*26.56*	283.26	*22.36*	171.02	*27.22*	285.78	*22.94*	177.02	*26.37*	279.80	*22.22*	196.01	*27.83*	291.73
43 < School Size ≤ 68	*16.16*	176.01	*23.88*	287.00	*14.72*	174.10	*24.78*	292.50	*13.07*	180.65	*23.50*	286.12	*14.00*	198.05	*25.06*	296.84
68 < School Size ≤ 101	*12.54*	182.98	*14.37*	291.25	*11.32*	180.47	*15.71*	295.64	*10.36*	189.23	*17.01*	288.02	*9.98*	205.96	*16.13*	300.49
101 < School Size	*26.32*	192.55	*9.04*	299.79	*27.27*	192.80	*8.39*	302.07	*22.87*	198.08	*8.06*	289.93	*23.24*	218.06	*7.95*	310.14

The averages of the school mean scores in each size category are presented in Table 4.21, and the distributions are shown in Figure 4.23 for each school sector. Schools seem not to differ much in achievement according to their sizes.

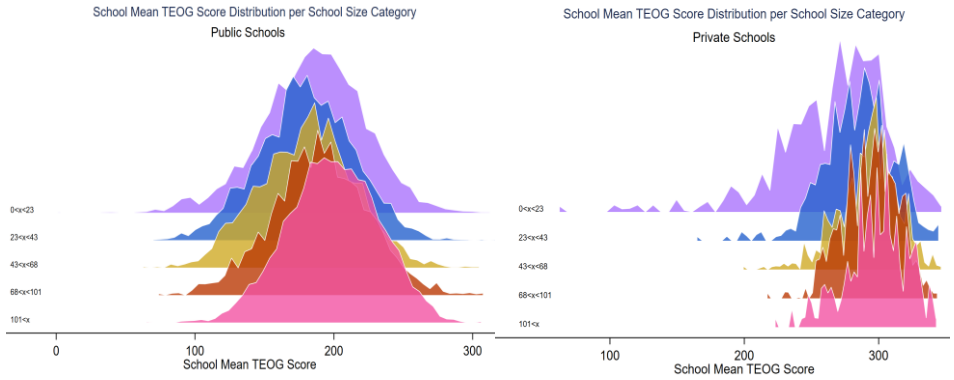

Figure 4.23 School Size Category and School Mean TEOG Scores

Table 4-22 Province development statistics by school sector

	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
Province Development Level	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score
1 - Most Developed	*24.74*	196.83	*58.26*	289.87	*22.87*	195.90	*58.63*	292.42	*22.64*	199.43	*58.33*	283.72	*22.74*	219.08	*61.29*	298.03
2	*17.33*	197.25	*13.50*	294.03	*16.93*	197.16	*13.56*	296.37	*16.74*	201.75	*14.00*	291.08	*16.71*	220.74	*12.75*	302.63
3	*14.43*	187.46	*11.50*	289.81	*14.66*	187.88	*11.12*	295.31	*14.59*	193.42	*10.38*	289.25	*14.60*	212.70	*10.43*	301.85
4	*12.56*	194.18	*6.03*	287.67	*12.72*	195.33	*6.15*	293.77	*12.88*	200.45	*6.56*	287.89	*12.77*	219.89	*6.08*	301.29
5	*11.20*	189.30	*4.58*	294.22	*11.35*	190.85	*4.49*	297.27	*11.53*	196.14	*4.23*	293.07	*11.43*	215.55	*3.38*	304.05
6 - Least Developed	*19.67*	157.89	*6.14*	272.91	*21.37*	157.48	*6.05*	275.64	*21.50*	164.78	*6.49*	267.06	*21.63*	183.86	*6.08*	284.69

Table 4.22 and 4.23 display statistics related to provincial and district development levels for each school type.

Table 4-23 District development statistics by school sector

Distrcit Development Level	2014				2015				2016				2017			
	Public		Private		Public		Private		Public		Private		Public		Private	
	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score	%	Score
1 - Most Developed	*12.42*	199.49	*44.75*	289.98	*11.15*	198.84	*44.39*	293.24	*11.04*	202.36	*43.99*	285.48	*11.23*	222.40	*46.36*	299.41
2	*30.54*	194.97	*40.07*	291.05	*30.03*	194.42	*40.59*	293.70	*29.38*	198.97	*40.44*	286.17	*29.37*	217.91	*40.44*	299.33
3	*21.39*	186.40	*11.83*	287.77	*21.29*	186.75	*11.61*	290.67	*21.25*	192.41	*11.34*	282.73	*21.15*	211.50	*9.60*	297.57
4	*11.19*	185.42	*1.56*	274.52	*11.30*	186.50	*1.56*	286.27	*11.60*	191.40	*1.71*	279.68	*11.54*	211.88	*1.50*	293.29
5	*12.26*	167.25	*1.12*	270.65	*12.63*	167.98	*1.07*	278.74	*12.95*	175.96	*1.37*	275.77	*12.85*	184.88	*1.20*	287.88
6 - Least Developed	*12.13*	148.27	*0.67*	245.99	*13.50*	147.75	*0.78*	237.85	*13.69*	153.24	*1.16*	224.78	*13.75*	172.54	*0.90*	241.15

Private schools are significantly more concentrated in provinces and districts which are the most developed. The averages of school mean test scores increase as the development level improves.

Figure 4.24 displays the distributions for provincial and district development levels. In terms of provincial development, only the distribution of the least developed category differentiates from the other categories. For district development levels, at least two developed categories differentiate as such.

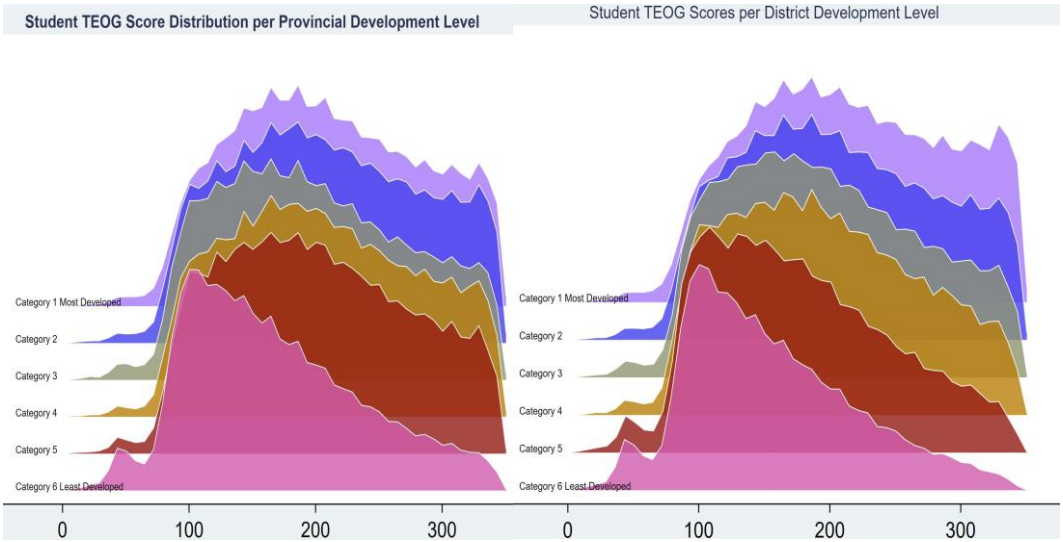

Figure 4.24 Province and District Development Levels and Test Score Distributions

4.4 Conclusion of the Data Chapter

Descriptive analysis carried out in this chapter had the objective of presenting information in the raw data by describing variables and their sub-categories of variables in the data at hand, their relevance to the research questions of the study, and reducing data to a more understandable form via the use of tables, graphs, numerical summaries (mean, range, etc.) as well as data visualizations for whole distributions for each category.

This chapter made it clear that the statistical relationships exist between TEOG test scores and various student, family, and school characteristics are strong. However, a multivariate approach is required if the impact of each unique factor is to be estimated (Bradley & Taylor, 2004, p.286). Descriptive statistics explored in this chapter help us understand the dataset. Inference statistics to be investigated in the next chapters are aimed to help us draw conclusions from this dataset.

CHAPTER 5

THE METHODOLOGY

5.1 Introduction

In this chapter, I will define and justify my methodological choices to investigate private school effects on the human capital outcomes of students and schools.

The methodology chapter will first present the general aspects of the methodological approach, which include the subjects, the measures, the sample and sampling technique, and underlying philosophical assumptions. Then I will revisit the research questions declared in the introduction chapter. Next, I will describe the steps of the data analysis (research designs, identification strategies and estimating models) for each of the research questions. In the following, I will state the methodological and data-related limitations.

5.2 General Research Design Choices

This subsection presents the overarching methodological choices made for the study.

Research context

The research takes place in Turkey, which is an emerging economy, and it covers the period from the 2013-2014 academic year to the 2016-2017 academic year, which corresponds to a period of drastic changes in private schooling in Turkey led by the institutional and policy-based interventions described in chapter 3.

The switch to the TEOG examination scheme was the most prominent institutional change in terms of this study. The TEOG exam provided the unique opportunity to observe private school students to take the same nationally held, standardized, high-

stakes test as their public school counterparts. Comparable exams preceding or succeeding the TEOG exam do not provide such an opportunity since taking a standardized test is not mandatory for those exams. In effect, test scores of almost all private school students have been observed throughout the TEOG period and provide a measure of the human capital for this study. The processes and procedures of the TEOG exam have been summarized in a separate chapter for institutional and political background (chapter 3).

Research philosophy

Research philosophy refers to the worldview about how to collect and explore data about a phenomenon. As a quantitative study, the research philosophy adopts positivism, implying that reality and the data representing reality can be collected and explored objectively.

Research approach

Among the two broad approaches to research, inductive research is exploratory, and data is used to generate the theory. Deductive research, on the other hand, sets off on an existing theory and expands upon it with collected data. As the data analysis will be founded on the human capital theory, the research type will be deductive.

Specifically, I will utilize the workhorse empirical model of the economics of education literature, which is the education production function approach. It analyzes the productivity relationship between inputs and outputs in the context of the education production process during the schooling period. This approach explicitly takes into account the contributions of family inputs, peer inputs, and genetic endowments (for mental capacity) of students to the human capital formation of students, in addition to school inputs. Structural characteristics of schools constitute a subcomponent of the school inputs, and the type of control of the school (i.e., being a private or a public school) is among the important structural characteristics of schools.

Types of Research Questions:

Todd & Wolpin (2003, p. F8) underlines that there are two distinct types of research questions that can be explored empirically. Depending on the specific research objectives, both types of questions might be relevant:

Type-1: How would an exogenous change in an input, holding all other inputs constant, affect achievement?

Type-2: What would be the total effect of an exogenous change in an input on achievement, that is, not holding other inputs constant?

Note that the objective of the first question is to learn about the ceteris paribus effect of an input of interest (which is private school status, for instance, in this study). Knowledge about production technology is sufficient to estimate this effect. Hence, a primary objective of empirical research is to understand the technology for combining inputs to produce achievement outcomes, i.e., to learn about the structural parameters of the education production function. Even though observational data reflects the nonrandom assignment of inputs due to the purposeful choices made by parents or schools, it can be utilized to learn about features of the production technology. We can obtain an answer to the first question as long as we can observe all relevant inputs into the production process and the student endowments. However, not having data on some of the (historical and/or current) inputs and the unobservable nature of student endowments creates important challenges for estimation that require methodological remedies. These remedies include invoking assumptions in the production technology and input determination rules as well as the implementation of fixed-effects approaches to control unobserved but fixed inputs.

The second question aims to learn about the total effect, which includes both the ceteris paribus effect of a change in the input of interest plus, the indirect effects that operate through changes in the levels of other inputs due to the change in the input of interest. In addition to the knowledge of production technology, knowledge about the

family decision rule is required, too. A controlled experiment or a natural experiment that randomly assigns the input of interest delivers an answer to the second question and identifies a treatment effect (policy effect) in the program evaluation literature (Imbens and Rubin, 2015). However, it generally cannot provide an answer to the first question; i.e., experimental studies do not estimate production function parameters (Todd & Wolpin, 2003, p. F9).

In this study that depends on observational data, I will be answering Type-1 questions.

Sample and the sampling technique:

The data of the whole student population in grade 8 (i.e., in the graduating year of lower-secondary school) of each successive student cohort from the 2013-2014 academic year to the 2016-2017 academic year is obtained, as well as the data of all of the school population that inhabit these students. Hence, the study sample equals the population, and no sampling technique is applied.

Data collection method:

The administrative register data is provided by the Turkish Ministry of National Education (MoNE) from their existing administrative databases. Student data is provided anonymously by the MoNE. I did not directly collect data from any students or other human subjects. Note that this is an ex-post evaluation of past data. This data is then augmented by other complementary school information, which is open access on the internet. The details of the data are provided in Chapter 4.

The data represent the most comprehensive version of TEOG exam data and complementary student and school-level data. Although I utilized pre-existing data, it was not open-access data. This study creates an important contribution to data availability for research purposes.

Burgess (2016) argues that the survey data, especially the data from labor force surveys, were the mainstay of the economics of education. However, important new

learning comes from the utilization of other types of data that include administrative or registry data. He underlines the importance of recruitment and use of such data and calls for research funding to support such initiatives. This study follows the lines of this suggestion.

As an important limitation, sample sizes for private school students generally remain small in most of the studies, which leads to substantially wide confidence intervals for the estimates from data analysis (Jerrim et al., 2014, p.292). The population data to be used in this study will prevent such limitations.

Levels of the analysis:

The empirical models will also address the challenges and exploit the opportunities provided by the nested structure implied by schooling (i.e., the fact that students are nested within schools and schools are nested within neighborhoods).

Analysis Tools:

Statistical analysis programs Stata (version 17) is used for the data analysis.

Missing Data Strategy:

Missing values are ubiquitous in all disciplines and particularly common in the social sciences. Missing values in student observations are more likely when the input measures are derived from questionnaires, as students fail to complete questionnaires fully. As presented in chapter 4, although the missing observations for outcome measures are quite limited for public school students and almost non-existing for private school students, some of the input measures have missingness comparable to what we see in international tests like the PISA. Addressing missingness differs for the case where missingness is in the dependent variable and for the case where missingness is in predictor variables. Allison (2012, p.7) argues that the best option is generally omitting the cases with missing data on the dependent variable, while methodological options are larger when missingness is in the predictor variables if

the analysis is carried out in a regression context. The cases with valid observations on the dependent variable but missing values on one or more predictor variables contain potentially valuable information that will be completely lost.

Little and Rubin (2019) distinguish the missingness pattern, indicating which values are missing in the data matrix, from the missingness mechanism(s), which is about the relationship between missingness and the values of variables observed in the data matrix. The widely-adopted framework to think about missingness is Rubin (1976)'s classification of the missingness mechanisms.

For the purposes of this analysis, I will focus on the regression under missing data in the context of education production function analysis. Let X_1 be an input variable (for instance, parental education) entering into the education production function. A missing X_{1i} means that the observation for X_1 of student i is missing. For students with missing values for X_1, if missingness on X_{1i} is unrelated to the observed values of $X_{2i},\ldots, X_{pi}$ (other inputs into education production), to the observed value of the Y_i (output of the education production function) and to the unobserved value of X_{1i} itself, then, according to the classification of Rubin (1976), missingness mechanism implies that missing values for X_{1i} are missing completely at random (MCAR). Again, for students with missing values for X_1, if missingness on X_{1i} is uncorrelated with the unobserved value of X_{1i} after adjusting for the observed values of $X_{2i},\ldots, X_{pi}$ and Y_i; then, the missingness mechanism implies that missing values for X_{1i} are missing at random (MAR). Finally, if the missingness on X_{1i} is correlated with the unobserved value of X_{1i} even after controlling for the observed $X_{2i},\ldots, X_{pi}$ and Y_i, then missingness mechanism implies that the missing values for X_{1i} are missing not at random (MNAR).

Analysis of missing data is problematic since the majority of the statistical methods (and also the majority of the statistical software) assume that all variables are measured for all units. Almost all statistical software, by default, deletes the complete entry for a student even if that student has missing value on just a single variable. In

other words, even if only the value of X_{1i} is missing for the student i, all observed values of $X_{2i},\ldots, X_{pi}$ as well as the value of Y_i are discarded from the dataset. Only the observations of students that don't have even a single missing value for any of the input or output variables are allowed to remain in the analysis. Hence, this method is referred to as complete case analysis or listwise deletion. If a substantial amount of data is discarded, then the loss of statistical power can pose a problem for further analysis with this approach.

Alternative methods to the complete-cases approach includes pairwise deletion, dummy-variable adjustment, imputation based on least-squares, multiple imputation and full information maximum likelihood (FIML) estimation.

In the pairwise deletion, means or variances/covariances are estimated using all available data for each of the variables at hand. In other words, unlike listwise deletion, we do not discard a student's whole entry even if only one or some variable values are not observed for that student. Then the aggregate summaries (mean, variance, etc.) are used in the analysis, as in the case of regression of mean test scores in each school on mean values of several inputs in each school. Pairwise deletion is consistent in the case of MCAR but is likely to be biased in the case of MAR (Allison, 2009, p. 76). The method performs poorly when input variables are correlated to a high degree. Pairwise deletion is generally regarded among the inferior approaches to the missingness problem.

Dummy-variable adjustment (Cohen and Cohen, 1985) adds a dummy to the regression function that indicates that data is missing on the associated predictor. It aims incorporating all the available information into the regression function. This works especially well when dealing with categorical data, where the missing value is just added as its own category and then you run the model as normal. This is a common approach in machine learning applications. The strength of the dummy variable method is that it distinguishes the missing from non-missing observations and partially relates missingness to other inputs. However, its disadvantage is that it

ignores the differences between different observations with missing data (Nick Huntington-Klein, 2021), which may create problems even with MCAR data (Jones, 1996).

In the case of a traditional imputation method that substitutes the means for missing values or in the case of the more advanced method of least-squares imputation, the statistical software cannot address the uncertainty associated with imputed data. The variances related to the variables with missing values are underestimated, and bias is propagated into the regression coefficients (Allison, 2009, p.77).

Multiple imputation is executed in three steps. In the first step, the missing values are replaced with multiple sets (at least five sets) of simulated values to obtain a complete dataset. In the second step, standard data analysis methods are applied to each imputed dataset. In the third and final step, pooling is applied to adjust the obtained parameter estimates for uncertainty. Note that obtaining imputed values that are as similar as possible to the missing values is not the direct aim of this approach. What is aimed is to conduct valid statistical inference (Rubin 1996).

The full information maximum likelihood (FIML) approach adjusts the likelihood function so that each case (student) contributes information on its variables with observed values (Arbuckle, 1996). FIML just analyzes everything that is present in the data and assumes multivariate normality for the model. This is a model-based method. FIML does not need any additional assumptions for missingness in outcome variables but requires joint normality assumptions when input variables have missingness. As dummy variables do not have a normal distribution for sure, FIML seems, at first sight, to be problematic to handle them. However, both simulation-based evidence and implementation experience suggest that even with the missingness in dummy variables, it performs well (Schaffer, 1997).

Both multiple imputation and FIML methods provide valid inference (i.e., they provide unbiased parameter estimates, and their dispersion term estimates account for the increased variability due to missing values) as long as the missing values are

MCAR or MAR. These model-based estimation methods use all the data (including all available observations in predictor variables, plus, the observations for the outcome variable) and are based on established principles of statistical inference. Hence, they are regarded to have significantly better statistical properties over more traditional methods like pairwise deletion, dummy variable adjustment, and least squares imputation. For instance, simulation studies have shown that FIML is superior to pairwise deletion even when normality assumptions required by FIML are violated (Little, 1992, p.1231).

Comparison of Multiple Imputation and FIML:

Multiple imputation and FIML depend on similar assumptions and possess similar statistical properties.

One advantage of multiple imputation is that if the chained equation approach is used, a multivariate normality assumption is not needed. MI generally makes it easier to include auxiliary variables. An important challenge with multiple imputation is that the imputation model has to be coherent with the data analysis model. For instance, if the analysis model has interactions or some type of transformation in the variables, then the imputation model needs to have these aspects as well. This poses a problem if imputation is conducted by another analyst or the study explores lots of exploration with different specifications, then this leads to inflexibility in the data analysis part. FIML does not have these limitations as everything happens under a single model, and all of the variables in the data analysis model will be utilized to account for the missingness (Allison, 2009).

Moreover, the imputation part bears significant computation time, which makes working with big data sets impractical (Allison, 2012, p.6). My own experience with a data set with more than 4 million base units also verifies this limitation in terms of the imputation stage. Little (1992, p.1235) also confirms that the multiple imputation method works well in small samples, while the FIML method works well in large samples.

Another attractive property of FIML is that it produces a deterministic output. In contrast, multiple imputation produces a different result in each run since it crucially depends on random draws (Allison, 2012, p.5). Finally, the implementation of multiple imputation is more accessible as most of the mainstream statistical software allows the straightforward application. FIML, on the other hand, requires specialized software to implement it. For linear models with missing data on the predictor variables (which is the focus of this study), the SEM command in STATA and the PROC CALIS suit in SAS are the two alternatives. Considering the trade-offs between the two methods, I choose FIML.

Comparison of FIML and Complete-Cases Approach:

Complete case analysis provides a useful baseline reference method. It is similar to OLS, but it minimizes the sum of squares of residuals with respect to not only the parameters but also the missing values. FIML (and multiple imputation), on the other hand, treats the missing values as random variables (Little, 1992).

In particular, for the regression set-up, complete case analysis is valid if missingness of X_{1i} depends on X_{1i}, X_{2i},…, X_{pi} but not on Y_i after controlling for X_{1i}, X_{2i},…, X_{pi} (Glynn and Laird 1986). Complete case analysis under these conditions is highly robust to MCAR, to MAR and even to MNAR mechanism for missing input variable(s), as long as the regression function is correctly specified (Allison, 2009, p.75) and the missingness the input variable(s) does not depend on the output variable (i.e., to the test score achievement or the GPA). The complete-cases method will yield approximately unbiased estimates for regression coefficients even if the data are not missing at random. Little (1992, p. 1229) underlines that this useful property is not present even in more sophisticated approaches like multiple imputation or FIML. This method allows for the missingness of X_{1i} depending on itself. This is important in the sense that we never have a chance to assess if missingness in a certain input variable depends on itself (conditional on other input variables) as we cannot observe it. We would require knowledge of those missing values. Moreover, Allison (2014) argues

that complete case analysis is more advantageous over FIML (or multiple imputation) for sample sizes over 500,000 to 1 million as the loss of power is much less a problem for those sample sizes.

FIML is valid when the data are MAR and the missingness of X_{1i} depends on $X_{2i},\ldots,X_{pi}$ and Y_i but not on X_{1i} (Little & Rubin, 1987). As multiple different variables can be jointly missing for the same students, the information provided by Y_i, which is (almost) not missing in this dataset, can be beneficial to a good extent. Note that the probability of missingness on a specific variable may depend on the missingness of another variable without even violating the MCAR assumption (Allison, 2009, p.73). Missingness of two or more variables may always occur together. While joint missingness may not pose a problem in this sense, each of these variables would also not provide information to each other, either. However, if Y_i and the other observed predictor values provide sufficient information, then FIML will work well.

Despite substantive loss of power, complete case analysis is the safest method since it is not prone to Type-I errors. Conventional imputation methods of pairwise deletion, dummy-variable adjustment and traditional imputation methods, on the other hand, often lead to substantially underestimated standard errors and p-values (Allison, 2009, p. 87).

FIML is more efficient than complete-cases analysis under MCAR. FIML is both more efficient and less biased than complete-cases under MAR. Both FIML and complete-cases are prone to bias under MNAR. However, complete-cases are more robust than the maximum likelihood under MNAR. Joint use of FIML next to complete case analysis is a crosscheck for the assumptions of complete case analysis under regression set-up, which requires (i) missingness cannot depend on the dependent variable (test score outcomes or GPAs) and (ii) the regression model for the education production function is correctly specified.

Based on the above considerations, I will use both FIML and Complete-Cases approaches next to each other and assess if they produce close or different results. If they agree with each other, that would be reassuring.

5.3 Research Questions

Here I will reiterate the research question before moving on with data analysis.

Research Question 1: Are private schools more effective than public schools in producing human capital?

Research Question 2: Does competition from private schools increase public school effectiveness in producing human capital?

Research Question 3: Does the gender achievement gap differ by school sector?

5.4 Data Analysis

5.4.1 Preliminary Assessment for Missing Data at the Student Level

I will first compare the model outputs of the two methods suggested above to handle missingness in student level data.

My strategy is to estimate the education production function that incorporates *within school-by-cohort* fixed effects separately for each of Full Information Maximum Likelihood (FIML) and complete-cases methods. FIML utilizes all available observed data, whereas complete-cases utilize only the observations without any missing values for all covariates. As described in the previous subsections, each has its own strengths and required assumptions. Complete-cases can be seen as a reference method, for which FIML provides a sound comparison. It would be reassuring if these two methods did not display very different parameters.

I estimate below the school-by-cohort fixed effects model, where Y_{ijt} is the TEOG test score of student *i* who studied in school *j* in year *t*[35], $\boldsymbol{X_{ijt}}$ is the vector of student level covariates corresponding to parents' and students' inputs, $\bar{\boldsymbol{X}}'_{\boldsymbol{jt}}$ is the vector of school-by-cohort means of covariates, e_{ijt} is the random error term at the student level, and $\bar{e}_{.jt}$ is the school-by-cohort means of covariates error terms. In order to be able to implement FIML, I utilize the school-by-cohort mean-differenced representation of a fixed-effects model.

$$Y_{ijt}\text{-}\bar{Y}_{.jt} = (\boldsymbol{X'_{ijt}}\text{-}\bar{\boldsymbol{X}}'_{\boldsymbol{.jt}})\beta_x + (e_{ijt} - \bar{e}_{.jt}) \quad (1)$$

Note that FIML depends on the assumption of joint normality between different explanatory variables. Mean centering also help as it transforms even the binary variables (which are, for certain, not normal-distributed) to approximately normal-distributed variables, to make them more suitable to the joint normality assumption[36].

How the variables are demeaned (within each school-by-cohort) in the context of each alternative approach requires care. In both approaches, I will use the observed means for variables that have no missing values (for instance, gender or number of siblings) as they correspond to the true means in the population. As FIML uses all available information, I will use all observed values (whether or not they have missingness in values of some other variables for that specific student) in each school-by-cohort when calculating the covariate-means within each school-by-cohort. For the Complete-Cases approach, I will only use observations of students who are in the Complete-Cases group. After calculating the means as such, I will check in a graph how the distribution of means from each approach compares. The other reason for comparing the parameter estimates from these two different approaches is that while

[35] As I am analyzing just the last year's performance of each student (when TEOG is held), cohort and year coincide for each specific student.

[36] Note that a joint normality assumption is not a too strict requirement as, even in the presence of binary variables, the method often works satisfactorily. The same is valid for the alternative method of multiple imputation, too.

FIML can only be used for single-level analysis, I will utilize Complete-Cases set of observations for multilevel analysis in successive sections. Whenever I do more *within analysis* in the following sections, I will implement both approaches together.

I will run the above mean-differenced fixed effects model separately for public school students, private school students, and all students. By the utilization of school-by-cohort fixed effects, I will be able to control all school differences, including the time-variant school factors. In effect, I will be only comparing the girls and boys who studied in the same school and in the same cohort.

5.4.2 Private School Effects

In this section I will explore the following first research question:

Research Question-1: *Are private schools more effective than public schools in producing human capital, where human capital is measured by standardized test scores from TEOG exam held between 2013-14 and 2016-2017 academic years?*

The core task in the literature on private school effects is ideally specifying the production function that will lead to unbiased estimates of the relationship between student outcomes and private school status.

I posit the following conceptual model based on the education production function framework:

$$Y_{ijt} = \alpha + f_t + \delta D_j + \boldsymbol{X'_{ijt}}\beta_w + \boldsymbol{\overline{X}'_{\cdot jt}}\beta_c + \boldsymbol{S'_{jt}}\beta_s + f_i + u_{jt} + e_{ijt} \quad (2)$$

Above, the human capital measure Y_{ijt} stands for TEOG score (standardized to a mean of 0 and a standard deviation of 1; jointly for all 4 cohorts), for student *i* enrolled in school *j* during the year *t*, which is the year (s)he took the TEOG exam. Overall intercept α corresponds to the population's mean TEOG score. f_t is the cohort (year) fixed effect. D_j is the dummy for private sector status and δ is the difference between the mean private school effect and the mean public school effect on the human capital

development of the student population. The vector $\boldsymbol{X_{ijt}}$ represents the student level covariates corresponding to family (parents' and students') inputs into human capital production. The vector $\overline{\boldsymbol{X}}'_{\boldsymbol{jt}}$ corresponds to the school-by-cohort (*jt*) means of family covariates, which represent peer inputs. The combined specification $\boldsymbol{X}'_{\boldsymbol{ijt}}\beta_w$ and $\overline{\boldsymbol{X}}'_{\boldsymbol{jt}}\beta_c$ implies that the β_w is the within school-by-cohort effects and β_c is the contextual effects of family covariates. The vector $\boldsymbol{S}'_{\boldsymbol{jt}}$ contains observed school inputs. f_i is the fixed student contribution, which corresponds to any fixed differences that affect learning. As mentioned before, this term primarily covers innate cognitive ability and personality traits, but also involves any aspect of the student that affects the achievement formation that do not change during the school years being considered. A random school term u_{jt} collects all unobserved school inputs, and a random student term e_{ijt} represents a idiosyncratic stochastic term plus all input history.

The education production function framework promises unbiased estimation if all of the past inputs into human capital, plus, the innate ability and personality traits are observed and put into the function[37]. It should be noted that the above described conceptual model differs from the standard education production function due to the fact that (as in most of the literature) it depends on cross-sectional data that involves solely the contemporaneous measures of the inputs. This requires to make assumptions (based on Todd & Wolpin, 2003, p. F16) that "(i) *only contemporaneous inputs matter to the production of current achievement* or (ii) *inputs are unchanging over time, so that current input measures capture the entire history of inputs;* and (iii) *contemporaneous inputs are unrelated to endowed mental capacity*". Assumption (i) or (ii) are to an extent tenable in the TEOG examination context. It was the first standardized nation-wide exam for those students that involved questions only from last year's curriculum. A catch-up program would cover the held-back content before the last year, and for the period before the TEOG exam year, it can be argued that

[37] When that is the case, e_{ijt} above will be equal only to stochastic idiosyncratic term.

input was more or less stable. The accompanying assumption of the relation between contemporaneous inputs and the endowed mental capacity is a stronger assumption. Parents, for instance, might increase their input to compensate for less-able children or push even more the more-able children so that they get even higher score points. In terms of school inputs, Hanushek (1979, p. 364) argues that the association between innate abilities of students and school inputs, after allowing for family inputs, is likely to be small.

Moreover, again, as in most of the previous studies, direct family and school input measures are not available. For instance, actual verbal conversation between the child and the parent, or the actual amount of homework done, which are direct measures of the quality of the home learning environment, is not in the data. Instead, these direct measures are proxied by other factors such as parental education and the existence of the student's own room. Similarly, instead of direct measures of the quality of the school learning environment (instructional effectiveness, classroom management, disciplined and focused students, etc.), proxy inputs like teacher experience, teacher credentials, and expenditure per pupil are utilized. What we learn from literature is that the proxies for family inputs work considerably better than proxies for school inputs (mainly teacher inputs). Direct measures of innate ability, on the other hand, are out of the reach of researchers in economics. I acknowledge these limitations, and I will aim to reduce the impact of these limitations by using empirical modeling strategies as much as possible. Moreover, despite these limitations, the education production function framework provides important conceptual assistance in terms of the specification of control variables, as described in the literature review section.

One important problem in directly using the model (2) is that the family inputs, especially family socioeconomic measures like parental education or income, are correlated with unobserved school factors collected in u_{jt} due to non-random sorting of students into schools (due to selection by the parents, and in the case of private schools, the school administrators) and also due to residential segregation. For instance, better educated parents act better in locating their children in schools with a

better reputation and more effective teachers. As mentioned before, while parental education acts as a good proxy for family inputs, teachers' education level (master, Ph.D.) is found to be a poor proxy in the literature. Moreover, families may better know about school reputation via their social networks, etc., compared to analysts like us. Thus, utilizing family input proxies but avoiding their endogeneity due to unobserved school variables (Level 2 endogeneity) is required.

Poor proxy problem with schools can be addressed by specifying school-fixed effects, which allow to control for all fixed school inputs. However, we want to estimate δ, the parameter for private school status, and when the fixed effects specification is used, the private school status dummy D_j (with its parameter) would be dropped from the estimation. The fixed effects specification does not allow the estimation of school-level (Level-2) variables.

Castellano, Rabe-Hesketh, & Skrondal (2014), on the other hand, argue that two distinct types of private school effects can be conceived[38]. The first is named as Type-A private school effect and equals the total of the effects of School Inputs and Peer Inputs[39]. Type-B private school effects, on the other hand, only include School Inputs.

5.4.2.1 Type-A private school effects:

Type-A effects are considered to be more relevant to the parents' decisions about private school choice. This is the total added value of a private school, and parents are expected to be more interested in how much the achievement potential is for their children, be it predominantly due to the teachers or to the peers in the school.

On the other hand, it would be unfair to assess the performance of a disadvantaged school where the peer input is limited, but teachers exert the utmost effort in terms of school inputs. Hence, for accountability purposes, the Type-B effect is the more

[38] They iterate Raudenbush & Willms (1995)'s conceptualization of individual schools into school type variables.
[39] Raudenbush & Willms (1995) call these as School Practice and Contextual Effects, respectively.

appropriate type of effect (Castellano, Rabe-Hesketh, & Skrondal, 2014; Raudenbush & Willms, 1995).

The estimation of Type-A private school impact is less challenging and only requires a good set of family input measures. Type-B private school impacts are considerably more challenging to estimate.

I will start with estimating Type-A effects. I will follow the steps of the Two-Step Estimator[40], originally belonging to Raudenbush & Willms (1995) and adapted by Castellano, Rabe-Hesketh, & Skrondel (2014) for estimation of private school effects.

I first write equation (2) in the form of (3) and estimate the following school-by-cohort fixed effects model[41] where the new term f_{jt} corresponds to the school-by-cohort fixed-effects:

$$Y_{ijt} = \alpha + \boldsymbol{X'_{ijt}}\beta_w + f_i + f_{jt} + e_{ijt} \quad (3)$$

Remember the limitation mentioned above due to a potential correlation between f_i and $\boldsymbol{X'_{ijt}}$ (i.e., between innate ability and contemporaneous family resource inputs). Equation (3) suffers from Level-1 endogeneity due to the correlation between $\boldsymbol{X'_{ijt}}$ and f_i, (Papageorge & Thom, 2020; Todd &Wolpin, 2003). In the spirit of Chamberlain (1982), and Falck, Mang, & Woessmann (2017), I model this dependency explicitly:

$$f_i = \boldsymbol{X'_{ijt}}\pi_f + \omega_i \quad (4)$$

where by construction ω_i is uncorrelated with observed variables $\boldsymbol{X'_{ijt}}$. Now re-writing the outcome equation:

[40] Not related to the 2SLS.

[41] To compare with regular models with time and school fixed effects; for instance, let's assume we have 100 schools and 4 cohorts, then there we will have a total of 104 fixed effects. In the school-by-cohort fixed effect strategy, we have 100x4=400 fixed effects. For a similar strategy, see for instance Rivkin, Hanushek & Kain (2005, p.418).

$$Y_{ijt} = \alpha + \boldsymbol{X'_{ijt}}(\beta_w+\pi_f) + \omega_i + f_{jt} + e_{ijt} \quad (5)$$

The above model will produce unbiased estimates for the vector of parameters $\hat{\beta}_w+\pi_f$ as there is no Level-1 endogeneity problem now. By running the above model, I obtain predictions for and deduct them from Y_{ijt} to obtain quasi-residuals (Ballou et al., 2004):

$$\hat{Q}_{ijt} = Y_{ijt} - \boldsymbol{X'_{ijt}}(\hat{\beta}_w+ \hat{\pi}_w) \quad (6)$$

Note that $\boldsymbol{X'_{ijt}}(\hat{\beta}_w+\hat{\pi}_w)$ corresponds to the Compositional Effect of a school (Duncan, Jones, & Moon, 1998), which is solely due to variations in the prevalence of students with favorable background characteristics in a school. Even if school inputs or peer inputs do not contribute to human capital achievement (i.e., β_c=0 and β_s= 0), we would still have this compositional effect. By this deduction, we equalize the playing field between public schools and private schools (indeed, between all schools) in terms of compositional differences.

As achievement is the total of compositional inputs, contextual (peer) inputs and school inputs, and we took away the majority of the compositional part (i.e., other than ω_i), the remaining is the total of contextual and school inputs and remaining student effect ω_i from the compositional part.

The right-hand side of quasiresiduals now becomes[42]:

$$\hat{Q}_{ijt} = \alpha_{\cdot t} + \delta D_j + \boldsymbol{S'_{jt}}\beta_s + \omega_i + u_{jt} + e_{ijt} \quad (7)$$

Note that by taking out $\boldsymbol{X'_{ijt}}$ from the equation we take out its counterpart $\boldsymbol{\bar{X}'_{jt}}$, too.

[42] Following Rivkin, Hanushek, & Kain (2005), I allow for changes in tests over time through inclusion of a separate fixed effect for each year.

Student covariates have simultaneous effects at both levels. When they enter into level 1 as $\boldsymbol{X'_{ijt}}$, they automatically enter at level 2 as $\boldsymbol{\bar{X}'_{\cdot jt}}$, too[43]. Indeed, this latter Level 2 presence of student covariates leads to Level-2 endogeneity with pure school variables like private school status. It is still not possible to single out the peer effect from the school input effect. However, as the explicit peer term $\boldsymbol{\bar{X}'_{\cdot jt}}$ is gone, peer effects are now injected into School Input variables D_j, $\boldsymbol{S'_{jt}}$, and u_{jt}[44]. Thus, the estimate of δ from equation (7) is called Type-A private school effect. It represents private schools' combined contribution, which involves providing (potentially) better school inputs and also recruiting better peer groups of students compared to public schools. An estimate of δ will tell us about that.

Consistent estimates of δ, β_s, and var(u_{jt}) can be achieved by fitting equation (7) with Maximum Likelihood (Gourieroux & Monfort, 1995). I will specify a two-level multilevel model for quasi-residuals:

$$\begin{aligned} \hat{Q}_{ijt} &= \alpha_{jt} + \omega_i + e_{ijt} \\ \alpha_{jt} &= \alpha_{\cdot t} + \delta D_j + \boldsymbol{S'_{jt}}\beta_s + u_{jt} \end{aligned} \tag{8}$$

In the above model (8), the remaining compositional input ω_i, (which is the remaining part of the fixed student contribution majorly representing residual innate ability not correlated with other inputs) may still create Level-1 endogeneity problem for the parameter of private school dummy. Hanushek (1979, p. 364) argues that the innate abilities of students can be associated with school inputs and peer inputs because higher socioeconomic status families are sorted in certain neighborhoods and at the

[43] And we take $\boldsymbol{X'_{ijt}}$ off, its counterpart $\boldsymbol{\bar{X}'_{\cdot jt}}$ goes away, too.

[44] Note that in the context of the education production function framework, private school status is a school input. School inputs are formed by three pillars: School structure, school resources, and school processes. School structure refers to school characteristics, mainly represented by four factors: school location (urban, suburban, rural), school size, type of control (public, private), and religious affiliation. School resources include the quantity and quality of teachers, amenities, material resources (books, computers, etc.) and fiscal resources. School processes include school organization and management, teaching practices, and school climate (Rumberger & Palardy, 2004).

same time they manage to put their children into higher quality schools. However, associations, when family input differences are controlled, are likely to be substantively small. He also underlines that the if the students are put into classrooms based on ability-tracking, then the associations between innate abilities and school inputs are inclined to be higher if there are systematical differences in terms of school inputs across those classrooms. Tracking during lower-secondary school is not common, or if that is the case, school inputs are not expected to vary by much. Moreover, we are comparing schools to schools here, not classrooms in the same school.

Papageorge & Thom (2020)'s study into "human capital related genes" (measured by the *polygenic score*), on the other hand, implies that $\boldsymbol{X}'_{\boldsymbol{ijt}}\hat{\pi}_w$ takes a large portion away from the original f_i term, which would reduce the harm caused by ω_i significantly.[45] Their findings indicate that family inputs (especially) parental education interact strongly with f_i. Heckman and Mosso (2014) also underline that the impact of genes is also a function of socioeconomic status since the *expression* of genes is affected strongly by environmental factors.

I will also estimate a further two-level multilevel model that involves district-by-cohort random effects (v_{kt}) and district-by-cohort means of school input variables.

$$\hat{Q}_{ijkt} = \alpha_{jkt} + \omega_i + e_{ijkt}$$

$$\alpha_{jkt} = \alpha_{.t} + \delta D_j + \ell \overline{D}_k + \boldsymbol{S}'_{\boldsymbol{jkt}}\beta_s + \overline{\boldsymbol{S}}'_{\boldsymbol{..kt}}\beta_n + u_{jkt} + v_{kt} \qquad (9)$$

In the above specification, as district-by-cohort means of school inputs are included, school input variables (including the private school dummy) cannot be correlated with the district-by-cohort random term. Hence, v_{kt} will act as a district fixed effect term, and we will be controlling all district-level factors, including time-varying ones, as

[45] Van Dam (2018)'s reporting article related to the study of Papageorge & Thom (2020) further suggests that the human capital related genes are almost evenly distributed in the population.

we have not only district random terms but district-by-cohort random effects. Note that the parameters of meaned input variables (i.e., ℓ and β_n) may not be consistent, but δ and β_s will be consistent, which are the parameters of primary interest.

5.4.2.2 Type-B private school effects:

Next, I will estimate Type-B private school effects. Type-B effects are only the effects of School Inputs (school resources, school structure, and school processes) in isolation from Peer Inputs. In my conceptual model described at the beginning of the section and re-written below, $\bar{\boldsymbol{X}}'_{\boldsymbol{\cdot jt}}\beta_c$ represents Peer Inputs defined in the form of contextual effects. Contextual effect, in an example, can be expressed as the influence of studying (for a specific student i) in a school having a mean parental education $\bar{X}_{\cdot j}$ after controlling for that student's own parents' education status. Students from more educated parents are generally of higher ability and are more disciplined. Their parents tend to pressure teachers and the principal for more effort. These provide a context that facilitates learning[46].

$$Y_{ijt} = \alpha + f_t + \delta D_j + \boldsymbol{X}'_{\boldsymbol{ijt}}\beta_w + \bar{\boldsymbol{X}}'_{\boldsymbol{\cdot jt}}\beta_c + \boldsymbol{S}'_{\boldsymbol{jt}}\beta_s + f_i + u_{jt} + e_{ijt} \qquad (10)$$

The above model is a version of the correlated random effects model (also known as Mundlak Device due to Mundlak, 1978; and contextual effects due to Raudenbush, 1989). Castellano, Rabe-Hesketh, & Skrondal (2014) state that the inclusion of school means of student level covariates in addition to student covariates addresses adequately the contextual effects (peer inputs).

If there was no level 2 endogeneity, $\hat{\beta}_c$ would provide a consistent estimate of contextual effects via the parameter vector β_c. However, under the level 2

[46] Interpretation of the parameters belonging to the school-means of input variables as contextual effects is a mainstay in school effectiveness literature that heavily uses multilevel modeling (Kreft, Leeuw, & Aiken (1995), Rabe-Hesketh & Skrondal, 2022; Raudenbush & Bryk, 2001; Snijders & Bosker, 2012). It also has its use (similar to the context here) in the studies of economists such as Palardy & Rumberger (2008), Wooldridge, 2010; Rockoff et al. (2011); Hanchane & Mustafa (2012), Urquoila (2016), Arkangelksy & Imberns, 2019; and Yang (2022).

endogeneity problem described in the previous subsection, these parameters cannot be consistently estimated.

On the other hand, Altonji & Mansfield (2018) further argue, based on their theoretical general equilibrium model, that controlling for a rich set of meaned student covariates will also control for the means of remaining unobserved student covariates, too, including the problematic school-by-cohort means of f_i, especially when there is access to population-level data[47].

Even more, if the correlation between the means of student covariates $\bar{\boldsymbol{X}}'_{\boldsymbol{.jt}}$ and the school's random contribution u_{jt} is positive; and additionally, if the parameter of the contextual effect β_c is positive; then β_c will be overestimated. It is in general the case that schools that attract better peer groups also have better teachers (corr($\bar{X}'_{.jt}$, u_{jt})>0) and peer effects of these schools are expected to be positive (β_c>0)[48].

Overestimation of the contextual effect parameter, in turn, will lead to underestimation in school input variables, including the parameter of the private sector dummy $\hat{\delta}$. Thus, $\hat{\delta}$ can only be interpreted as a lower bound for the Type-B effect of private school effect. Moreover, in the same conditions, Type-A private school effect would serve as an upper limit for the Type-B private school effect (Castellano, Rabe-Hesketh, & Skrondal, 2014, p. 354).

Hence, acknowledging the impossibility of discerning contextual effects (representing peer effects) from school input effects (endogenous to family inputs), I will be able to only provide a lower limit for the Type-B private school effect estimate and discuss the rationale of adopting Type-a private school effect as the upper limit.

[47] Also, remember the arguments discussed above that innate ability is less an issue when estimating school input variables.

[48] It may also be the case that specialized programs in some disadvantaged schools imply better teachers operate in held-back schools. But first, this is not prevalent; second, these disadvantaged schools, by their background, not in the group of public schools to be compared with private schools under econometric adjustment.

5.4.3 Voucher Program Competition Effects

I will explore the second research question in this part:

Research Question 2: *Does competition from private schools increase public school effectiveness in producing human capital?*

There are three potential effects of voucher programs. First is the participation effect, which corresponds to the mean effect on the students who received the voucher. The second effect is the competition effect, which indicates whether more choice in school selection and associated higher availability of private schools increase the competitive pressures on public schools. It is assumed that this will increase, in turn, the public school quality. The third effect is related to the second one. Higher availability of private schools may also lead to cream-skimming of better public students by private schools, which may put public schools even in a more disadvantaged position. In some other cases, students who face problems in public schools may switch to private schools to a higher degree, which would lead to a negative sorting effect for private schools. As it is very hard to single out the competition effect from the cream-skimming (sorting) effect, generally, their net effect on public school outcomes is explored. Learning about the net competition effect and simultaneously elaborating on the design of the voucher program at hand will inform the design of similar programs in the future.

This research question ideally begs variation across markets in the size of private enrollment at the aggregate market level. This type of variation is present in the cross-section in many countries (Hsieh & Urquiola, 2006). Cross-market variation in private school attendance is expected to emerge if households and schools react differently against the voucher program across localities (Urquiola, 2016). Variations in the size of the supply or the mode of implementation are also expected to increase the mentioned cross-market variation. Experiments are not suitable for assessing the competition effect (or the competition effect net off cream-skimming effect) as it

would be almost impossible to manipulate private sector entry across randomly selected markets. Instead, Urquiola (2016) states that the literature has resorted to large-scale, nationwide voucher reforms that have induced large changes in private enrollment, at least in some markets. Current research so far has produced mixed results. Challenging aspects that beg further research are that the voucher programs are context dependent and generally endogenous to how they are designed (Hanushek et al., 2013).

I will estimate the competition effect of the voucher program between the 2014-15 and 2016-17 academic years. As the measure of competition, I will use the percentage of private school students in a province. My data started one year before the introduction of the voucher program. Hence, the academic year of 2013-2014 will constitute a baseline. The details of the program are described in section 3.3 of chapter 3. The large-scale voucher program had an intensive start much suited to the "voucher shock" type of design mentioned in the literature review chapter.

The criteria of the program were different at the province, school, and student levels. At the province level, the allocation rule was announced to be based on the socio-economic development level of the province (for instance: Official Gazette, August 12, 2018, no:30507)[49]. The weights for each group were as follows from level one to level six: {0.95, 0.95, 1.00, 1.00, 1.20, 1.30}. Hence, less developed provinces received more weight.

Based on the weights and the number of students in each province, the Ministry of National Education declared in each year the quotas for each province before the start of the applications by the schools and the students each year (see MoNE, 2014, p.4-6). The announced quotas each year are reported in Kahyaoğlu & Karataş (2018, p.600-613). The actual quota numbers also reflect the size of the student population. If the demand did not cover the allocation, the Ministry of National Education also stated they would make adjustments later on.

[49] Başbakanlık Mevzuatı Geliştirme ve Yayın Genel Müdürlüğü (resmigazete.gov.tr)

The quota was shared as 40 percent for the first-year students and 20 percent for each remaining years. A student could benefit from the program at any grade. None of the students in the population I analyze could enter in their first year as my scope of data ends in 2017, and the program started in 2014. My 2014 graduating cohort was leaving when the program started. The 2015 cohort could get 20 percent of the 2015 allocation, while the 2017 cohort could get the total of 20 percent from 2015, 20 percent from the 2016 quota, and 20 percent from 2017 quota allotted to the province according to the allocation rule. Moreover, a total of 75,000 subsidies were made available in 2014-15, 50,000 in 2015-16, and 15,000 in 2016-17. So each rule and yearly allocation created variation. I calculated the amount of quota that each cohort was eligible for according to the initial allocation. Conditional on the socioeconomic development level of the province, my hypothesis is that it will qualify as a good instrument to address the endogeneity of private sector concentration in a province to the quality of the private sector[50]. I will also control for the selection factors that private school entrepreneurs take into attention when making their new school location decisions. Finally, I will use as an additional instrument the vote share received by the religion-focused Welfare Party (*Refah Partisi*) in the 1995 elections for the religious school demand basis. This instrument type is the most implemented one in the literature, as described in the literature review chapter.

Addressing the endogeneity of private school size with public school quality calls for sources of supply-side variation, and religious private schools are supplied not predominantly for achievement per se but for religious-values based education. So the rationale here is not purely exogenous variation but variation due to reasons different than the achievement level and quality of incumbent public schools. Perhaps different than many other countries, religious public schools (Imam Hatips) were

[50] Gallego (2002) and Auguste & Valenzuela (2003) also use the student population at county level (market size), which is an implicit part of my instrument. As the theoretical basis, Bresnahan & Reiss (1987, 1991) suggest that per capita demand for a good is a function of total population and demographic characteristics, both of which vary across markets. Accordingly, different thresholds are formed for new entries in different markets.

increasing rapidly in numbers over the voucher program period, while the largest network of religious private schools with 22 percent of total private school enrollment exogenously abolished right in the middle of summer of 2016. These all create deep fluctuations in the share of the private sector as well as the composition of the school types in each province over such a short period of time. Table 5.1 below represents the yearly changes in the numbers of different school types.

Table 5-1 Number of Schools per School Type per year

	2014	2015	2016	2017
Ordinary public middle school	14,370	13,463	12,394	12,854
Boarding public middle school	365	353	335	325
Imam Hatip middle school	98	84	1,338	1,648
Mid. schools for students with special needs	74	125	117	120
Private middle schools (closed group)	233	241	287	0
Private middle schools (other)	663	784	1,177	1,333
Total	15,803	15,050	15,648	16,280

I will specify the following schools in a market model often used in the literature:

$$Y_{jkt} = \alpha + \delta \overline{D}_{.kt} + \boldsymbol{S'_{jkt}}\beta_w + \boldsymbol{\overline{S}'_{.kt}}\beta_c + \boldsymbol{P'_{kt}}\beta_p + u_{jkt} \quad (11)$$

Y_{jkt} is the mean TEOG score achievement of the school j in province k in year t. The variable $\overline{D}_{.kt}$ is the share of enrollment in percentage points in province k in year t. The vector $\boldsymbol{S'_{jkt}}$ includes school background variables, and $\boldsymbol{\overline{S}'_{.kt}}$ are the province-by-cohort meaned school background variables. $\boldsymbol{P'_{kt}}$ collects province-level variables, u_{jkt} is the error term that collect school and provincial level unobserved factors.

Note that the above model, in part, corresponds to the human capital production function of each school but also aims to control for the factors that are important for private school entrepreneurs' school location and operational enlargement decisions. The control variables are specified according to the factors highlighted in the literature review chapter part 2.4.2, pp. 39-40.

5.4.4 Private Schools and Gender Achievement Gap

In this section I will focus on the students' gender gap in human capital formation and answer the third and final research question:

Research Question 3: *Does the gender achievement gap differ by school sector ceteris paribus?*

I will estimate the below joint model, where Y_{ijt} is the TEOG test score of student *i* who studied in school *j* in year *t*[51], W_{ijt} is the cumulative GPA (covering the last 3 years) of the same student, G_{ijt} is the gender variable (G=1 for girls), $\bar{G}_{.jt}$ is the school-by-cohort mean of gender (i.e., girl ratio in a school-by-cohort), $\boldsymbol{X_{ijt}}$ is the vector of student level covariates corresponding to parents' and students' inputs other than gender input, $\boldsymbol{\bar{X}'_{.jt}}$ is the vector of school-by-cohort means of covariates, e_{ijt}, ε_{ijt} are the random error terms at student level other than mean gender input, and $\bar{e}_{.jt}$, $\bar{\varepsilon}_{.jt}$ are the school-by-cohort means of covariates error terms. e_{ijt} and ε_{ijt} are assumed to be correlated with the correlation term ρ. I will fit this joint model with FIML approach in the Structural Equation Modeling (SEM) modeling framework, and alternatively fit it under a Seemingly Unrelated Regression (SUREG) framework.

$$
\begin{aligned}
Y_{ijt}\text{-}\bar{Y}_{.jt} &= \tau(G_{ijt}\text{-}\bar{G}_{.jt}) + \vartheta(G_{ijt}\text{-}\bar{G}_{.jt})*D_j + (\boldsymbol{X'_{ijt}}\text{-}\boldsymbol{\bar{X}'_{.jt}})\beta_w + f_i\text{-}\bar{f}_{.j} + (e_{ijt} - \bar{e}_{.jt}) \\
W_{ijt}\text{-}\bar{W}_{.jt} &= \gamma(G_{ijt}\text{-}\bar{G}_{.jt}) + \varphi(G_{ijt}\text{-}\bar{G}_{.jt})*D_j + (\boldsymbol{X'_{ijt}}\text{-}\boldsymbol{\bar{X}'_{.jt}})\theta_w + f_i\text{-}\bar{f}_{.j} + (\varepsilon_{ijt} - \bar{\varepsilon}_{.jt}) \\
&\text{Corr}(e_{ijt} - \bar{e}_{.jt}, \varepsilon_{ijt} - \bar{\varepsilon}_{.jt}) = \rho \qquad (12)
\end{aligned}
$$

Above is a type of difference-in-differences model with school-by-cohort fixed effects[52] fitted jointly to two different measures. The estimate of τ indicates the

[51] As I am analyzing just the last year's performance of each student (when the TEOG exam is held), cohort and year coincide for each specific student.

[52] All variables, including the output variable, is mean differenced to translate the model into a fixed effects model. Fitting joint models with explicit fixed effect terms is not possible, so I absorb them with mean differencing.

average gender achievement gap in public schools, and $\tau+\vartheta$ informs on the average gender achievement gap in private schools for TEOG scores. γ and $\gamma+\varphi$ indicate the same for teachers' grades. School-by-cohort fixed effects are implemented to eliminate bias due to differential sorting of each gender, i.e., if one of the genders is sorted to better schools. For teachers' grades, school-fixed effects also control for differences in grading tendencies (including score boosting, etc.). As we observe students just once in the data set and as we are comparing students in the same cohort and same school, all school factors are being controlled. Fixed-by-cohort school effects address the unobserved school inputs if we can assume that all the students within the same school receive the same school inputs. If this assumption is valid, then any bias on the parameter estimates of the gender achievement gap should be due to unobserved family factors, which involve fixed student contribution or the error term that implicitly involve past inputs.

Note that one of the rationales for joint modeling above is that if a set of skill measures does not correspond to the same set of student abilities, then a joint model of those measures better predicts each student outcome compared to a separate model for each outcome (Jackson, 2018, p.2074). The other rationale is to assess their correlation in the error terms. The prior belief is that both measures evaluate principally the same content and cognitive skills, but teachers' grades also reflect their assessment of the non-cognitive skills, stereotyping bias, and any other factor. Thus, the expectation for the correlation coefficient ρ is that it is not very close to one, but also, it is not a low value, either.

Note that I also allow for different parameters for family inputs in each equation (but also allow for correlation in their errors) instead of the common specification mostly used in the literature, which stacks scores from teacher-scored and blindly-scored exams together (see the seminal study of Lavy, 2008). This is reasonable for the case of the same type of exams that are evaluated by the biased own teacher of the student and an unbiased external evaluator. However, when comparing test scores and teacher's grades, there could be differences not only due to the evaluator who does

the grading but also due to the evaluation method itself. For instance, school grades may depend more on homework, which implies different inputs for which different genders have different skills. For example, the existence of an own room or an own desk at home for the child may be more important for the teacher's grades than the TEOG test scores. Similarly, Fiorini & Kane (2014) states that while the time inputs from parents are highly influential for cognitive skills, non-cognitive skills tend to be insensitive to parental time inputs. Latter skills are found to depend more on the parenting style, which combines mild discipline with parental warmth.

As teacher-grades are more dependent on non-cognitive skills, allowing for different production function parameters for TEOG scores and teachers' grades will address these measure-specific aspects.

After calculating and checking with the value of ρ, I will deduct the first equation from the second.

$$Q_{ijt} - \bar{Q}_{.jt} = (\gamma - \tau)(G_{ijt} - \bar{G}_{.jt}) + (\varphi - \vartheta)(G_{ijt} - \bar{G}_{.jt}) * D_j + (\boldsymbol{X'_{ijt}} - \bar{\boldsymbol{X}}'_{.\boldsymbol{jt}})(\theta_w - \beta_w) + \epsilon_{ijt} - \bar{\epsilon}_{.jt} \quad (13)$$

In the above, the outcome now becomes a difference term. Note that the student fixed contribution term $f_i - \bar{f}_{.j}$ is eliminated now. As implied by its name, the contribution of this term to two outcomes that measure the same content and cognitive skills shall be the same (i.e., fixed). Moreover, we can expect that past inputs to affect both cognitive skills measures approximately the same in terms of cognitive skills and be eliminated by differencing. The estimate for $(\gamma - \tau)$ will indicate a residual gender achievement gap in the public school sector; $(\gamma - \tau) + (\varphi - \vartheta)$ is the residual gender achievement gap in the private school sector.

As cognitive aspects are eliminated, the residual gender gap (or *gender grading gap (gender bias)* as it is defined by Terrier, 2020) within each school sector should be related to non-cognitive aspects (Cornwell et al., 2013; Golsteyn & Schils (2015), which I will discuss in depth in the next chapter.

5.4.5 Methodological Limitations and Fixes

As with every study, this study has limitations. Indeed, the essence of the education production function framework is to make clear some of the most important limitations if they exist. As mentioned above, the production of human capital is cumulative and ideally requires direct measures of inputs and endowments. Like most of the studies, despite having access to population-level data, the data involves only the observed contemporary inputs. Unobserved contemporary inputs, availability of proxies instead of direct input measures, missingness in student data, and potential measurement errors are the challenges to be addressed. Fortunately, education production function literature also provides a set of methodological fixes for these challenges.

One fix is to explicitly recognize that there are unobserved inputs (Todd & Wolpin, 2007). School fixed effects, for instance, address the unobserved school inputs if we can assume that all the students within the same school receive the same school inputs. This assumption may not hold for all schools, but in the context of middle schools in Türkiye this assumption can be regarded as not strong[53]. I utilized this fix in addressing missing data, singling out compositional effects, and exploring gender achievement gap differences between private and public schools. However, I could fully benefit from implementing fixed effects only when I am utilizing only the variation within school-by-cohorts. An important limitation when estimating school level (level-2) variables like school sector or school type is that these variables do not change over time. If a school is a private school, it remains as a private school throughout all of the study period. When school fixed effects are applied, the school sector variable is dropped from estimation and its parameter cannot be estimated. To overcome this limitation, I used an *augmented fixed effects* approach, which enables

[53] However, for upper-secondary schools, this would be a stronger concern comparatively. Students in high schools are segregated in terms of their specialization, for instance (math track, science track, literature track, different sub-specializations in vocational schools). Inputs may differ accordingly. For this study, there is evidence for the possibility that it does not hold in terms of gender, but for other inputs it can be assumed to hold.

the estimation of the parameters of school-level variables while controlling for unobserved contemporary school inputs.

In order to control past inputs and endowments (innate ability and personal traits), a value-added approach is mainly utilized (Todd & Wolpin, 2003). However, as the students in this study did not have a centralized exam score before, I could not use this fix. Another strategy to control student fixed effects is the between-subject strategy, i.e., differencing education production functions of two different subjects (for instance, math and science). I could not use this strategy either because I only got the *overall* TEOG score and overall GPA score provided by the Ministry of National Education. I also did not have data on within-family indicators such as sibling composition, which would have helped me to single out the family fixed contribution (Dayıoğlu, Kırdar, & Tansel, 2009).

In order to fix the student unobserved variable problem, in the first research question, I discarded the correlated parts of the student fixed contribution via the fixed effects approach[54] and argued from evidence in the literature that the correlation between a school-level factor of private school status and the remaining student error is not expected to be extensive (but still be present to a degree). One rationale for this is that the individual students' endowments and past inputs vary more than school-average inputs and endowments. So it is a bigger issue for parents to make investment decisions endogenous to a child's endowments.

Teachers, however, focus on average student endowments, so their endogenous decision has a smaller scope for being altered drastically in the final year scompared to previous years. For instance, while the standard deviance of student scores is 1, the standard deviation for the school-average of the test scores is 0.54. Roughly, similar difference in variability can be envisaged for endowments.

[54] For instance, just controlling for parental education reduces the effect of human capital genes by approximately 25 percent in Papageorge & Thom (2020, table 2).

The impacts of the past inputs and the endowment are expected to reduce substantively when we control for both student and school background variables. However, the remaining part of the school error term might still be problematic. Hence, a bold causal statement will not be appropriate. However, we can argue that the estimate is informative as it controls for an important part of the endogenous effect. Moreover, most of the estimates from observational studies also prone to these limitations, so comparison with them will still be informative.

I also acknowledged that singling out Peer Inputs from School Inputs is not possible with the observational data at hand, so I estimated a lower bound estimate. To single out peer inputs, we need an exogenous variation in school inputs. Controlling for school inputs like teacher experience, expenditures-per student, vs., won't be a complete solution even if we can observe them fully as they are shown by the literature as poor proxies of direct school inputs.

In the second research question, the competition measure and instrumental variable are at the level of province, which might be seen as a high level. These measures and instrumental variables start from the school-district level or counties and might be at the level of states or countries in the literature. Province is a sweet spot (in terms of scale) for many of the mid-tier cities in Türkiye, which can be seen as the target group of the voucher program at hand. The number of private schools is too many in large cities and comparatively modest in mid-tier cities. My main instrument is driven by the preliminary allocation rule, and it is at the level of the province. This level is where we can have a conditional exogenous (but not random) variation. There are 81 provinces in Türkiye, and at the minimum, cross-sectional analysis among markets (here market = province) is sufficient for this type of analysis (Urquiola, 2016).

Here I also have four years of data (368 data instances in total) where substantial variation in the data is seen in the numbers of schools due to the political dynamics of the period, as depicted in Table 5.1 (on page 118). The outcome measure, on the other hand, is on the level of school. Hence, this is a multilevel model utilizing

information from all levels jointly, where provinces correspond to level 2, and schools to level 1.

In the third research question, I implemented a "differencing between measures" (i.e., between TEOG test scores and GPA) strategy to address student-level unobserved inputs, which include past inputs and endowments. I exploit the specific property of the TEOG exam, which implies that the TEOG questions are based on the same content and cognitive achievements that within-class examinations assess.

Missingness in student data is something out of my control for the reason of the anonymity of students. To address missingness in student data, I surveyed the missing data handling methods, chose the two most effective strategies and implemented them in conjunction with a fixed school effects approach. To account for anomalies in data, I provided detailed data summaries for each year and graphs that show whole distributions. I conducted a comprehensive digital document search to obtain info on critical school-level variables like school type, school size, location of the school, etc. I did not have the refugee status in my data, so I provided Table 3.1 (in page 48) that displays the number of refugee students in grade 8 in Turkish schools and show that the numbers of such students were limited for the period of this study.

In this study, I delimit my exploration only to effectiveness, which corresponds to human capital outcomes when inputs are held constant between public schools and private schools. However, efficiency can be another criterion for private public school comparisons. Production (or productive) efficiency delineates the conditions under which goods are produced at the lowest unit cost possible. Public schools are assumed to not operate on the production possibilities frontier and are presumed that they can produce more human capital without increasing the inputs. Public schools are believed to be not efficient. Are private schools, under more competitive pressures than public schools, efficient? As I don't have the cost information in private schools, I cannot analyze this dimension. Entrepreneur's initial investment costs (building, etc.) and profits confound how much is invested each year for actual education

provision. So, here I will provide two clues about the cost structure. The first clue is provided by the voucher program's regulatory clauses. It says that a voucher up to 1.5 times the cost of the student to the public education system can be subsidized, and the monetary value of the subsidy is announced each year (changing between 3 to 4 monthly minimum wages for a school year). The second clue is about the teacher wages in the private sector. I have been involved in quite a number of job interviews as a recruiter or as an observer with private sector teachers, generally the leading ones, for project work, and they require to disclose their current wage structure; and it was their benefit to prove the most amount as possible. My experience indicates that the experienced, better (rated) teachers in better (rated) private schools are *expected* to earn comparable wages to public school teachers of similar experience. But it should be noted that private schools recruit more teachers per students. Average of pupil-to-teacher ratio for public schools is 18.21, and for public schools is 9.63. Moreover, private school teachers are not required to serve in remote regions for part of their career in contrast to public school teachers, which may convince them to work in their preferred location at a less than their normal wage rate. Hence, I can only provide these revealed costs for a limited number of cases, and leave in-depth exploration for a future study.

CHAPTER 6

FINDINGS AND DISCUSSION

6.1 Comparative Analysis of Methods in Handling Data Missingness

In this subsection, I will compare the parameter estimates of student level variables under within-school analysis between the full information maximum likelihood (FIML) and complete cases methods. As discussed in methodology chapter, FIML utilizes all observed data and produces consistent and efficient estimates under missing at random (MAR) assumption. Complete-cases analysis only uses the data of students who have no missing data in each of the input variables. It is consistent under MAR and also more robust to missing not at random (MNAR) mechanism of data missingness compared to FIML, but it is not as efficient due to dropped observations. Table 6.1 presents the summary of student-level missingness[55].

Table 6-1 Summary Table for Missingness in Student Level Variables

	Missingness by each variable		Overall Missingness	
Student Variables:	***Number***	***%***	***Net Missing***	***Cumulative %***
TEOG Score	81,999	1.68	81,991	1.68
GPA	906	0.00	205	1.68
Mother Education	718,222	14.67	688,632	15.75
Father Education	700,747	14.31	140,777	18.62
Mother Occupation	711,911	14.54	40,596	19.45
Father Occupation	684,272	13.98	15,875	19.78
Own Room at Home	264,519	5.40	25,428	20.29
Home Status	266,748	5.45	4,967	20.40
Family Income	525,358	10.73	160,386	23.67
Cumulative Missing:	-	-	**1,158,857**	**23.67**
Data Available for Complete Cases:			**3,736,627**	**76.33**
Total Number of Observations:			**4,895,484**	**100.00**

[55] Note that school level variables to be utilized in this study do not have missingness problem.

The net missing column in Table 6.1 above represents an additional number of students who have missing values as we go down each cell of the column and provides information about the extent of joint-missingness. The cumulative missing column, on the other hand, indicates how many students in total are discarded by the addition of each new input variable.

The following school-by-cohort demeaned model is utilized separately for FIML and Complete-Cases to obtain parameter estimates for student-level variables.

$$Y_{ijt}\text{-}\bar{Y}_{\cdot jt} = (\boldsymbol{X'_{ijt}}\text{-}\boldsymbol{\bar{X}'_{\cdot jt}})\beta_x + (e_{ijt} - \bar{e}_{\cdot jt}) \quad (14)$$

Before running the models, the means for each variable needs to be calculated. As described in the Methodology chapter, I use the observed means for variables with no missing values (for instance gender or number of siblings) as they correspond to the true means in the population. As FIML use all available information, all observed values (whether or not they have missingness in values of some other variables for that specific student) in each school-by-cohorts will be used when calculating the covariate means within each school-by-cohort. For the Complete-Cases approach, I will only use observations of students who are in the Complete-Cases group.

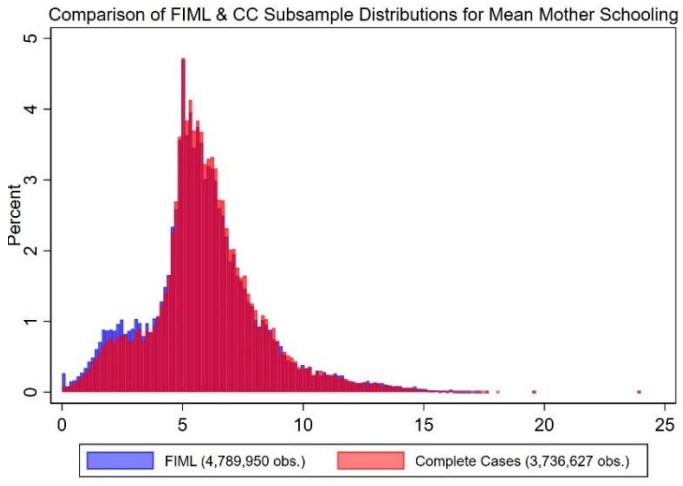

Figure 6.1 Comparison of Mean Distributions under FIML & CC Approaches

In Figure 6.1 above, the distribution of means for the *School-by-Cohort Mean Mother Schooling* (years spent in education) variable is displayed as an example. The distributions are highly close despite the difference of over 1 million observations (that are used for calculating means) between the two methods.

In the below table, the parameter estimates from FIML and Complete-Cases approaches are reported. For FIML part, I utilized the Structural Equation Modeling (SEM) suite of STATA 17, which allows FIML specification for single-level models. For complete-cases analysis, I used regression with robust errors. Note that both the outcome and the control variables are entered into analysis with their school-by-cohort de-meaned forms to eliminate school-level differences across them.

Table 6-2 Within School-by-Cohort Model Results for FIML & Complete-Cases[56]

Estimation method:	**FIML**						**Complete Cases**					
School-by-cohort	**Public Schools**		**Private Schools**		**All Schools**		**Public Schools**		**Private Schools**		**All Schools**	
centered controls:	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*
Girl	**0.311**	*0.000*	**0.127**	*0.000*	**0.312**	*0.000*	**0.312**	*0.000*	**0.125**	*0.000*	**0.312**	*0.000*
Girl*Private School	-	-	-	-	**-0.191**	*0.000*	-	-	-	-	**-0.195**	*0.000*
Age of the student	-0.099	*0.000*	-0.009	*0.052*	-0.098	*0.000*	-0.101	*0.000*	-0.024	*0.000*	-0.095	*0.000*
Age data is mismeasured	-0.009	*0.035*	-0.055	*0.111*	-0.009	*0.027*	-0.002	*0.634*	-0.030	*0.449*	-0.004	*0.464*
Mother Schooling	0.034	*0.000*	0.029	*0.000*	0.034	*0.000*	0.034	*0.000*	0.030	*0.000*	0.034	*0.000*
Father Schooling	0.048	*0.000*	0.033	*0.000*	0.048	*0.000*	0.049	*0.000*	0.034	*0.000*	0.048	*0.000*
Employed Mother	0.039	*0.000*	-0.029	*0.000*	0.037	*0.000*	0.037	*0.000*	-0.027	*0.000*	0.035	*0.000*
Employed Father	0.053	*0.000*	0.018	*0.425*	0.055	*0.000*	0.054	*0.000*	0.036	*0.000*	0.056	*0.000*
Public Official Mother	0.177	*0.000*	0.060	*0.000*	0.138	*0.000*	0.147	*0.000*	0.071	*0.000*	0.113	*0.000*
Public Official Father	0.171	*0.000*	0.116	*0.000*	0.168	*0.000*	0.173	*0.000*	0.114	*0.000*	0.168	*0.000*
Family Income	0.082	*0.000*	-0.036	*0.000*	0.080	*0.000*	0.086	*0.000*	-0.032	*0.000*	0.083	*0.000*
Siblings	-0.030	*0.000*	-0.005	*0.011*	-0.030	*0.000*	-0.034	*0.000*	-0.006	*0.004*	-0.033	*0.000*
Parents not live together	-0.169	*0.000*	-0.176	*0.000*	-0.169	*0.000*	-0.186	*0.000*	-0.153	*0.000*	-0.180	*0.000*
Public Lodging Home	0.079	*0.000*	0.054	*0.000*	0.077	*0.000*	0.083	*0.000*	-0.004	*0.000*	0.077	*0.000*
Own Home of Family	0.061	*0.000*	-0.005	*0.164*	0.059	*0.000*	0.065	*0.000*	0.050	*0.260*	0.061	*0.000*
Own Room of Student	0.111	*0.000*	0.007	*0.230*	0.111	*0.000*	0.109	*0.000*	0.002	*0.796*	0.107	*0.000*
Constant	0.000	*0.000*	0.002	*0.000*	0.000	*0.000*	0.016	*0.000*	0.013	*0.000*	0.016	*0.000*
No of school-by-cohorts:	58,603		4,718		63,321		57,433		4,666		62,099	
No of unique students:	4,657,700		237,784		4,895,484		3,569,250		167,377		3,736,627	

*Coefficients indicate change in terms of standard deviations in standardized TEOG score for one unit increase in a control variable.

[56] The variable "*Age data is mismeasured*" belongs to a group of (on average 9,000 students per year) whose age is extremely young to enter TEOG exam. Age info is calculated from birth info coming from Ministry of Interior and normally, there should be no error. If there is an error, this signals that this child is most probably from a disadvantaged group. She can be a refugee or a Turkish child from a remote place whose birth certificate was not registered by birth. Generally, these children's birth certificate is activated when they start school by the school principal. Hence there are TEOG participants who are at the age of 8. When the database recruits info from the birth certificate data, these type of error can happen. I control this with a dummy. It is seen that the overall effect of this is not large.

One presumed disadvantage of the Complete-Cases approach is the loss of precision. However, it is observed from Table 6.2 that this does not bear much of a challenge due to the large data size, as suggested by Allison (2014). Moreover, Figure 6.2 below reports the distributions of predicted values from both models. The non-overlap region is quite limited. I will utilize multilevel modeling techniques, and they are only compatible with the complete-cases approach. Hence, I will continue with that subsample until the third research question. I will utilize both FIML and seemingly unrelated regression approaches at that section.

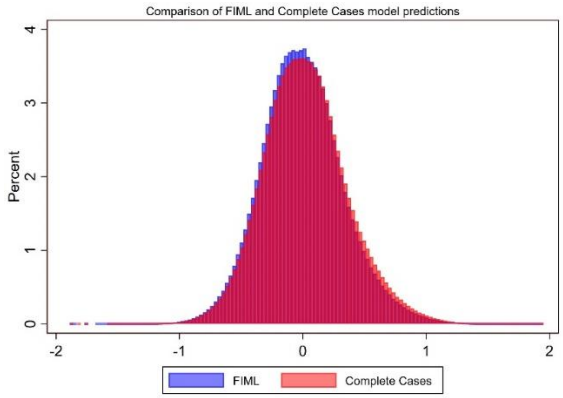

Figure 6.2 Predictions from FIML and Complete-Cases

6.2 Private School Effects

In Table 6.3, I report the fixed effects regression findings. This time I use all categories of each family input variable, so the graduation of effects can be observed in more detail. The parameter estimates are, as expected, in line with the estimates of the previous subsection. Note that these parameter estimates are after controlling for school-by-cohort fixed effects, so they are adjusted for time-fixed and time-varying school differences. Parental education and gender inputs have the major effect sizes in relation to the outcome measure. Parental occupation and family income are next. Single parent and own room are close to them in terms of influence. The signs of the effects all along are as expected.

Table 6-3 School-by-Cohort Fixed Effects Model Parameter Estimates

Estimation method:	**School-by-Cohort Fixed Effects Regression**					
School-by-cohort centered	**Public Schools**		**Private Schools**		**All Schools**	
controls:	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*
Girl	0.312	*0.000*	0.125	*0.000*	0.304	*0.000*
Age of the student	-0.099	*0.000*	0.024	*0.000*	-0.096	*0.000*
Age is problematic dummy	-0.002	*0.648*	-0.031	*0.396*	-0.001	*0.786*
Mother Education dummies						
Illiterate	base		base		base	
Some Primary School	0.058	*0.000*	0.099	*0.000*	0.059	*0.000*
Primary School	0.133	*0.000*	0.116	*0.000*	0.134	*0.000*
Lower-secondary school	0.172	*0.000*	0.129	*0.000*	0.173	*0.000*
High School	0.357	*0.000*	0.257	*0.000*	0.355	*0.000*
Institute	0.552	*0.000*	0.392	*0.000*	0.535	*0.000*
Undergraduate	0.567	*0.000*	0.423	*0.000*	0.543	*0.000*
Masters	0.560	*0.000*	0.445	*0.000*	0.546	*0.000*
PhD	0.305	*0.000*	0.470	*0.000*	0.428	*0.000*
Father Education dummies						
Illiterate	base				base	
Some Primary School	0.049	*0.000*	-0.007	*0.924*	0.050	*0.000*
Primary School	0.108	*0.000*	-0.029	*0.676*	0.108	*0.000*
Lower-secondary school	0.205	*0.000*	0.004	*0.958*	0.206	*0.000*
High School	0.390	*0.000*	0.133	*0.051*	0.391	*0.000*
Institute	0.617	*0.000*	0.304	*0.000*	0.612	*0.000*
Undergraduate	0.672	*0.000*	0.357	*0.000*	0.660	*0.000*
Masters	0.686	*0.000*	0.391	*0.000*	0.676	*0.000*
PhD	0.666	*0.000*	0.417	*0.000*	0.688	*0.000*
Mother Occupation dummies						
Unemployed	base				base	
Other occupation	-0.033	*0.000*	0.015	*0.000*	-0.031	*0.000*
Private sector worker	-0.068	*0.000*	-0.015	*0.115*	-0.066	*0.000*
Self-employed	-0.011	*0.062*	0.007	*0.656*	-0.014	*0.020*
Public sector worker	-0.035	*0.001*	0.160	*0.506*	-0.032	*0.001*
Retired	0.052	*0.000*	0.025	*0.062*	0.045	*0.000*
Public sector official	0.036	*0.000*	0.034	*0.001*	0.029	*0.000*
Army/Internal affairs officer	0.028	*0.000*	0.024	*0.000*	0.170	*0.000*
Teacher/Lawyer/Prosecuter	0.094	*0.000*	0.062	*0.000*	0.066	*0.000*
Father Occupation dummies						
Unemployed	base				base	
Other occupation	0.050	*0.000*	0.019	*0.350*	0.050	*0.000*
Private sector worker	0.075	*0.000*	0.051	*0.014*	0.075	*0.000*
Self-employed	0.120	*0.000*	0.012	*0.587*	0.117	*0.000*
Public sector worker	0.096	*0.000*	0.114	*0.000*	0.098	*0.000*
Retired	0.185	*0.000*	0.114	*0.000*	0.185	*0.000*
Public sector official	0.195	*0.000*	0.128	*0.000*	0.194	*0.000*
Army/Internal affairs officer	0.167	*0.000*	0.106	*0.000*	0.167	*0.000*
Teacher/Lawyer/Prosecuter	0.261	*0.000*	0.118	*0.000*	0.252	*0.000*
Family Income dummies						
Very low	base				base	
Low	0.097	*0.000*	0.035	*0.128*	0.097	*0.000*
Medium	0.215	*0.000*	0.034	*0.297*	0.217	*0.000*
High School	0.267	*0.000*	0.034	*0.769*	0.266	*0.000*
Very high	0.233	*0.000*	0.034	*0.141*	0.227	*0.000*
Sibling dummies						
No siblings	base				base	
One sibling	0.005	*0.001*	0.022	*0.000*	0.007	*0.000*
Two siblings	-0.052	*0.000*	0.001	*0.811*	-0.048	*0.000*
Three siblings	-0.095	*0.000*	-0.013	*0.132*	-0.091	*0.000*
Four or more siblings	-0.110	*0.000*	-0.020	*0.112*	-0.107	*0.000*
Parents not live together	-0.170	*0.000*	-0.146	*0.000*	-0.167	*0.000*
Home status dummies						
Rented house	base				base	
Public Lodging	0.074	*0.000*	0.046	*0.000*	0.073	*0.000*
Family's Own home	0.058	*0.000*	-0.002	*0.588*	0.566	*0.000*
Own room at home	0.102	*0.000*	-0.002	*0.685*	0.102	*0.000*
Constant term	0.550	*0.000*	0.239	*0.015*	0.545	*0.000*
No of School-by-cohorts:	57,433		4,666		62,099	
No of Students:	3,569,250		167,377		3,736,627	

I obtained the predictions for $\boldsymbol{X'_{ijt}}(\beta_w+\pi_f)$ obtained from presented in Table 6.3, and discarded them from student standardized TEOG scores to obtain quasi-residuals. In Figure 6.3 below, how the variation has become less wide is presented:

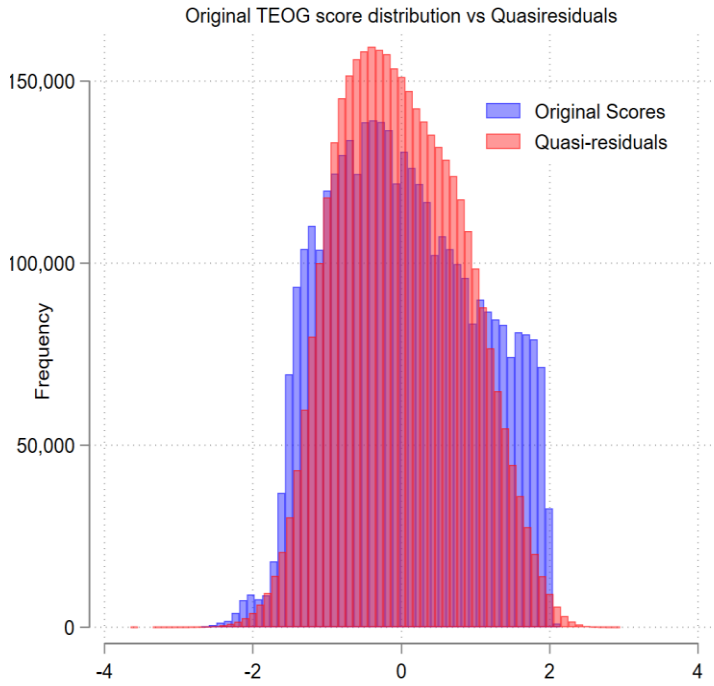

Figure 6.3 TEOG Score Distribution versus Quasiresidual Distribution

The findings from two-level multilevel analysis is reported in Table 6.4. The estimates are provided both for a single private school dummy and for six different school types. The unadjusted parameter estimate of the private school dummy is **1.31** standard deviations. We observe from the first column that about half of this is accounted for the compositional effects that arise due to sorting. Further adjusting with other school-level inputs doesn't make much difference. The last column also employs district-by-cohort random effects[57], which leads to a slight increase in the parameter estimates.

[57] The combined use of district-by-cohort random effects and district-by-cohort means of pure school variables (school type, school size, etc.) approximately replicates using district-by-cohort fixed effects in terms of consistent estimation of the parameter(s) of "private school status" or "school type". The replication is almost exact when number of schools in a district is more than 50. The typical number of schools in a district is in the range of 30-40 in Türkiye.

In terms of findings, using standardized test score measure (to mean zero a standard deviation 1), a positive effect of between 0.58-0.62 standard deviation for Type-1 private school effect, which corresponds to the combined effect of school inputs and the peer inputs. This type of effect is not estimated in Economics literature; however, it is the least problematic estimate and is relevant for parental school choice decisions.

Table 6-4 Type-A School Effects (School Inputs + Peer Inputs)

	Coef.	*p-val.*	Coef.	*p-val.*	Coef.	*p-val.*
Private school dummy	-		-			
School type dummies						
Ordinary public middle school	base	base	base	base	base	base
Boarding public middle school	0.088	*0.000*	0.121	*0.000*	0.059	*0.000*
Imam Hatip middle school	0.012	*0.036*	-0.002	*0.759*	0.005	*0.397*
Mid. schools for students with special needs	-0.817	*0.000*	-0.850	*0.000*	-0.803	*0.000*
Private middle schools	**0.581**	*0.000*	**0.585**	*0.000*	**0.623**	*0.000*
Cohort dummies:						
2013-2014	base		base		base	
2014-2015	-0.005	*0.121*	-0.025	*0.000*	-0.006	*0.000*
2015-2016	0.054	*0.000*	0.020	*0.000*	0.055	*0.000*
2016-2017	0.292	*0.000*	0.258	*0.000*	0.293	*0.000*
District's socioeconomic development						
Category 1 (most developed)	-		base		-	
Category 2	-		-0.053	*0.258*	-	
Category 3	-		-0.024	*0.000*	-	
Category 4	-		-0.023	*0.000*	-	
Category 5	-		-0.072	*0.000*	-	
Category 6 (least developed)	-		-0.207	*0.000*	-	
Provincial mean Teacher-Pupil ratio gap between public and private schools	-		-0.010	*0.000*	-	
Location type						
Village	-		base		base	
Ville	-		-0.009	*0.000*	0.007	*0.324*
City	-		0.024	*0.000*	0.056	*0.000*
School size category						
x<=23	-		base		base	
23<x<=43	-		-0.040	*0.000*	-0.036	*0.000*
43<x<=68	-		-0.033	*0.000*	-0.024	*0.000*
68<x<=101	-		-0.004	*0.000*	0.012	*0.030*
101<x	-		0.043	*0.000*	0.074	*0.000*
Double sessions school dummy	-		-0.507	*0.000*	-0.022	*0.000*
District mean of private school	NO		NO		NO	
District means of school types	NO		NO		YES	
District mean of location type	NO		NO		YES	
District mean of school size	NO		NO		YES	
District mean of double session school	NO		NO		YES	
Constant term	-0.094	*0.000*	0.050	*0.162*	0.006	*0.949*
Random intercept var. (District-by-Cohort)	-		-		0.030	
Random intercept var. (School-by-Cohort)	0.065		0.057		0.029	
Residual variance (Student)	0.605		0.605		0.666	
Number of individual schools						
Number of students						

Estimates remain robust for adding other structural school inputs. Correlational random effects at the level of district-by-cohorts slightly increase the effect estimate. This formulation is analogous to adding district-by-cohort effects and controls for all district-by-cohort and higher-level fixed influences. A size of 0.62 is a medium to high effect size in general and a quite high effect size for education research.

Table 6.5 below reports parameter estimates for the Type-B private school effects.

Table 6-5 Type-B School Effects (School Inputs) –Part 1

	School Types	
	Coefficient	***p-value***
Private School	-	-
School type:		
Ordinary public middle school	base	base
Boarding public middle school	0.188	*0.000*
Imam Hatip middle school	-0.033	*0.000*
Mid. schools for students with special needs	-0.755	*0.000*
Private middle schools	**0.256**	*0.000*
Girl	0.303	*0.000*
School-mean Girl	0.104	*0.000*
Age	-0.093	*0.000*
School-mean Age	-0.034	*0.000*
Mother's schooling	0.034	*0.000*
School mean mother's schooling	0.010	*0.000*
Father's schooling	0.048	*0.000*
School mean father's schooling	0.033	*0.000*
Working mother	-0.036	*0.000*
School-mean working mother	0.047	*0.000*
Working father	0.057	*0.000*
School mean working father	-0.111	*0.000*
Public official mother	0.137	*0.000*
School mean public official mother	0.071	*0.014*
Public official father	0.167	*0.000*
School mean public official father	0.170	*0.000*
Family income	0.082	*0.000*
School mean family income	-0.045	*0.000*
Siblings	-0.032	*0.000*
School mean siblings	-0.072	*0.000*
Parents not live together	-0.173	*0.000*
School mean parents not living together	-0.418	*0.000*
Family living in public housing	0.077	*0.000*
School mean family living in public house	-0.182	*0.000*
Family own their home	0.060	*0.000*
School mean family own their home	-0.045	*0.000*
Own room at home	0.107	*0.000*
School mean own room at home	0.141	*0.000*

Table 6-6 Type-B School Effects (School Inputs) – Part 2

Cohort dummies:		
2013-2014	base	
2014-2015	-0.010	*0.000*
2015-2016	0.039	*0.000*
2016-2017	0.266	*0.000*
District's socioeconomic dev.		
Category 1 (most developed)	base	
Category 2	-0.004	*0.258*
Category 3	0.067	*0.000*
Category 4	0.078	*0.000*
Category 5	0.095	*0.000*
Category 6 (least developed)	0.065	*0.000*
Provincial mean Teacher-Pupil ratio gap between public and private schools	-0.003	*0.000*
Location type		
Village	base	
Ville	-0.043	*0.000*
City	-0.024	*0.000*
School size category		
x<=23	base	
23<x<=43	-0.042	*0.000*
43<x<=68	-0.054	*0.000*
68<x<=101	-0.048	*0.000*
101<x	-0.042	*0.000*
Double sessions school dummy	-0.023	*0.000*
Constant term	0.800	*0.000*
Random intercept var. (School-by-cohort)	0.042	*0.000*
Residual variance (Student)	0.609	*0.000*
Number of Students	3,669,768	
Number of Schools-by-Cohorts	60,409	

As isolating the effects of school inputs from peer inputs is not possible with observational data, I estimate a lower bound of 0.254 standard deviations for the Type-2 effect, which corresponds only to the private school effect due to school inputs. This effect is the most estimated effect in the literature, and the typical estimate is around 0.20 standard deviations (see page 35 in this study).

As we discussed in the methodology chapter, if schools that attract better peer groups also have better teachers, and at the same time, if peer effects of these schools are expected to be positive, then we can set the Type-A effect as an upper bound for Type-B effect. Jackson (2010) points out that schools acquire their reputations over longer time frames. Hence, variations in peer inputs across schools observed in cross-

sectional data are, in fact, associated with the account of historical variations in all accumulated school inputs over time. He underlines that this is important and cannot be created in a randomized experiment. All else equal, he argues that teachers are expected to prefer to teach within high-achieving, more elite secondary schools. Hence, it is expected to observe large differences in teacher characteristics across schools.

In the Turkish school system, teachers gain service points over time that they can use during their appointment to a new school and this process is egalitarian. However, there are ways to move to the school you want. One of them is the centralized school vice-principal examination. If you are successful in the exam and appointed to any school, you can withdraw from the post and choose a school of your interest as a standard teacher. There are some other strategies similar to this. School principals also have a say to a certain extent and can collect better teachers into their schools. Better schools are also located in higher socioeconomic quarters, and for the more skillful teachers who are in the private-tutoring business after school hours, these schools are economically more advantageous. This is expected to motivate them to move to those schools.

One interesting thing in Table 6.4 and Table 6.5 is the contrasting performance of boarding middle schools (*pansiyonlu bölge ortaokulları*). In terms of Type-A effects they are not much different than ordinary public schools. However, the lower limit of Type-B effect is almost on par with private schools. The reason behind this is their disadvantaged backgrounds, which make the Peer Inputs in these schools weak. It is seen that school inputs have compensated for the disadvantaged background. Let's have a look at the background variables mostly used in literature (Sirin et al., 2005) to define socioeconomic status:

Table 6-6 Average Socioeconomic Status Indicators per School Type

School Type:	Mother Schooling	Father Schooling	Employed Mother	Employed Father	Family Income
Ordinary public middle school	5.67	7.21	0.21	0.95	1.89
Boarding public middle school	3.80	5.49	0.21	0.90	1.65
Imam Hatip middle school	5.84	7.84	0.17	0.96	1.99
Mid. schools for students with special needs	5.02	6.43	0.19	0.91	1.53
Private middle schools (closed group)	10.76	12.53	0.45	0.99	2.76
Private middle schools (other)	11.86	12.86	0.56	0.99	2.76

Boarding school students almost always have the most disadvantaged families in terms of all of the socioeconomic status indicators in Table 6.6 and demographic indicators in Table 6.7. Only for the employed mother criterion, Imam Hatip middle schools have a lower employment ratio[58] compared to Boarding middle schools. I had discussed the mismeasured age problem before and expressed that this may indicate disadvantaged status. Again boarding middle schools seem to face this problem the most, almost ten times more than other schools. It is likely that their primary school teachers guided these students into boarding middle schools.

Table 6-7 Average Demographics for each School Type

School Type:	Age Problem	Girl	Single Parent	No of Siblings	District SEDI*
Ordinary public middle school	0.008	0.48	0.09	1.91	2.61
Boarding public middle school	0.026	0.45	0.07	2.79	4.69
Imam Hatip middle school	0.002	0.52	0.07	1.86	2.42
Mid. schools for students with special needs	0.002	0.39	0.11	1.91	1.70
Private middle schools (closed group)	0.003	0.46	0.06	1.21	1.94
Private middle schools (other)	0.002	0.46	0.11	0.94	1.65

*** Lower value indicates higher socioeconomic development level**

6.3 Voucher Program Competition Effects:

In the Table 6.9 below, the results from the model that relates school mean TEOG test scores on private school concentration on each province.

[58] Interestingly, in Table 6.7 below, Imam Hatip schools also have the highest girl-to-boy ratio. This is in line with the tendency in Islamic countries that religious schools increase girls' schooling. As discussed before, Imam Hatips are vocational-religious schools and the most direct vocations (working as an *imam* or *muezzin* in a mosque) they train is not suitable to be carried by women.

Table 6-8 Findings for (net) Private Competition Effects on Public Schools

School-mean of standardized student TEOG score:	**OLS**		**Single IV**		**Two IVs**	
	Coefficient	***p-value***	**Coefficient**	***p-value***	**Coefficient**	***p-value***
Private school ratio in the province	**-0.015**	*0.000*	**-0.010**	*0.000*	**-0.012**	*0.000*
Public school type:						
Ordinary middle schol	base		base		base	
Boarding middle school	0.117	*0.000*	0.117	*0.000*	0.114	*0.000*
Imam Hatip middle School	-0.040	*0.000*	-0.040	*0.000*	-0.039	*0.000*
School for specialized-needs students	-0.988	*0.000*	-0.988	*0.000*	-0.978	*0.000*
Year 2015 dummy	-0.029	*0.000*	-0.032	*0.000*	-0.033	*0.000*
Year 2016 dummy	-0.045	*0.000*	0.035	*0.000*	0.038	*0.000*
Year 2017 dummy	-0.266	*0.000*	0.259	*0.000*	0.261	*0.000*
School size:						
School size 23<x<=43	-0.057	*0.000*	-0.057	*0.000*	-0.054	*0.000*
School size 43<x<=68	-0.069	*0.000*	-0.070	*0.000*	-0.066	*0.000*
School size 68<x<=101	-0.060	*0.000*	-0.061	*0.000*	-0.057	*0.000*
School size >101	-0.044	*0.000*	0.046	*0.000*	-0.040	*0.000*
Double sessions school	-0.039	*0.000*	-0.039	*0.000*	-0.039	*0.000*
School-mean mother years of schooling	0.083	*0.000*	0.083	*0.000*	-0.084	*0.000*
School-mean father years of schooling	0.097	*0.000*	0.097	*0.000*	0.097	*0.000*
School mean family income	0.107	*0.000*	0.108	*0.000*	0.108	*0.000*
Province SEDI category 2	-0.068	*0.000*	-0.058	*0.000*	-0.066	*0.000*
Province SEDI category 3	-0.030	*0.000*	-0.016	*0.035*	-0.032	*0.000*
Province SEDI category 4	0.028	*0.000*	0.041	*0.000*	0.029	*0.001*
Province SEDI category 5	0.006	*0.000*	0.021	*0.029*	0.002	*0.862*
Province SEDI category 6 (least dev.)	0.055	*0.000*	0.054	*0.000*	0.025	*0.103*
Gap between public and private schools in teacher-pupil-ratio (province level)	-0.010		-0.010	*0.000*	-0.010	*0.000*
Province mean of mother schooling	0.095	*0.000*	0.085	*0.000*	0.080	*0.000*
Province mean of father schooling	-0.144	*0.000*	-0.139	*0.000*	-0.141	*0.000*
Province mean of family income	0.160	*0.000*	0.190	*0.000*	0.188	*0.000*
Province variance of mother schooling	-0.054	*0.000*	-0.046	*0.000*	-0.045	*0.000*
Province variance of father schooling	0.206	*0.000*	0.181	*0.000*	0.190	*0.000*
Province variance of family income	-0.303	*0.000*	-0.278	*0.000*	-0.315	*0.000*
Constant	-0.603	*0.000*	-1.433	*0.000*	-1.381	*0.000*
No of school-by-cohorts:	46,129		45,893		45,893	
R-square:	0.64		0.67		0.67	

First, I shall note that the control variables reflect both the education production function for schools and the location selection decision factors of private school entrepreneurs documented in the literature (see pages 38-39 of this study). Comparison is restricted to urban areas and public schools.

I find a competition effect of -0.045 standard deviations on the mean of a typical school from one standard deviation increase in the share of private school students in the same province (Table 6.6 reports the effect of 1 percentage point increase in the private school concentration. The standard deviation of private sector concentration itself is 3,003). The instrumental variable estimate of the same effect is -0.036. The size of the effect is in line with the modest impact sizes seen in the literature. The larger part of the literature finds modest positive competition effects. However, some of the most influential contributions like Dee (1998) found -0.023 sd, Husted & Kenny (2000) document -0.060 sd, and Geller, Sjoquist & Walker (2001) detect between -0.018 to -0.137 sd competition effects.

The tests for two-IVs indicate a Cragg-Donald Wald statistic of 23,000 (shall be bigger than 10); a Sargan statistic of 0.55 (it is the overidentification test for both instruments with Ho: at least one of the instruments is valid, so it must not be rejected and shall be larger than 0.05), and a Hausman statistic for endogeneity test for endogenous regressors is 0.08 (shall be rejected if it the regressor is endogenous). The statistics indicate that at least one of the instruments is valid, but there is no evidence for the case that private schools' location decisions are endogenous to public school quality. It might be that the period covered in the study has large public interventions that overcome market dynamics.

The negative impact estimate is likely related to the "voucher shock" design of the voucher, which did not have the complementary accountability component. It was likely that the voucher program was not planned to be long-lasting, but instead, it was aiming for a fast transition of public school students into private schools. As discussed in the literature review section (2.4.2), this design often produces negative outcomes in terms of net competition effects. The institutional and political backdrop summarized in chapter 3 provides several insights to this extent.

Moreover, as I also have actual voucher assignments to each province for the 2015-16 and 2016-17 cohorts, I instrumented the actually assigned voucher concentration with

the conditionally exogenous preliminary allocation for these two years. I find an effect of -0.029 sd. This time endogeneity test was not rejected. Hence, evidence suggests that the assigned voucher concentration were endogenous with public school quality.

It should be acknowledged that the duration of the study period is comparatively short, so it is not possible to analyze medium to longer term impacts. However, the immediate effects of voucher programs also bear importance as most of the shuffling in the school system happens in this period. High negative impacts from the beginning bear the risk for negative competition effect for public schools, which corresponds to the case that the public schools reduce their effort after losing their higher achieving students.

6.4 Private Schools and Gender Achievement Gap

Before starting the discussion about the gender achievement gap, we shall remind you that the gender *attainment* gap was present both in the overall school system and in private schools. The girl-boy ratio was 0.934 instead of close to 1, which shows that a significant proportion of girls were not self-selected into the middle school population. The girl-boy ratio in private schools was even lower at 0.869. This is concerning in terms of attainment, but also poses a concern for the exploration of gender achievement gap. While gender is randomly determined at birth, sample selection and consequent school selection are not random. When we look at how boys and girls compare in the in-school population, we observe from Table 6.10 that girls have a slightly more advantageous background. However, this slight difference provides more of evidence that the sample selection of girls is due to their family backgrounds or household endowments. Moreover, it could also be argued to a certain extent that there is not much room for sample selection due to inherent ability when all other family inputs are so close to each other among the two gender groups. The similarity of family inputs between each gender group also means that the regression approach is well suited to capture the differences between the two groups even if the functional form assumptions on the education production function are strong.

Table 6-9 Comparison of Boys and Girls by Family Inputs

Family Inputs	Boys	Girls
Mother Schooling	5.896	5.959
Father Schooling	7.437	7.496
Employed Mother	0.224	0.222
Employed Father	0.950	0.950
Public Official Mother	0.032	0.033
Public Official Father	0.091	0.092
Family Income	1.929	1.928
Siblings	1.848	1.914
Single Parent	0.088	0.090
Family Owns the House	0.664	0.660
Public Housing	0.017	0.017
Student Has Own Room	0.430	0.441
N	2,394,938	2,233,753

In terms of the gender achievement gap, which compares TEOG scores or GPAs of students across both genders, we have observed from Figures 4.6 and 4.7 (from data description chapter, pp. 73-74) that girls' score and grade distributions stochastically dominated boys' distribution in both measures for public schools. For private schools, only in the highest score range do boys have higher propensity for TEOG score distribution. In terms of GPA, girls' distribution stochastically dominates boys' distribution in the private school population, too. While these are descriptive, they still hint the existence of a gender achievement gap even after controlling other inputs in an education production framework.

Golsteyn & Schils (2014) states that the determinants of the gender achievement gap have remained largely unexplored. Burgess et al. (2004) find that mostly explored observable school inputs such as size, selection regime, and religious denomination or mode of financing do not predict the gender achievement gap. Bharadwaj et al. (2016) explored, via PISA (2006 and 2009) data, the roles played by parents, classroom environments, and individual characteristics of the students. He concludes that none of these inputs seem to be able to account for this gap. Lai (2009) also finds that girls generally receive more parental inputs, and exert higher effort compared to boys. However, a substantial proportion of the gender achievement gap remains unexplained

even after controlling for those factors. Incorporating the school (or even classroom) fixed effects does not explain the gap, either, which implies sorting across schools shall not be strongly related to the gender achievement gap.

I contribute to this literature by providing new evidence from Türkiye, an important emerging economy. Bedard & Cho (2010) had to discard at that time Türkiye from their survey for OECD countries on early gender test score gaps[59] due to the sizeable numbers of girls who drop out of school before grade 8. So, the findings here shall also compensate for that gap. Moreover, by also focusing on the differences of the gender achievement gap in public and private school sectors, I explore a new dimension of the gender achievement gap that previously was not explored. My findings indicate that there is an observable school characteristic that predicts the gender achievement gap, which is the private school status.

The findings here also add to the small but fast-growing literature on *comparing different measures* (i.e., standardized tests vs. teachers' grades) of the same learning content and the same set of skills, which started with the seminal study by Lavy (2008) on grade discrimination; and expanded by (Burgess & Greaves, 2013; Cornwell, Mustard, & Van Parys, 2013; Falch & Naper, 2013; Rangvid, 2015; Jackson, 2018; Lavy & Sand, 2018; Terrier, 2020, Graetz & Karimi, 2022). Exploring this research area is important since grading differences is one of the important mechanisms by which non-cognitive skills affect important life outcomes. Brian (2002) and Jackson (2018) argue that the coefficient of a non-cognitive skill is comparable in size to the coefficients of socioeconomic status or cognitive ability.

I will highlight some of the methodological advantages due to institutional setup and the data availability in this study. Among them are: (i) in some of the studies, only a selective sample of students take the standardized cognitive test, which brings strong challenges for unconfounded estimation (Graetz & Karimi, 2022). As all students had to take the TEOG exam, the selective sample problem is not present here. (ii) Content

[59] The study utilized data of period between 1995 and 2003.

and skill-set differences between the standardized test and the teachers' teaching programs complicate the comparison and differencing between measures. TEOG and school exams test the same achievements coming from the same official curriculum, so this challenge is not present in this study. (iii) small sample sizes and low detecting power are highlighted as a problem in the literature (Terrier, 2020); the population level data here eliminates this limitation, (iv) I employ a joint modeling framework that allows for more flexibility in terms of functional forms, which at the same time addresses missing data problem that is often not addressed in this literature, (v) Gneezy, Niederle, and Rustichini (2003) and Shurchkov (2012) suggest that males perform relatively better in competitive environments and girls struggle when they are subject to high-stakes tests. The design of the TEOG exam involved a lot of breaks during testing of different subjects to get rest and calm down, involved application of the test in the student's own classroom, and offered a make-up option. This help alleviate such disadvantages of girls in high-stakes exams (Falch & Naper, 2013).

The results of the joint models of TEOG scores and GPAs under FIML and seemingly unrelated regression frameworks are presented in Table 6.11 below. As observed in Table 6.11 when the human capital measure is TEOG achievement score, findings indicate 0.30 standard deviations of the gender achievement gap in public schools, while this reduces to 0.12 standard deviations in private schools. When the measure is GPA, the gender gap is 0.44 standard deviations in public schools, and this reduces to 0.13 standard deviations in private schools. The errors of these two measures have a correlation coefficient of 0.91.

When these findings are compared to the literature findings, we also see that girls systematically outperform boys in grade point averages (GPA) in the literature. Although it is documented that there is a girls' advantage in overall GPAs from school, these girls and their boy peers are found to perform on par on standardized achievement tests. Gratez & Karimi (2022) find that boys are outperformed by girls by about 0.34 standard deviation in school grades, just the opposite is observed in terms of SAT scores with an equal effect size. Reardon et al. (2019, pp. 2475-2476)

documents 0.23 sd for English language arts, but no gap in math, and the authors also list several references in which test score gap lies in the range of 0.15 sd to 0.20 sd. Dickerson, McIntosh, & Valente (2015) documents an average gap of 0.10 standard deviations (sd) for 19 African countries (estimates for individual countries vary between 0.06 sd to 0.34 sd).

Table 6-10 Findings from the joint model of TEOG test scores and GPAs

Estimation method:	**Full Information Maximum Likelihood**						**Seemingly Unrelated Regressions**					
School-by-cohort	**Public Schools**		**Private Schools**		**All Schools**		**Public Schools**		**Private Schools**		**All Schools**	
mean centered:	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*
TEOG Score (stand.):												
Girl	**0.298**	*0.000*	**0.127**	*0.000*	**0.298**	*0.000*	**0.312**	*0.000*	**0.125**	*0.000*	**0.312**	*0.000*
Girl*Private School	-	-	-	-	**-0.180**	*0.000*	-	-	-	-	**-0.194**	*0.000*
Age of the student	-0.156	*0.000*	-0.013	*0.004*	-0.154	*0.000*	-0.095	*0.000*	0.024	*0.000*	-0.094	*0.000*
Age data is mismeasured	-0.019	*0.000*	-0.055	*0.116*	-0.019	*0.027*	-0.003	*0.634*	-0.030	*0.409*	-0.003	*0.564*
Mother Schooling	0.035	*0.000*	0.029	*0.000*	0.034	*0.000*	0.034	*0.000*	0.029	*0.000*	0.034	*0.000*
Father Schooling	0.049	*0.000*	0.034	*0.000*	0.049	*0.000*	0.049	*0.000*	0.034	*0.000*	0.048	*0.000*
Employed Mother	0.041	*0.000*	-0.031	*0.000*	0.039	*0.000*	0.037	*0.000*	-0.027	*0.000*	0.035	*0.000*
Employed Father	0.062	*0.000*	0.015	*0.500*	0.064	*0.000*	0.054	*0.000*	0.036	*0.074*	0.056	*0.000*
Public Official Mother	0.173	*0.000*	0.060	*0.000*	0.135	*0.000*	0.177	*0.000*	0.059	*0.000*	0.138	*0.000*
Public Official Father	0.171	*0.000*	0.117	*0.000*	0.168	*0.000*	0.170	*0.000*	0.115	*0.000*	0.167	*0.000*
Family Income	0.086	*0.000*	-0.036	*0.000*	0.083	*0.000*	0.084	*0.000*	-0.033	*0.000*	0.082	*0.000*
Siblings	-0.030	*0.000*	-0.005	*0.011*	-0.030	*0.000*	-0.033	*0.000*	-0.005	*0.022*	-0.033	*0.000*
Parents not live together	-0.176	*0.000*	-0.178	*0.000*	-0.175	*0.000*	-0.176	*0.000*	-0.156	*0.000*	-0.174	*0.000*
Public Lodging Home	0.079	*0.000*	0.054	*0.000*	0.078	*0.000*	0.078	*0.000*	0.051	*0.000*	0.077	*0.000*
Own Home of Family	0.064	*0.000*	-0.005	*0.164*	0.062	*0.000*	0.062	*0.000*	-0.004	*0.245*	0.060	*0.000*
Own Room of Student	0.115	*0.000*	0.006	*0.230*	0.115	*0.000*	0.107	*0.000*	0.002	*0.688*	0.107	*0.000*
Constant	-0.026	*0.000*	0.001	*0.000*	-0.024	*0.000*	0.012	*0.000*	0.012	*0.000*	0.012	*0.000*
GPA (standardized):												
Girl	**0.438**	*0.000*	**0.137**	*0.000*	**0.438**	*0.000*	**0.448**	*0.000*	**0.136**	*0.000*	**0.448**	*0.000*
Girl*Private School	-	-		-	**-0.311**	*0.000*		-		-	**-0.321**	*0.000*
Age of the student	-0.152	*0.000*	0.026	*0.000*	-0.149	*0.000*	-0.085	*0.000*	0.046	*0.000*	-0.083	*0.000*
Age data is mismeasured	-0.095	*0.047*	0.023	*0.670*	-0.009	*0.058*	0.001	*0.035*	-0.011	*0.642*	0.001	*0.892*
Mother Schooling	0.032	*0.000*	0.020	*0.000*	0.031	*0.000*	0.031	*0.000*	0.020	*0.000*	0.030	*0.000*
Father Schooling	0.045	*0.000*	0.022	*0.000*	0.044	*0.000*	0.044	*0.000*	0.022	*0.000*	0.042	*0.000*
Employed Mother	0.055	*0.000*	-0.003	*0.000*	0.055	*0.000*	0.050	*0.000*	-0.002	*0.000*	0.050	*0.000*
Employed Father	0.075	*0.000*	0.030	*0.000*	0.078	*0.000*	0.066	*0.000*	0.043	*0.000*	0.069	*0.000*
Public Official Mother	0.110	*0.000*	0.038	*0.000*	0.074	*0.000*	0.114	*0.000*	0.039	*0.000*	0.138	*0.000*
Public Official Father	0.144	*0.000*	0.065	*0.000*	0.139	*0.000*	0.143	*0.000*	0.066	*0.000*	0.078	*0.000*
Family Income	0.091	*0.000*	-0.009	*0.000*	0.090	*0.000*	0.088	*0.000*	-0.007	*0.000*	0.087	*0.000*
Siblings	-0.033	*0.000*	-0.015	*0.000*	-0.034	*0.000*	-0.036	*0.000*	-0.015	*0.000*	-0.035	*0.000*
Parents not live together	-0.198	*0.000*	-0.126	*0.000*	-0.192	*0.000*	-0.207	*0.000*	0.035	*0.000*	-0.201	*0.000*
Public Lodging Home	0.080	*0.000*	0.037	*0.000*	0.074	*0.000*	0.077	*0.000*	0.014	*0.000*	0.073	*0.000*
Own Home of Family	0.075	*0.000*	0.013	*0.000*	0.073	*0.000*	0.071	*0.000*	0.030	*0.000*	0.070	*0.000*
Own Room of Student	0.011	*0.000*	0.027	*0.000*	0.111	*0.000*	0.102	*0.000*	0.004	*0.000*	0.103	*0.000*
Constant	-0.026	*0.000*	0.001	*0.000*	0.000	*0.000*	0.039	*0.000*	0.001	*0.000*	0.038	*0.000*
Error correlation:	**0.910**		**0.855**		**0.908**		**0.909**		**0.857**		**0.907**	
No of school-by-cohorts:	58,603		4,718		63,321		57,433		4,666		62,099	
No of unique students:	4,657,700		237,784		4,895,484		3,569,250		167,377		3,736,627	

*Coefficients indicate change in terms of standard deviations in standardized TEOG score for one unit increase in each control variable.

The potential bias from innate ability & personal traits, as well as past inputs, is still present in above estimates. One supportive aspect of arguing that the bias shall be limited is the similarity in the background distributions of girls and boys. Despite sample selection due to dropped out girls, these are very similar. The linear specification will not bear functional form concerns when these distributions are similar, and it is more plausible to think that unobserved variables would be similarly distributed as every other input is similarly distributed.

When I use the differencing strategy, as presented in Table 6.12 below, the unobserved effects from innate ability & personality traits and the past human capital inputs cancel and drop out of estimation. In the remaining variation (called *the grading gap*), the gap is 0.14 sd for public schools and 0.01 sd for private schools. The remaining gap in public schools is in congruence with previous findings. The seminal study by Lavy (2008) found a grading gap of 0.10 sd, while Rangvid (2015) finds between 0.10 sd and 0.20 sd., Cornwell et al. (2013) obtains 0.22 sd, Falch & Naper (2013) finds 0.09 sd, Bharadwaj et al. (2013) detect 0.09 sd, Fryer & Levitt (2010) obtain 0.20 sd., and Terrier (2020) comes up with 0.30 sd.

Table 6-11 The Differenced Model Results

Estimation method:	**Full Information Maximum Likelihood**						**Seemingly Unrelated Regressions**					
School-by-cohort	**Public Schools**		**Private Schools**		**All Schools**		**Public Schools**		**Private Schools**		**All Schools**	
mean centered:	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*	Coef.	*p-value*
Difference (GPA-TEOG)												
Girl	**0.141**	*0.000*	**0.010**	*0.000*	**0.141**	*0.000*	**0.136**	*0.000*	**0.011**	*0.000*	**0.136**	*0.000*
Girl*Private School	-	-	-	-	**-0.132**	*0.000*	-	-	-	-	**-0.127**	*0.000*
Age of the student	0.008	*0.000*	0.039	*0.000*	0.008	*0.000*	0.010	*0.000*	0.022	*0.000*	0.011	*0.000*
Age data is mismeasured	0.009	*0.000*	0.045	*0.022*	0.010	*0.027*	0.004	*0.000*	0.019	*0.000*	0.004	*0.464*
Mother Schooling	-0.003	*0.000*	-0.009	*0.000*	-0.004	*0.000*	-0.003	*0.000*	-0.009	*0.000*	-0.004	*0.000*
Father Schooling	-0.005	*0.000*	-0.012	*0.000*	-0.005	*0.000*	-0.004	*0.000*	-0.012	*0.000*	-0.005	*0.000*
Employed Mother	0.013	*0.000*	0.027	*0.000*	0.015	*0.000*	0.013	*0.000*	0.026	*0.000*	0.014	*0.000*
Employed Father	0.012	*0.000*	0.015	*0.218*	0.013	*0.000*	0.012	*0.000*	0.007	*0.000*	0.013	*0.000*
Public Official Mother	-0.064	*0.000*	-0.052	*0.000*	-0.060	*0.000*	-0.062	*0.000*	-0.020	*0.000*	-0.059	*0.000*
Public Official Father	-0.028	*0.000*	-0.022	*0.000*	-0.030	*0.000*	-0.027	*0.000*	-0.049	*0.000*	-0.029	*0.000*
Family Income	0.005	*0.000*	0.028	*0.000*	0.006	*0.000*	0.004	*0.000*	0.026	*0.000*	0.005	*0.000*
Siblings	-0.003	*0.000*	-0.010	*0.011*	-0.004	*0.000*	-0.003	*0.000*	-0.011	*0.004*	-0.003	*0.000*
Parents not live together	-0.021	*0.000*	0.052	*0.000*	-0.016	*0.000*	-0.030	*0.000*	0.034	*0.000*	-0.027	*0.000*
Public Lodging Home	0.001	*0.421*	-0.170	*0.000*	-0.002	*0.000*	-0.001	*0.000*	-0.016	*0.000*	-0.003	*0.000*
Own Home of Family	0.010	*0.000*	0.018	*0.164*	0.011	*0.000*	0.009	*0.000*	0.018	*0.000*	0.010	*0.000*
Own Room of Student	-0.005	*0.000*	0.021	*0.230*	-0.004	*0.000*	-0.005	*0.000*	0.022	*0.000*	-0.004	*0.000*
Constant	0.030	*0.000*	0.000	*0.280*	0.029	*0.000*	0.027	*0.000*	0.009	*0.000*	0.025	*0.000*
No of school-by-cohorts:	58,603		4,718		63,321		57,433		4,666		62,099	
No of unique students:	4,657,700		237,784		4,895,484		3,569,250		167,377		3,736,627	

*Coefficients indicate change in terms of standard deviations in standardized TEOG score for one unit increase in a control variable.

The motivation for this analysis followed Elder & Jepsen (2014, p.37), who proposed a future research to further explore the private school effect on non-cognitive outcomes, with a focus on whether such effects vary across demographic subgroups. I explored how private schooling impacts non-cognitive outcomes differently for boys and girls. The findings in Table 6.12 indicate that private schools affect the non-cognitive skills of boys and girls more evenly compared to public schools.

On the other hand, Burgess (2016) raised the need for a deeper understanding of and a set of more informative measures for non-cognitive skills. Different versions of differencing strategies are implemented in the literature. First is the value-added modeling, which is implemented by subtracting the test scores of the past from the current test score. A more recent strategy is "within subject" differencing (suggested first by Dee, 2005), which employs differencing (standardized) scores or grades from two different subjects of the same student; for instance, subtracting the score of the math exam from the science exam. Both strategies allow for eliminating students' fixed contributions as here. Time-varying unobserved determinants of student achievement create concern in the value-added models. On the other hand, the challenge for the between-subject approach is that allocating students to teachers of different subjects might be associated with unobserved student inputs that vary across subjects (Clotfelter, Ladd & Vignor, 2010). Other than threats to estimation, they also do not isolate the non-cognitive dimension. But they would provide potentially more robust estimates for the overall gender achievement gap. The non-existence of a prior nationally held standardized exam eliminates the possibility for utilization of value-added approach. On the other hand, data at hand does not include outcome measures divided into subjects. So I cannot additionally utilize those strategies.

My strategy of "between measure differencing (standardized TEOG score vs. GPA)" is more similar to between-subject strategy. However, in my case, the set of teachers is the same. My strategy is a version of deducting blindly and non-blindly scored exams (Lavy, 2008). Special characteristics of TEOG exam, especially being "achievement based" just like the courses taught by the teachers make the comparison

highly relevant. There are several studies showing that both the content covered in class and the questions asked in TEOG are in harmony with the achievements declared in the official curriculum prepared by the Board of Education[60] (*Talim ve Terbiye Kurulu*) (see for instance, Ocak & Kalender, 2016; Öner & Bahadırtaş, 2021; Topçu, 2017). Another assumption is that the TEOG test and GPA reflect the same cumulative achievement level. GPA is formed by the last three years of data, so I implicitly assume that there is no decay in the past inputs. As discussed in the literature review part, this assumption is not a strong assumption (Harris, 2010). In essence, the achievements in the curriculum are cumulative, and rehearsals of past achievements are provided before moving into more advanced ones.

As cognitive aspects are eliminated, the residual gender gap (or *gender grading bias* as it is defined by Terrier, 2020) within each school sector shall be related to non-cognitive aspects as follows (Cornwell et al., 2013; Golsteyn & Schils (2015):

(i) A specific teacher may be incorporating non-cognitive skills of children into grades/marks without any personal gender bias, where girls' marks are amplified due to their stronger non-cognitive skills (self-sufficiency, etc.) and boys' marks are deducted due to their weaker non-cognitive skills (unruly behavior, etc.) (Jacob, 2002).

(ii) A specific teacher may have *grading bias* for one of the gender types due to gender stereotyping; and this reflected as negative bias when giving the mark, but otherwise the teacher does not behave differently during teaching (Stoet & Geary, 2012). Stereotype threat is found to be higher in localities where traditional gender roles are stronger. Bonesronning (2008), Lavy (2008), Lindahl (2007), Xu & Lee (2018) find that girls are more generously rewarded in teachers' grades. Dee (2007) finds that girls achieve more when their teachers are females, and boys achieve better when they have male teachers. Sand & Lavy (2018) find that as the proportion of own daughters

[60] Expert views: Eğitimcilere göre TEOG soruları nasıldı? | PervinKaplan.com

of the teacher increases, the grading bias to the advantage of the girls increases. Lindahl (2007) argues that teachers are perhaps afraid to discriminate girls for reputation purposes and so give them higher grades.

(iii) A specific teacher may have a *behavioral bias* against one of the genders, i.e., the teacher discriminates against one of the gender types in the allocation of schooling inputs (Dickerson, McIntosh, & Valente, 2015). The attention of the teacher within the classroom may vary by the gender of students based on the gender match between the teacher and the student (Carrell et al., 2010; Dee, 2007). The number and intensity of interactions, on the other hand, are in general found to be to the advantage of girls (Falch & Naper, 2013).

Jackson (2018, p.2081) also states that standardized test scores and grade marks provide almost the same measure in terms of cognitive skills. However, he argues that when teachers give their marks, they also incorporate non-cognitive factors such as effort, class behavior, punctuality, and progress of the student during the course period. So grades given by teachers are expected to reflect more skills compared to the skills measured by standardized test scores.

Based on the above literature findings, one reason might be that discipline and school climate in private schools are better in the sense that boys do not delve into unruly behavior or do not have the mindset that they should not study much in order to get more respect from boy peers (Lazear, 2001). Also, lower class sizes in private schools might come to mind, as we have discussed that more crowded classrooms are associated with more disruptive behavior, which affects boys more. Another explanation might be that better peers are together in private schools. However, as we are employing school fixed effects, and any school level factor, including peers, is already controlled for.

As I eliminated the student level (past input and endowment) differences via differencing, the reason for the difference in gaps between the two sectors must be due

to either the gender-specific factors or the school factors, but not due to student factors. Note that the remaining variation after differencing is also expected to measure non-cognitive skill differences reflected as a part of grade marks.

The question is how school factors can have an effect when we control for school fixed effects. Fixed effects approach eliminates *average* differences *between* schools. As there is still a lot of gap remaining in public schools after differencing, different school inputs might have been applied to students from different genders within public schools. In other words, boys may receive fewer inputs in public schools due to the differential behaviour of teachers. Even if there is no input differential, teachers might be just incorporating weak noncognitive skills of students into the grades or even discriminate grades of different genders based on the stereotypes in his/her head, perhaps without even knowing it.

On the other hand, as there is not much gap remaining after differencing in private schools, the comparable school inputs might be applied to both genders in private schools, or grade discrimination is considerably less. I cannot single out each mechanism based on available data and leave this for future research. But I will discuss the consequences of these gaps and potential policy proposals.

Teachers' gender biases can be highly influential on the gender achievement gap. These biases may lead boys to fall behind at school and lead girls to stay distant from scientific courses in high school. GPA entered with a weight of 0.3 into high school admission during the TEOG period, hence these biases had a substantive impact on which high school students are assigned in the centralized allocation system. Hence, grading gaps can affect human capital accumulation (Lavy & Sand, 2018). In terms of a policy proposal, an intervention into teachers' evaluation methods and behaviors in public schools might be useful to reduce gender achievement gaps in this sector.

Golsteyn & Schils (2015) underlines that understanding why there are achievement differences by gender in schools may assist in laying out better school processes that prevent such gaps from occurring. Their findings indicate that differences in non-

cognitive skill endowments play a strong role. This implies a role for better school processes since in contrast to the innate ability, which stays constant after primary school, non-cognitive skills remain malleable throughout K12 education. Jackson (2018) also puts great emphasis on *teachers' value added* for non-cognitive skills, not just *teacher grading bias*. Hence a double-tier strategy is recommended, which increases the human capital of teachers to discriminate less over and contribute more to the development of non-cognitive skills of students.

CHAPTER 7

CONCLUSION

Hanushek (1986) suggested that educational production is about transforming an individual into one with different and better qualities. I explored if private schools do this more effectively compared to public schools. Moreover, I investigated whether higher concentrations of private schools induce more public school quality. Finally, I explored if the gender achievement gap differs between the public and private sectors.

The studied period 2013-2017 included the standardized exam (TEOG) taken by all private and public school students. It also witnessed the implementation of a large-scale voucher program and a rapid expansion of the private sector enrolment by up to 78 percent. On top of that, a failed coup attempt led to the sudden abolishment of 20 percent of private schools in 2016, which provided an exogenous variation for analysis. All these led to a large dataset with a lot of variation.

In terms of findings, using standardized test score measure (to mean zero a standard deviation one), I find a positive effect of around 0.62 standard deviations (sd) for the Type-1 private school effect, corresponding to the combined effect of school inputs and the peer inputs. As the unadjusted difference is 1.31 sd, half of this is accounted for by the compositional differences between the two school sectors. I find a lower bound of 0.25 sd for the Type-2 effect, which corresponds only to the private school effect due to school inputs. The latter is the effect mostly estimated in the literature, and the typical estimate is 0.20 sd.

I find a competition effect of -0.045 sd on the mean of a typical school from one standard deviation increase in the share of private school students in the same province. The instrumental variable estimate of the same effect is -0.036 sd. These

effect sizes align with the modest impact sizes seen in the literature. The majority of the recent literature finds modest positive competition effects. However, some of the most influential previous contributions like Dee (1998) found -0.023 sd, Husted & Kenny (2000) document -0.060 sd, and Geller, Sjoquist & Walker (2001) detect between -0.018 sd to -0.137 sd competition effects. The actually assigned voucher concentration per province has an impact of -0.028 sd on average on the public school achievement outcomes. This finding highlights the importance of better voucher program design in the future.

Regarding gender achievement gaps, when the human capital measure is the TEOG achievement score, I find a 0.30 sd gender achievement gap in public schools. In comparison, this reduces to 0.12 sd in private schools under a school-by-cohort fixed effects setup, where the differences between schools cannot be a reason for the differences. When the measure is GPA, the gender gap is 0.44 sd in public schools, which reduces to 0.13 sd in private schools. The errors of these two measures have a correlation coefficient of 0.91. When I use the differencing strategy, the unobserved effects from innate ability and historical human capital inputs cancel. In the remaining variation, which can be regarded to reflect non-cognitive skills based on the previous literature, the gap is 0.14 sd for public schools and 0.01 sd for private schools. These findings on the gender achievement gap indicate the need for focused case studies to explore why such sizeable differences exist between the two school sectors.

REFERENCES

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, vol. 25, no. 1, pp.95-135.

Abdulkadiroglu, A., Pathak, P., & Walters, C. (2018). Free to choose: can school choice reduce student achievement? *American Economic Journal*: *Applied Economics* 10(1): 175-206.

Abdulkadiroglu, A., Pathak, P., Schellenberg, P., & Walters, C. (2021). Do parents value school effectiveness? *American Economic Review*, vol. 110, pp. 1502-1539.

Akyüz, Y. (2012). *Türk Eğitim Tarihi (M.Ö. 1000 – M.S. 2012),* Ankara, Pegem Akademi.

Alderotti, G., Rapallini, C., & Traverso, S. (2023). The Big Five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, vol. 94, 102570.

Allison, P.D. (2009). Missing Data. In Millsap, R.E. and Maydeu-Olivares (eds.) *The SAGE Handbook of Quantitative Methods in Psychology*. SAGE Publications.

Allison, P.D. (2012). Handling missing data by maximum likelihood. SAS Global Forum 2012, Statistics and Data Analysis, Paper 312-2012

Allison, P.D. (2014). Listwise deletion is not evil. Statistical Horizons.

Almlund, M. et al. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, and L. Wößmann (Eds.), *Handbook of the Economics of Education, Volume 4*, pp. 1–181. Amsterdam: Elsevier.

Altınok, S. K. (2019). *Eğitim ve Öğretim Desteği'nin Özel Okullaşmaya Etkisinin Sosyolojik Analizi*. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Sosyoloji Anabilim Dalı Yüksek Lisans Tezi.

Anand, P., Mizala, A., & Repetto, A. (2009). Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile. *Economics of Education Review*, 2009, vol. 28, issue 3, pp. 370-381.

Anderson, M. R., Guzman, T., & Ringquist, E. J. (2013). Evaluating the effectiveness of educational vouchers. In E. J. Ringquist (Ed.), *Meta-analysis for public management and policy* (pp. 311-351), San Francisco, CA: Jossey-Bass

Angrist, J., Bettinger, E., Bloom, E., King, E. & Kremer, M. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review*, 92 (5): 1535-1558.

Angrist, N., Djankov, S., Goldberg, P. K., & Patrinos, H. A. (2021). Measuring human capital using global learning data. *Nature*, 592, 403–408.

Arbuckle, J.L. (1996). Full information estimation in the presence of incomplete data. In Marcoulidesi G. A. and Schumacker, R.E. (eds), *Advanced Structural Equation Modeling: Issues and Techniques*. Mahvah, NJ: Lawrence Erlbaum Associates.

Arenas, A. (2004). Privatization and vouchers in Colombia and Chile. *International Review of Education*, vol. 50, pp. 379-395.

Atherton, P. Appleton, S. & Bleaney, M. (2012). International school test scores and economic growth. *Bulletin of Economic Research*, vol. 65, issue 1, pp. 82-90

Broer, M., Bai, Y., & Fonseca, F. (2019). Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMMS: IEA Research for Education and Springer Open.

Balart, P., Oosterveen, M. & Webbink, D., (2018). Test scores, noncognitive skills and economic growth, *Economics of Education Review*, 63, 134-153.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background variables in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, vol. 29, pp. 37-65.

Barnov, B.S., Cain, G.L., & Goldberger, A.S. (1980). Issues in the analysis of selection bias. Discussion Paper, University of Wisconsin-Madison, Institute for Research on Poverty.

Barro, R. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106, 407–443.

Barro, R. (1997). *Determinants of economic growth: A cross-country empirical study*. Cambridge, MA: MIT Press.

Barro, R.J. & Lee, J.W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics* 32. No 3. (December): 363-394

Barro, R.J. & Lee, J.W. (2001) International data on educational attainment: Updates and implications. *Oxford Economic Papers* 53, no. 3 (July): 541-563

Barro, R.J. & Lee, J.W. (2013). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics* 104: 184-198.

Barrow, L. (2006). Private School Location and Neighborhood Characteristics. *Economics of Education Review*, vol. 25 (6), pp. 633–645.

Bassanini, A. and Scarpetta, S. (2001). The driving forces of economic growth: Panel data evidence for the OECD countries. OECD Economic Studies, 33, 9–56.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, New York: National Bureau of Economic Research

Becker, G.S., & Lewis, H.G. (1973). On the Interaction Between the Quantity and Quality of Children. *Journal of Political Economy*, Part II, march-april, pp. 279-88

Becker, G.S. (1993) *A Treatise on the Family*. Cambridge, MA: Harvard University Press

Bedard, K. & Cho I. (2010). Early gender test score gaps across OECD countries. *Economics of Education Review*, vol. 29, pp.348-363.

Behrman, J.R. & Birdsall, N. (2003). The quality of schooling: Quantity alone is misleading. *American Economic Review*, 73, 928-946.

Behrman, J.R. (2010). Investment in Education – Inputs and Incentives. In D. Rodrik M.R. Rosenzweig (eds). *Handbook of Development Economics*, Vol. 5, Elsevier.

Belfield, C., & Levin, H. (2002). The effects of competition on educational outcomes. *Review of Educational Research*, 72, 279-341.

Bell, A., & Jones, K. (2015). Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods*, 3 (1): 133–53.

Benhabib, J. & Spiegel, M.M. (1994). The role of human capital in economic development - evidence from aggregate cross-country data. *Journal of Monetary Economics*, Vol. 34, Issue 2, October 1994, pp. 143-173

Benhabib, J. & Spiegel, M.M. (2005). Human capital and technology diffusion. In *Handbook of Economic Growth*, edited by P. Aghion and S. N. Durlauf. Amsterdam: North Holland: 935-966.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75, no. 4, pt. 1:352–65.

Bernal, K. & Keane, M.P. (2011). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics*, vol. 29, no:3, pp. 459-512.

Bertrand, M., Goldin, C. & Katz, L.F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3): 228-255.

Betts, Julian R. (1995). "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics*, 77 (2): 231–50.

Beuermann, D., & Jackson, C. K. (2022). "The short and long-run effects of Attending the schools that parents prefer", *Journal of Human Resources*, 57, 725–746.

Beuermann, D., Jackson, C. K., Navarro-Sola, L., & Pardo, F. (2023). What is good school, and can parents tell? Evidence on the multidimensionality of school output. *The Review of Economic Studies*, vol. 90, issue. 1, January 2023, pp. 65-101.

Blair, J.P. and Staley, S. (1995). Quality competition and public schools; further evidence. *Economics of Education Review*, 14, 193–208.

Bonesronning, H. (2008). The effect of grading practices on gender differences in academic performance. *Bulletin of Economics Research*, vol. 60, issue 3, pp. 245-264.

Borghans, L.; Golsteyn B.H.; Heckman J.J. & Humphries J.E. (2016). What grades and achievement tests measure? *Proc Natl Acad Sci (PNAS) USA*, 113(47):13354-13359

Borland, M.V. and Howsen, R.M. (1992). Student academic achievement and the degree of market concentration in education. *Economics of Education Review*, 11, 31–9.

Bowen, D.H. & Trivitt, J.R. (2014). Stigma without sanctions: The (lack of) impact of private school vouchers on student achievement. *Education Policy Analysis Archives*, vol. 22, no. 87, pp.1-22.

Bowles, S. (1970). Towards an Educational Production Function. In Hansen, W.L. (Ed.), *Education, Income, and Human Capital* (pp. 11-70). National Bureau of Economic Research (NBER).

Bowles, S. and H. Gintis (1975). The Problem with Human Capital Theory: A Marxian Critique. *The American Economic Review*, 65 (2): 74–82.

Bradley, S. and J. Taylor (2002). The effect of the quasi-market on the efficiency–equity trade-off in the secondary school sector. *Bulletin of Economic Research*, 54, 295–314.

Bradley, S., & Taylor, J. (2004). The economics of secondary schooling. *International Handbook on the Economics of Education*, 368.

Brown, B.W. & Saks, D.H. (1987). The microeconomics of the allocation of teachers' time and student learning. *Economics of Education Review*, vol. 6, no. 4, pp. 310-332.

Burgess, S. (2016). Human Capital and Education: The State of the Art in the Economics of Education. IZA Discussion Papers 9885, Institute of Labor Economics.

Burgess, S., McConnell, B., Popper, C. & Wilson, D. (2004). Girls rock, boys roll: an analysis of the age 14–16 gender gap in English schools. *Scottish Journal of Political Economy*, 51(2), pp. 209–29.

Cain, G.G., & Goldberger, A.S. (1983). Public and private schools revisited. *Sociology of Education,* vol. 56, pp.208-218.

Canbolat, Y. (2021). The long-term effect of competition on public school achievement: Evidence from the Indiana choice scholarship program. *Education Policy Analysis Archives*, vol 29, no: 97.

Caner, A., Guven C., Okten C., & Sakalli, S.O. (2016). Gender Roles and the Education Gender Gap in Turkey.*Social Indicators Research,* 129 (3):1231–54.

Card, D. (1995). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. In: Louis, N., Christofides, E.K.G., Swidinsky, R. (Eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderka*mp. University of Toronto Press, Toronto, pp. 201–222.

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, *3*, 1801-1863.

Card, D. & Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States? *Journal of Political Economy*, vol. 199, no. 1, pp.1-40.

Carneiro, P. & J.J. Heckman (2003). Human capital policy. In Heckman, J.J., Krueger, A.B., & Friedman, B.M. (Eds.), *Inequality in America: What role for human capital policies?* Cambridge, MA: MIT Press.

Carr, M. (2011). The impact of Ohio's EdChoice on traditional public school performance. *Cato Journal*, 31(2), 257-284.

Castellano, K. E., Rabe-Hesketh, S., & Skrondal, A. (2014). Composition, Context, and Endogeneity in School and Teacher Comparisons. *Journal of Educational and Behavioral Statistics*, vol. 39, no. 5, pp. 333-367.

Chattopadhyay, S. (2012). *Education and Economics: Disciplinary Evolution and Policy Discourse*, OXFORD University Press

Chakrabarti, R. (2008). Impact of voucher design on public school performance: Evidence from Florida and Milwaukee voucher programs. FRB of New York Staff Report, (315) 45.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. *Quarterly Journal of Economics*, 126(4), 1593–1660

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633–2679

Chiswick, B.R. (1988). Differences in Education and Earnings Across Racial and Ethnic Groups: Tastes, Discrimination and Investments in Child Quality. *Quarterly Journal of Economics*, vol. 103, pp. 571-597.

Chung, H.J. (2018). Gender equality and the gender achievement gap: A crossnational study using PISA 2015. Ph.D. Dissertation, The Pennsylvania State University.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2010). Teacher credentials and student achievement in high school: A cross subject analysis with student fixed effects. *Journal of Human Resources*, vol. 45, pp. 655-681.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum

Cohen, J. & Cohen, P. (1985). *Applied Multiple Regresion and Correlation Analysis for the Behavioral Sciences (2nd edn.)*. Mahvah, NJ: Lawrence Erlbaum Associates.

Cohen-Zada, D. (2009). An alternative instrument for private school competition. *Economics of Education Review*. 28:29-37

Coleman, J., Hoffer, T. & Kilgore, S. (1982). Cognitive outcomes in public and private schools. *Sociology of Education*, vol. 55, pp.65-76.

Collins, A.I. (2007). *Social studies for schools*. Ibadan: University Press.

Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, 48(1), 236–264.

Crain, R.L., & Ferrer, R.L. (1982). Achievement prediction with school level equations: A non-technical example using the public and private schools data. Report no: 323. John Hopkins University: Center for Social Organization of Schools.

Cullen, J.B., Jacob, B. A., & Levitt, S.D. (2005). The impact of school choice on student outcomes: an analysis of the Chicago public schools. *Journal of Public Economics*, vol. 89, pp. 729-760.

Cunha, F., Heckman, J.J. & Schennach, S.M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78 (3): 883–931.

Currie, J. & Thomas, D. (1999). Early test scores, socioeconomic status and future outcomes. NBER Working Papers, No: 6943, National Bureau of Economic Research, Inc.

Çelik, S., Kul, Ü., & Çalık Uzun, S. (2018). Ortaokul matematik dersi öğretim programındaki kazanımların yenilenmiş Bloom taksonomisine göre incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(2), ss. 775-795.

Dal, N. (2018) Bugünkü büyümede en büyük unsur teşvik olmuştur. *Öğretmenim Dergisi*, yıl:10, sayı:113, s.6

Das, J., Singh, A. & Chang, A.Y. (2022). Test scores and educational opportunities: Panel evidence from five low- and middle-income countries. *Journal of Public Economics*, vol. 206, Feb. 2022, 104570

Dauchy, E.P. (2013). Private School Competition and Public School Performance in Arizona. Available at: SSRN: https://ssrn.com/abstract=2599123

Dayioğlu Tayfur, M.; Kirdar, M., & Tansel, A. (2009). Impact of sibship size, birth order, and sex composition on enrolment in Urban Turkey. *Oxford Bulletin of Economics and Statistics*, vol. 71, no. 3, pp. 399-426.

DeAngelis, C.A. & Wolf, P.J. (2020). Private School Choice and Character: More Evidence from Milwaukee. *Journal of Private Enterprise*, The Association of Private Enterprise Education, vol. 35, pp. 13-48.

Dee, T.S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources,* 42(3), 528–554

Dee, T. & M.R. West (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis,* vol. 33, no.1, pp. 23-46

Deming, D.J., Hastings, J.S., Kane, T.J. and Staiger, D.O. (2014). "School Choice, School Quality, and Postsecondary Attainment", *The American Economic Review,* 104, 991–1013.

Denison, F. (1962). *The Sources of Economic Growth in the United States and the Alternatives Before Us*, New York: Committee for Economic Development, pp. 297.

Dickerson, A.; McIntosh, S., & Valente, C. (2015). Do the maths: An analysis of the gender gap in mathematics in Africa. *Economics of Education Review,* 46, 1-22.

DiPrete, T.A. & Buchmann, C. (2013). *Rise of the Women, The Growing Gender Gap in Education and What It Means for American Schools*. Russell Sage Foundation.

Ding, W. & Lehrer, S.F. (2007). Do peers affect student achievement in China's secondary schools? *Review of Economics and Statistics*, 81(3), 543-552.

Dougherty, C. (2003). Numeracy, literacy, and earnings: Evidence from the National Longitudinal Survey of Youth. *Economics of Education Review*, 22(5), 511–521.

Downes, T.A. and Greenstein, S.M. (1996). Understanding the supply decisions of nonprofits: Modelling the location of private schools. *RAND Journal of Economics*, vol. 27 (2), pp. 365–390

Duckworth, A.L., C. Peterson, M.D. Matthews & D.R. Kelly (2007). Grit: perseverance and passion for long-term goals." *Journal of Personality and Social Psychology*, 92 (6): 1087–1101.

Duncan, C., Jones, K., & Moon, G. (1998). Context, composition, and heterogeneity: Using multilevel models in health research. *Social science & Medicine*, vol. 46, pp. 97-117.

Elder, T., & Jepsen, C. (2014). Are Catholic primary schools more effective than public primary schools? *Journal of Urban Economics*, vol. 80, pp. 28-38.

Egalite, A.J. (2013). Measuring competitive effects from school vouchers programs: A systematic review. *Journal of School Choice*, vol. 7, issue 4.

Egalite, A.J., & Mills, J.N. (2021). Competitive impacts of means-tested vouchers on public school performance: Evidence from Louisiana. *Education Finance & Policy*, vol. 16(1): 66-91.

Englander, S. and Gurney, A. (1994). Medium-term determinants of OECD productivity. OECD Economic Studies, 22, 49–109.

Epple, D., & Romano, R. (2008) Educational vouchers and cream skimming. *International Economic Review*, vol. 49, no.4.

Epple, D., Romano, R. E., & Urquiola, M. (2017). School vouchers: A survey of the economics Literature. *Journal of Economic Literature*, 55(2), 441-492.

Evans, W.N., Oates, W.E., & Schwab, R.M. (1992). Measuring peer group effects: A study of teenage behaviour. *Journal of Political Economy*, 100s, 966-91.

Falch, T., & Naper, L.R. (2013). Educational evaluation schemes and gender gaps in student achievement. *Economics of Education Review*, vol. 36, pp.12-25.

Feinstein, L., & Symons, J. (1999). Attainment in Secondary School. *Oxford Economic Papers*, vol. 51, no. 2, pp. 300-321.

Figlio, D. and Stone, J. (1999). 'School choice and student performance: are private schools really better?', *Research in Labor Economics*, 18, 115–40.

Figlio, D., Hart, C.M.D., & Karbownik, K. (2020). Effects of scaling up private school choice programs on public school students. National Bureau of Economic Research. (No.26758)

Fiorini, M. & Keane, M. P. (2014). How the allocation of children's time affects cognitive and noncognitive development? *Journal of Labor Economics*, vol. 32, no. 4.

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *Quarterly Journal of Economics*, 128(1), 249–285

Freedman, M. (1962). *Capitalism and freedom*. Chicago. University of Chicago Press.

French, R., Sariaslan, A., Larsson, H., Kneale, D. & Leckie, G. (2022). Estimating the importance of families in modeling educational achievement using Swedish administrative data. *Journal of Research on Educational Effectiveness*, published online: 26 April 2022, DOI: 0.1080/19345747.2022.2054480

Frenette, M., & Chan, P.W.C. (2015) Academic outcomes of public and private high school students: What lies behind the differences? Catalogue no. 11F0019M, Statistics Canada.

Fryer, R.G., & Levitt, S.D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210–240.

Gallego, F., "Voucher-School Competition, Incentives, and Outcomes: Evidence from Chile." unpublished manuscript, MIT, 2006.

Gamoran, A., (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis* 18 (1), 1 – 18

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel / Hierarchical Models*. Cambridge University Press.

Geller, C.R., Sjoquist, D.L. & Walker, M.B. (2006). The effect of private school competition on public school performance in Georgia. *Public Finance Review*, 34, 4.

Gemello, J. M., and Osman, J.W. (1983). The Choice for Private and Public Education: An Economist's View," in Thomas James and Henry M. Levin (eds.*), Public Dollars for Private Schools: The Case of Tuition Tax Credits*. Philadelphia: Temple University Press, 196-209.

Genc, M. & Colakoglu, O. M. (2021). Modeling the Effects of Instructional Quality on Mathematical Literacy Performance from the Students' Perspective: PISA 2012 Turkey Sample. *Education and Science*, vol. 46, no. 206, pp. 1-26

Gevrek, Z.E., Gevrek, D., & Neumeier, C. (2020). Explaining the gender gaps in mathematics achievement and attitudes: The role of societal gender equality. *Economics of Education Review*, 76, 101978.

Gintis, H. (1971). Education, technology and the characteristics of worker productivity. *American Economic Review*, vol. 61, pp. 266-279.

Glynn, R.J., Laird, N.M., and Rubin, D.B. (1986). Selection modeling versus mixture modeling with nonignorable nonresponse. In Wainer, H (ed.): *Drawing Inferences from Self-Selected Samples*, 115–142. New York: Springer.

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118, 1049–1074.

Gok, F. (2004). Eğitimin Özelleştirilmesi. In Balkan, S. & Savan, S. (eds) *Neoliberalizmin Tahribatı* – Türkiye'de Ekonomi, Toplum ve Cinsiyet, s. 94-110

Goldberger, A.S., & Cain, C.C. (1982). The causal analysis of cognitive outcomes in the Coleman, Hoffer and Kilgore report. *Sociology of Education*, vol. 55:103-22.

Goldhaber, D. (1996). Public and private high schools: is school choice an answer to the productivity problem? *Economics of Education Review*, 15, 93–109.

Golsteyn, B.H.H. & Schils, T. (2014). Gender gaps in primary school achievement: a decomposition into endowments and returns to IQ and non-cognitive factors. *Economics of Education Review*, vol. 41, pp. 176-187.

Gong, G., & Samaniego, F. J. (1981). Pseudo maximum likelihood estimation: Theory and applications. *The Annals of Statistics*, 9, 861–869.

Gourieroux, C., & Monfort, A. (1995). *Statistics and Econometric Models* (Vol. 2). Cambridge, England: Cambridge University Press.

Gradstein, M. & Justman, M. (2000), 'Human capital, social capital and public schooling', *European Economic Review*, 44, 879–90.

Graetz, G. & Karimi, A. (2022). Gender gap variation across assessment types: Explanations and implications. *Economics of Education Review*, 91, 102313.

Gray, N. L., Merrifield, J. D., & Adzima, K. A. (2016). A private universal voucher program's effects on traditional public schools. *Journal of Economics and Finance*, 40(2), 319-344.

Green, F. (2020). Private schools: choice and effects. In Bradley, S. & Green, C. (eds.). *The Economics of Education, A Comprehensive Overview, Second Edition*, Academic Press.

Gümüş, A. (2015). Küreselleşme bağlamında değişen eğitim politikaları. In Gümüş, A. (ed.) *Türkiye'de Eğitim Politikaları*. Ankara, Nobel Akademik Yayıncılık, pp.168-169

Feinstein, L., & Symons, J. (1999) Attainment in secondary school. *Oxford Economic Papers*, vol. 51, pp. 300-321.

Ferman, B. & Fontes, L.F. (2022). Assessing knowledge or classroom behavior? Evidence of tteachers' grading bias. *Journal of Public Economics*, volume 216, 104773.

Hamilton, L., & Guin, K. (2006). Understanding how families choose schools. In Betts, J. and Loveless, T. (eds.) *Getting Choice Right*, 40–61. Washington, DC: Brookings.

Hanchane, S. & Mostafa, T. (2012). Solving endogeneity problems in multilevel estimation: An example using education production functions? *Journal of Applied Statistics*, vol. 39, pp. 1101-1114.

Hansen, W.L. (1970). Introduction to "Education, Income, and Human Capital". In Hansen, W.L. (ed.). *Education, Income, and Human Capital*. NBER.

Hanushek, E.A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, vol. 14, pp. 351-388.

Hanushek, E.A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, vol. 24, pp. 1141-1177.

Hanushek, E.A. (2003). The failure of input-based schooling policies. *The Economic Journal*, vol.113, pp. F64-F98.

Hanushek, E.A. (2008). Education Production Functions. In: The New Palgrave Dictionary of Economics. Palgrave Macmillan, London

Hanushek, E.A. (2015). Economics of Education. *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, Vol. 7

Hanushek, E.A. (2020). Education production functions. In Bradley, S. & Green, C. (eds.) *The Economics of Education, A Comprehensive Overview, Second Edition*, Academic Press

Hanushek, E.A., Piopiunik, M., & Widerhold, S. (2019). The value of smarter teachers: International evidence on teacher skills and student performance. *Journal of Human Resources*, vol. 54, pp. 857-899.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73, 103–130

Hanushek, E. A. & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607–668.

Hanushek, E. A and Woessmann, L. (2012). Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation, *Journal of Economic Growth*, 17: 267-321

Hanushek, E. A. & Woessmann, L. (2015). *The Knowledge Capital of Nations: Education and the Economics of Economic Growth*. Cambridge, MA: MIT Press.

Hanushek, E. & Woessmann, L. (2019). *The economic benefits of improving educational achievement in the European Union: an update and extension: executive summary*. European Commission, Directorate-General for Education, Youth, Sport and Culture Publications Office.

Hanushek, E. A & Woessmann, L. (2021). *Education and Economic Growth*, Oxford Research Encyclopedia of Economics and Finance, Oxford: Oxford University Press

Harmon, C.; Oosterbeek, H., & I. Walker (2003) The returns to education: Microeconomics. *Journal of Economic Surveys*, 17(2): 115-55

Harris, D.N. (2010) Education Production Functions: Concepts. *International Encyclopedia of Education*, pp.402-406

Heckman, J. J. (1992). Randomization and social policy evaluation. In C. Manski and I. Garfinkel (eds.), *Evaluating Welfare and Training Programs*, Cambridge, MA: Harvard University Press, pp. 201–30.

Heckman, J.J. (2000). Causal parameters and policy analysis in economics: a twentieth century retrospective'. *Quarterly Journal of Economics*, vol. 115, pp. 45–97.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annual. Review of Economics*, *6*(1), 689-733.

Heckman, J.J., Pinto, R. & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103 (6): 2052-86.

Heckman, J. J.; Stixrud, J. & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behaviour, *Journal of Labor Economics,* Vol 24, No. 3.

Heckman J., Vytlacil E. (2001) Identifying the role of cognitive ability in explaining the level of change in the return to schooling. *The Review of Economics and Statistics*, 83(1)

Heller-Sahlgren, G. & Jordahl, H. (2022). Test Scores and Economic Growth: Update and Extension. *IFN Working Paper* No. 1408.

Henriquez, F., Lara, B., Mizala, A. & Repotto, A. (2012). Effective schools do exist: low-income children's academic performance in Chile. *Applied Economics Letters*, vol. 19(5), pp. 445-451.

Heyns, B., & Hilton, T.L. (1982). The cognitive tests for High School and Beyond: An assessment. *Sociology of Education*, vol. 55, pp. 89-102

Hsieh, C.T. and Urquiola, M. (2006). The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's Voucher Program. *Journal of Public Economics*, vol. 90 (8-9), pp. 1477– 1503

Hoxby C.M. (1994). 'How school choice affects the achievement of public school students', in Hill, P. (ed.), *Choice with Equity*, Stanford: Hoover Institution Press

Hoxby, C.M. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper No. 7867.

Howell, W.G., Peterson, P.E., Wolf, P.J., & Campbell, D.E. (2002). *The Education Gap: Vouchers & Urban Schools*. Washington, D.C.: Brookings Institution Press.

Hungerman, D.M., & Rinz, K. (2016) Where does voucher funding go? How large-scale subsidy programs affect private-school revenue, enrollment, and prices? *Journal of Public Economics*, vol. 136, pp. 62-85.

Imbens, G.W. (2022) Causality in econometrics: choice vs chance. *Econometrica*, vol 90, no. 6 (November 2022), 2541-2566.

Jabbar, H., Fong, C. J., Germain, E., Li, D., Sanchez, J., Sun, W.-L., & Devall, M. (2019). The competitive effects of school choice on student achievement: A systematic review. *Educational Policy*, 1-35

Jackson, C. K. (2010). "Peer quality or input quality?: Evidence from Trinidad and Tobago." NBER (National Bureau of Economic Research) Working Paper 16598

Jackson, C.K. (2018) What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy*, 2018, vol. 126, no 5.

Jackson, C. K., Johnson, R.C., & Persico, C. (2016). The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economics*, vol. 131 (1), pp. 157–218.

Jacob, B.A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, 21, 589–598.

Jepsen, C. (2000) The private schooling market and its effects on student achievement. Ph.D. dissertation. Northwestern University, Evanston, Il.

Jencks, C., & Phillips, M. eds. (1998) *The Black-White Test Score Gap*. Brookings Institutiom Press, Washington, DC.

Jencks, C., & Phillips, M. (1999). Aptitude or achievement: Why do test scores predict educational attainment and earnings? In Mayer, S. & Peterson, P.E. (Eds.), *Earning and learning: How schools matter* (pp. 15–47). Brookings Institution Press.

Jerrim, J. & Micklewright, J. (2014). Socio-economic gradients in children's cognitive skills. Are cross-country comparisons robust to who reports family background? *European Sociological Review*, 30(6), 766-781

Jerrim, J., Parker, P.D., Katyn Chmielewski, A. & Anders, J. (2016). Private schooling, educational transitions, and early labour market outcomes: Evidence from three Anglophone countries. *European Sociological Review*. 32(2), 280-294

Kablan, Z., Baran, T, & Hazer, Ö. (2013). İlköğretim matematik 6-8 öğretim programında hedeflenen davranışların bilişsel süreçler açısından incelenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAT),* cilt 14, sayı 1, pp. 347-366.

Kautz, T; Heckman, J.J.; Diris, J.; Ter Weel, B.; & Borghans, L. (2017). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. NBER Working Paper 20749.

Kahyaoğlu, R. B., & Karataş, S. (2018). Milli Eğitim Bakanlığının Özel Okullara Teşvik Yardımına İlişkin Mevcut Durum Analizi. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, no. 16, s. 585-626.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross country evidence on the effect of gender inequality in education on economic development. *World Bank Economic Review*, 16, 345-373.

Kleitz, B.; Weiher, G.R.; Tedin, K. & Matland, R. (2000). Choice, charter schools, and household preferences. *Social Science Quarterly,* vol. 81, pp. 846-854.

Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In von Davier, M., Gonzalez, E., Kirsch, I. & Yamamoto, K. (Eds.), *The role of internationally large-scale assessments: Perspectives from technology, economy, and educational research* (pp. 115-147) Dordrecht: Springer Science + Business Media

Knowles, S., Lorgelly, P.K. & Owen, P.D. (2002). Are educational gender gaps a brake on economic development? Some cross-country empirical evidence. *Oxford Economic Papers*, vol. 54, pp. 118-149.

Koedel, C., & Mihaly, K., & Rockoff, J.E. (2015). Value-added modeling: A review. Economics of Education Review, vol. 47, pp. 180-195.

Komatsu, H. & Rappleye, J. (2017). A new global policy regime founded on invalid statistics? Hanushek, Woessmann, PISA, and economic growth. *Comparative Education*, 53:2, 166-19

Kreft, I. G. G., de Leeuw, J., & Aiken, L.S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research* 30: 1–21

Krueger, A.B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, vol. 114, pages. 497-532.

Krueger, A. B. (2003). Economic considerations and class size. *Economic Journal*, 113(485), F34–F63

Krueger, A.B. and Lindahl, M. (2001). Education for Growth: Why and for Whom? *Journal of Economic Literature*, 39, 1101-1136.

Krueger, A.B., & Whitmore, D. (2002). Would smaller classes help close the black-white achievement gap? In J.E. Chubb, & T. Loveless (Eds.), *Bridging the Achievement Gap* (pp. 11-46). Washington, DC: Brookings Institution Press.

Kuzu, Y.; Kuzu, O. & Gelbal, S. (2019). TEOG ve LGS sistemlerinin öğrenci, öğretmen, veli ve öğretmen velilerin görüşleri açısından incelenmesi. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, cilt 5, sayı 1, ss. 112-130.

Küçükçayır, G. & Cemaloğlu, N. (2017) Türkiye'de geçmişten günümüze özel okullar ve uygulanan politikalar üzerine bir araştırma. In Akbaba Altun, S., Örücü, D., Beycioğlu, K., Kondakçı, Y., & Koşar, S. (eds.) *Eğitim Yönetimi Araştırmaları*. Ankara, Pegem Akademi, s. 87-99

Lai, F. (2010). Are boys left behind? the evolution of the gender achievement gap in Beijing's middle schools. *Economics of Education Review*, 29(3), 383-399.

Lang, K. (1994). Does the human capital / educational sorting debate matter for development policy? *American Economic Review*, 84 (1), 353–8

Lankford, H., &Wyckoff, J. (1992). Primary and secondary school choice among public and religious alternatives. *Economics of Education Review*, 11, 317–37

Lavy, V. (2008). Do gender stereotypes reduce girls' and boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 91(10–11), 2083–2105

Lavy, V. & Sand, E. (2018). On the origins of gender gaps in human capital: Short- and long-term consequences of teachers' biases. *Journal of Public Economics*, vol. 167, pp. 263-279.

Lavy, V. & Schlosser, A., (2007). Mechanisms and impacts of gender peer effects at school. NBER Working Paper No. 13292.

Lazear, E.P. (2001). Educational Production. *Quarterly Journal of Economics*, vol. 116, pp. 777-803.

Leckie, G., & Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society:* Series A (Statistics in Society), 172, 835–851.

Leckie, G. (2018). Avoiding bias when estimating the consistency and stability of value-added school effects. *Journal of Educational and Behavioral Statistics*, vol. 43, no. 4, pp. 440-468.

Leibowitz, A. (1974). Home investments in children. *Journal of Political Economy*, 82, no. 2, pt. 2: S111–S13

Leuven, E. & Oosterbeek, H. (2011). "Overeducation and mismatch in the labor market", in E. A. Hanushek, Machin, S. and Woessmann, L. (Eds.), *Handbook of the economics of education, Vol. 4*. Amsterdam: North Holland.

Levin, H.M. (1976). Concepts of economic efficiency and educational production. NBER Chapters, in *Education as an industry*, pp. 149-198, National Bureau of Economic Research, Inc.

Lin, D., Lutter, R., & Ruhm, C. J. (2018). Cognitive performance and labour market outcomes. *Labour Economics*, 51, 121–135

Lindahl, E. (2007). Comparing teachers' assessments and national test results—Evidence from Sweden. Institute for Labour Market Policy Evaluation (Working Paper 2007:24).

Lindahl, M. (2015). *Review of The Knowledge Capital of Nations: Education and the Economics of Growth*. By Hanushek, E.A. and Woessmann, L. CESifoBook Series. Cambridge and London: MITPress

Little, R.J.A. (1992). Regression with missing X's: a review. *Journal of the American Statistical Association,* 87: 1227–1237.

Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd edn.). Wiley.

Long, J.E., & Toma, E.F. (1988). The determinants of private school attendance, 1970-1980. *The Review of Economics and Statistics*, vol. 70, no. 2, pp. 351-357.

Lubienski, S.T., & Lubienski, C. (2006). School sector and academic achievement: a multilevel analysis of NAEP mathematics data. *American Educational Research Journal* 43 (4), 651–698.

Lubienski, C., Corinna, C., & Lubienski, T.S. (2008). What do we know about school effectiveness? Academic gains in public and private schools. *Phi Delta Kappan* 89 (9), 689–695.

Lucas, R. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, 22.

Mankiw, N.G. (1995). The growth of nations. *Brookings Papers on Economic Activity*, 1: 275-326

Mankiw, G., D. Romer and D. Weil (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107, 407–37

Marks, G. & M. O'Connell (2021). Inadequacies in the SES – Achievement model: Evidence from PISA and other studies. *Review of Education*, British Educational Research Association, vıl. 9, issue 3, e3293

Martinez-Vazquez, J. and Seaman, J (1985). Private Schooling and the Tiebout Hypothesis. *Public Finance Quarterly*, 13, 293-318.

Mayer, S.E. (1999). From Learning to Earning. In S. Mayer & P. E. Peterson (Eds.), *Earning and learning: How schools matter* (pp. 3-17). Brookings Institution Press.

McEwan, P.J. (2000). Comparing the effectiveness of public and private schools: A review of evidence and interpretation. Occasional paper 3, National Center for the Study of Privatization in Education, Teachers College, Columbia University, New York.

McEwan, P. (2004). The potential impact of vouchers. *Peabody Journal of Education*, 79(3), 57–80.

McMillan (2001). Parental involvement, competition, and public school performance: An empirical analysis using individual data. *Proceedings, Annual Conference on taxation and Minues of the Annual Meeting of the National Tax Association*, vol. 92, pp. 150-155.

Meghir, C., & Rivkin, S. (2011). Econometric methods for research in education. *Handbook of Economics of Education*, Volume 3, pp.1-87.

Milli Eğitim Bakanlığı (1961). *Özel Okullar Komitesi Raporu*. Milli Eğitim Basımevi, Ankara

Milli Eğitim Bakanlığı (2014) *5580 sayılı kanun kapsamındaki özel okullarda öğrenim görecek öğrenciler için verilecek eğitim ve öğretim desteği uygulama e-kılavuzu*. Milli Eğitim Bakanlığı Özel Öğretim Kurumları Genel Müdürlüğü 25031903_eitimveretimdestei_klavuz25082014.pdf (meb.gov.tr)

Miller, R. (2022) The effect of private schools on measures of socioemotional development in adolescence: Evidence fro0m India. *Journal of Human Capital*, vol. 16, no. 2.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66 (4), 281–302.

Mincer, J. (1974). *Schooling, Experience, and Earnings*, New York: National Bureau of Economic Research.

Monk, D.H. (1989) The education production function: Its evolving role in policy analysis. *Educational Evaluation and Policy Analysis*, vol.11, no.1, pp.31-45.

Mosteller, F. (1995). The Tennessee Study of Class Size in the Early School Grades. *The Future of Children*, vol. *5*(2), pp. 113–127.

Morgan, C., Petrosino, A. & Fronius, T. (2015). The impact of school vouchers in developing countries: A systematic review. *International Journal of Educational Research*, vol. 72, pp. 70-79.

Morgan, S.L. (2001). Counterfactuals, causal effect heterogeneity, and the Catholic school effect on learning. *Sociology of Education*, 74 (4), 341-374.

Mulligan, C. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy*, 107(6), pp. 184-224.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, vol. 46 (1), pp. 69–85.

Murnane, R.J., Newstead, S., & Olsen, R.J. (1985). Comparing public and private schools: The puzzling role of selectivity bias. *Journal of Business & Economic Statistics*, vol.3, no.1

Murnane, R J., Willett, J. B. & Levy, F. (1995). The growing importance of cognitive skills in wage determination, *The Review of Economics and Statistics*, vol. 99, no. 2, pp. 251-166

Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4), 547–568.

Neal, D. (1997). The effects of Catholic secondary schooling on educational achievement. *Journal of Labor Economics*, 15, 98-123.

Neal, D. (1998). What have we learned about the benefits of private schooling? *Federal Reserve Bank of New York Policy Review*, 4(1), 79-86.

Nelson, R. and Phelps, E. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 56, 69-75.

Nguyen, A., Taylor, J. Bradley, S. (2003). The effect of Catholic schooling on educational and labour market outcomes: further evidence from NELS', mimeo, Department of Economics, Lancaster University.

Niederle, M. & Vesterlund, L. (2010). Explaining the gender gap in math test scores: the role of competition. *Journal of Economic Perspective*, 24(2): 129-144.

Noell, J. (1982). Public and Catholic schools: A reanalysis of "Public and private schools." *Sociology of Education*, *55*(2-3), 123–132.

Nomi, T., & Raudenbush, S. W. (2016). Making a success of "algebra for all": The impact of extended instructional time and classroom peer skill in Chicago. *Educational Evaluation and Policy Analysis*, 38(2), 431–451.

Ocak, G., & Kutlu Kalender, M.D. (2017). Ortaokul 5. ve 6. sınıf fen bilimleri öğretim programının kazanım ve içerik ilişkisinin değelendirilmesi. *Bayburt Eğitim Fakültesi Dergisi*, vol. 12, no. 23, pp. 209-225.

OECD (1998). Returns to investment in human capital. Ch. 4 in *Human capital investment — An international comparison*, Paris: OECD, Centre for Educational Research and Innovation

OECD (2011). PISA in Focus: Private schools: Who benefits?, 2011/7 (August)

OECD (2017). Schoolwork-related anxiety: In PISA 2015 results (volume III): students' well-being. OCED publishing, Paris.

OECD (2013). *Skilled for Life? Key Findings from the Survey of Adult Skills*, OECD Publishing, Paris

OECD (2019). *Turkey – Country Note – PISA 2018 Results,* OECD2019 – Volume I-III

OECD (2019). *Skills Matter: Additional Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris

OECD (2020). *Education Policy Outlook: Turkey*, June 2020, OECD Publishing, Paris

OECD (2020). Private schools and school choice. In *PISA 2018 results (volume V): Effective Policies, Successful Schools*. OECD Publishing. Paris.

OECD (2021). *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies*. OECD Skills Studies, OECD Publishing, Paris

Oosterbeek, H. (1992). Study, duration and earnings – a test in relation to the human capital versus screening debate. *Economics Letters*, 40, 223–8.

Ozawa S, Laing S.K., Higgins C.R., Yemeke T.T., Park C.C., Carlson R., Ko, Y.E., Guterman L.B., & Omer S.B. (2022). Educational and economic returns to cognitive ability in low- and middle-income countries: A systematic review. *World Development,* 149:105668

Ozkaya, N. (1986). Türkiye'de özel öğretim kurumları ve özel öğretimden beklenenler. *Din Öğretimi Dergisi*, Ankara, Milli Eğitim, Gençlik ve Spor Bakanlığı Yayınları, Sayı 8-9, s. 36-41

Ozturk, A. (2017). *Essays on quasi-experimental studies in labor economics*. Ph.D. thesis. Middle East Technical University (METU), Türkiye

Page, M. E. (2010). *Signaling in the Labour Market*, in D. J. Brewer and P. J. McEwan (eds), Economics of Education, UK and USA: Elsevier and Academic Press

Pak, S.Y. (2004). Articulating the boundary between secularism and Islamism: The imam-hatip schools of Turkey. Anthropology & Education Quarterly, no.3:324

Paloyo, A.R. (2020). Peer effects in education: recent empirical evidence. In Bradley, S. & Green, C. (eds.). *The economics of education, a comprehensive overview, second edition*, Academic Press

Patrinos, H. A. (2007). Demand-side financing in education. In Education policy series. Paris: UNESCO.

Pepinsky, T. (2018). A note on listwise deletion versus multiple imputation. *Political Analysis*, 26(4), 480-488.

Prichett, L. (2001). Where has all the education gone? The World Bank Economic Review, vol. 15, no. 3, pp. 367-391.

Psacharopoulos, G. (1994). 'Returns to investment in education: a global update', *World Development*, 22 (9), 1325–43.

Sanayi ve Ticaret Bakanlığı (2019). *İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıramalası Araştırması SEGE-2017*. Kalkınma Ajansları Genel Müdürlüğü yayını, sayı 3.

Sacardote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? Chapter 4 in Handbook of Education, Volume 3, pp.249-277, Elsevier.

Schafer, J.L. (1997). *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.

Sen, S., Yildirim, İ. & Karacabey, M.F. (2020). The implementation of the value-added assessment to determine school effectiveness. *Education and Science*, vol. 45, pp. 79-91

Shakeel, M.D., Anderson, K.P. & Wolf, P.J. (2021). The participant effects of private school vouchers around the globe: a meta-analytic and systematic review. *School Effectiveness and School Improvement*, vol. 32, issue 4, pp. 509-542.

Quintini, G. (2014). Skills at work: How skills and their use matter in the labour market. *OECD Social, Employment, and Migration Working Papers*, No. 158

Papageorge, N. W., & Thom, K. (2019). Genes, education, and labor market outcomes: Evidence from health and retirement. *Journal of the European Economic Association*, vol. 18, issue 3, pages 1351-1399

Pianta, R.C. & Ansari, A. (2018). Does attendance in private schools predict student outcomes at age 15? Evidence from a longitudinal study. *Educational Researcher*, vol. 47, issue 7., pp. 419-434.

Psacharopoulos, G. and Patrinos, H. (2018). Returns to investment in education: A decennial review of the global literature, *Education Economics*, Vol. 26, Issue 5, pp. 445-458.

Reardon, S.F., Fahle, E.M., Kalogrides, D., Podolsky, A., & Zarae, R.C. (2019). Gender achievement gaps in U.S. school districts. *American Educational Research Journal*, vol. 56, pp. 2474-2508.

Reardon, S.F., Cheadle, J.F., Robinson, J.P. (2009). The effects of catholic schooling on math and reading development in kindergarten through fifth grade. *Journal of Research on Educational Effectiveness*, 2(1), 45–87.

Ritzen J.M. & Winkler, D.R. (1977). The production of human capital over time. *The Review of Economics and Statistics*, vol. 59, no.4, pp. 427-437.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica* 73 (2), 417–458.

Robertson, D., & Symons, J. (2003). Do peer groups matter? Peer group versus schooling effects on academic attainment. *Economica*, vol. 70, no. 227, pp. 31-53.

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*., vol. 94 (2), pp. 247–252.

Romer, P. (1990). Endogenous technical change. *Journal of Political Economy*, 89, S71–S102.

Rose, H. (2006). Do gains in test scores explain labor market outcomes? *Economics of Education Review*, volume 25, issue 4, pp. 430-446.

Rose, H. & Betts, J. R. (2004). The Effect of High School Courses on Earnings. The *Review of Economics and Statistics*, vol 86., no: 2, pp. 497-513

Rubin, D.B. (1976). Inference and missing data. *Biometrika* 63: 581–592.

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Rumberger, R.W. & Palardy, G.J. (2004). Multilevel models for school effectiveness research. In Kaplan, D. (ed.) *The SAGE handbook of quantitative methodology for the social sciences*. SAGE Publications.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics*, 6(1), 253-272.

Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review*, 94(4), 813–835.

Sanayi ve Teknoloji Bakanlığı (2019). İllerin ve Bölgelerin Sosyo-Ekonomik Gelişmişlik Sıralaması Araştırması SEGE-2017.

Sander, W. (1996). Catholic grade schools and academic achievement. *The Journal of Human Resources*, vol. 31, no. 3, pp. 540-548.

Sander, W. (1999). Private schools and public school achievement. *Journal of Human Resources*, vol. 34. pp. 697-739.

Sander, W., & Cohen-Zada, D. (2020). The economics of catholic schools. In Bradley, S. & Green, C. (eds.). *The Economics of Education, A Comprehensive Overview, Second Edition*, Academic Press.

Savvides, A. & Stengos, T. (2009). *Human Capital and Economic Growth*. Stanford Economics and Finance, Stanford University Press, Stanford, California

Saygın, P. O. (2019). Gender bias in standardized tests: evidence from a centralized college admissions system. *Empirical Economics*, vol. 59(2), pages 1037-1065.

Schafer, J.L. (1997). *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall.

Schanzenbach, D.W. (2020). The economics of class size. In Bradley, S. & Green, C. (eds.). *The Economics of Education, A Comprehensive Overview, Second Edition*, Academic Press.

Schultz, T.W. (1960). Capital Formation by Education, *The Journal of Political Economy*, 68 (6): 571–83.

Schultz, T.W. (1961). Investment in Human Capital. *The American Economic Review*, 1(2), 1-17.

Schwerdt, G. & Woessmann, L. (2020). Empirical methods in the economics of education. In Bradley, S. & Green, C. (eds.). *The Economics of Education, A Comprehensive Overview, second edition*, Academic Press.

Sen, S.; Yildirim, I. & Karacabey, M. (2020). The implementation of the value-added assessment to determine school effectiveness. *Education and Science*, vol. 45, no: 202, pp. 79-91.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*,10(5), 1189–1213.

Sianesi, B., & Van Reenen, J. (2003). The returns to education: Macroeconomics. *Journal of Economic Surveys*, 17(2), 157–200.

Snijder, T.A.B. & Bosker, R.J. (2012). *Multilevel Analysis, 2nd Edition, An Introduction to Basic and Advanced Multilevel Modeling*. SAGE.

Sobel, I. (1982). Human Capital and Institutional Theories of the Labour Market: Rival and Components?, *Journal of Economic Issues*, 16 (1): 255–72.

Solow, R. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70, 65–95.

Somers, M. A., McEwan, P. J., & Willms, D. (2004). How effective are private schools in Latin America? *Comparative Education Review*, 48, 48– 69

Spence, M.A. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87 (3): 355–74.

Spence, M.A. (2002). Signalling in retrospect and the informational structure of markets. *American Economic Review*, 92 (3), 434–59.

State Planning Organization (1979). Fourth Five-Yearly Development Plan.

Stiglitz, J.E. (1975). 'The theory of "screening", education, and the distribution of income', *American Economic Review*, 65 (3), 283–300.

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology, 16*(1),93–102.

Strom, B., & Falch, T. (2020). The role of teacher quality in education production. In Bradley, S. & Green, C. (eds.). *The Economics of Education, A Comprehensive Overview, Second Edition*, Academic Press.

Swinton (2010). document empirical evidence for the effect of effort on learning and performance. *Economics of Education Review*, vol. 29, pp. 1176-1182.

Taubman, P.J. and T.J. Wales (1973). Higher education, mental ability and screening. *Journal of Political Economy*, 81, 28–55.

Taşkın, P. (2010). 5580 sayılı özel öğretim kurumları kanunu ve getirdiği yenilikler. *Ankara Barosu Dergisi*, yıl:68, sayı:2010/4

Terrier, C. (2020). Boys lag behind: How teachers' gender biases affect student achievement. *Economics of Education Review*, 77(5): 101981

ter Weel, B. (2008). The noncognitive determinants of labor market and behavioral outcomes: Introduction to the symposium. *Journal of Human Resources*, 43(4), pp. 729–737.

Tiebout, C.M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64:416-24.

Todd, P.E. & Wolpin, K. (2003) On the specification and estimation of the production function for cognitive achievement. *The Economic Journal,* 113, pp. F3-F33.

Todd, P.E., and Wolpin, K.I. (2007). The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital* 1, no. 1:91–36.

Tonello, M. (2016). Peer effects on non-native students on natives' educational outcomes: Mechanisms and evidence. *Empirical Economics*, 51(1), 383-414.

Tooley, J. (1996). *Education Without the State*, Institute of Economic Affairs, London: IEA.

Tumen, S. (2018). The Impact of Low-Skill Refugees on Youth Education. IZA Discussion Papers, No. 11869, Institute of Labor Economics (IZA), Bonn.

Tumen, S. (2019a). The effect of refugees on native adolescents' test scores: Quasi-experimental evidence from PISA. *Journal of Development Economics*, vol. 150, May 2021, 102633..

Tumen, S. (2019b). Refugees and native flight from public to private schools. *Economic Letters*, vol 181, August 2019, pp. 154-159.

Tunç & Baydar (2022). TEOG, LGS ve TIMMS matematik sorularının MATH taksonomisine göre incelenmesi. *Bayburt Eğitim Fakültesi Dergisi*, cilt: 17, sayı: 33, ss. 20-53.

UNDP (United Nations Development Programme). (2022). *Human Development Report 2021-22: Uncertain Times, Unsettled Lives: Shaping our Future in a Transforming World*. New York.

UNESCO (United Nations Educational, Scientific and Cultural Organization). (2022). *Gender Repor – Deepening the debate on those still left behind*. Global Education Monitoring Report. UNESCO.

Urquiola, M. (2016). Competition among schools: Traditional public and private schools. In Hanushek, E.A., Machin, S., & Woessmann, L. (pp. 209-237) *Handbook of the Economics of Education, Volume 5.*

Uygun, S. (2003). Türkiye'de dünden bugüne özel okullara bir bakış (gelişmeler ve etkileri). *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, cilt 36, sayı 1-2, s.107-120.

Valenzuela, J. P. (2005). Do students benefit from school competition? Evidence from Chile. In *Essays in economics of education*. PhD thesis.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16, 219–242.

Van Dam, A. (2018). It's better to be born rich than gifted. The Washington Post, October 9, 2018. Reached from: It's better to be born rich than gifted - The Washington Post

Watts T.W. (2020). Academic achievement and economic attainment: Reexamining associations between test scores and long-run earnings. *AERA Open*, American Educational Research Organization, April-June 2020, vol. 6, no. 2, pp. 1-16.

Weiss, Y. (2015). Gary Becker on human capital. *Journal of Demographic Economics*, vol. 81, issue 1, pp. 27-31.

West, E. and Palsson, H. (1988). Parental choice of school characteristics: estimation using state-wide data. *Econ. Inquiry* 26, 725-740.

Westphal, A., Vock, M. and Kretschmann, J. (2021). Unraveling the Relationship Between Teacher-Assigned Grades, Students Personality, and Standardized Test Scores. *Frontiers in Psychology*, Volume 12, Article 627440.

Willingham, W. W., Pollack, J. M., and Lewis, C. (2002). Grades and test scores: accounting for observed differences. *J. Educ. Measur*. 39, 1–37.

Winkler, D.R. (1987) Screening models and education. In Psacharopoulos, G. (ed.) *Economics of Education: Research and Studies*. New York: Pergamon Press.

Woessmann, Ludger. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *Journal of Economic Perspectives*, 30 (3): 3-32.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, 47(1), 21–41.

Yang, Y. (2022). A correlated random effects approach to the estimation of models with multiple fixed effects. *Economic Letters*, vol. 213, 110408.

Yıldırım, S. (2019). Predicting mathematics achievement: the role of socioeconomic status, parental involvement, and self-confidence. *Education and Science*, vol. 44, no: 198, pp. 99-113.

Zanzig, B.R. (1997). Measuring the impact of competition in local government education markets on the cognitive achievement of students. *Economics of Education Review*, 16, 431–41.

Zhan, C. (2018). School choice programs and location choices of private schools. *Economic Inquiry*, vol 56, issue 4, pp. 1622-1645.

Xu, D. & Li, Q. (2018). Gender achievement gaps among chinese middle school students and the role of teachers' gender. *Economics of Education Review*, vol. 67, 82-93.

A. CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Özbaylanlı, Bilgehan
Nationality: Turkish

EDUCATION

Degree	Institution	Year of Graduation
Ph.D.	METU, Economics	2023
Pre-Doc.	Oxford University, Sustainable Urban Development (with Jean Monnet Scholarship)	2019
MPA/ID	Harvard University, Kennedy School of Government, International Development	2012
BS	METU, Industrial Engineering	2001
BS	METU, Environmental Engineering	2001
High School	Bornova Anadolu High School, İzmir	1997

WORK EXPERIENCE

Year	Place	Enrollment
2018 - Present	Turkish Ministry of National Education	World Bank Project Manager

2017 - 2018	Ministry of Development	Senior Specialist
2013 - 2016	Ministry of Development	Head of IT Department
2003 - 2013	Prime Ministry – State Planning Organization	Specialist
2001- 2003	Ministry of Environment and Forest	Junior Specialist

PUBLICATIONS

1. Gender Achievement Gap in Turkey, Jean Monnet Research Report (at Oxford University under supervision of Ass. Prof. Nihan Akyelken), 2020

2. Impact of School Type on Student Academic Achievement (with M. Cansiz and M. H. Colakoglu), *Journal of Education and Science*, vol. 44, no:97, pp. 275-314, 2019

3. Türkiye'de Kültürel Sermayenin Öğrenim Başarısı Üzerine Etkisi (with M. Cansiz and M. H. Colakoglu), *Journal of Economy, Culture and Society*, Istanbul University Press, Issue:58:127-152, Dec. 2018

4. Benefits of Technoparks for Innovative and Technology-Based Entrepreneurs, (with M. Cansiz), *Journal of Productivity* (Verimlilik Dergisi), ISSN: 1013-1388, 2018/3

5. Teknoparkların Ar-Ge ve Yenilik Fikirlerine Katkıları, (with M. Cansiz) *Journal of Productivity* (Verimlilik Dergisi) 2017/3

B. TURKISH SUMMARY / TÜRKÇE ÖZET

Giriş:

Eğitim, uzun bir süre boyunca sadece oynadığı sosyal rol itibariyle değerlendirilmiştir. İnsan sermayesi teorisinin öncüleri Schultz (1960) ve Becker (1964), öğrencilerin eğitim süreçleri ile mesleki kariyerlerine değer katan üretken beceriler arasındaki kritik bağlantıyı ortaya koymuştur. Bu anlamda eğitim, sosyal rolünün yanı sıra, ekonomik değere de haiz bir yatırımıdır. Kişinin kendisine yapılan bir sermaye yatırımı olması nedeniyle insan sermayesi olarak adlandırılmıştır (Hansen, 1970).

O dönemden bugüne, eğitim bireysel boyutta ve ülke genelinde geliri artıran ve gelirin dağılımını iyileştiren bir yatırım unsuru olarak görülmeyi sürdürmektedir. İnsan sermayesi teorisinden iki ana araştırma alanı doğmuş ve bu iki alan birlikte *eğitim iktisadı* alt bilim dalını oluşturmuştur. İlk araştırma alanı eğitimle gelir arasındaki ilişkiye odaklanmış ve bunu hem bireysel düzeyde (Mincer, 1974), hem de ulusal düzeyde (Lucas, 1988; Mankiw, Romer & Weil, 1992) ele almıştır. İkinci araştırma alanı ise eğitim süreçlerinde insan sermayesinin nasıl üretildiğine odaklanmıştır (Bowles, 1970; Hanushek, 1979). Her iki alan da bir birini tamamlayıcı niteliktedir, zira eğitim faaliyetleri sonucu üretilen insan sermayesi, bireysel gelirlerin ve milli gelirin üretimine de bir girdi olarak katkı vermektedir. Özel okulların insan sermayesi üretimi üzerindeki etkilerine odaklanan bu çalışma, eğitim ekonomisi literatürünün ikinci ana araştırma alanında yer almaktadır.

İnsan sermayesi, bir bireyin sahip olduğu bilişsel ve bilişsel-olmayan becerilerin, doğuştan gelen genetik yeteneğin, kişilik özelliklerinin ve edinmiş olduğu bilgilerin toplamına karşılık gelmektedir. Mikro düzeyde, insan sermayesinin dağılımı kazançların dağılımını belirler. Makro düzeyde ise, bir ülkenin toplam insan sermayesi stoku ekonomik büyüme ve sosyoekonomik kalkınma için kritik önem taşımaktadır (Hanushek & Woessmann, 2015). İnsan sermayesi hem okullarda hem de hanelerde

yürütülen eğitim faaliyetleri sonucu üretilmektedir. Ancak, okullardaki eğitim faaliyetleri, politika araçlarının doğrudan uygulanması açısından daha elverişlidir.

Aileler ve devlet kurumları insan sermayesine ne miktarda ve nasıl yatırım yapacaklarına dair bir karar süreci yaşamaktadır. İktisat bakış açısı ve analiz yöntemleri söz konusu karar süreçlerine önemli katkı sağlama potansiyeli taşımaktadır. İnsan sermayesi yatırımları açısından önemli kararların önde gelen birini de özel okullara yönelik alınan yatırım kararları oluşturmaktadır. Özel okulların devlet okullarına kıyasla daha iyi bir eğitim kalitesi sağladığı varsayımı yaygın bir olgudur. Bunun başlıca bir nedeni özel okulların daha fazla rekabet etkisi altında olmalarıdır. Aynı zamanda özel okullar, başvuran adaylar arasından öğrencilerini bizzat seçmektedir. Bu da özel okullarda daha nitelikli bir akran grubu oluşmasına imkan sağlayabilmektedir. Özel okullar sağladıkları bu avantajlara karşılık velilerden okul ücreti talep etmektedir.

Aileler okul sektörüne (özel/devlet) yönelik kararlarında fayda ve maliyetleri göz önünde bulundurmaktadır. Anne babalar maliyet yönünden özel okul ücretlerine ilişkin piyasa fiyatlarını inceleyerek fikir sahibi olabilmektedir. Ancak faydaların hesaplanması açısından benzer bir net gösterge bulunmamaktadır. Gerek akademik kazanımlar gerekse diğer çıktılar açısından okulda geçen sürelerin ne kadar fayda sağlamış olduğu bilgisi ebeveynlerin ve kamu yetkililerinin daha sonraki kararlarına ışık tutacaktır.

Öte yandan kamu karar vericileri, ailelerin okul seçimini genişletmeyi hedefleyen teşvik programlarına yatırım yapmaktadır. Amaçlanan doğrudan fayda, *katılım etkisi* olarak ifade edilen, sağlanan teşvikten dolayı öğrencinin kendisine daha uygun bir okulda öğrenim görmesinden kaynaklı insan sermayesi artışıdır. Katılım etkisinin ortaya çıkabilmesi için, teşvik nedeniyle okul kararının özel okuldan devlet okuluna dönüşmesi veya zaten özel okula yazılacaklar için farklı bir özel okula kayıt olunması gereklidir. Teşvik olsa da olmasa aynı özel okulda eğitim görecek öğrenciler için katılım etkisi söz konusu olmamaktadır. Beklenen dolaylı fayda ise, *rekabet etkisi*

olarak adlandırılan, bir bölgede özel okul yoğunluğunun artmasından dolayı o bölgedeki devlet okullarının da eğitim kalite ve verimliliğini artırmaları beklentisidir. Ancak, daha becerili öğretmen ve öğrencilerin, özel okul teşvikinin etkisiyle özel okullara daha yoğun geçiş yapması sonucunda, devlet okullarında ortalama öğretmen ve akran kalitesi düşebilmekte, bu da *kaymağını-sıyırma* şeklinde ifade edilen bir etkiye yol açabilmektedir. Söz konusu mekanizmanın devlet okullarındaki eğitimin kalitesi ve verimliliği üzerinde negatif bir etki oluşturması beklenir.

Özel okullara yönelik teşvik programlarının net etkilerinin belirlenmesi hangi etkinin daha güçlü olduğunu ortaya koymaktadır. Net etkilere ilişkin bulgular, teşvik programının tasarım detayları ve teşvikin uygulandığı dönemdeki kurumsal arka plan hep birlikte göz önüne alındığında, gelecekte hem daha iyi teşvik programlarının tasarımı için hem ulusal hem de uluslararası düzeyde dersler çıkarma imkânı doğabilmektedir (Egalite & Mills, 2021).

Bu çalışma aşağıda yer alan üç araştırma sorusuna cevap aramaktadır. İlk iki araştırma sorusu aynı zamanda özel okul etkilerine ilişkin literatürde cevabı aranan ana iki soruya karşılık gelmektedir. Urquiloa (2016) eğitim ekonomisi alanındaki el kitaplarının beşincisinde söz konusu iki sorunun da cevaplarının hala tam olarak ortaya konamadığı ifade etmektedir. Aynı zamanda, bu politika alandaki kamu müdahaleleri ülke bağlamıyla yakından ilişkilidir ve literatürdeki bulgular Amerika Birleşik Devletleri, Birleşik Krallık, Şili, İsveç ve Hindistan şeklinde sıralanabilecek sayılı ülkeyle sınırlı kalmaktadır.

Üçüncü araştırma sorusu ise yakın dönemde öne çıkan ve literatür birikimi oluşmaya devam eden (hem bilişsel ve hem de bilişsel-olmayan becerilere ilişkin) cinsiyet kazanım farkının özel okullar ve devlet okulları bağlamında farklılık gösterip göstermediğini incelemektedir. Cinsiyet kazanım farkına ilişkin okul etkenleri çeşitli çalışmalarda ele alınmış olsa da, okul sektörü açısından söz konusu farkların nasıl değişim gösterdiği daha önce araştırılmamıştır.

Araştırma Soruları:

Araştırma Sorusu 1: Özel okullar insani sermaye üretiminde devlet okullarından daha mı etkilidir?

Araştırma Sorusu 2: Özel okullardan kaynaklanan rekabet, devlet okullarının insani sermaye üretimindeki etkinliği artırmakta mıdır?

Araştırma Sorusu 3: Cinsiyet kazanım farkı okul sektörüne göre farklılık göstermekte midir?

Her üç soru kapsamında da insani sermaye ölçütü olarak Türkiye'de 2013-14 ve 2016-2017 akademik yılları arasında gerçekleştirilen Temel Eğitimden Ortaöğretime Geçiş (TEOG) sınavından elde edilen standartlaştırılmış test skorları kullanılacaktır. Üçüncü soru için, son üç yılda okuldaki derslerde elde edilen not ortalaması ek bir insani sermaye ölçütü olarak ele alınacaktır.

Literatür Taraması ve Teorik Çerçeve:

Bu bölüm, niceliksel ve niteliksel insan sermayesi ölçütleri, insan sermayesinin nasıl üretildiği, özel okul etkilerinin neler olduğu ve bu etkileri nasıl öğrenebileceğimiz ile ilgili teorik çerçeveyi ve literatür bulgularını ele almakta ve dört alt bölümden oluşmaktadır. İlk alt bölümde, insan sermayenin nicel ölçümlerini tanımlanmaktadır. Bu bölüm aynı zamanda insan sermayesi teorisinin ortaya çıkışını, gelişimini ve hem mikroekonomik hem de makroekonomik araştırmalar üzerindeki güçlü etkisini özetlemektedir. İkinci bölüm, genellikle standartlaştırılmış test puanlarıyla ölçülen bilişsel becerilerin nitel ölçümlerine odaklanmaktadır. Ayrıca, gelecekteki eğitim ve işgücü piyasası sonuçları için çok etkili olduğu tespit edilen bilişsel olmayan becerilere de değinilmektedir. Üçüncü alt bölüm, eğitimin girdileri ve çıktıları arasındaki üretim ilişkilerini analiz etmek için kullanılan eğitim üretim fonksiyonu yaklaşımını ortaya koymaktadır. Son alt bölüm, özel okul etkilerini ve özel okul seçimini teşvik eden politikaların genel çerçevesini ve literatürdeki bulguları özetlemektedir. Dört alt

bölümün tümü birbiriyle yakından ilişkili olup, analiz modellerinin oluşturulmasında da birbirini tamamlayan niteliktedir.

İnsan sermayesinin nicel ölçütleri, okullarda eğitimde geçirilen yılları veya elde edilen en yüksek derece veya diplomayı içermektedir. Bunlar *edinim* ölçütleri olarak tanımlanmaktadır. Mikro perspektiften bakıldığında, bu bireylerin edinimine karşılık gelmekteyken, makro-perspektife göre de ülkedeki ortalama eğitim süresi veya belirli bir dereceye / diplomaya sahip vatandaşların payı ile ölçülen toplam insan sermayesi stokuna işaret etmektedir.

İnsan sermayesine yapılan yatırımları bireysel düzeyde analiz etmek için Mincer modeli (Mincer, 1958, 1975) kullanılmaktadır. Mincer modeli, temel olarak ücret göstergelerini insan sermayesi girdileriyle ilişkilendiren bir modeldir. Bu model en yaygın olarak kullanılan ampirik modellerden biri olup, eğitimde geçen her ilave sene ile ilişkili ücretlerdeki yüzde artış, 5 ila 15 arasındadır (Hanushek & Woessman, 2008, s.615; Chattopadhyay, 2012). Bu değerler, ilköğretim düzeyinde, kızlar açısından ve daha az gelişmiş ülkelerde daha yüksek bulunmuştur (Psarcharopoulos & Patrinos, 2018).

Rakip bir teori olan sinyal/tasnif hipotezi ise eğitimin, insan sermayesini arttırdığı için değil, doğuştan gelen içsel becerilere işaret ettiği için önemli olduğunu öne sürmektedir (Spence, 1973; Stiglitz, 1975). Bu yaklaşıma göre eğitimin esas fonksiyonu, daha düşük yetenek ve motivasyona sahip adaylarla, daha yüksek yetenek ve motivasyona sahip adayları birbirinden ayırabilmesidir. İşverenler, potansiyel çalışanlarının doğuştan gelen becerileri ve motivasyonları hakkında bilgi asimetrisi ile karşı karşıyadır. Öğretmenler ve okul sistemi söz konusu bilgi asimetrisini gidermeleri nedeniyle değer yaratmaktadır. Bu bakış açısının bir yansıması ise okulda geçen süreler ile ilgilidir. Öğrencinin doğuştan gelen beceri ve motivasyon seviyesini ortaya koymaya yetecek en kısa eğitim süresi ideal eğitim süresi olmasıdır. Bu anlamda, eğitimin insan sermayesine katkısını eğitimde geçen süre ile ölçen insan sermayesi teorisinin nicel ölçütlerinin tam tersi bir çıkarım söz konusudur.

Literatürde oluşan konsensus, eğitimin sadece bir sinyal ve tasnif aracı olmadığı, öğrencilerin insan sermayelerini bizzat artırdığı yönündedir. Card (1999), Oosterbeek (1992), Woessmann (2016) ve diğer birçok çalışma, doğuştan gelen beceriler kontrol edildiğinde dahi, okul döneminde geçen sürenin insan sermayesin pozitif katkı yaptığını ortaya koymaktadır. Sinyal / tasnif teorisinin öncülerinden Spence (2002) de her iki mekanizmanın birlikte fonksiyon gösterdiğini ifade etmektedir.

Makro perspektiften bakıldığında ise insan sermayesini ilk kez açıkça ele alan model Mankiw, Romer & Weil (1992) tarafından ortaya konan genişletilmiş Solow modelidir. Bu model işgücünü eğitilmiş ve eğitilmemiş işgücü olarak iki grup olarak ele almaktadır. Bu modelde eğitim bir üretim faktörü haline gelmektedir. Büyümeye girdi sağlayan insan sermayesi daha fazla eğitimli işgücü birikiminden kaynaklanmaktadır. İnsan sermayesine vurgu yapan diğer bir model ise Lucas (1998)'ın içsel büyüme modelidir. Bu modelde iki ayrı sektör bulunmaktadır. Sektörlerden biri insan sermayesini kullanarak çıktı üretirken, diğer sektör ise mevcut insan sermayesini yeni insan sermayesi yaratmak için kullanmaktadır. Genişletilmiş Solow modelinden farklı olarak, insan sermayesi stoku arttıkça, çıktının limitsizce artışı söz konusu olabilecektir. Benzer bir içsel model de Romer (1990) tarafından önerilmiş olup, söz konusu modelde insan sermayesi araştırma geliştirme sektörünün girdisi olarak ele alınmaktadır. İnsan sermayesinin temel rolü ekonominin yenilik kapasitesinin artırılmasıdır. Üçüncü bir model ise Benhabib & Spiegel (1994) ve Nelson & Phelps (1996) tarafından ortaya konan teknoloji difüzyonu modelidir. Bu yaklaşımda insan sermayesi toplam faktör verimliliğine girdi sağlaması ile büyümeye katkı sağlamaktadır. Eğitim ile artırılan insan sermayesi, toplam faktör verimliliğinde lider ülkeleri yakalamak için gerekli itici gücü oluşturmaktadır. Bu da diğer ülkelerde geliştirilen teknolojilerin transferi ve ülke içinde difüzyonu ile gerçekleşmektedir. Bunu gerçekleştirebilmek için iyi eğitimli insan gücüne ihtiyaç duyulmaktadır.

İnsan sermayesinin makro etkilerine ilişkin Sianesi ve Van Reenen (2003, s.177-180)'de incelenen on üç çalışma kapsamında, ülkedeki ortalama eğitim düzeyinin bir yıl artması durumunda, genişletilmiş Solow modeli kişi başına düşen çıktı seviyesini

yüzde 3 ila 6 oranında artmasını tahmin etmektedir. İçsel büyüme teorilerine dayanan modellerde ise aynı durumda yüzde bir puanlık daha hızlı büyüme tahmin edilmektedir.

Brown ve Saks (1987, s.319) ve Hanushek (2012), eğitim yoluyla insan sermayesi üretiminin ölçülmesinde sadece eğitimin niceliğinin (eğitimde geçen süre) ötesinde eğitimin niteliği (kazanım) üzerinde de durulması gerektiğini öne sürmektedir. Kazanımlar standartlaştırılmış test skorları ile ölçülmektedir (Hanushek, 2017). Uluslararası Öğrenci Değerlendirme Programı (PISA) ve Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMMS) gibi uluslararası sınavların yanı sıra Amerika Birleşik Devletlerinde gerçekleştirilen Eğitiminde İlerlemenin Ulusal Değerlendirmesi (NAEP) ve ülkemizdeki TEOG gibi ulusal çapta sınavlar kazanım düzeylerinin belirlenmesinde kullanılmaktadır. Bu sınavların ortak özelliği okullarda öğrenim gören öğrencileri kapsamasıdır. Bunlardan farklı olarak, Uluslararası Yetişkin Becerilerinin Ölçülmesi Programı (PIAAC) tarafından üretilen skorlar işgücü piyasasında becerilerin önemini doğrudan ölçmektedir. Okuryazarlık ve matematiksel beceriler tüm bu sınavlarda ortak olarak test edilmektedir ve PIAAC bulguları bu iki beceri başlığının ekonomik getiriler anlamında en fazla katkıyı sağladığını ortaya koymaktadır (Hanushek ve ark., 2015).

Standartlaştırılmış test skorlarını ve gelecekteki gelir düzeyini birbiri ile ilişkilendiren geniş bir literatür bulunmaktadır. Bu kapsamda, Chetty, Friedman & Rockoff (2014), Dougherty (2003), Hanushek (2009), Heckman, Stixrud & Urzua (2006), Lin, Lutter & Ruhm (2019), Rose (2006), Ozawa et al. (2022) ve Watts (2020)'da yer alan literatür özetleri, standartlaştırılmış test puanlarının, gelecekte işgücü piyasasında aktif olarak yer alma ve kazanılan ücret düzeyiyle doğrudan ilişkili olduğunu ve bunun hem gelişmiş hem de gelişmekte olan ülkelerde geçerli olduğunu ortaya koymaktadır.

Cinsiyet kazanım farkı, yani kızlar ve erkekler arasındaki ortalama test puanı farkları, gelecekte işgücü piyasasındaki ücret farklarını öngörmesi ve eğitim alanındaki eşitsizliğin bir yansıması olması nedeniyle ekonomistler tarafından da gittikçe daha

fazla araştırılan bir alan konumundadır (Lai, 2010). Bu çalışmalardaki bulgular, kadınların eğitimin birçok alanında önemli ilerlemeler kaydettiğini ve hatta hemen her alanda erkeklerden daha iyi bir ortalama performans gösterdiğini göstermektedir (DiPrete & Buchmann, 2013; Gül, 2006; UNESCO, 2022).

Öte yandan, Gintis (1971) ve Kautz ve ark. (2014), bilişsel olmayan becerilerin önemini vurgulamakta ve gelecekteki eğitim ve işgücü piyasası çıktılarını en az bilişsel insan sermatesi ölçütleri kadar tahmin ettiklerini iddia etmektedir. Daha da önemlisi, bilişsel olmayan beceriler aileler, okullar ve sosyal ortamlar tarafından tüm K12 dönemi boyunca geliştirilebilir beceriler olarak değerlendirilmektedir.

Öte yandan, öğretmenler genellikle öğrencilerin davranışlarına yansıyan bilişsel-olmayan becerileri dikkate almaktadır. Bu beceriler arasında okula devam durumu, sınıfta derse konsantre olma, ödevleri zamanında yapma ve sınıfta disiplin kurallarına uygun şekilde davranma gibi davranış ve tutumlar yer almaktadır. Ayrıca, öğretmenlerin gözlemledikleri söz konusu davranış ve tutumları öğrencilerinin notlarına yansıttıkları görülmektedir (Jackson, 2018, Willingham, Pollack & Lewis, 2002). Bu nedenle, öğretmenler tarafından verilen notlar ve standartlaştırılmış testlerden alınan skorlar, eğitim sonuçlarının farklı göstergeleri olarak kabul edilmektedir. Borghans, Golsteyn, Heckman & Humphries (2016), ortak testlerin varoluş nedeninin, öğretmenlerce verilen notlarda öğrencilerin kişilik özellikleri ve tutumlarının da nota yansıması olduğunun altını çizmektedir. *Cinsiyet notlandırma farkı* olarak bilinen, ortak test skorları ve öğretmenin kendisi tarafından verilen notlar arasındaki farklar, siyah ve beyaz öğrenciler arasındaki başarı farkını (Jencks & Phillips, 1998), cinsiyete dayalı ücret farkını (Bertrand, Goldin & Katz, 2010) ve sosyal sınıfla ilgili boşlukları (Kautz ve ark., 2014) anlamak için araştırılmıştır. Heckman, Pinto ve Savelyev (2013) bu farkların gelecekteki işgücü katılımı ve ücretler için güçlü belirleyiciler olduğunu saptamaktadır.

Niteliksel insan sermayesi ölçütlerine makro perspektiften bakıldığında ise, Hanushek ve Kimko (2000), öncü çalışmalarında, ülkelerin ortalama test skorları arasında

uluslararası karşılaştırmanın yapılabileceği yöntemleri ortaya koymuştur. Hanushek & Woessmann (2008, 2015)'ın takip eden çalışmaları, standartlaştırılmış test puanlarında bir standart sapma iyileşmenin ekonomideki büyüme oranını 2 yüzdelik puan artırdığını göstermektedir. Hanushek & Woessmann (2015), aynı zamanda, düşük başarılı öğrencilere mi yoksa yüksek başarılı öğrencilere mi odaklanılacağı konusuna dikkat çekmektedir. Yüksek başarı gösteren öğrenciler, iş liderlerinin, yenilikçilerin veya yıldız bilim adamlarının rollerini üstlenme potansiyeline sahip oldukları için önemlidir. İçsel büyüme modelleri (Lucas, 1998; Romer, 1990), bu öğrenciler tarafından oluşturulacak Ar-Ge işgücünün rolünü vurgulamaktadır. Öte yandan, akademik açıdan çok başarılı olmasalar da, belli bir başarı düzeyini karşılayan çok daha geniş öğrenci gruplarına odaklanma, hâlihazırda var olan teknolojilerin farklı endüstrilerde kullanımını sağlamak için değerlidir. Genişletilmiş Solow modeli (Mankiw ve ark., 1992), temel becerilerin, özellikle de okuryazarlık becerilerinin birikimini vurgulamaktadır. Hanushek & Woessmann (2015), her iki gruba da odaklanmanın ekonomik büyümeye katkıda bulunacağını ve her iki öğrenci grubunun birbirini tamamlayıcı olduğunu ifade etmektedir.

Eğitim üretim fonksiyonu modeli, insan sermayesinin eğitim yoluyla üretimini incelemek için geliştirilen temel teorik ekonomik çerçevedir. Eğitim üretim fonksiyonu, okul girdileri ile (genellikle standartlaştırılmış bir test puanı ile ölçülen) okul çıktılarının arasındaki verimlilik ilişkisini ortaya koyar (Bowles, 1970, s.12; Todd & Wolpin, 2003, s.F3). Eğitim ekonomistleri, okullardaki ve hanelerdeki öğrencilerin insan sermayesi birikim süreci ile bir firmanın üretim süreci arasında benzerlik kurmaktadır. Nihai insan sermayesi seviyesini belirleyen her şey, üretim sürecine girdi olarak kabul edilir. Bu girdiler öğrenci girdilerini, ebeveyn girdilerini, akran girdilerini, okul girdilerini ve diğer çevresel girdileri içerir. Söz konusu *girdi-çıktı yaklaşımı*, eğitim süreçleri sonucu beceri oluşumunun incelenmesine yönelik ekonomik yaklaşımın belirgin özelliğidir (Becker, 1999). Söz konusu girdiler, öğrencinin doğuştan gelen beceri ve kişilik özellikleri üzerine yıllara sari olarak eklenen aile ve okul girdilerinin kümülatif toplamına karşılık gelmektedir. İdeal olan,

bu girdilerin eğitim üretim fonksiyonu modelinde doğrudan kontrol edilmesidir. Ancak "evde ebeveynlerin çocuklarına eğitim desteği sağlamak üzere ayırdıkları süre" gibi direk girdi verilerine erişim çoğunlukla mümkün olmadığı için, "ebeveynlerin eğitim durumu (edinimi)" gibi vekil değişkenler analiz aşamasında kontrol değişkeni olarak kullanılmaktadır.

Özel okulların kazanımları artırmada devlet okullarına göre daha etkili olduğu genel bir öngörü olarak değerlendirilebilir. Özel okulların, devlet okullarına göre daha yüksek seviyede bir rekabet baskısıyla karşı karşıya kalmalarından dolayı, daha kaliteli eğitim verme yönünde daha fazla motivasyona sahip olmaları beklenen bir durumdur (DeAngelis & Wolf, 2020). Eğer özel okullar söz konusu öngörüleri doğrulayacak şekilde öğrencilerin insan sermayelerine daha fazla katkı sağlıyorsa, bu durumda hem aileler hem de teşvik programı yürüten kamu yetkilileri için özel okul kararı daha anlamlı hale gelecektir. Aileler için akademik başarı tek başına önemli olmayabilir zira okul güvenliği, spor ve sanat altyapıları gibi farklı faktörler özel okul kararında etkin rol oynayabilir. Ancak, diğer her şey sabitken, daha yüksek kazanım düzeyinin bu kararı daha çekici hale getirmesi beklenir. Kamu yetkilileri de ülkenin insan sermaye stokunu artırmak, öğrencileri kendilerine en uygun okulla eşleştirmek, eğitimde fırsat eşitliğini artırmak ve özel okulların devlet okulları için rekabet ortamı sağlayacak olması gibi gerekçelerle özel okul teşvik programlarını tasarlamaktadır. Hem ebeveynlerin hem de kamu yetkililerin gelecekteki kararlarına girdi sağlamak üzere, geçmiş dönemlerdeki özel okul kararlarının veya özel okullara yönelik teşviklerin ne kadar fayda sağladığının ortaya konması önem taşımaktadır.

Kurumsal Arka Plan:

Kurumsal ve politik arka plan, bu çalışmada ele alınan dönem süresince oldukça dinamik bir yapı sergilemektedir. Kurumsal arka planın anlaşılması, çalışmanın bulgularını yorumlanması açısından da önem taşımaktadır. Söz konusu dönem, özel okulların insan sermayesi üretimi üzerindeki etkilerini incelemek için oldukça uygun ve diğer dönemlerde görülmeyen özellikler taşımaktadır.

İlk olarak, TEOG sınavı hem özel okul öğrencileri hem de devlet okulu öğrencilerinin zorunlu olarak girmesi gereken bir sınavdı. Ortaokulun son yılı her iki dönemde birer kez gerçekleştirilen iki adımlı TEOG sınavı, bu dönemlerdeki okul sınavlarının da birinin yerine geçmekteydi. TEOG, kendisinden önceki ve sonraki sınav sistemlerindeki gibi öğrencilerin veya velilerin kararına bağlı olarak girilmesi tercihe kalmış bir sınav özelliği taşımamaktaydı. Yine bu dönemde okulların ortalama TEOG skoru, okul başarısını doğrudan temsil eden bir gösterge olarak görülmekteydi ve bu da TEOG sınavının okul yöneticileri ve öğretmenlerce de önemli görülmesini sağlamaktaydı. Özel okul etkilerinin araştırılmasında önemli kısıtlardan birini tüm öğrencilerin objektif olarak karşılaştırılabileceği çıktı ölçütlerinin olmaması oluşturmaktadır. Bu anlamda, bu çalışma ile söz konusu kısıtlılık aşılmaktadır.

İkincisi, çalışmanın kapsadığı dönemde özel okul sayısı ve tüm okullar içindeki oranının ciddi anlamda artmasına yardımcı olan, ülke çapında geçerli bir özel okul teşvik programının yürütülüyor olmasıdır. Teşvik programının amaçları Milli Eğitim Bakanlığı tarafından şu şekilde belirlenmiştir: (i) eğitim kalitesine odaklı rekabetçi bir eğitim sistemi kurmak, (ii) özel okulların olanaklarından yararlanmak isteyen tüm öğrencilerin yararlanmasını sağlamak ve (iii) özel sektörün eğitim yatırımlarını teşvik etmek ve tetiklemek. Sadece birkaç yıl içinde, özel ortaokullardan mezun olan öğrenci sayısı yüzde 63,2 ve özel okul öğrencilerinin toplam öğrenci sayısı içindeki oranı yüzde 78,9 artış göstermiştir. Bu değerler çok uzun dönem boyunca oldukça yatay bir seyir izledikten sonra söz konusu artışların olması dönemi farklı kılmaktadır ve istatistiki analizler için gerekli veri oynaklığının sağlanmış olması yine bu çalışmanın önemini artırmaktadır. Özelikle ikinci araştırma sorusunun etkin şekilde analiz edilmesi için ülke genelinde gerçekleştirilen ve uygulanmasında bölgeler arası farklar olan bir teşvik programının bulunması büyük önem taşımaktadır.

TEOG sınavı bilişsel becerileri ölçen standartlaştırılmış bir test sınavıdır. Ortaokul eğitiminin sadece son (dördüncü) yılında, iki yarıyıla dağıtılarak gerçekleştirilmiştir. Her dönemde, öğrenciler iki gün boyunca 6 farklı öğrenme alanında 6 farklı teste tabi tutulmuştur. Bu testler, matematik, fen bilimleri, okuma-anlama, yakın tarih, dini

bilgiler ve İngilizce yeterliliğini ölçmüştür. Her test 20 soru içermiş ve test başına 40 dakika süre tanınmıştır. Öğrencilerin dinlenebilmeleri için testler arasında yarım saatlik molalar verilmiştir. Geçerli nedenlerle sınavlara giremeyen öğrenciler için ikinci bir mazeret sınavı imkânı tanınmıştır. Ortalama olarak, kızlar yüksek riskli testlere tabi tutulduklarında erkeklerden daha fazla olumsuz yönde etkilenmektedir (OECD, 2017; Niederle & Vesterlund, 2010; Saygın, 2018). TEOG sınavı yukarıda sayılan operasyonel özellikleri sayesinde, öğrencinin cinsiyetine bağlı sınav kaygısı faktörünün etkilerini sınırlamıştır.

TEOG testi, devlet okullarını seçecek tüm öğrencileri merkezi olarak tahsis ettiği için bu öğrenciler için oldukça belirleyici bir sınavdı. Bir puanlık fark, evden uzakta bir liseye kayıtlı olmak anlamına gelebilmekteydi. Bir öğrenci evinin yakındaki seçkin devlet okulu için gerekli notu alamadığında, ebeveynleri önemli bir maliyet farkı yaratacak olan özel bir okul alternatifine yönelebilmekteydi. Bu, TEOG'u önemli sonuçlar doğuran bir sınav statüsüne taşımaktaydı. Öte yandan, özel liseler de öğrenci seçerken TEOG sınav sonuçlarını göz önüne almaktaydı. Bu nedenle, daha itibarlı bir özel bir liseye girmek için TEOG sınavında da yüksek bir puan gerekmekteydi. TEOG sınavı tasarımı gereği okuldaki derslerde görülen müfredatla aynı kazanımları ölçmekteydi. TEOG ayrıca ortaokulun son yılında her iki dönem için sınıf içi sınavlardan birinin yerine yapılıyordu. Bu nedenle, her iki sınavda da aynı bilişsel içerik test edilmekteydi.

Kullanılan Veri:

Hem Burgess (2016) hem de Hanushek (2020), okul faktörlerinin etkilerinin araştırılmasında, son zamanlardaki en dikkat çekici yeniliğin ulusal, büyük ölçekli idari veritabanlarının kullanılması olduğunu vurgulamaktadır. Bu çalışmanın önemli bir katkısı, 4,8 milyondan fazla öğrenciyi ve birbirini izleyen dört kohorta yayılmış 18 bin özgün okulu kapsayan, öğrenci ve okul popülasyonlarının tamamına karşılık gelen çok geniş kapsamlı verileri literatüre tanıtmasıdır.

Çalışma verileri birçok çalışmada genellikle metodoloji başlıklı kısımda özetlenmektedir. Bu çalışmada, popülasyon düzeyindeki bu veri setinin ilk kez kullanılıyor olması nedeniyle, detaylı veri özetlerinin sunulduğu ve teorik çerçeve ile bağların kurulduğu bir veri sunumu, metodoloji bölümünün hemen öncesinde ayrı bir bölüm dahilinde sunulmuştur.

Verinin ham halinde özel okul ve devlet okulu ortalama başarı farkının 1.31 standart sapma (ss) gibi oldukça yüksek bir düzeyde olduğu görülmektedir. Tüm öğrencilerin ortaokula devam ettiği durumda bir olması gereken kız öğrencilerin erkek öğrencilere oranının devlet okullarında 0,93, özel okullarda ise 0,84 olduğu gözlemlenmektedir. Bu anlamda kızlar aleyhine bir *cinsiyet edinim farkı* söz konusudur. Ancak yine verinin ham halinde TEOG skorları açısından 0,32 ss, not ortalamaları açısından 0,44 ss'lik kızların avantajına bir *cinsiyet kazanım farkı* bulunmaktadır. Girdi değişkenler arasında ebevynlerin eğitim durumları, çalışma statüleri ve ayrı yaşayıp yaşamadıkları hem TEOG skorunu hem de not ortalamasını güçlü şekilde etkileyen aile girdileri olarak öne çıkmaktadır. Öğrencinin kaç tane kardeşi olduğu ve evde kendisine ait bir odası olup olmadığı da diğer önemli aile girdilerini oluşturmaktadır. Literatüre kıyasla okulun öğrenci sayısının büyüklüğü, okulun yer aldığı lokasyonun gelişmişlik seviyesi, okulun ikili eğitim yapıp yapmadığı gibi okul faktörler ham veride öğrenci başarısıyla çok ilgili gözükmemektedir. Genel olarak, özel okul öğrenci grubunun hem aile girdileri hem de okul girdileri açısından hissedilir derecede daha avantalı olduğu görülmektedir.

Metodolojik Yaklaşım:

Bu çalışmanın metodolojisi, ana hatlarıyla, eğitim üretim fonksiyonu teorik çerçevesi bağlamında, gözlemsel verilere dayalı nicel araştırma yöntemlerinin uygulanmasını kapsamaktadır.

Eğitim üretim fonksiyonu yaklaşımı ile cevaplanan soru "*girdi değişkenlerde dışsal bir değişiklik sonucu, diğer tüm girdi değişkenler sabitken, kazanım çıktıları nasıl etkilenir*?" şeklindedir (Todd & Wolpin, 2003). Bu soruyu yanıtlamak için üretim

teknolojisi hakkında bilgi sahibi olmak, diğer bir ifadeyle eğitim üretim fonksiyonun yapısal parametrelerini tahmin etmek yeterlidir. Üretim teknolojisinin özellikleri hakkında bilgi edinmek için gözleme dayalı veriler kullanılabilir. Geçmiş ve son dönemdeki girdilerin tümünün gözlemlenebilmesi bunun için yeterli olmaktadır. Söz konusu girdilerden bir kısmının eksik olması durumunda ekonometrik çözüm yöntemleri uygulanarak yansız tahminler elde edilmeye çalışılmaktadır.

Metodolojik yaklaşımın önemli bileşenlerinden biri, ayrıntılı bir eksik veri stratejisinin benimsenmesidir. Öğrenci verilerinin bir kısmı, PISA gibi benzeri diğer sınavlarda da görülen makul bir oranda veri eksiklikleri barındırmaktadır. Bu eksiklikler TEOG skoru veya not ortalamaları gibi çıktı değişkenlerinden ziyade, açıklayıcı değişkenler bazında ve bu değişkenlerin bir kısmında görülmektedir. Eksik verilerin hangi değişkenlerde görüldüğü öğrenciden öğrenciye değişiklik göstermektedir. Eksik veriyi ele alma adına farklı yöntemler bulunmaktadır. Bu yöntemler arasından, veri seti büyüklüğü ve veri eksikliklerinin açıklayıcı değişkenlerde yoğunlaşması göz önüne alındığında, tam bilgi en çok olabilirlik tahmincisi (full information maximum likelihood) ile tam gözlemler/liste bazında silme (complete-cases/listwise deletion) yaklaşımları en uygun metodlar olarak öne çıkmaktadır. Tam bilgi en çok olabilirlik tahmincisi tüm öğrencilerin gözlemlenebilen ne kadar değişken değerleri varsa tümünü kullanmaktadır. Tam gözlemler metodunda ise yalnızca değişkenlerinin tümünde değerler gözlemlenen öğrencilere analizde yer verilmektedir. İlk metot, veri setinde yer alan tüm bilgiyi kullanması açısından avantajlıyken, eğer veri eksikliği, eksik değerin olduğu değişkenden kaynaklanıyorsa, bu durumda tam gözlemler yöntemi daha güvenilir sonuçlar verebilmektedir. Farklı güçlü ve zayıf yönlere sahip iki metodun karşılaştırmalı kullanımı ile bulguların dirençliliğinin kontrol edilmesi amaçlanmıştır.

İlk araştırma sorusu kapsamında iki farklı özel okul etkisi araştırılmaktadır. İlki, Tip-1 özel okul etkisi olarak adlandırılmakta ve okul girdileri ile akran girdilerinin etkilerinin birleşimden oluşmaktadır. Bu etkiyi tahmin etmek üzere İki-Adımlı tahminci (Castellano, Rabe-Hesketh, & Skrondal, 2014) kullanılmaktadır. İlk adımda

sabit etkiler modellemesi ile okul ve yıl sabit etkileri birlikte kontrol edilerek, sadece aynı yıl ve aynı okuldaki öğrencilerin aile ve öğrenci girdi değişkenlerinin parametre değerleri yansız şekilde tahmin edilmektedir. Akabinde, söz konusu parametre tahmin değerleri ve aile değişken değerleri kullanılarak her okul kapsamında hane girdilerinin TEOG skoruna katkıları tahmin edilmekte ve öğrencinin TEOG sınav notundan çıkarılmaktadır. Böylelikle kalan dağılımda, öğrencilerin aile girdilerinden kaynaklı olarak okullara yanlı tasnifinin etkisi ayrıştırılmıştır. İkinci adımda, kalan değerler iki seviyeli regresyon yöntemiyle okul değişkenlerine karşı modellenmektedir.

İkinci etki ise Tip-2 özel okul etkisi olarak adlandırılmakta olup, sadece okul girdilerin etkilerinden oluşmaktadır. Bu etkiyi tahmin etmek üzere *ilişkili rastsal etkiler modeli* (Wooldridge, 2010; Snijders & Bosker, 2012) uygulanmaktadır. Bu yöntemin temel özelliği aile girdi değişkenlerine (ebeveynlerin eğitim durumu, öğrencinin evde kendine ait bir odasının olması, vb.) ilişkin okul ortalamalarının da modele eklenmesidir. Bu, Eğitim üretim fonksiyonu teorik çerçevesinde de akran girdilerine karşılık gelmektedir. Altonji & Mansfield (2018) kurdukları denge modeli ile, zengin bir gözlemlenebilen aile girdileri değişken setine analizde yer verildiğinde gözlemlenemeyen aile değişkenlerinin de kontrol edildiğini göstermektedir. Bu çerçevede, *ilişkili rastsal etkiler modeli* akran etkilerini kontrol edebilmekte, ancak, bunu yaparken okul etkilerinin bir kısmını da yutmaktadır. Burada hala devam eden temel sorun, gözleme dayalı veri ile analiz yapılırken, okul girdileri ile akran girdilerinin birbirinden net olarak ayrıştırılmasının mümkün olmamasıdır. Bu nedenle ilişkili rastsal etkiler modeli ile tahmin edilen değer, özel okul etkisinin okul girdilerinden kaynaklı kısmı için ancak bir alt sınıra tekabül etmektedir. Eğer aile girdilerine ilişkin değişkenlerin okul ortalamaları ile (gözlemlenemeyen okul girdilerine karşılık gelen) okul rastsal terimlerinin arasında pozitif bir ilişki varsa ve yine okul ortalamalarının parametre değerleri pozitif ise (yani akran etkileri pozitif ise), bu durumda Tip-2 özel okul etkisinin üst sınırı, İki-Adımlı tahminciyle tahmin edilen Tip-1 özel okul etkisine karşılık gelecektir. Pratik ifadeyle, eğer iyi öğrencileri çeken okullar aynı zamanda daha iyi yönetici ve öğretmenlere sahipse ve aynı

zamanda bu okullardaki akran etkileri pozitif ise, bu durumda Tip-2 özel okul etkisi gözleme dayalı veriyle dahi alt ve üst sınırlar dahilinde tahmin edilebilmektedir.

İkinci araştırma sorusu kapsamında, her ildeki özel okul öğrencisi sayısının o ildeki toplam öğrenci sayısına oranı ile yine o ildeki ortalama devlet okulu TEOG performansı karşılaştırılarak özel okul teşvik programının *rekabet etkileri* araştırılmaktadır. Teşvik programlarının etkilerinin ölçülmesinde ulusal düzeyde yürütülen bir teşvik programı kapsamında bölgeler, iller veya ilçeler bazındaki tahsis varyasyonundan faydalanılmaktadır. Bu tür destekleri deneysel yöntemlerle analiz etmek mümkün olmadığı için, ulusal çapta yürütülen ve iller bazında tahsis farklılıkları içeren teşvik programlarının varlığı önem taşımaktadır.

Diğer taraftan, teşvik programlarının rekabet etkilerinin araştırılmasında duyulan bir kaygı, özel okul yatırımcılarının okul açacakları yeri seçerken veya özel okulu tercih eden ailelerin evlerinin lokasyonunu seçerken oradaki ortalama devlet okulu kalitesini göz önünde bulundurmalarıdır. Bu da içsellik (endojenite) problemine işaret etmektedir. Özellikle kapsamlı bir teşvik programının yürürlükte olmadığı durumda, piyasa mekanizmasının bu şekilde işlemesi beklenir. Ancak kapsamlı ve cezbedici bir teşvik programının etkin olduğu dönemde, gerek eğitim alanındaki girişimcilerin gerekse ailelerin kararlarının durağan dönem piyasa dinamikleri yerine teşvik programının dinamiklerinden etkilenmesi beklenebilir. Bu çalışmanın kapsadığı dönemde kapsamlı bir teşvik programının yürürlükte olmasından dolayı içsellik etkilerinin daha az olması beklenebilir. Yine de, bu riske karşı *araç değişken* olarak teşviğin illere, illerin gelişmişlik durumuna göre farklı kotalar dahilinde tahsis edilmesi kullanılmıştır (gelişmişlik seviyelerinin de kontrol değişkeni olarak eklenmek suretiyle). Eğitim üretim fonksiyonu yine baz çıktı modeli olarak kurgulanmış, ancak bu sefer özel okul yatırımcılarının ve ailelerin lokasyon kararlarını etkilediği literatür tarafından ortaya konan açıklayıcı değişkenlere de modelde yer verilmiştir.

Üçüncü araştırma sorusu kapsamında, öğrencilerin TEOG skorları ve okul not ortalamaları bütünleşik bir şekilde modellenmiştir. Yapısal eşleştirme modeli ve

görünüşte ilişkisiz regresyon modelleri karşılaştırmalı olarak kullanılarak veri eksikliklerinin etkileri de göz önünde bulundurulmuştur.

Bulgular:

İlk önce, diğer tüm unsurlar sabitken özel okulların devlet okullarına kıyasla daha fazla insan sermayesi üretip üretmediğini araştırılmıştır. Özel okulların sağladığı okul girdileri ve akran grubu girdileri bir arada değerlendirildiğinde, özel okulların devlet okullarına göre 0,62 standard sapma (ss) daha fazla insan sermayesi ürettiği tespit edilmiştir. Bu etki tipi literatürde genellikle tahmin edilmemektedir, ancak ebeveynlerin özel okul kararı için anlamlı bir tahminci olarak değerlendirilmektedir. Ebeveynler için ister daha iyi okul girdilerinden isterse daha iyi akran girdilerinden kaynaklansın, çocuklarının sınav başarılarındaki ve insan sermayelerindeki toplam artışın onlar için anlamlı bir gösterge niteliğinde olması beklenir.

Tip-2 ortalama özel okul etkisi dahilinde ise 0,25 sd'lik fark alt sınırı saptanmıştır. Bu etki sadece okul girdileri açısından özel okul ve devlet okulu farklılıklarından kaynaklanmaktadır. Okul performansını (özellikle de okul idarecileri ve öğretmenlerin performanslarını) ölçerken akran girdilerindeki farklılıkları kontrol etmek okullar arası adil bir karşılaştırma için gereklidir. Literatürde bu etki genellikle 0.20 ss düzyinde tahmin edilmekte olup, bu çalışmanın bulgusu da literatürle uyumlu gözükmektedir.

Literatürdeki daha güvenilir tahminler genellikle rastsal olarak tahsis edilen öğrenci teşviklerine dayanmaktadır. Bu çalışmada teşviklerin rastsal olarak tahsis edilmemesi nedeniyle okul girdileri ile akran girdilerini tam olarak birbirinden izole etmek mümkün olmamıştır. Ancak, kullanılan yöntem sayesinde bir alt sınır tahmin edilmiştir.

İkinci olarak, iller bazında özel okul öğrenci oranının, o ildeki ortalama devlet okulu kazanımı ile ilişkisi modellenmiştir. İlk olarak sıradan en küçük kareler yöntemi uygulanmıştır. İl bazında özel okul öğrenci oranındaki bir standart sapmalık artışın, o ildeki devlet okulu ortalama TEOG performansını -0.045 ss azalttığı tahmin edilmiştir.

Araç değişken kullanımı sonucunda -0.036 ss düzeyinde benzer bir etki tahmin edilmiştir. Literatürde, özellikle de yakın dönem literatürde, genellikle küçük boyutta pozitif etkiler tahmin edildiği görülmektedir. Bununla birlikte, literatürde öne çıkan çalışmalardan Dee (1998) -0,023 ss, Husted & Kenny (2000) -0,060 ss ve Geller, Sjoquist & Walker (2001) -0.018 ve -0.137 ss arası (yine küçük boyutta) negatif etkiler tahmin etmişlerdir. Hem etkinin büyüklüğü hem de negatif veya pozitif olması teşvik programının tasarımıyla ve uygulama dönemindeki ülke bağlamıyla yakından ilişkilidir. Bu nedenle yukarıdaki tahminin yanı sıra teşvik programının detaylarına ve uygulama dönemindeki ülke bağlamına ilişkin ayrıntılı bilgilere çalışmada yer verilmiştir. Çalışmanın bulguları aynı zamanda daha iyi teşvik programı tasarımının önemini vurgulamaktadır.

Üçüncü olarak, *cinsiyet kazanım farkının* okul sektörüne göre farklılık gösterip göstermediği, test skorları ve ağırlıklı not ortalamalarını birlikte ele alan ortak bir modelle araştırılmıştır. Devlet okullarında TEOG skorları açısından 0,30 ss ve ağırlıklı not ortalamaları açısında 0.44 ss'lik bir fark tespit edilmiştir. Bu fark özel okullar için sırasıyla 0,13 ss ve 0,12 ss'e düşmektedir. Her iki insan sermayesi ölçütünün hata terimleri arasında 0,91'lik bir koreleasyon katsayısı tespit edilmiştir.

Literatürde daha önce cinsiyet kazanım farkı açısından anlamlı bir fark yaratan hiçbir okul faktörü bulunamamıştır (Burgess ve ark., 2004; Bharadwaj ve ark., 2016). Okul sektörünü cinsiyet kazanım farkı analizlerinde ilk kez ele alan bu çalışmanın bulguları, Türkiye bağlamında cinsiyet kazanım farkının özel okullar ve devlet okulları arasında farklılık gösterdiğini ortaya koymaktadır.

Diğer taraftan, TEOG skorunun üretim fonksiyonu ve okul notlarının üretim fonksiyonu birbirinden çıkarıldığında, kalan *cinsiyet notlandırma farkının* (Terrier, 2020) devlet okullarında 0.14 ss, özel okullarda 0.01 ss olduğu saptanmıştır. Bu strateji ile, gözlemleyemediğimiz beceriler olan doğuştan gelen öğrenme kapasitesi ve önceki dönemlerde edinilen becerilerin yanlılık doğurabilecek etkilerinden de aranılmaktadır. Literatürde, cinsiyet notlandırma farkının temel olarak cinsiyetler arası bilişsel-

olmayan beceri farklarından kaynakladığı öne sürülmektedir ve bu farklar cinsiyet eşitliği açısından kaygı yaratan farklar olarak görülmektedir. Öğretmenleri yanlı notlama yapmamaya yönlendirecek farkındalık yaratma, izleme ve teşvik etme yaklaşımlarının hayata geçirilmesi önemli görülmektedir. Diğer taraftan, öğretmenlerin sadece yanlı notlandırma farkı perspektifinden değerlendirilmemesi, onların bilişsel-olmayan becerilere katabilecekleri değere de önem verilmelidir (Jackson, 2018). Bu nedenle, bilişsel-olmayan becerileri öğrencilerine nasıl kazandıracakları konusunda öğretmenlerin profesyonel gelişimlerinin desteklenmesi önemli bir politika enstrümanı olarak değerlendirilmektedir.

C. TEZ İZİN FORMU / THESIS PERMISSION FORM

ENSTİTÜ / INSTITUTE

Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences ☐

Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences ☒

Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics ☐

Enformatik Enstitüsü / Graduate School of Informatics ☐

Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences ☐

YAZARIN / AUTHOR

Soyadı / Surname : ÖZBAYLANLI
Adı / Name : BİLGEHAN
Bölümü / Department : İktisat / Economics

TEZİN ADI / TITLE OF THE THESIS (**İngilizce** / English): PRIVATE SCHOOL EFFECTS ON HUMAN CAPITAL PRODUCTION – EVIDENCE FROM TÜRKİYE

TEZİN TÜRÜ / DEGREE**:** **Yüksek Lisans** / Master ☐ **Doktora** / PhD ☒

1. **Tezin tamamı dünya çapında erişime açılacaktır. /** Release the entire work immediately for access worldwide. ☒
2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two years. *** ☐
3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. ***** ☐

***** *Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

Yazarın imzası / Signature ..... **Tarih** / Date .....
(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)
(Library submission date. Please fill out by hand.)

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.

Procurement Of Interactive Boards