BETA
Bureau
d'économie
théorique
et appliquée

Documents de travail

«From Access to Excellence: The Dual Growth of Türkiye's Education System»

Auteurs

Nadir ALTINOK, Claude DIEBOLT

Document de Travail n° 2025 – 13

Mars 2025

Bureau d'Économie
Théorique et Appliquée
BETA

https://www.beta-economics.fr/

Contact :
jaoulgrammare@beta-cnrs.unistra.fr

Université de Strasbourg · cnrs · Université de Lorraine · INRAE · AgroParisTech · Université Haute-Alsace

From Access to Excellence:
The Dual Growth of Türkiye's Education System

Nadir ALTINOK[1] and Claude DIEBOLT[2]

Abstract: Türkiye experienced rapid growth in completion rates in both primary and secondary education over the last two decades. The completion rate at the secondary level increased from less than 60% in 2000 to almost 90% in 2023. We question to what extent this school expansion has impacted the quality of education in Türkiye. To do this, we use a new database of schooling quality from Altinok and Diebolt (*Cliometrica*, 2024) that provides comparable data from 1970 to 2020 for more than 130 countries around the world. In addition to this quality dimension, we also explore equity issues. In order to provide a macro analysis, we use a counterfactual approach by comparing the relative performance of Türkiye to the most similar developing countries. While Türkiye is one of the most challenging countries among the OECD members for ensuring access and retention of young people in formal education, we show that the recent democratization of education did not lead to a significant reduction in the quality of its schooling system. On the contrary, in parallel with this expansion of schooling access, Türkiye experienced significant progress on indicators evaluating the quality of education (mainly math and reading scores). However, as in other countries like France and Germany, inequality in access, completion, and skill levels remain high in Türkiye and will likely be among the most pressing education policy challenges in the coming years.

Keywords: PISA, Türkiye, Schooling Quality, Economics of Education, Development, Edumetrics, Cliometrics.

JEL codes: C8, I2, N3, O1.

[1]Nadir ALTINOK, BETA/CNRS & University of Lorraine (France). Email: nadir.altinok@univ-lorraine.fr (corresponding author)

[2]Claude DIEBOLT, BETA/CNRS & University of Strasbourg (France). Email: cdiebolt@unistra.fr

Genişletilmiş Özet

Son TIMSS 2023 (*Trends in International Mathematics and Science Study*) araştırmasında Türkiye dikkat çekici bir başarı göstererek 58 katılımcı ülke arasında 8. sırada yer aldı ve 553 puanlık ortalama skoru ile uluslararası ortalamanın oldukça üzerinde bir performans sergiledi (Koca ve diğerleri, 2024). Ancak, OECD'nin PISA (*Programme for International Student Assessment*) araştırması, bu önemli başarıyı tam olarak doğrulamamaktadır. Bu çelişki, bir eğitim sisteminin performansını ölçmede kullanılan verilerin kalitesinin belirleyici rolünü ortaya koymaktadır.

Bu çalışmada, çok sayıda ülkenin eğitim sistemlerinin kalitesine ilişkin kapsamlı bir veri seti kullanıyoruz. Sadece PISA gibi uluslararası sınavlardan elde edilen ortalama puanları değil, aynı zamanda aşağıdaki iki temel boyutu içeren ek ölçümleri de dikkate alıyoruz:

1. Evrensel eğitim kalitesi, öğrencilerin okuma, matematik ve fen bilimleri gibi temel becerileri kazanma yeteneğini değerlendirmektedir.
2. Gelişmiş yetkinlikler, yenilikçi bir ekonomik büyümeyi teşvik etmek için gerekli olan ileri düzey becerileri kapsamaktadır.

Bu iki boyutu birleştirerek, eğitim sistemlerinin etkinliğini ölçen özgün bir gösterge öneriyoruz. Bunun yanı sıra, eğitimde eşitliği değerlendirmek için iki ek gösterge daha sunuyoruz. Eşitlik, burada uluslararası sınavlarda en yüksek ve en düşük performans gösteren %10'luk dilimler arasındaki göreceli fark olarak tanımlanmıştır ve eğitim sistemlerinin seçiciliğini hesaba katmak amacıyla okullaşma oranları ile düzeltilmiştir.

İki farklı çalışmanın verilerini kullanarak, Türkiye'nin eğitim performansının benzer ekonomik gelişmişlik seviyesine sahip ülkelere kıyasla daha iyi olup olmadığını değerlendiriyoruz. Bu veriler, OECD ülkelerinin çoğunu ve birçok gelişmekte olan ülkeyi kapsamaktadır. Karşılaştırmalı analiz ve karşı-gerçeksel metodolojisinden faydalanarak, Türkiye'nin eğitim performansındaki değişimin benzer ekonomik gelişmişlik seviyesindeki ülkelerden anlamlı şekilde farklı olup olmadığını belirlemek için çifte farklar yöntemini kullanıyoruz.

Analizlerimiz, Türkiye'nin öğrenci başarı testlerindeki olağanüstü performans artışını kısmen doğrulamaktadır; özellikle 2000'li yılların başından itibaren önemli bir yükseliş görülmektedir. Çoğu benzer ülke eğitim performansında iyileşme kaydetmiş olsa da, Türkiye'nin performans artışı belirgin şekilde daha yüksektir. Bu bulgu, alternatif performans eşikleriyle de doğrulanmaktadır: 2020 yılında Türkiye'de ortaöğretimde minimum başarı seviyesine ulaşan öğrencilerin oranı %88 iken, karşılaştırma grubundaki ülkelerde bu oran yalnızca %73'tür. 2000 ve 2020 yılları arasındaki çifte farklar analizimiz de bu pozitif eğilimi doğrulamaktadır; Türkiye için fark +%3,6 olarak hesaplanmıştır.

Bununla birlikte, Türkiye eğitim sisteminin nispeten yüksek etkinliğine rağmen, eşitlik konusunda önemli zorluklarla karşı karşıya olduğu görülmektedir. Eğitimde eşitlik göstergemiz, en yüksek başarı gösteren %10'luk dilim ile en düşük başarı gösteren %10'luk dilim arasındaki farkı ölçmektedir. Bu fark ne kadar büyükse, eğitim sistemi öğrencilere eşit fırsatlar sunma konusunda o kadar başarısızdır. Sonuçlarımız, Türkiye'de bu eşitlik göstergesinin son 50 yılda düştüğünü göstermektedir (1970-2020): 589 puandan 540 puana gerileyerek her on yılda yaklaşık 10 puanlık bir düşüş yaşamıştır. Buna karşılık, okullaşma oranı ile düzeltilmiş eşitlik göstergesi, özellikle ortaöğretimde anlamlı şekilde artmıştır. Bu eğilimler, Türkiye'nin kitlesel eğitimde kaydettiği ilerlemeyi yansıtmaktadır: 2000 yılında %60 olan ortaöğretim tamamlama oranı, 2023 itibarıyla %90'a yükselmiştir. Bu gelişmeler büyük ölçüde, zorunlu eğitimi 12 yıla çıkaran 2012 reformu (*4+4+4* reformu) ile mümkün olmuştur.

Ancak, karşı-gerçeksel analizimiz, Türkiye ile karşılaştırılabilir ülkelerin eğitimde eşitlik açısından daha iyi performans gösterdiğini ortaya koymaktadır. Bu durum, insan sermayesi kalitesinin Türkiye genelinde nasıl dağıldığı konusunda önemli sorular doğurmaktadır. Ayrıca, Türkiye'deki eğitim eşitsizliklerinin yalnızca bölgeler arasında (Uysal & Gelbal, 2018) değil, aynı zamanda okullar arasında da önemli ölçüde farklılık gösterdiği görülmektedir (Polat ve diğerleri, 2024).

Son olarak, eğitim kalitesinin iyileştirilmesine odaklanan bir eğitim reformunun potansiyel uzun vadeli kazanımlarını değerlendirmek için ekonomik bir öngörü çalışması yürütüyoruz. Bir tahmin modeli kullanarak (Gust ve diğerleri, 2024), kalite iyileştirme ve eğitimde eşitliğin birlikte artırılmasının Türkiye için 2100 yılına kadar önemli ekonomik kazançlar sağlayacağını gösteriyoruz. Karışık bir eğitim politikası senaryosunda, 2100 yılı itibarıyla GSYH'nin %111 oranında artacağı öngörülmektedir. Buna karşılık, yalnızca kalite artışına odaklanan bir senaryoda bu artış %68,1 ile sınırlı kalmaktadır. Bu bulgular, Türkiye'nin uluslararası sınavlardaki başarılarını artırmanın ötesinde, *Vizyon 2023* programında (MEB, 2018) vurgulandığı gibi, eğitim sisteminde bölgesel ve okul bazında daha fazla eşitliği sağlama hedefini de önceliklendirmesi gerektiğini ortaya koymaktadır.

1. Introduction

In 2023, Türkiye achieved a remarkable performance in the international TIMSS (Trends in International Mathematics and Science Study) assessment, ranking 8th out of 58 participants with an average score of 553, well above the international average (Koca et al., 2024). This progression is part of a trend of improving results for Turkish students in mathematics and science, which has been evident throughout successive cycles of TIMSS, particularly since the 2000s. Compared with previous assessments, Türkiye has consolidated its position among OECD countries, showing notable progress at both the 4th and 8th grade levels. However, while these results illustrate a significant improvement in Türkiye's education system, questions remain as to the representativeness of the 2023 sample and the structural challenges still facing the country.

According to Polat et al. (2024), the TIMSS 2023 survey in Türkiye was not carried out in some provinces due to the earthquake of February 6, 2023. Thus, students from nine provinces, including Adıyaman, Gaziantep, Hatay, and Malatya, were not included in the final sample. This exclusion raises questions about the representativeness of the results obtained, given that these regions, heavily affected by the disaster, generally experience difficulties in terms of access to quality education. Furthermore, previous studies, such as those by Spaull (2018), have already shown that sample composition can influence international rankings. In this respect, the performance observed in 2023 may not reflect the reality of the Turkish education system as a whole.

At the same time, Türkiye's results in the latest OECD PISA assessment survey do not confirm its advantageous position found in TIMSS 2023. Its mathematics score is below the OECD average (453 points compared to an average of 472). Results in reading and science confirm this central position. The question is whether these contradictions will persist over the long term and to what extent Türkiye's performance has actually improved since the advent of international surveys of student achievement.

Improving the quality of education is therefore a central issue if Türkiye is to avoid falling into the "middle-income trap." As various researchers have pointed out (Suna & Özer, 2021), the level of qualification of the working population is a decisive factor for sustained, inclusive economic growth. Despite progress in access to education, imbalances persist, particularly in terms of the quality of teaching and performance gaps between schools and regions.

According to Sarıer (2020), performance in mathematics and science is strongly influenced by students' socio-economic status, reflecting persistent inequalities in access to quality education. To avoid this pitfall and foster sustainable development, it is essential to reduce these disparities and improve the quality of education overall. Indeed, studies have shown that countries that have successfully overcome this trap have implemented education policies that promote research, innovation, and the development of advanced skills (Yediyıldız & Ustun, 2024).

Overall, the trend in Türkiye's school enrolment rate is positive, with increasing participation in secondary and tertiary education. According to the OECD, between 2003 and 2018, Türkiye doubled the number of students eligible for the PISA test, and the country significantly reduced the number of young people not in education or training (OECD, 2021). The 4+4+4 reform has played a crucial role in this progress, extending the duration of compulsory education and promoting greater access to education for girls. In addition, the "2023 Vision" project has also helped to improve the quality of the Turkish education system by focusing on pedagogical innovation, teacher training, and the reduction of educational inequalities (MEB, 2018).

Türkiye's school enrolment rate has undergone significant change over the decades, but the country remains marked by significant regional disparities. Despite this expansion in access to education, significant gaps remain between provinces. According to Çelik & Gür (2013), in 2012, provinces such as Bilecik, Rize, Artvin, and Bolu had secondary school enrolment rates above 90%, while regions such as Bitlis, Şanlıurfa, Şırnak, and Siirt remained around 40%. This situation highlights the persistent challenges of educational equity and the need for targeted policies to improve access to education in disadvantaged areas. In addition, the study by Uysal & Gelbal (2018) indicates that regional economic development directly influences access to education, with wealthier provinces benefiting from better school infrastructure and higher-quality educational supervision.

Furthermore, Türkiye's performance in various international assessments, such as PISA, PIRLS, TIMSS, and ABIDE, reveals discrepancies depending on the studies and disciplines assessed. According to T. Aydın & Çilek (2024), Türkiye's performance in mathematics increased from 2012 to 2022 (+15.0). However, according to Sarıer (2020), in 2018, Türkiye reached its highest average scores in mathematics after a gradual increase since 2003, although a decline was observed in 2015. On the other hand, according to Suna & Özer (2021), more than 60% of the variation in mathematics scores in Türkiye is attributable to

differences in achievement between schools—more than double the rate observed in OECD countries. These discrepancies show that, while some progress has been made, significant disparities persist, particularly between schools. It also appears that students' performance is strongly influenced by their access to educational resources, including the availability of qualified teachers, teaching materials, and an environment conducive to learning (Polat et al., 2024).

Finally, educational inequalities in Türkiye are particularly visible in the PIRLS, TIMSS, and PISA surveys. According to Yediyıldız & Ustun (2024), 64% of 15-year-old students in Türkiye are classified as socio-economically disadvantaged, placing the country among the most unequal in the OECD on this indicator. Furthermore, according to , socio-economic status directly influences performance in math, science, and reading. Other studies, such as those by Suna & Özer (2021), point out that socio-economic characteristics explain a significant proportion of the variance in school results. These findings show that improving results on international assessments cannot be separated from a broader reflection on equity and the redistribution of educational resources in Türkiye. According to Köseleci (2015), the performance gap between the most and least developed regions is exacerbated by the concentration of elite schools in large cities, accentuating inequalities between students.

All in all, Türkiye's results in international surveys show significant progress, particularly in TIMSS 2023, but also reveal potential problems with comparability of results over time and between countries, as well as major challenges in terms of equity and equal opportunities.

We thus propose to use a methodology to aggregate Türkiye's results in international assessments, from a dual perspective of measuring both Turkish educational performance and aspects of educational inequality. We use several indicators and various methodologies measuring both average quality and inequalities to analyze the state of knowledge and enhance the robustness of the results obtained. In doing so, we also seek to address the problems of measurement error inherent in international surveys.

While longitudinal analyses capture fixed effects and trace school career paths, cross-sectional approaches allow for a more comparative reading and, in turn, include many countries. With this in mind, we propose a hybrid, counterfactual approach. Using cross-sectional data spread over time, we analyze Türkiye's educational performance using a database of learning outcomes for nearly 130 countries between 1970 and 2020. In so doing, we combine the advantages of cross-sectional data (comparability with other education systems) with those of longitudinal analyses (the possibility of tracing the evolution of

inequalities over time). We also use the double-difference method to correct exogenous effects similar to all developing countries and detect structural changes in the Turkish education system.[3]

In the remainder of this article, we first present the original data produced on the quality of education. The methodology used to construct the data is explained, followed by an analysis of Türkiye's performance. The latter focuses on a number of dimensions to confirm or refute the trends observed. Finally, we use an economic projection model to estimate the economic impact of a better education policy in Türkiye, using a historical macro-simulation model based on a counterfactual approach.

2. Data and methodology

The originality of our work lies in the use of a database containing comparable scores in several areas of school competence, at different levels and over a period covering around five decades (1970-2020). [4] This database presents not only school performance scores, but also measures of performance in certain benchmarks and scores for sub-samples, enabling us to gain a better understanding of the equity of education systems, with a view to international comparison.

2.1. Assessments of student learning

Our assessment of the quality of education systems is based on the dimension of learning achievement, measured through an exercise that standardizes international surveys. The international database on student learning compiles results from various surveys (TIMSS, PIRLS and PISA)[5]. Other assessments focus on adult skills (IALS, PIAAC[6]). We present these surveys very briefly. More information can be found in the respective reports for each assessment (Mullis et al., 2020, 2023; OECD, 2023b; Paccagnella, 2016).

The International Association for the Evaluation of Educational Achievement (IEA) was the first organization to conduct international surveys of student achievement. After a pilot survey in the 1960s (Foshay et al., 1962), several assessments tested primary and secondary

[3]In a way, our approach can be seen as a natural, quasi-experimental experiment, with Türkiye as the control group and other developing countries as the control group.

[4]For a detailed description, see (Altinok & Diebolt, 2024).

[5]*Trends on International Mathematics and Science Study*, *Progress on International Reading Literacy Study* and *Programme for International Student Assessment*, respectively.

[6]Respectively *International Adult Literacy Survey* and *Programme for the International Assessment of Adult Competencies*.

school students in mathematics and science, but it was above all the TIMSS assessment that made its mark on the international community. This survey began in 1995 and focuses on mathematics and science. Since then, several waves of tests have been carried out every four years.[7] The IEA has also assessed students' reading skills since the 1970s, but it is undeniably the PIRLS survey that has highlighted young people's reading skills. In parallel with the IEA surveys, the PISA assessment has attracted a great deal of media attention, particularly in Germany (Waldow, 2009). PISA is an assessment developed by the OECD. Launched in 2000, it tests 15-year-old students in three skill areas (mathematics, science and reading). The PISA assessment has become the survey covering the most countries in the world. Organized every three years, eight different waves have been carried out up to 2022, with over 70 countries participating in the latest wave. Türkiye began participating in PISA in the second wave, in 2003.

Alongside assessments of student achievement, surveys of adult skills (IALS and PIAAC) can also be very useful for better assessing student skills. Using a specific methodology, some authors have, in fact, obtained panel data relating to the skill levels of young adults (Schwerdt & Wiederhold, 2019). All the surveys used to obtain the database can be found in Table 1.

Table 1: International surveys of student or adult learning achievement

No	*Years*	*Organization*	*Abbrev.*	*Material*	*Number of countries or regions*	*Grade*
1	1959-60	IEA	Pilot input	M,S,L	12	7/8
2	1964	IEA	FIMS	M	12	7/FS
3	1970-71	IEA	SSS-RC	L	15	4/8/FS.
4	1970-72	IEA	FISS	S	19	4/8/FS.
5	1980-82	IEA	SIMS	M	19	8/FS
6	1983-84	IEA	SISS	S	23	4/8/ FS
8	1990-91	IEA	RLS	L	32	3-4/7-8
9	1995-2019	IEA	TIMSS	M,S	45-38-26-48-66-65-64	3-4/7-8/ FS
11	1997-2019	UNESCO	LLECE	M,S,L	13-16-15-16	3-6
12	1999-2007	UNESCO	SACMEQ	M,L	7-15-16	6
13	2000-2019	CONFEM	PASEC	M,L	22-22-10-14	2/5 then 3-6
14	2001-2021	IEA	PIRLS	L	35-41-55-50-65	4
15	2000-2022	OECD	PISA	M,S,L	43-41-57-74-65-71-79-81	15 years
16	2010-2019	USAID/R	EGRA	L	29-40-5	1 to 6
17	2008-2019	ASER	ASER	L	2	1 to 6
19	2011-2017	OECD	PIAAC	L	39	Adults

Note: IEA = "International Association for the Evaluation of Educational Achievement", NCES = "National Center for Education Statistics", CONFEMEN = "Conférence des Ministres de l'Education des Etats et Gouvernements de la Francophonie", OECD = "Organisation for Economic Co-operation and Development", USAID = "United States Agency for International Development", RTI = "Research Triangle Institute", ASER = "Australian Council for Educational Research", FIMSS = "First International Mathematics Study", SSS-RC = "Six Subject Survey: Reading Comprehension", FISS = "First International Science Study", SIMS = "Second International Mathematics Study", IAEP = "International Assessment of Educational Progress", RLS =

[7]The second wave dates back to 1999, then 2003, 2007... The latest wave with available data dates from 2019. Although TIMSS 2023 data have recently been released, they could not be included in our analyses.

"Reading Literacy Study", TIMSS = "Trends on International Mathematics and Science Study", LLECE = "Latin American Laboratory for Evaluation of the Quality of Education", SACMEQ = "The Southern and Eastern Africa Consortium for Monitoring Educational Quality", PASEC = "Programme d'Analyse des Systèmes Educatifs de la Confemen", PIRLS = "Progress in International Reading Literacy Study", PISA = "Programme of International Student Assessment", EGRA = "Early-Grade Reading Assessment", PIAAC = "Programme for International Assessment of Adult Competencies". Subjects: M= mathematics; S= science; L= reading. FS = "end of upper secondary".

2.2 Methodology

Our analysis includes several recently published original datasets measuring multiple dimensions derived from student achievement surveys (Altinok & Diebolt, 2024, 2025). In addition to traditional average scores from surveys such as PISA, we also use alternative measures such as the proportion of students reaching specific thresholds, differences between deciles, or even gaps between socioeconomic levels. By mixing these dimensions with traditional measures used in education, it is possible to obtain indices of efficiency or equity.

2.2.1. Education quality database

Before preparing data on specific sub-samples or thresholds, it is essential to have an international database on the quality of education systems. The detailed methodology of this approach can be found in Altinok & Diebolt (2024). Below, we briefly present the general approach to obtaining average scores before continuing with the dimension of educational inequalities.

Altinok & Diebolt (2024)'s approach is based on grouping the various international achievement tests (presented in the previous section) onto comparable scales. Since these surveys are diverse in nature, the authors developed a methodology for adjusting them to each other in order to avoid any potential bias. The surveys were adjusted to enable comparisons over time and space. The general idea behind this approach is to use the results of countries that have participated simultaneously in several surveys over the same period (Angrist et al., 2021).

In fact, by combining all the countries that have participated in several assessments, it is possible to calculate conversion rates, as one would for an exchange rate between currencies, for example. This method is used to adjust the scores of surveys carried out from the 1990s onwards and has already been implemented by previous research (Angrist et al., 2013, 2021; Gust et al., 2024). An additional approach was also implemented to anchor the surveys to each other over time, particularly before 1990. Following the pioneering approach of Hanushek & Kimko (2000), it is thus possible to anchor international survey data by referring

to the U.S. results in the national survey of student achievement, namely the National Assessment of Educational Progress (NAEP).[8]

As for PISA and TIMSS, the calculation of a general conversion rate is sufficient to obtain a comparable database in terms of learning outcomes.[9] What is new, however, is the implementation of the multiple imputation method used in Lim et al. (2018). By using data that are highly correlated with educational attainment, Altinok & Diebolt (2024) explain that this approach makes it possible to extend both the period covered by attainment assessments and to predict scores for countries that have not participated in all these assessments.

This imputation procedure uses the "Amelia II" package (Honaker et al., 2011) and is useful for obtaining panel data for almost half a century (1970-2020).[10] Indeed, the database prepared by Angrist et al. (2021) only covers the period 2010-2020. The combination of multiple imputation and the earliest surveys, anchored through conversion rates and NAEP evaluation, allows Altinok & Diebolt (2024) to generate comparable scores for key OECD and developing countries from 1970 to 2020. This database will serve as the foundation for our analysis.

Furthermore, in order to combine the quantity and quality of education, Altinok & Diebolt (2024) propose associating the average number of years of schooling with students' scores on achievement assessments. Filmer et al. (2020) were the first to perform such a conversion. This combination thus provides a relative measure of the quantity of education, taking into account the educational quality differential between countries. The main hypothesis is based on the idea that one year of education in a given country does not necessarily yield the same academic performance in terms of learning outcomes as in another country. Some countries prove to be more efficient than others, as highlighted by assessments such as PISA. With this in mind, the hybrid indicator we are implementing is the "Learning-Adjusted Years of Schooling" (Filmer et al., 2020). To calculate this relative indicator, the score obtained is

[8]Details of the methodology used to obtain a database of over 130 countries are presented in Altinok & Diebolt (2024). The present article focuses on the results for Türkiye, comparing this country with a sample of 20 developing countries with an economic level close to Türkiye's in 1970.

[9]Assuming that the average mathematics scores of countries participating in TIMSS 1999 and PISA 2000 are equal to 553 in TIMSS and 533 in PISA, then the conversion rate between PISA and TIMSS will be $553/533 = 1{,}0375$. PISA scores will thus be revised upwards by around 3.75%.

[10]Details of the multiple imputation procedure can be found in Altinok & Diebolt (2024). In contrast to Lim et al. (2018), we employ this multiple imputation method using different steps to minimize measurement error. Specifically, eight steps are implemented, ranging from predicting scores for countries with the most observations over a short period (between 1995 and 2020, step number 1) to countries with very little data on education quality and over a longer period (between 1970 and 2020, step number 8). This imputation procedure uses explanatory factors previously selected via multiple regressions and includes several explanatory variables.

divided by 700 (considered a theoretical ceiling), giving a value between 0 and 1. This ratio is then multiplied by the average number of years of schooling, taken from the database of Barro & Lee (2013). The new database thus obtained provides information about three complementary variables: the quality of education systems (via scores on various achievement surveys and performance thresholds), the quantity of education (via the average number of years of schooling), and the Learning-Adjusted Years of Schooling (LAYS).

2.2.2. Extending the base to school inequalities

The use of anchoring methods provides a panel database on the quality of education. In practical terms, data on the quality of education are spread over several school levels (primary, secondary) and skill areas (mathematics, reading, science) and are available between 1970 and 2020 for most OECD countries (including Türkiye) , as well as a large number of developing countries.[11]

Analyzing the performance of an education system solely in terms of average scores can overlook the issue of equity. It is therefore appropriate to propose measures to assess the inequitable dimension of educational achievement. As school enrollment is not universal in all countries, we adjust inequality indicators based on school completion rates. This issue is more prevalent in developing countries like Türkiye.

Three main dimensions are used to analyze performance differences within countries. Firstly, we reproduce the data anchoring methodology almost identically, focusing on performance thresholds and scores across the different deciles. The approach used is very similar to that for average scores, with the difference that we prefer to use anchoring by the so-called "equipercentile" method instead of the "conversion rate" method in order to calibrate each performance threshold across countries and over time (Kolen & Brennan, 2014). The main advantage of this method is that it takes into account the distribution of scores within each database. In concrete terms, instead of anchoring average scores with conversion rates, we anchor each percentile of a reference survey with the same percentile of another survey when calculating conversion coefficients. Around a hundred of these coefficients are calculated for a pair of surveys (each referring to a particular percentile). This allows us to take into account any differences in score distribution within each survey. Indeed, as we wish to calculate the proportions of students exceeding performance thresholds, it is important to

[11]It should be noted that we use the term "panel" in the sense that the cross-sectional data are grouped together without longitudinal follow-up. However, as the data are statistically representative of the different countries included in the surveys, we can consider them as a panel for monitoring the performance of education systems.

take into account the distribution of scores within each assessment before anchoring the surveys to each other.

We calculate the proportions of students reaching different competency thresholds. More specifically, we are interested in two levels. The performance thresholds are based on those developed by the IEA, and these thresholds are also quite similar to those defined by the OECD in the PISA test. The first level refers to a "Minimum Proficiency Level" (MPL), whose main aim is to highlight the proportion of a population that achieves a certain level of competence, in order to obtain minimum skills for everyday activities. The Sustainable Development Goals explicitly refer to this minimum threshold (UNESCO Institute of Statistics, 2019). The minimum proficienct threshold refers to the IEA's minimum benchmark, the "Low International Benchmark", which corresponds to a threshold of 400 points. Students who reach this threshold have basic skills in mathematics, reading and science.[12] At primary level, they can perform basic arithmetic operations on one- and two-digit numbers as well as solve problems involving a small number of parameters. Students can handle fractions and common geometric shapes. Statistically, they can read and complete bar graphs and tables.

Alongside this minimum threshold, it is interesting to look at the proportion of students reaching the Advanced Proficiency Level (APL). The main purpose of this threshold is to measure the proportion of the population achieving a high level on tests of student achievement, and thus identify the proportion of elites within each country. We use the IEA's reference threshold, the Advanced International Benchmark, which is set at a level of 625 points. [13] At this level, primary school pupils can apply their knowledge and understanding in a variety of complex situations, and explain their reasoning rigorously. It is no longer a question of having basic knowledge, but rather of being able to use it in complex situations while knowing exactly which skills to mobilize, and mobilizing more sophisticated tools (such as the use of fractions or decimals). The use of multidimensional tools is essential to reach this advanced level. The set of skills acquired for each skill area and threshold is presented in more detail in OECD (2023b) and Mullis et al. (2020). In particular, we analyze the proportion of students achieving the Minimum Proficiency Level (MPL) against the

[12]This threshold is very close to the level 2 defined in the PISA assessment (OECD, 2023b).

[13]It is important to note that this threshold is very close in secondary education to level 5 of the PISA study. The thresholds are close to 625 points for the 3 skill areas, bearing in mind that we use approximately the same threshold in reading. More precisely, the threshold is 607 points in mathematics, 633 points in science and 626 points in reading. See OECD (2023b) , pages 92, 9-100 and 103 for the definition of thresholds according to the PISA assessment.

proportion achieving the Advanced Proficiency Level (APL). A country that manages to get almost its entire population to reach the minimum threshold can be considered egalitarian. If the same country also manages to get a proportion of students to exceed the advanced skills threshold, it will be both egalitarian and successful.

In addition to average scores, we also calculate scores for each decile. The gap between deciles may prove to be a means of measuring educational inequality, following the approach developed by Piketty (2013) in the context of income inequality. In doing so, we calculate scores at deciles 1 and 9, i.e. D1 & D9, respectively, the average score of the lowest and highest performing 10%. An egalitarian country should logically obtain an interdecile ratio close to 1, meaning a very small gap in performance between the two extreme deciles. On the other hand, an unequal country that manages to get the best performers to do better, while neglecting the worst performers, will have a high interdecile ratio. Measuring the interdecile ratio will therefore enable us to assess the level of inequality in Türkiye in comparison with other countries.

An additional way of measuring educational inequalities concerns the potential differences in performance that may exist between populations with specific characteristics, such as whether or not they come from a family with high socio-economic capital. This brings us back to the classic analysis of inequalities in the sociology of education (Coleman, 1966). It is important to note that the initial anchoring is done on the scores of each sub-population, but the use of ratios is only carried out during the multiple imputation procedure (Altinok & Diebolt, 2025). The approach adopted here is to replicate the initial anchoring methods on sub-populations, then use ratios marking the absolute differences between sub-populations to perform data imputation.

Here, we use only the variability of scores relating to differences in socio-economic level. For this purpose, Altinok & Diebolt (2025) calculated socio-economic level scales within all surveys, using available data. It is impossible to have a standardized and similar scale for all surveys. However, as the authors use the interquartile range for each index (i.e. splitting the distribution of scores into four equal parts), comparison between different periods should not suffer from the lack of standardization of scales for this index between surveys. If students with the highest socio-economic levels have significantly higher scores than those from families without social and economic capital, the country can be considered unequal in terms of educational achievement. We use an index of socio-economic level, obtained from virtually all the assessments included in our database. As the measurement of this index is quite

heterogeneous between achievement tests, it is primarily the absolute gap between quartiles that will be weighed in our analysis. The bottom quartile (Q1) will measure the average score of students from the 25% poorest families in terms of socio-economic status. We then measure the level of inequality by calculating the difference in performance between the top quartile (Q4) and the bottom quartile (Q1). The greater the difference, the greater the inequality.[14]

2.2.3. Synthetic performance indices

The indicators presented above each measure a particular dimension and may relate to measures of effectiveness, efficiency, or equity. First, we propose to construct an indicator of educational effectiveness. This indicator is based on the assumption that a system can be considered efficient if it succeeds in getting all its students to reach the minimum performance threshold (MPL), without sacrificing those who could also reach the excellence threshold (APL). However, it is virtually impossible to have the entire population reach the threshold of excellence. This is why we weight the value relative to this threshold to obtain our efficiency index:

$$efficacy_{country,t} = \frac{(MPL+2\times APL)}{2} \times 10 \quad (1)$$

with MPL = Minimum Proficiency Level & APL = Advanced Proficiency Level

So let's assume a country A where 80% of the population reaches the minimum threshold and only 5% manage to reach the excellence threshold. Conversely, in country B, almost the entire population reaches the minimum threshold (95%), with almost a fifth reaching the excellence threshold (20%). The efficiency index is equal to 450 in country A compared to 675 for country B, which represents a difference of 50% in favor of the latter country.

Beyond efficiency, a system can also be measured by its degree of equity, i.e. the extent to which it enables everyone to succeed, regardless of their initial resources. The notion of equity is fairly subjective but often refers to the principle of equal opportunities to succeed, irrespective of individual characteristics. Equity does not imply perfect equality between all students, but rather equal access and success justified by factors independent of socio-economic characteristics or based on various dimensions (such as gender, immigrant status, or religion). It is impossible to combine all possible dimensions. The most commonly used one, given the ease of obtaining comparable data, refers to the socio-economic status of parents.

[14]It remains possible that if the structure of the population changes over time, such as the arrival of large numbers of poor immigrants in a country, the interquartile range may be modified exogenously.

However, the measure of socio-economic level is quite heterogeneous between the different assessments and may be interpreted differently depending on the economic level of the country. A more neutral measure refers to the analysis of deciles. More specifically, the gap between the bottom and top deciles can measure a degree of inequality, somewhat similar to the Gini index. We therefore propose to measure the equity of education systems by combining the extreme deciles in terms of student scores. The bottom 10% decile is calculated as the average score of students in the bottom 10%. In contrast, the top 10% score refers to the average score obtained by the top 10% of students. The greater the deviation from the median score, the more inequitable an education system is considered to be. More generally, the formula we use is as follows:

$$equity_{absolute,t} = \left(1 - \frac{Top_{10\%} - Bottom_{10\%}}{Median\_score}\right) \times 1\,000 \quad (2)$$

where $Bottom_{10\%}$ is the average score obtained by the bottom 10% of students, $Top_{10\%}$ is the average score obtained by the top 10% and $Median\_score$ is the median score.

This equity index thus represents the standardized score of the gap between the highest and lowest performers, relative to the median score. As Spaull (2018) points out, the incomplete schooling of a young population can call into question the interpretation of results within surveys such as PISA. If an education system is selective and does not include the entire school-age population, it is possible for the equity index to be biased. Let's take the example of a country where only half the population aged 15 actually attends school. In this case, the calculated index will tend to overestimate the level of equity in the country under the assumption that the population excluded from the education system would have performed less well if they had attended school. To take account of this exclusion bias, we calculate an adjusted equity index in which we include the school completion rate:

$$equity_{adjusted,t} = \left(1 - \frac{Top_{10\%} - Bottom_{10\%}}{Median\_score}\right) \times completion \times 1\,000 \quad (3)$$

where $completion$ represents the completion rate for the school level in question.

Descriptive statistics are presented in Table 2. Türkiye's position falls between the two groups of countries in terms of quality of education (respectively OECD and developing countries groups). However, the indicator relating to the number of years of schooling highlights the clear gap between developed countries and Türkiye. On average, OECD countries record around 10 years of schooling over the period 1970-2020, compared with 5.6 years for Türkiye. Even the group of developing countries has a higher level of schooling than

Türkiye (6.9 years). This difference is confirmed by the hybrid indicator of years of schooling adjusted by the quality of education, which is around half as high in Türkiye as in OECD countries.

Although the equity index is relatively close between Türkiye and the group of OECD countries, once the completion rate is taken into account, adjusted equity is significantly lower in Türkiye. This disparity can be explained by the low enrollment of young people between 1970 and 2010. Finally, the efficiency index falls between the average for developing countries and that for OECD countries. Türkiye thus occupies a special position here, appearing to hold an average position among developing countries but lagging behind when OECD countries are used as a benchmark. It therefore seems important to compare Türkiye with similar countries, as the process of universal schooling was only implemented later in Türkiye, unlike in developed countries.

Table 2: Education database (1970-2020)

	Average	Standard deviation	Minimum	Maximum
Türkiye				
Quality of education	462.2	14.2	445.4	501.3
Years of schooling	5.6	1.9	2.6	8.1
Mixed indicator (LAYS)	3.7	1.3	1.7	5.8
Equity indicator	535.6	24.4	499.4	589.5
Adjusted equity indicator	394.2	102.0	284.2	537.5
Efficiency indicator	454.9	27.9	424.1	526.4
OECD-20				
Quality of education	499.2	37.6	379.7	599.2
Years of schooling	9.9	2.2	2.6	13.8
Mixed indicator (LAYS)	7.2	1.9	1.7	10.8
Equity indicator	537.4	43.9	340.4	633.6
Adjusted equity indicator	478.0	87.7	150.1	630.3
Efficiency indicator	534.7	84.7	288.9	754.9
Group of developing countries D20				
Quality of education	409.8	57.4	268.6	545.7
Years of schooling	6.9	2.7	1.6	13.0
Mixed indicator (LAYS)	4.1	2.0	0.6	9.8
Equity indicator	473.5	99.5	24.2	639.8
Adjusted equity indicator	338.3	142.3	21.1	597.7
Efficiency indicator	341.3	112.0	140.5	616.3

Notes: The data include all observations available between 1970 and 2020 for a selection of 20 OECD countries, 20 developing countries and Türkiye. A total of 220 observations are available for the full sample of developing countries and 11 observations for Türkiye (data available every 5 years). Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

2.3 Analysis

2.3.1. Double-difference analysis

In order to analyze the evolution of Türkiye's performance rigorously, we employ a counterfactual approach known as the double-difference method. It is possible that the recent increase in Türkiye's performance is mainly due to a catch-up effect with countries that were at the top in the 1970s. In this case, this evolution cannot be considered exceptional or as the result of educational policies specific to Türkiye. To test this hypothesis, we calculate the gap between Türkiye's performance and that of countries with roughly the same characteristics in the 1970s, i.e. a panel of 20 developing countries with roughly the same GDP per capita as Türkiye.[15]

We then calculate the performance gap between Türkiye and the selected countries in subsequent years, using the previously defined performance indicators. If the gap is significant, then we can deduce that the trajectory followed by Türkiye results from the country's own actions and is not the consequence of a simple catch-up phenomenon and/or of factors external to Türkiye.

Let's assume the following performance indicator (denoted *educ*) for Türkiye and a panel of countries similar to Türkiye, which we denote DEV_{20}:

$$edu_{TUR,1970} = X \; et \; edu_{DEV20,1970} = Y \tag{4}$$

We can calculate the variation in Türkiye's performance between two periods:

$$\Delta edu_{TUR,t\rightarrow t+10} = edu_{TUR,t+10} - edu_{TUR,t} \tag{5}$$

Similarly, it is possible to calculate the variation in performance for the control group (here the 20 developing countries in our panel, noted):

$$\Delta edu_{DEV20,t\rightarrow t+10} = edu_{DEV20,t+10} - edu_{DEV20,t} \tag{6}$$

A significant effect of Türkiye's performance can be observed by calculating the difference between the variation in Türkiye's performance and that of the selected panel of countries (using the double-difference method):

$$\Delta edu_{TUR-net,t\rightarrow t+10} = \left(edu_{TUR,t+10} - edu_{TUR,t} \right) - \left(edu_{DEV20,t+10} - edu_{DEV20,t} \right) \tag{7}$$

[15]We looked for countries with a GDP per capita close enough to Türkiye's in 1970 to form the control group sample. The countries in this group are as follows: Algeria, Brazil, Chile, Colombia, Costa Rica, Czech Republic, Ecuador, El Salvador, Guatemala, Iran, Malaysia, Mauritius, Mexico, Nicaragua, Peru, South Africa, Vietnam, Slovenia, Türkiye and Uganda.

The main drawback of this double-difference approach is the potential presence of measurement errors that could bias the net effect of the variation in Türkiye's performance. For example, if there were a measurement error that overestimated the indicator in a given year, then the variation calculated between two periods would mainly reflect this measurement error.

To partially mitigate this estimation bias, we propose two alternatives. First, we calculate the absolute deviation of Türkiye's performance relative to the panel considered for each period:

$$edu_{TUR-relative,t} = edu_{TUR-abs,t} - edu_{DEV20-abs,t} \quad (8)$$

Then, by calculating the average of these indicators, it is possible to determine whether the school performance observed in Türkiye significantly differs from that of other comparable countries:

$$edu_{mean,TUR-relative,t} = \frac{\sum_{t=1}^{n} edu_{TUR-relative,t}}{n} \quad (9)$$

Another method is to extend the time intervals used to calculate the variation, to reduce potential estimation biases. Instead of using a 10-year interval, we could use intervals of 10, 15 or 20 years:

$$\Delta edu_{TUR-dif,t\rightarrow t+15} = \left(edu_{TUR,t+15} - edu_{TUR,t} \right) - (edu_{DEV20,t+15} - edu_{DEV20,t}) \quad (10)$$

2.3.2. Model for projecting an improvement in the quality of education

Furthermore, in order to assess the potential economic benefits of improving the quality of education in Türkiye, we use a forecasting model developed in several studies by economists such as Eric Hanushek and his co-authors (Gust et al., 2024; Hanushek et al., 2017b, 2017a; Hanushek & Woessmann, 2012).

In these forecasting models (the key features of which are presented in Appendix B), a macro-simulation of the impact of an educational reform is conducted using the results of growth models estimated in a previous study (Hanushek & Woessmann, 2012). We apply this model and attempt to measure the trajectory that Türkiye could experience if it implemented an effective policy to improve the quality of education.

Our first scenario estimates the economic benefits of a policy aimed at improving the test scores of the entire population by a quarter of a standard deviation, or around 25 points on the

PISA-comparable scale. By simulating this increase, we can estimate the potential economic growth Türkiye could experience by 2100.

In these models, the skills of each cohort differ from those of the current workforce. We simulate a situation where an educational policy is introduced, gradually affecting the entire population over the period 2020-2100. Initially, the workforce maintains its pre-reform skill level. This workforce is then partially, and eventually fully, replaced by individuals who have experienced the reform. This transition takes several years, considering that the duration of working life is 40 years (W = 40). Thus, each year, 2.5% of Türkiye's total workforce comprises students who have experienced the educational reform.

We calculate the skill level of the workforce for each year between 2020 and 2100 by replacing the oldest workers with those from the new cohorts (i.e. $\bar{A}_t$) weighted as $1/W$, where W is the working life. In calculating the gain in cognitive skills from such a reform, we consider 4 different phases:

1. School reform ($t = 1, \dots, R$): during the *R-term* reform, workers with an initial skill level are gradually replaced by more skilled workers.

2. Main replacement ($t = R + 1, \dots, W$): workers with the initial skill level (i.e. pre-reform) are replaced by new workers for the next ($W - R$) years.

3. Quality reinforcement ($t = W + 1, \dots, W + R$): for the following R years, some workers who have only partially experienced the reform are replaced by workers who have fully experienced the reform.

4. Population with complete reform ($t = W + R + 1$, ...): The entire workforce concerned has a skill level at the level of the desired reform.

In these models, a reform scenario assumes a linear progression path taking 15 years before the reform is fully operational (parameter *R)*. This means that the educational levels of each of the first cohorts following the educational reform will have different (and better) levels of educational completion. For each year of the simulation, we calculate the GDP growth rate resulting from the educational reform as follows:

$$g_t^{reform} = p + \gamma \bar{A}_t \quad (11)$$

GDP with or without the reform changes as follows:

$$GDP_t^{\Delta} = \left(1 + g_t^{\Delta}\right) \times GDP_t^{\Delta} \quad (12)$$

where $\Delta \in (reform, without\ reform)$.

The total value of the reform is given by the sum of the discounted annual GDP differences:

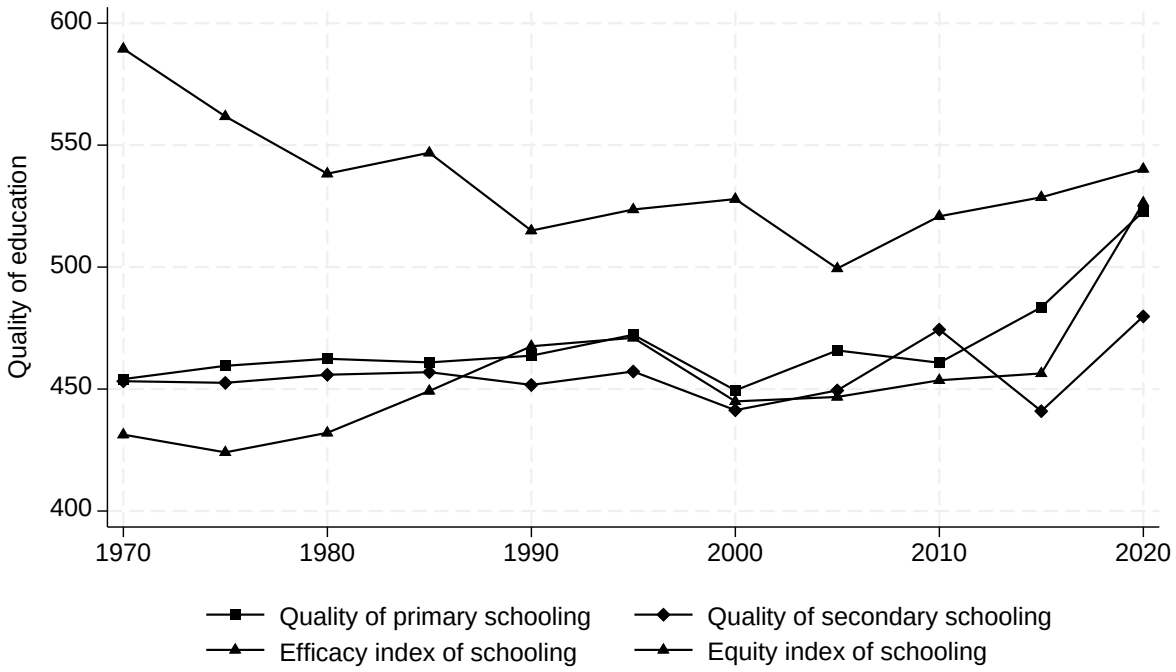
$$Value\ of\ the\ reform = \sum_{t=2020}^{t=2100}(GDP^{t}_{reform} - GDP^{t}_{without\ reform}) \times (1+d)^{-(t-2020)} \quad (13)$$

where *d* is the discount rate.

3. Results

We explore the overall performance of the Turkish education system by analyzing the evolution of the various indicators used in our study (Figure 1). The quality of primary and secondary education appears to have fallen below the OECD average in the 1970s (i.e. set at around 500 points). However, an upward trend appears to have begun in the 2000s. The efficiency indicator of the education system, which combines the proportions of students reaching the two performance thresholds, seems to show an increase between 1970 and 1990, followed by a decline until the early 2000s. Thereafter, the index rises significantly, which can be explained by the educational policies outlined in the introduction. However, equity in Türkiye's education system seems to be eroding almost continuously, except during the last decade.

Figure 1: Performance indicators for the Turkish education system (1970-2020)



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

As indicated in the introduction to this article, the main innovation of our work lies in the availability of comparable data for the quality and equity of education over almost half a century (1970-2020). Restricting ourselves to a single country can conceal general trends and attribute exogenous improvements to one country. Our method is comparative and focuses here on countries similar to Türkiye, developing countries that we call DEV-20. One advantage of restricting ourselves to these countries is the availability of long-term data, and the fact that Türkiye is positioned among countries that have followed a similar trajectory in terms of school enrolment, particularly at the secondary level where enrolment was not universal in the 1970s. In this way, we can avoid comparing Türkiye with OECD countries where universal secondary schooling was achieved by the end of the 1970s, which would tend to overestimate Türkiye's performance in the various skill areas.[16] Looking at the quantitative dimension of education alone, we see a significant gap between Türkiye and OECD countries (Figure A.1). Secondary school completion rates are significantly lower in the developing countries in our sample than in OECD countries. At the same time, there has been very strong growth in student enrolment, especially since 1995. Beyond the quantitative dimension, a purely qualitative approach also highlights the dominance of OECD countries (Figure A.2). Indeed, the quality of education has always appeared superior for OECD countries, even if Türkiye seems to have caught up in the last decade. An analysis of our efficiency index confirms our approach of comparing Türkiye with developing countries. Although efficiency is higher in Türkiye than in developing countries, the country's singular position suggests that it should be placed at the level of emerging countries such as Argentina and Indonesia (Figure A.3). The equity of Türkiye's education system appeared quite high in the 1970s, especially in comparison with other developing countries (Figure A.4). However, it seems to have gradually declined over the years. It is therefore necessary to analyze all dimensions of education in order to gain a comparative perspective on Türkiye's performance.

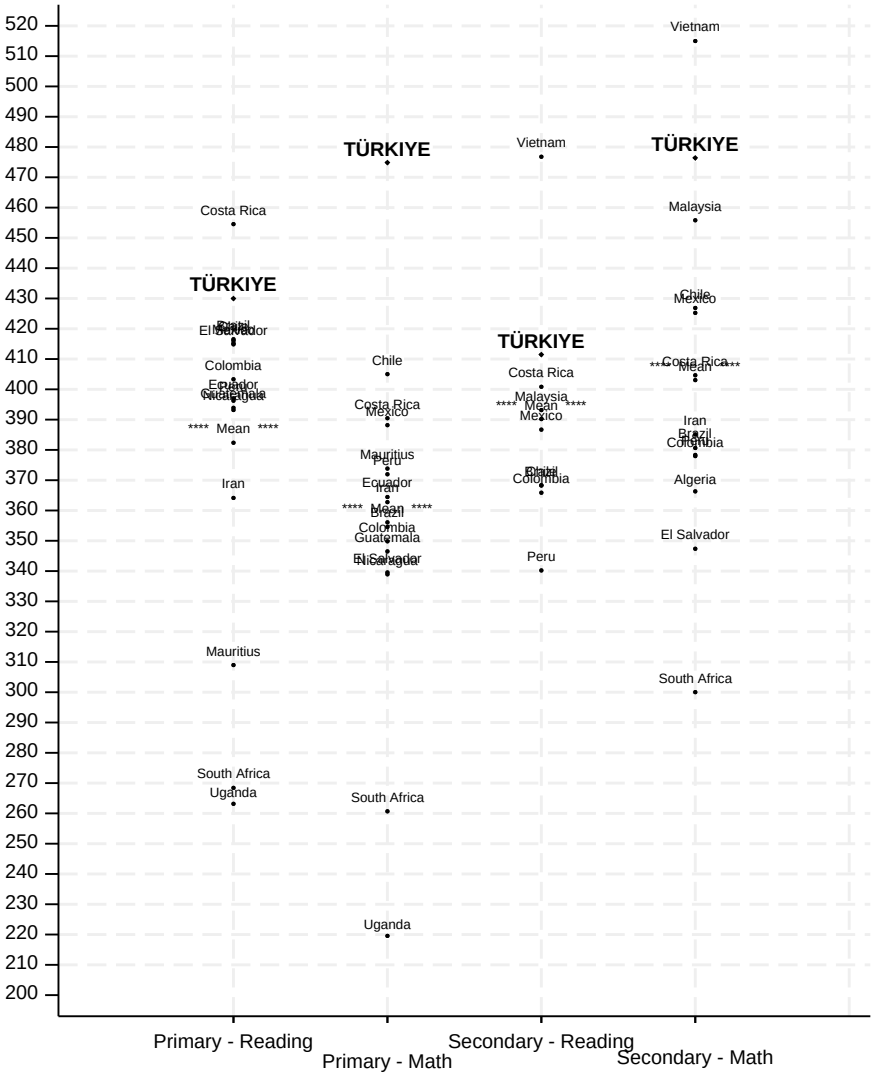
3.1. High performance compared with other developing countries

We begin by comparing the quality of education in Türkiye in a cross-section for the year 1970. The skill level we find for Türkiye is one of the highest among the developing countries in our panel - called *DEV-20* - for both grade levels, in reading and mathematics (Figure 2). While Türkiye's average score is close to 500 points in mathematics at the primary level, its average level is only 360 points for the countries in our panel. Assuming that one year of

[16]By way of example, Spaull (2018) has shown that Türkiye's results in the PISA survey appeared to be overestimated in the early 2000s due to a lower enrolment of 15-year-olds than in other OECD countries.

achievement is equivalent to around 40 points for a developing country (Avvisati & Givord, 2023), this would mean that Türkiye was more than four years ahead of the other countries. It is worth noting that only Vietnam outperforms Türkiye, particularly at the secondary level, while Malaysia scores around 20 to 30 points lower than Türkiye. Moreover, in 2020[17], Türkiye's position remains high, despite lower performance than countries such as Slovenia, the Czech Republic and Vietnam (Figure A.5). On the other hand, African countries such as South Africa and Uganda have significantly lower scores than the other countries in the sample.

Figure 2: Performance of DEV-20 countries in 1970



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

[17]It should be noted that the 2020 data exclude the Covid-19 crisis because our projections stop at 2019 and we extrapolate the 2020 results by reproducing the trend observed over the last period (for example between 2015-2018 according to PISA or 2015-2019 according to TIMSS).

Table 3 shows the evolution of performance for each skill area (mathematics, reading and science) and each school level analyzed (primary and secondary). The period studied spans 50 years, from 1970 to 2020. Based on an international average in 2000 of 500 and a standard deviation of 100, it appears that, since 1970, school performance in Türkiye has increased, even if it was below the OECD average in 2000. To assess the significance of the variation in performance, we calculated the measurement error for each mean score.[18] The reading level for primary school rose from 430 to 496 points between 1970 and 2020 (see Figure A.6). This represents an increase of around 70 points, or 14 points per decade, which is the equivalent of two school years of learning over half a century. These variations are almost all significant in terms of the standard errors calculated.[19] If we look at the column relating to long-term standardized variation (column 8), the gains are all positive over the long term: the Turkish education system has therefore globally improved its ability to teach students since 1970. The average level in secondary education tended to fall between 1970 and 2000, while a clear rise was observed over the last two decades (see also Figure A.7).

Table 3: Trends in education quality and productivity effects

	(1)		(2)		(3)		(4)		(5)		(6)		(7)			(8)			(9)	(10)	(11)	(12)
	Mean performance in score points												Trends over time						Productivity effects			
	1970		1980		1990		2000		2010		2020		Short term			Long Term			Short term		Long term	
	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Var	s.e.		Var	s.e.		L.B.	U.B.	L.B.	U.B.
Mean	454	(5.7)	459	(4.9)	458	(4.9)	445	(4.9)	468	(4.8)	501	(2.0)	28	(2.6)	*	10	(1.2)	*	1.9	2.8	1.6	2.4
Primary	454	(5.6)	462	(4.7)	464	(4.7)	449	(4.7)	461	(4.7)	523	(4.8)	37	(3.4)	*	14	(1.5)	*	2.4	3.7	2.3	3.4
Reading	430	(4.4)	452	(3.5)	456	(3.5)	445	(3.5)	472	(3.1)	496	(3.1)	26	(2.4)	*	13	(1.1)	*	1.7	2.6	2.2	3.3
Math	475	(6.2)	469	(4.8)	473	(4.7)	454	(4.8)	460	(4.7)	534	(4.8)	40	(3.4)	*	12	(1.6)	*	2.7	4.0	2.0	3.0
Science	457	(5.3)	466	(4.7)	461	(4.7)	450	(4.7)	451	(4.7)	538	(4.5)	44	(3.2)	*	16	(1.4)	*	3.0	4.4	2.7	4.0
Secondary	453	(5.7)	456	(4.9)	452	(4.9)	441	(4.9)	474	(4.8)	480	(2.0)	19	(2.6)	*	5	(1.2)	*	1.3	1.9	0.9	1.3
Reading	411	(6.5)	435	(5.8)	439	(5.8)	432	(5.8)	468	(3.7)	461	(2.0)	15	(3.1)	*	10	(1.4)	*	1.0	1.5	1.6	2.5
Math	476	(6.0)	463	(4.9)	454	(4.9)	452	(4.9)	471	(4.8)	478	(2.0)	13	(2.6)	*	0	(1.3)		0.9	1.3	0.1	0.1
Science	472	(5.2)	469	(4.6)	462	(4.6)	441	(4.5)	485	(3.9)	501	(2.1)	30	(2.5)	*	6	(1.1)	*	2.0	3.0	1.0	1.4

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

In addition to the long-term analysis (50 years), we have also calculated short-term variations (over the last 20 years). Here, the results confirm and amplify the upward trend,

[18]The measurement error or "standard error" is calculated here in such a way as to encompass both the measurement error inherent in estimating *macro* performance from a *micro* survey, but also taking into account the measurement error arising from the multiple imputation method. Assuming that the two errors are independent, we can calculate the standard error as follows $error = measurement\ error_{survey} + measurement\ error_{imputation}$. The measurement error of the imputation is achieved through a *bootstrapping* method with 30 replications on the results from the imputation, which itself includes around 80 score predictions.

[19]Most short- and long-term variations are significant at the 5% level. To conduct the significance test, we combined the standard errors ("S.E.") and calculated the threshold using the following formula:$threshold = \sqrt{E.S._{France}^2 + E.S._{OCDE}^2} \times 1{,}96$. If the corresponding difference is greater than this value, we deduce that the difference is significantly different from 0 at the 5% threshold. Significant data are indicated with an asterisk (*).

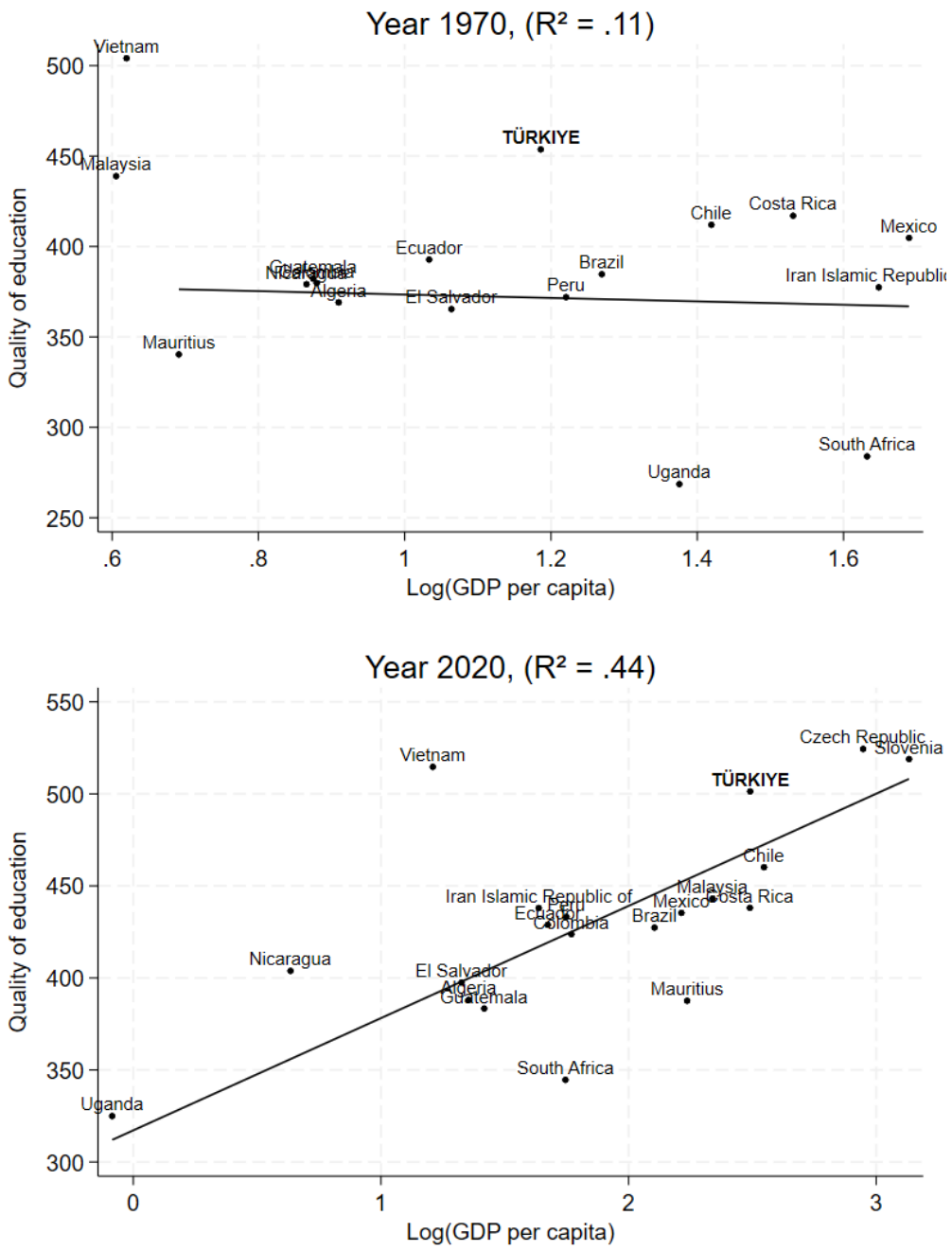
highlighting an acceleration in performance in all areas and levels since 2000 (column 7). Performance in mathematics at primary level has risen by around 40 points per decade over the last 20 years, while at secondary level it has risen by around 13 points per decade (although the latter is not significant). If we relate this variation to the equivalence in terms of years of acquisition, the effects are quite significant. While, over the long term, Türkiye has gained the equivalent of more than a year's acquisition in reading (range between 1.6 and 2.4 school years), the increase is just as significant over the short term (range between 1.9 and 2.8 school years). The increase at primary level is double that at secondary level (3 years vs. 1.6 school years respectively).

We know that the level of school skills is closely linked to the economic level of nations (Hanushek & Woessmann, 2015). Even if the relationship is less clear-cut, we might therefore logically expect better results within schools that are better endowed with financial resources (Hanushek, 2019). Moreover, in order to test the extent to which Türkiye's performance is in line with what might be expected given its economic level or the scale of its education spending, we carry out various tests. First, we carry out a graphical analysis of the correlations between each skill area and educational level and the economic wealth indicator. More specifically, we establish a correlation between educational performance and GDP per capita in dollars, adjusted by the purchasing power parity method.

Figure 3 illustrates the significant gap between Türkiye and other countries for the quality indicator, obtained by grouping all subjects and all levels. [20] While the expected score would be around 375 points, the actual score is around 450 points, which is 75 points higher than the score predicted by a simple prediction based on GDP per capita. The same comparison for the year 2020 shows a stronger correlation between economic level and school performance (the correlation coefficient is 0.44), while Türkiye manages to perform better than expected if we consider a linear relationship between economic level and education quality (501 vs. 475).

[20]This analysis was carried out for each skill area and grade level. Additional results can be found in Figures A.8-A.11.

Figure 3: Relationship between GDP per capita and quality of education (DEV-20 countries)



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

To gain a better understanding of this potential performance gap, we regress performance on the economic level of each country. Beyond the coefficient associated with this last variable, it is the residual attached to each country that attracts our attention (Table 4). In each regression, we associate a particular dimension of educational skills with GDP per capita. Since the data are panel data, this enables us to obtain country fixed effects that are not

explained by the economic wealth indicator. Of the 20 countries included in the sample[21], Türkiye's ranking is often in the top five, underlining its higher score than would be expected if its economic level were taken into account. In fact, it should be remembered that the residual ranking allows us to measure what is not taken into account in the model, namely economic level.

Table 4: School performance and economic level of countries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(GDP pc)	50.232***	12.696**	36.279***	37.437***	28.401***	34.483***	99.190***	41.456***	77.492***
	(2.404)	(6.454)	(5.649)	(6.105)	(7.481)	(3.622)	(3.152)	(8.214)	(4.271)
Intercept	294.023***	383.119***	344.413***	317.415***	368.244***	328.309***	99.457***	94.317***	124.756***
	(8.770)	(12.309)	(12.475)	(13.285)	(11.745)	(7.596)	(16.608)	(28.489)	(15.541)
Level	Primary	Secondary	Both	Both	Both	Both	Primary (adj.)	Secondary (adj.)	Both (adj.)
Skill	All	All	Reading	Maths	Sciences	All	All	All	All
Adjusted	No	No	No	No	No	No	Yes	Yes	Yes
R-squared	0.53	0.10	0.31	0.31	0.22	0.30	0.68	0.45	0.56
Observations	502	433	293	337	305	935	502	427	929
Nb. countries	17	17	20	19	20	20	17	17	20
1970-2020									
Ranking	3	5	5	5	5	5	4	4	4
Residuals	-16.10	-19.82	-4.90	-28.89	-18.86	-17.99	39.17	123.97	81.50
Best ranked	Czech Rep.	Czech Rep.	Czech Rep.	Czech Rep.	Vietnam	Czech Rep.	Uganda	Vietnam	Vietnam
2000-2020									
Ranking	2	4	4	4	4	4	6	4	3
Residuals	10.39	22.48	30.30	6.71	13.19	16.33	13.15	87.21	50.62
Best ranked	Czech Rep.	Vietnam	Vietnam	Vietnam	Vietnam	Vietnam	Nicaragua	Vietnam	Vietnam

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

In almost all regressions, the Czech Republic and Vietnam show the highest residuals. The latter country, which is often highlighted in PISA studies, performs much better than it should, taking into account its economic level (Boman, 2022; Parandekar & Sedmik, 2016; World Bank, 2011). For Türkiye, the ranking highlights a higher-than-expected performance in all cases over the short term, while a certain lag can be noted in the long-term data (see lines "Period 1970-2020" and "Period 2000-2020" for the long and short term, respectively). These results confirm those recently published for TIMSS 2023 (Polat et al., 2024) and PISA 2022 (OECD, 2023b). By way of illustration, the results of the regression between reading and economic level suggest an "over-performance" of around 30 points in the short term (column 3), once economic level is taken into account. As pointed out earlier (Avvisati & Givord, 2023), this difference is equivalent to a year of schooling gained for Türkiye. It should be noted, however, that even if Türkiye's "over-performance" appeared in the 2000s, it seems to go beyond a simple catch-up effect, since the residuals are all positive in the short term. Moreover, these differences in residuals are based on fixed effects and controlled for

[21]For some regressions, the number of countries is less than 20, due to missing values.

economic level. It therefore seems highly likely that the explanations for Türkiye's gap lie in structural rather than cyclical causes, and that they can be explained in particular by the educational reforms carried out since the early 2000s. Over the long term, however, the results are more mixed. Residuals are mostly negative (columns 1 to 6).

As previously mentioned, using the quality of education indicator alone can hide disparities in enrolment rates. A country may perform well on student achievement surveys mainly due to an educational policy aimed at selecting only the best students (Spaull, 2018). To circumvent this potential selection bias, we multiply the score in each domain and school level by the corresponding school completion rate. Thus, the average score at primary level is multiplied by the primary completion rate (column 7), while we use the total average score multiplied by the secondary completion rate thereafter (column 9). Türkiye's position among the residuals declines slightly, confirming our earlier findings of a certain lag in school completion (column 7). However, when all dimensions are taken into account, Türkiye's ranking returns to a rather good position (3rd place, column 9). This confirms that Türkiye's success is not solely due to a policy of selecting the best students and leaving the poorest out of the education system. The 2012 reform aimed at extending compulsory schooling to grade 12 may have had a strong impact on achieving universal enrolment. This reform restructured the Turkish education system into three stages of four years each: four years of primary education, four years of lower secondary education, and four years of upper secondary education (Köseleci, 2015).

A similar approach using an alternative indicator — educational expenditure as a percentage of GDP per capita — leads to similar conclusions (see Table A.1). Indeed, when comparing countries according to their level of educational expenditure, Türkiye is ahead of other developing countries in the short term, while the opposite is true in the long term.

3.2. A convergence of educational performance achieved and surpassed

The various results of international assessments all point to an increase in performance in mathematics and science over almost two decades (OECD, 2023a). In order to confirm or refute this trend, we have focused on the evolution of performance in the different skill areas and the two school levels since 1970, for countries with similar characteristics to Türkiye. We note that a clear upward trend does indeed seem to have taken place in Türkiye since the early 2000s, whereas its ranking was average in earlier periods (see Table 3). Previously, we observed a long-term upward trend in school performance in Türkiye, but also an acceleration over the last two decades for all skill domains (section 3.1). We feel it is important to compare

these variations with other developing countries (Table 5). To do this, we calculate for each skill area, and for each year, the difference between Türkiye's performance and that of the average of the 20 developing countries included in our sample (i.e. DEV-20). The difference is standardized and expressed in terms of standard deviation. Since our database has a standard deviation of 100, the calculation for standardization is done by dividing the performance gap by 100. While Türkiye had a primary math score of 475 in 1970, compared with an average of 356 points for DEV-20, the performance gap is 119 points, or 1.19 standard deviations in Türkiye's favor. Almost always, Türkiye's performance is superior to that of the DEV-20 average (i.e. the coefficients in columns 1-3 are all positive). The significance of deviations from the mean can be approximated by estimating the threshold at which the deviation becomes significant. The threshold is calculated by adding the standard errors and weighting them by the 5% error threshold.[22] Thus, in Table 5, relative differences in performance that are significant at the 5% threshold are indicated in bold and followed by an asterisk (*).

Table 5: Counterfactual analysis of Türkiye's performance
on average education quality scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Standardized difference			1970		2020		Short term trends		Long term trends		Short term	Long term
	1970	2000	2020	TUR s.e.	DEV s.e.	TUR s.e.	DEV s.e.	TUR s.e.	DEV s.e.	TUR s.e.	DEV s.e.	DinD s.e.	DinD s.e.
Mean	**0.69** *	**0.31** *	**0.70** *	454 (5.7)	385 (4.8)	501 (2.0)	431 (2.4)	**28.0** (5.3) *	8.2 (4.3)	9.5 (6.1)	9.2 (5.4)	**19.8** (3.3) *	0.3 (0.6)
Primary	**0.83** *	**0.44** *	**0.91** *	454 (5.6)	371 (4.6)	523 (4.8)	432 (2.5)	**36.7** (6.8) *	**13.2** (4.2) *	**13.7** (7.4)	**12.2** (5.2) *	**23.4** (3.9) *	**1.6** (0.7) *
Reading	**0.48** *	**0.22** *	**0.53** *	430 (4.4)	382 (4.7)	496 (3.1)	443 (2.4)	**25.8** (4.7) *	**10.0** (4.4) *	**13.3** (5.4) *	**12.1** (5.3) *	**15.8** (3.1) *	**1.2** (0.6) *
Math	**1.19** *	**0.68** *	**1.13** *	475 (6.2)	356 (4.7)	534 (4.8)	421 (2.5)	**39.9** (6.8) *	**17.6** (4.2) *	**11.8** (7.9)	**13.0** (5.4) *	**22.3** (3.9) *	**-1.2** (0.7) *
Science	**0.56** *	**0.18** *	**0.80** *	457 (5.3)	401 (4.9)	538 (4.5)	458 (2.4)	**44.4** (6.5) *	**13.2** (4.3) *	**16.2** (6.9) *	**11.3** (5.4) *	**31.2** (3.8) *	**4.9** (0.7) *
Secondary	**0.58** *	**0.16** *	**0.45** *	453 (5.7)	395 (4.9)	480 (2.0)	435 (2.6)	**19.2** (5.3) *	4.8 (4.5)	5.3 (6.0)	7.9 (5.5)	**14.4** (3.4) *	**-2.6** (0.6) *
Reading	**0.21** *	0.04	**0.22** *	411 (6.5)	390 (4.8)	461 (2.0)	439 (2.7)	**14.6** (6.1) *	6.0 (4.5)	9.9 (6.8)	9.8 (5.5)	**8.7** (3.7) *	**0.1** (0.7)
Math	**0.73** *	**0.21** *	**0.43** *	476 (6.0)	403 (4.9)	478 (2.0)	435 (2.6)	**13.3** (5.3) *	2.3 (4.5)	0.3 (6.3)	6.3 (5.5)	**11.0** (3.4) *	**-6.0** (0.7) *
Science	**0.63** *	0.01	**0.49** *	472 (5.2)	409 (5.1)	501 (2.1)	452 (2.7)	**29.8** (5.0) *	6.1 (4.7)	5.8 (5.6)	8.5 (5.8)	**23.7** (3.3) *	**-2.8** (0.6) *

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

We also calculate the variation in performance for the short (20 years) and long term (50 years). Columns 8-11 of Table 5 highlight that the variation in Turkish performance is fairly close to the DEV-20 country average over the long term (columns 10 and 11), even if progress is slightly higher for Türkiye. In the short term, however, the amplitude of the effects is different: while performance increases in almost all areas and school levels for the average DEV-20 country, the variation is more sustained in Türkiye. On average, primary school

[22] To conduct the significance test, we combined the standard errors ("S.E.") and calculated the threshold using the following formula: $hreshold = \sqrt{E.S._{Türkiye}^{2} + E.S._{DEV20}^{2}} \times 1{,}96$. If the difference between Türkiye and the DEV20 countries is greater than this value, we deduce that the difference is significantly different from 0 at the 5% threshold.

performance in DEV-20 countries improves by around 13 points per decade in the short term, while in Türkiye it rises by over 36 points. We note that Türkiye's performance gaps are greater at primary level than at secondary level. Explaining these differences by school level would require further research, which is beyond the scope of our study. However, it is possible to put forward a number of hypotheses that seem credible to us, such as the fact that the primary education system performs better than the secondary level, a more optimal allocation of primary school teachers (particularly in the field of mathematics), but also a "survey" effect specific to the data available: as Türkiye participates more in the PISA surveys than in the IEA assessments, the smaller gap observed at secondary level would suggest the difference in approach used in the surveys. PISA focuses more on skills, while the IEA emphasizes knowledge based on common curricula (Wu, 2010). It is also possible that the selective nature of secondary education in Türkiye reinforces inequalities and lowers average performance (Alacacı & Erbaş, 2010; Dincer & Uysal, 2010; Özdemir, 2015; Sarıer, 2021).

The results of relative differences provide arguments in favor of an improvement in Türkiye's performance in almost all levels and skill areas. The analysis of double differences tempers these results, particularly over the long term, and reinforces the idea of a catch-up phenomenon in terms of educational quality (column 13), particularly at secondary level. In fact, the progress observed in Türkiye is lower than that observed in the DEV-20 countries at this level of education, while the level of performance is higher in Türkiye than in these countries. In the short term, however, convergence does not seem to be taking place: Türkiye seems to be breaking away from the group of developing countries with a clear increase in its performance at both school levels. Over the last two decades, Türkiye appears to be moving closer to the group of OECD average performers.

In Figure 4, we refine the variation in school performance. We hypothesize that the countries in our sample with an initial competence deficit will have a more sustained rise in performance in subsequent decades. This convergence hypothesis can already be partially verified in Table 5, since Türkiye's relative lag is narrowing in most skill areas. [23] To verify this intuition, we cross-reference the initial level of educational measures with their variation in standardized form (i.e. expressed in standard deviations). The hypothesis of convergence is rather confirmed, as the slope of the regression line is negative in all quadrants. More

[23]This catching-up may initially be explained by a pure convergence effect, but other reasons could also explain it. The search for explanatory factors would require more detailed work on the variables involved in the educational production function, and the production of so-called multilevel models.

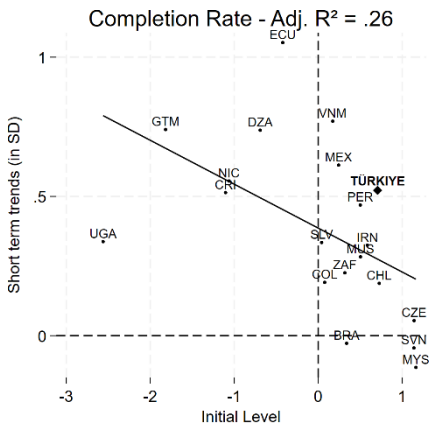
specifically, we are testing the convergence of three education indicators: a quantitative education indicator, i.e. the secondary school completion rate; a learning achievement indicator; and a hybrid indicator combining education quality with the secondary school completion rate.

Although Türkiye initially lagged behind in terms of access to education (it was on the negative side of the x-axis for completion rate in 1970), its position has improved markedly since then, confirming the convergence hypothesis for this indicator (first dimension: "completion rate"). This convergence is confirmed and even exceeded in the short term (2000-2020). For example, Türkiye has been catching up with other countries in terms of completion since the 2000s, while this indicator seems to be increasing by around 0.5 standard deviations in Türkiye, if we take into account the average variation of the DEV-20 countries (Quadrant A). Quantitative performance therefore appears to exceed that observed in other developing countries. Furthermore, the advance in learning achievement observed at the beginning of the period seems to be confirmed regardless of the period analyzed (positive ordinate values in Quadrants C & D. Moreover, this relative advance in educational attainment seems to have strengthened over time since the 2000s (positive ordinate in Quadrant C and a value close to 0.5 standard deviations). Our comparison between the initial level and variation on the hybrid quality-adjusted indicator reinforces these results: the low initial level tends to be reduced from the beginning of the 21st century (Quadrant F), and Türkiye's positive gap with other countries is clearly reinforced over the last two decades (Quadrant E).

Figure 4: Hypothesis of convergence of educational indicators in DEV-20 countries

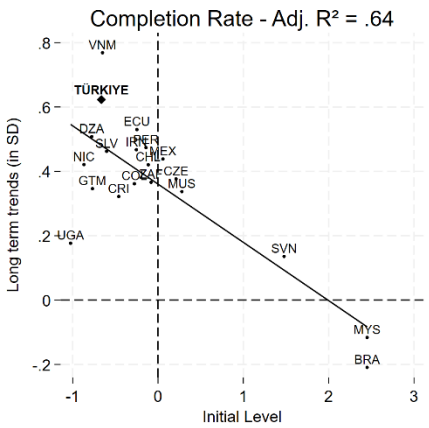
Short term trends (2000-2020)

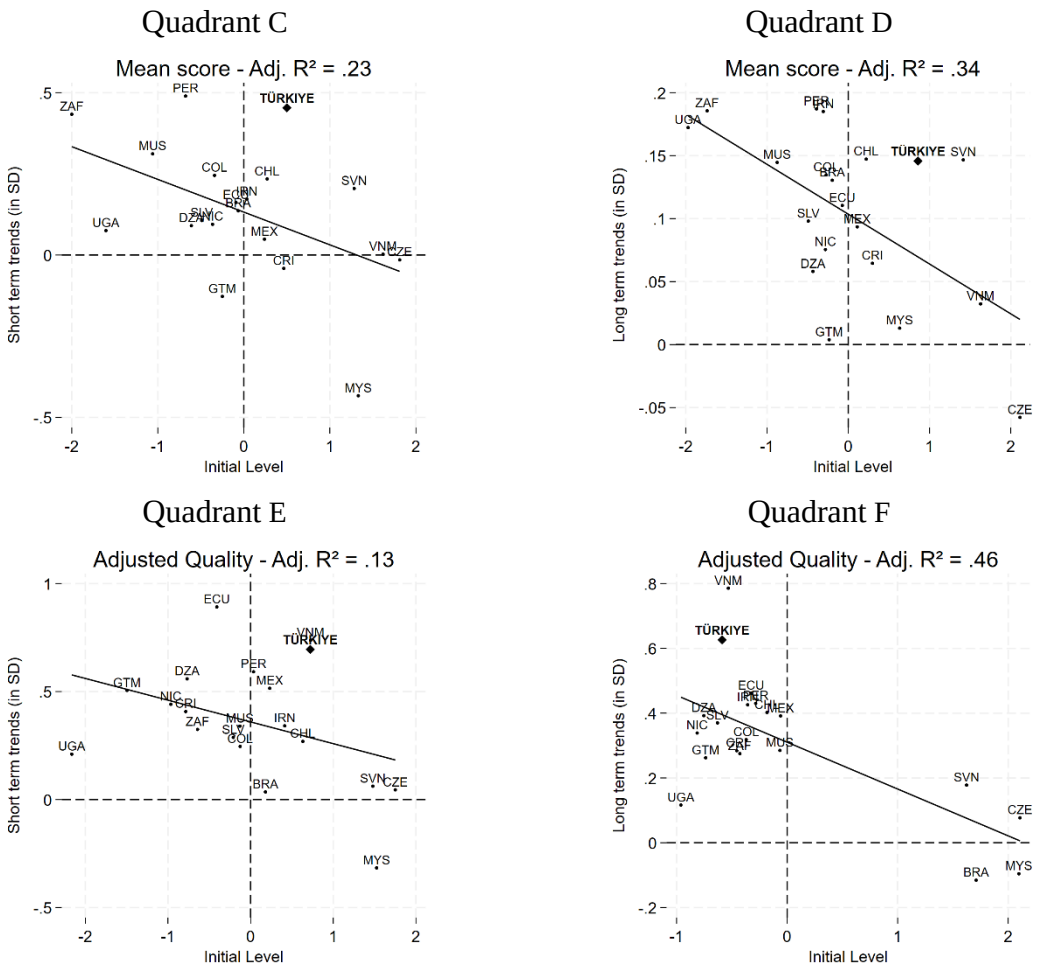
Quadrant A



Long term trends (1970-2020)

Quadrant B





Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

3.3. Robustness tests

The previous results point to a progression in educational performance in terms of quality in Türkiye, particularly over the last two decades, even after taking exogenous variations into account. In this section, we undertake calculations using the same methodology, but verifying the real potential progress of the Turkish education system across other dimensions.

Table 6 shows the trend in the proportion of Turkish pupils achieving the minimum and advanced proficiency levels (MPL and APL respectively). [24] The share of students reaching the minimum performance threshold rose from 72% to 89% on average between 1970 and 2020, i.e., a total increase of 17%, or an increase per decade of around 3%. This share accelerates in the short term, with the increase per decade rising from around 3% to an

[24]Only three dimensions of each measure are presented, in order to show the results of various indicators. However, these dimensions reflect averages and thus make it possible to synthesize the results obtained. The results for each dimension are available on request from the authors.

average of 4%. The results for the advanced skills threshold confirm those for the minimum skills threshold in terms of the acceleration observed. While average growth per decade is 0% over the long term, it has risen to 1.9% over the last two decades.

Beyond the performance thresholds, we also use the LAYS (learning-adjusted years of schooling) indicator, originally developed by Filmer et al. (2020). This indicator is calculated by combining the average education quality score and the number of years of schooling. The rise in the mixed indicator of school years adjusted by education quality shows a sharp increase since 1970: the average number of years has risen from 1.7 to 5.8 over the half-century studied. An additional measure concerns the school completion rate by level. The most significant change concerns the secondary school completion rate, which increased from 14% in 1970 to almost 100% in 2020. By crossing the completion rate and the average score, we obtain the adjusted measures of the quality of schooling. Adjusted quality thus rises from an average of 62 points to almost 501 points over the half-century.

Table 6: Robustness analysis, part 1.

Trends in performance thresholds and enrolment indicators

	(1)		(2)		(3)		(4)		(5)		(6)		(7)			(8)			(9)	(10)	(11)	(12)
	Mean performance in score points												Trends over time						Productivity effects			
	1970		1980		1990		2000		2010		2020		Short term			Long Term			Short term		Long term	
	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Var	s.e.		Var	s.e.		L.B.	U.B.	L.B.	U.B.
Minimum Prof. Level																						
Mean	72	(5.0)	74	(4.9)	82	(4.9)	80	(4.9)	82	(4.8)	89	(2.0)	4	(2.6)		3	(1.1)	*	0.3	0.4	0.5	0.8
Primary	71	(4.9)	70	(4.7)	81	(4.7)	79	(4.7)	76	(4.7)	89	(4.8)	5	(3.4)		4	(1.4)	*	0.3	0.5	0.6	0.9
Secondary	74	(5.0)	79	(4.9)	84	(4.9)	81	(4.9)	88	(4.8)	88	(2.0)	4	(2.6)		3	(1.1)	*	0.2	0.4	0.5	0.7
Advanced Proficiency Level																						
Mean	6.9	(4.9)	6.0	(4.9)	5.6	(4.9)	4.5	(4.9)	4.3	(4.8)	8.3	(2.0)	1.9	(2.6)		0.3	(1.1)		0.1	0.2	0.0	0.1
Primary	5.6	(4.7)	5.7	(4.7)	5.9	(4.7)	4.3	(4.7)	3.3	(4.7)	11.1	(4.8)	3.4	(3.4)		1.1	(1.4)		0.2	0.3	0.2	0.3
Secondary	8.2	(4.9)	6.4	(4.9)	5.2	(4.9)	4.7	(4.9)	5.4	(4.8)	5.4	(2.0)	0.4	(2.6)		-0.6	(1.1)		0.0	0.0	-0.1	-0.1
School attainment																						
Years of schooling	2.6	(0.0)	3.8	(0.0)	5.2	(0.0)	6.4	(0.0)	7.4	(0.0)	8.1	(0.0)	0.9	(0.0)	*	1.1	(0.0)	*	0.1	0.1	0.2	0.3
LAYS	1.7	(5.7)	2.5	(4.9)	3.4	(4.9)	4.1	(4.9)	5.0	(4.8)	5.8	(2.0)	0.9	(2.6)		0.8	(1.2)		0.1	0.1	0.1	0.2
Completion, pri.	91	(0.0)	88	(0.0)	95	(0.0)	97	(0.0)	99	(0.0)	99	(0.0)	1.1	(0.0)	*	1.6	(0.0)	*	0.1	0.1	0.3	0.4
Completion, sec.	14	(0.0)	24	(0.0)	35	(0.0)	79	(0.0)	100	(0.0)	100	(0.0)	10.3	(0.0)	*	17.2	(0.0)	*	0.7	1.0	2.9	4.3
Adjusted measures of quality of schooling																						
Mean	62	(5.7)	108	(4.9)	162	(4.9)	354	(4.9)	468	(4.8)	501	(2.0)	74	(2.6)	*	88	(1.2)	*	4.9	7.4	14.6	21.9
Primary	414	(5.6)	408	(4.7)	440	(4.7)	435	(4.7)	456	(4.7)	517	(4.8)	41	(3.4)	*	21	(1.5)	*	2.7	4.1	3.5	5.2
Secondary	62	(5.7)	107	(4.9)	160	(4.9)	350	(4.9)	474	(4.8)	480	(2.0)	65	(2.6)	*	83	(1.2)	*	4.3	6.5	13.9	20.9

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

It is now important to compare this trend with that observed in other developing countries (Table 7). The performance threshold indicators all point to higher performance in Türkiye. The double-difference method confirms these results, particularly in the short term. For

example, the share of students reaching the minimum performance threshold increased more in secondary school in Türkiye than in the DEV-20 countries, by around 3.6% to Türkiye's advantage per decade. At the primary level, the increase observed in Türkiye is the same as that observed in the other countries (i.e., the double-difference variation is then equal to 0, see column 12 and the line on the minimum performance threshold at the primary level). As expected, the strong growth in the secondary school completion rate underlines Türkiye's strong performance, which is, however, not yet visible through the LAYS indicator, which uses the number of school years. Finally, the analysis of the quality indicator adjusted by the completion rate shows Türkiye's higher relative performance at all levels, even if it is stronger at the secondary level. For example, education quality adjusted by completion rate at the secondary level increases by 37.6 points more per decade in Türkiye than in the other developing countries in our sample.

Table 7: Robustness analysis, part 2.

Counterfactual analysis of performance thresholds and enrolment indicators

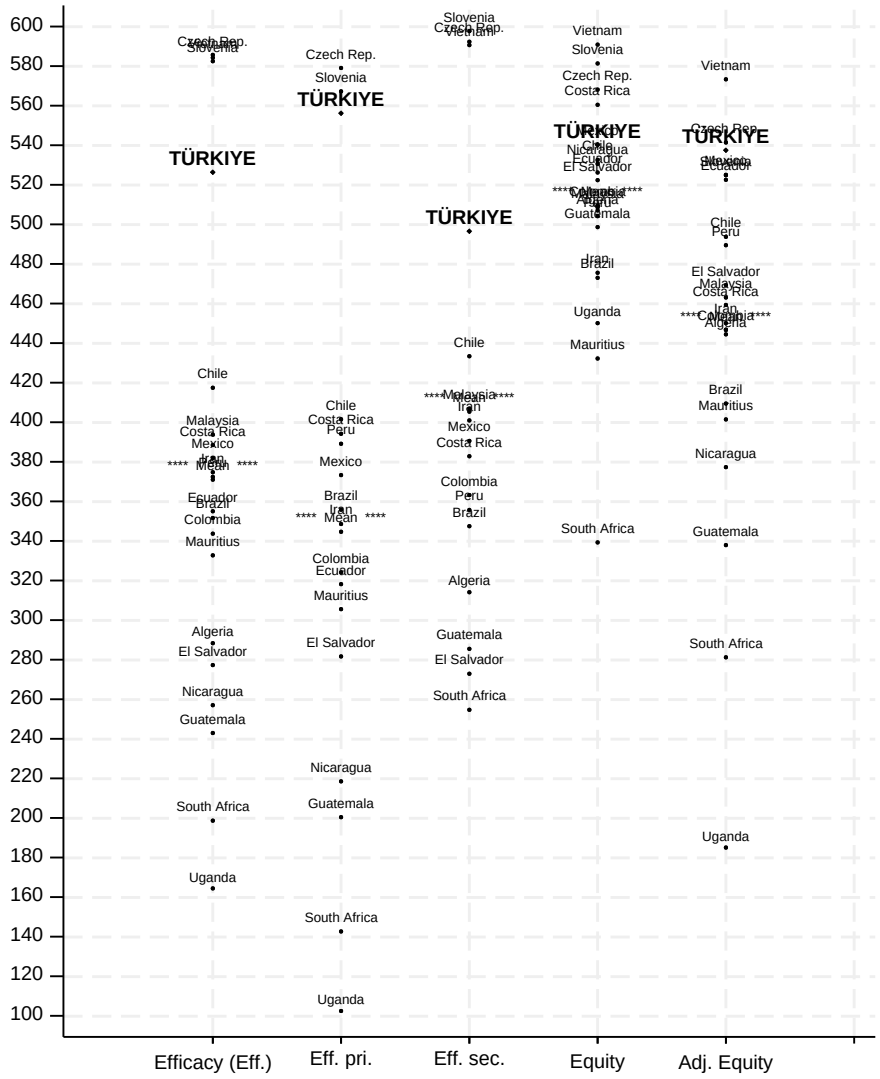
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Standardized difference			1970		2020		Short term trends (S.T.)		Long term trends (L.T.)		Difference in difference	
	1970	2000	2020	TUR	DEV	TUR	DEV	TUR	DEV	TUR	DEV	S.T.	L.T.
Minimum Profiency Level													
Mean	0.19 *	0.17 *	0.21 *	72 (5.0)	53 (3.9)	89 (2.0)	68 (2.4)	4.3 (2.6)	2.4 (2.1)	3.3 (1.1) *	2.9 (0.9) *	2.0 (1.7)	0.4 (0.3)
Primary	0.30 *	0.25 *	0.25 *	71 (4.9)	41 (3.6)	89 (4.8)	64 (2.5)	5.1 (3.4)	5.1 (2.1) *	3.6 (1.4) *	4.6 (0.9) *	0.0 (2.0)	-0.9 (0.3) *
Secondary	0.11	0.08	0.15 *	74 (5.0)	63 (4.2)	88 (2.0)	73 (2.6)	3.6 (2.6)	0.0 (2.3)	2.9 (1.1) *	2.0 (1.0) *	3.6 (1.7) *	0.9 (0.3) *
Advanced Proficiency Level													
Mean	0.03	0.01	0.05	7 (4.9)	4 (3.8)	8 (2.0)	3 (2.4)	1.9 (2.6)	0.0 (2.1)	0.3 (1.1)	-0.1 (0.9)	1.9 (1.7)	0.4 (0.3)
Primary	0.04	0.02	0.09	6 (4.7)	1 (3.4)	11 (4.8)	2 (2.5)	3.4 (3.4)	0.3 (2.1)	1.1 (1.4)	0.2 (0.8)	3.1 (2.0)	0.9 (0.3) *
Secondary	0.02	0.00	0.01	8 (4.9)	7 (4.1)	5 (2.0)	4 (2.6)	0.4 (2.6)	-0.3 (2.3)	-0.6 (1.1)	-0.5 (1.0)	0.7 (1.7)	-0.1 (0.3)
School attainment													
Years of sch.	-0.02 *	-0.01 *	-0.01 *	2.6 (0.0)	4.1 (0.0)	8.1 (0.0)	9.4 (0.0)	0.9 (0.0) *	0.9 (0.0) *	1.1 (0.0) *	1.1 (0.0) *	0.0 (0.0) *	0.0 (0.0) *
LAYS	0.00	-0.01	0.00	1.7 (5.7)	2.0 (4.8)	5.8 (2.0)	5.9 (2.4)	0.9 (2.6)	0.6 (2.2)	0.8 (1.2)	0.8 (1.1)	0.2 (1.7)	0.0 (0.3)
Compl., pri.	0.28 *	0.08 *	0.05 *	91 (0.0)	63 (0.0)	99 (0.0)	94 (0.0)	1.1 (0.0) *	2.3 (0.0) *	1.6 (0.0) *	6.1 (0.0) *	-1.2 (0.0) *	-4.5 (0.0) *
Compl., sec.	-0.17 *	0.14 *	0.19 *	14 (0.0)	31 (0.0)	100 (0.0)	81 (0.0)	10.3 (0.0) *	7.6 (0.0) *	17.2 (0.0) *	10.0 (0.0) *	2.7 (0.0) *	7.3 (0.0) *
Adjusted measures of quality of schooling													
Mean	-0.56 *	0.77 *	1.48 *	62 (5.7)	119 (4.8)	501 (2.0)	353 (2.4)	73.8 (2.6) *	38.2 (2.2) *	87.8 (1.2) *	47.0 (1.1) *	35.6 (1.7) *	40.8 (0.3) *
Primary	1.80 *	0.73 *	1.09 *	414 (5.6)	234 (4.6)	517 (4.8)	408 (2.5)	41.1 (3.4) *	22.8 (2.1) *	20.7 (1.5) *	34.8 (1.0) *	18.3 (2.0) *	-14.1 (0.4) *
Secondary	-0.78 *	0.39 *	1.14 *	62 (5.7)	141 (4.9)	480 (2.0)	366 (2.6)	64.7 (2.6) *	27.1 (2.3) *	83.5 (1.2) *	45.0 (1.1) *	37.6 (1.7) *	38.5 (0.3) *

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

It is possible to analyze the extent to which efficiency and equity have varied in Türkiye compared to other developing countries. Türkiye appears to be performing well on both dimensions in 2020 (Figure 5). The efficiency of an education system is measured by combining the shares of students reaching the minimum and advanced proficiency thresholds (see equation 1). School efficiency increases by around 100 points in Türkiye over half a century, from 431 to 526 points (Table 8). The increase is more marked at the primary level (146 points) than at the secondary level (44 points). This rise seems to have accelerated over

the last two decades (41 points per decade, column 7) compared with the long-term variation (19 points, column 8). The evolution of equity, measured by the gap relative to the median between the extreme deciles (see equation 2), highlights a decline in equity since 1970. Indeed, average equity fell from 589 points to 540 points between 1970 and 2020. The decline is even more marked in secondary schools, where it falls from 628 to 547 points over the same period. This can be explained by the rise in school enrolment rates among the least affluent populations. The democratization of education has increased school enrolment among the poorest and could therefore explain the rise in the gap between the extreme deciles. To test this, we adjust the equity indicator to take into account the school completion rate (equation 3). We then see that the adjusted equity index increases significantly in both educational levels. The increase is more marked over the recent period for the secondary level. It is possible that the rise in enrolment rates explains the increase in the gap between the extreme deciles.

Figure 5: Efficiency and equity in Türkiye's education systems and DEV-20 countries in 2020



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025). "Eff." = Efficacy, "pri." = primary ; "sec." = secondary, "Adj." = Adjusted.

Table 8: Robustness analysis, part 3.

Trends in education system efficiency and equity indices

	(1)		(2)		(3)		(4)		(5)		(6)		(7)			(8)			(9)	(10)	(11)	(12)
	Performance in score points												Trends over time						Productivity effects			
	1970		1980		1990		2000		2010		2020		Short term			Long Term			Short term		Long term	
	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Level	s.e.	Var	s.e.		Var	s.e.		L.B.	U.B.	L.B.	U.B.
Efficacy																						
Both levels	431	(0.0)	432	(0.0)	468	(0.0)	445	(0.0)	454	(0.0)	526	(0.0)	41	(0.0)	*	19	(0.0)	*	2.7	4.1	3.2	4.8
Primary	410	(0.0)	406	(0.0)	463	(0.0)	437	(0.0)	414	(0.0)	556	(0.0)	60	(0.0)	*	29	(0.0)	*	4.0	6.0	4.9	7.3
Secondary	453	(0.0)	458	(0.0)	472	(0.0)	453	(0.0)	493	(0.0)	497	(0.0)	22	(0.0)	*	9	(0.0)	*	1.4	2.2	1.5	2.2
Equity																						
Both levels	589	(0.0)	538	(0.0)	515	(0.0)	528	(0.0)	521	(0.0)	540	(0.0)	6	(0.0)	*	-10	(0.0)	*	0.4	0.6	-1.6	-2.5
Primary	551	(0.0)	498	(0.0)	479	(0.0)	460	(0.0)	491	(0.0)	533	(0.0)	37	(0.0)	*	-4	(0.0)	*	2.4	3.7	-0.6	-0.9
Secondary	628	(0.0)	578	(0.0)	551	(0.0)	596	(0.0)	551	(0.0)	547	(0.0)	-24	(0.0)	*	-16	(0.0)	*	-1.6	-2.4	-2.7	-4.1
Equity (adjusted)																						
Both levels	294	(0.0)	288	(0.0)	325	(0.0)	459	(0.0)	518	(0.0)	538	(0.0)	39	(0.0)	*	49	(0.0)	*	2.6	3.9	8.1	12.2
Primary	502	(0.0)	440	(0.0)	454	(0.0)	445	(0.0)	486	(0.0)	528	(0.0)	41	(0.0)	*	5	(0.0)	*	2.8	4.1	0.9	1.3
Secondary	87	(0.0)	136	(0.0)	195	(0.0)	473	(0.0)	551	(0.0)	547	(0.0)	37	(0.0)	*	92	(0.0)	*	2.5	3.7	15.4	23.0

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

We compare these variations with those observed in the other countries in our sample (Table 9). Almost all the coefficients are positive in the first three columns, suggesting that the variations observed in Türkiye are stronger than those in other developing countries. For example, while efficiency is equal to 526 points in Türkiye in 2020, it is only 371 points in the DEV-20 panel of countries. While for each decade, school efficiency increases by 40.7 points in Türkiye between 2000 and 2020, it is only 12 points for the DEV-20 countries, leading to a positive result for the double-difference method (+28.7%). The results for equity suggest a weaker performance for Türkiye if we disregard enrolment levels, particularly in secondary education (-36.7% gap in the short term). Once enrolment levels are taken into account (adjusted equity indicator), the gap remains negative, but its amplitude is divided by three (-11.7% versus -36.7%). This means that the issue of inequality in the Turkish education system is not solely due to mass schooling, but also to other factors specific to Türkiye, such as a policy of segregation between schools (Dincer & Uysal, 2010).

Table 9: Robustness analysis, part 4.

Counterfactual analysis of education system efficiency and equity indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Standardized difference			1970		2020		Short term trends		Long term trends		Short term	Long term
	1970	2000	2020	TUR	s.e. DEV s.e.	TUR	s.e. DEV s.e.	TUR	s.e. DEV s.e.	TUR	s.e. DEV s.e.	DinD	s.e. DinD s.e.
Efficacy													
Both levels	1.27 *	0.98 *	1.55 *	431	(0.0) 305 (0.0)	526	(0.0) 371 (0.0)	40.7	(0.0) * 12.0 (0.0) *	19.0	(0.0) * 13.3 (0.0) *	28.7	(0.0) * 5.7 (0.0) *
Primary	1.89 *	1.49 *	2.12 *	410	(0.0) 221 (0.0)	556	(0.0) 345 (0.0)	59.8	(0.0) * 28.2 (0.0) *	29.3	(0.0) * 24.8 (0.0) *	31.5	(0.0) * 4.5 (0.0) *
Secondary	0.72 *	0.41 *	0.91 *	453	(0.0) 381 (0.0)	497	(0.0) 405 (0.0)	21.7	(0.0) * -3.2 (0.0) *	8.8	(0.0) * 4.9 (0.0) *	24.9	(0.0) * 3.8 (0.0) *
Equity													
Both levels	1.42 *	0.56 *	0.31 *	589	(0.0) 447 (0.0)	540	(0.0) 510 (0.0)	6.2	(0.0) * 18.9 (0.0) *	-9.9	(0.0) * 12.5 (0.0) *	-12.7	(0.0) * -22.3 (0.0) *
Primary	1.26 *	0.08 *	0.31 *	551	(0.0) 425 (0.0)	533	(0.0) 502 (0.0)	36.7	(0.0) * 25.4 (0.0) *	-3.5	(0.0) * 15.4 (0.0) *	11.3	(0.0) * -18.9 (0.0) *
Secondary	1.59 *	1.04 *	0.30 *	628	(0.0) 469 (0.0)	547	(0.0) 517 (0.0)	-24.4	(0.0) * 12.4 (0.0) *	-16.2	(0.0) * 9.5 (0.0) *	-36.7	(0.0) * -25.7 (0.0) *
Equity (adjusted)													
Both levels	0.94 *	0.95 *	0.91 *	294	(0.0) 200 (0.0)	538	(0.0) 447 (0.0)	39.1	(0.0) * 41.2 (0.0) *	48.7	(0.0) * 49.2 (0.0) *	-2.0	(0.0) * -0.6 (0.0) *
Primary	2.42 *	0.38 *	0.53 *	502	(0.0) 259 (0.0)	528	(0.0) 475 (0.0)	41.3	(0.0) * 33.6 (0.0) *	5.2	(0.0) * 43.0 (0.0) *	7.6	(0.0) * -37.9 (0.0) *
Secondary	-0.55 *	1.52 *	1.29 *	87	(0.0) 141 (0.0)	547	(0.0) 419 (0.0)	37.0	(0.0) * 48.7 (0.0) *	92.1	(0.0) * 55.4 (0.0) *	-11.7	(0.0) * 36.7 (0.0) *

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

3.4. Projection model

In section 2.3.2, we presented the projection model initially developed by Hanushek & Woessmann (2012) and then updated in Gust et al. (2024). [25] We use this model to predict the growth Türkiye would experience if it succeeded in meeting targets for improving its education system. The parameters we use are those developed in Gust et al. (2024). We consider the gains obtained from an education policy over a period of 80 years (S), which leads to the year 2100. This period is roughly equivalent to the life expectancy of a child born in Türkiye at the start of the reform. The discount rate in our base model is set at 3%, a standard value in long-term projection models (see, for example, Borsch-Supan (2000)).

We first assume a growth in education quality of the order of 50 points, or 0.50 standard deviations (scenario 1). According to our estimates, and adopting the parameters presented in Table 10, we find that such a policy to improve the quality of Türkiye's education system should increase its GDP by 68% by 2100 (see Table 11). The gain from educational reform would thus be 736% of Türkiye's current GDP (i.e. around $16,500 billion over the entire 2020-2100 period). Even if such a scenario proves optimistic[26], it shows the extent to which

[25]We would like to thank S. Gust, E. Hanushek and L. Woessmann for providing us with the Stata codes to replicate and adjust the projection model.

[26]It is possible to draw up other, more realistic scenarios and compare the respective gains, but as the aim of our work is not to carry out simulations, we prefer to present only this policy for combating inequality. For example, in a recent publication, Gust et al (2024) show that a reform aimed at bringing the entire population up to at least the basic skills threshold would bring Türkiye $5,846 billion, or 261% of current GDP (see table A.6 on page 24). This reform would therefore be three times less effective in terms of growth than one aimed at intensively boosting the quality of Türkiye's education system.

an improvement in the quality of the education system can raise the economic level of a developing country like Türkiye.

Table 10: Projection model parameters

Parameter	Definition	Baseline value
R	Reform period	15
W	Length of work life	40
S	Simulation period (years)	80
d	Discount rate	3 %
p	Status quo growth rate	1,5 %
μ	Growth coefficient due to the reform	1,98 %
A*	Scores from achievement tests	500

Other scenarios are conceivable, in particular those that could focus more on school inequalities. For example, a so-called positive discrimination policy could aim to raise the performance of students with the lowest socio-economic capital to the level of students with the highest socio-economic capital (i.e., the average score of the poorest quartile would equal that of the richest quartile). This policy amounts to increasing the performance of a quarter of the population by 87 points. Thus, the total effect on the population is 22 points (obtained by dividing the targeted 87-point effect by four), i.e. a growth of 0.22 standard deviations. The total effect on GDP by 2100 would be an increase of 296% (or $6,638 billion discounted to 2020 dollars), with an average annual growth rate increasing by 1.42%, and a gain of 25.8% if the amounts are discounted.

A final scenario would aim to combine a policy of general improvement in the quality of education with one targeting pupils with the least socio-economic capital (a mix between scenario 1 and scenario 2). Thus, by raising the level of the entire population by 0.5 standard deviations, while at the same time boosting the performance of the poorest pupils (by making them converge toward that of the richest pupils), the total effect in terms of raising the quality of education is 72 points (i.e. 0.72 standard deviations, obtained by combining 50 points from scenario 1 and 22 points from scenario 2). The cumulative effect on GDP in 2100 would be a discounted increase of 111%. In the end, the cumulative gain from this reform would be $25,569 billion (or 1,141% of discounted 2020 GDP).

Table 11: Estimated gains from improving Türkiye's education system

	Scenario 1: Increasing the quality of education	Scenario 2: Reducing inequalities	Scenario 3: Combining quality and inequality
Value of reform (bn USD)	16,500	6,638	25,569
In % of current GDP	736 %	296 %	1,140.7 %
In % of discounted future GDP	15.7 %	6.3 %	24.4 %
GDP increase in year 2100	68.1 %	25.8 %	111.0 %

Notes: Methodology adapted from Gust et al. (2024). Discounted value of future increases in GDP until 2100 due to the reform scenario, expressed in billion USD, as a percentage of current GDP, and as a percentage of discounted future GDP.

4. Conclusion

Türkiye is a country that has simultaneously seen an increase in its results in international surveys of student achievement and a rise in access to education. This rather exceptional feature merits analysis. We have tried to understand whether these simultaneous trends can be explained by exogenous factors. Using a counterfactual analysis, we found that the progress observed in Türkiye over the last two decades is not mirrored in other countries with similar characteristics.

For example, while the average progression in terms of the quality of education has been 28 points per decade over the recent period, it is only 8.2 points for economically close countries. This suggests that Türkiye's progress is 20 points higher than that of developing countries with an economic level close to Türkiye's in the 1970s. This improvement is more marked at the primary level than the secondary level, although it remains significant at both levels.

In order to gain a better understanding of the performance of education systems, two indices have been calculated. The first refers to the efficiency of education systems and groups together the two indicators relating to performance thresholds. A system is said to be highly effective if it succeeds in bringing the entire population up to the minimum competency threshold while ensuring that a certain number can reach the threshold of excellence. The efficacy indices for Türkiye reinforce the idea that the Turkish education system is improving more than those of other developing countries. Recent efforts by the

Turkish government to extend compulsory schooling to the age of 16 appear not to have hampered the quality of the education system (Çelik & Gür, 2013; Köseleci, 2015; Özdemir, 2015). However, fears of rising inequality seem to be present, particularly at the secondary level. This issue seems to be the central problem of the Turkish education system (Alacacı & Erbaş, 2010; Sarıer, 2021; Yediyıldız & Ustun, 2024).

Indeed, equity at this level seems to have regressed since 1970 and especially since 2000. Using the double-difference method, we find that Türkiye has diverged in terms of equity. The other countries have improved their equity more than Türkiye, raising fears of a rise in inequality in this country. Further work is needed to understand the origins of these inequalities. Indeed, there appears to be a high degree of segregation between schools within the Turkish education system.

For example, the OECD has calculated a school inclusion index, which can be interpreted as a measure of the degree of equity and uniformity of educational outcomes between different schools in an education system. A low index suggests significant heterogeneity, where the quality of education and student achievement vary considerably from school to school. A low index of school inclusion, such as that implicitly suggested by the disparities observed in the Turkish education system, points to notable heterogeneity with significant differences in academic achievement between schools (T. Aydın & Çilek, 2024). This heterogeneity means that some schools achieve better results than others, which can be attributed to a combination of factors such as available resources, teacher quality, and students' socio-economic status (A. Aydın et al., 2012; Dolu, 2020; Sarıer, 2020).

In addition to inequalities between schools, there are also significant differences between Turkish provinces. Significant inequalities in educational achievement between Turkish provinces are evident in data from international assessments such as PISA and TIMSS. Uysal & Gelbal (2018), analyzing PISA data from 2009, 2012, and 2015, found significant differences between regions in math, science, and reading scores for all years studied, identifying Southeast and East-Central Anatolia as generally having the lowest scores. Other authors confirm these differences between Turkish provinces (Alacacı & Erbaş, 2010; Dinçer & Oral, 2013; Dolu, 2020; Köseleci, 2015). Drawing on PISA 2012 data, Köseleci (2015) points out that the percentage of 15-year-old students scoring below Level 2 varied considerably from region to region, ranging from 29% in Central Anatolia to 62% in Southeastern Anatolia region.

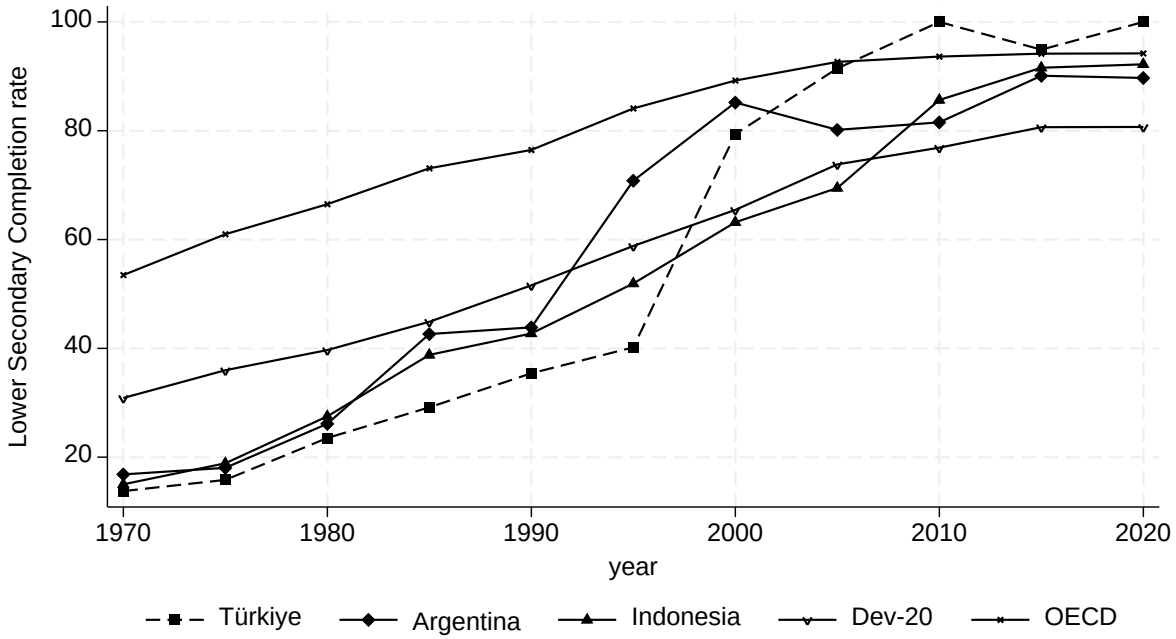
This shows that a policy aimed at improving the quality of education should also be accompanied by action to reduce educational inequalities. We show that such a mixed policy scenario could significantly improve Türkiye's economic performance. This acceleration in growth could then enable Türkiye to double its GDP by 2100 (+111%, see Table 11), in discounted terms. This growth could then undoubtedly lift Türkiye out of the developing country trap (Yilmaz, 2015) and into the status of a developed country.

The question remains, however, regarding the educational policies needed to achieve these objectives, and the sources of funding. Surveys of students' skills enable us to accurately measure the evolution of countries' school performance. However, the experiences of some countries cited as examples - such as Vietnam or the Netherlands - are difficult to replicate in other countries, and even more difficult in a country like Türkiye. Each country will have to find its own ways of targeting an effective education policy. A detailed analysis of the actions implemented in the past in Türkiye can help develop an effective policy.

In the context of this article, our research into the evolution of learning achievement in Türkiye has produced new evidence spanning half a century and in comparison with most emerging countries, leading to the conclusion that educational performance in reading and mathematics is declining, with average scores being significantly low. We have also shown that Türkiye's trajectory in terms of educational achievement tends to diverge from that of other developing countries. Generally speaking, our results enrich the conclusions drawn from national surveys, while providing new historical and comparative insights. Finally, they invite us to look beyond considerations centered on purely quantitative indicators and consider the major importance of the qualitative dimension of education.

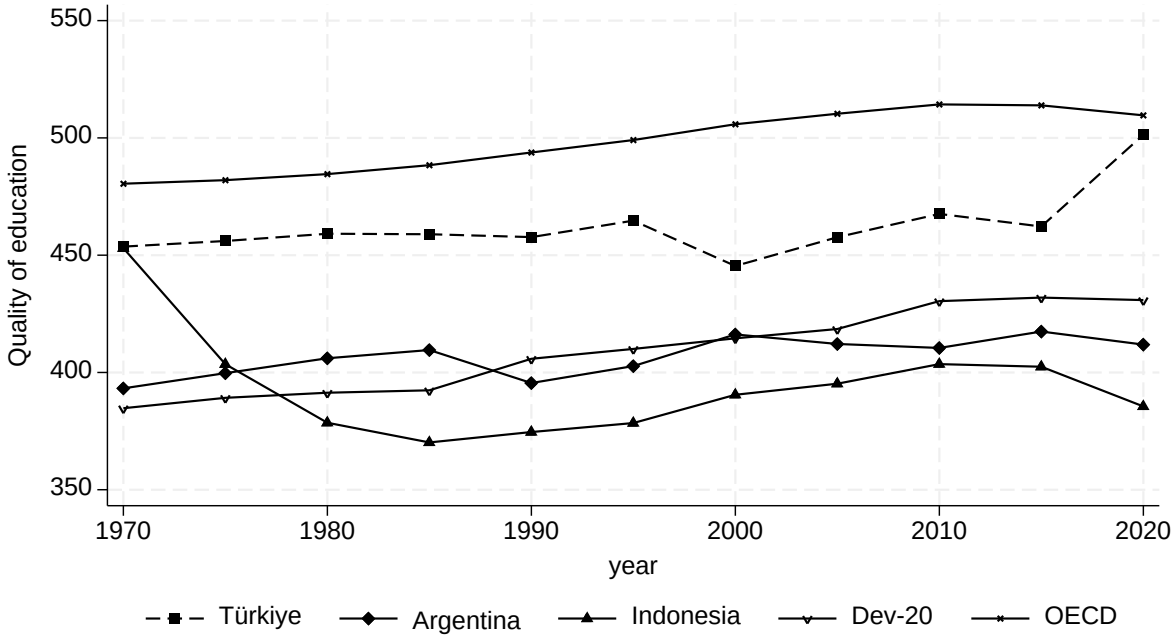
Appendix A

Figure A.1: Lower Secondary Completion Rate across different countries and regions (1970-2020)



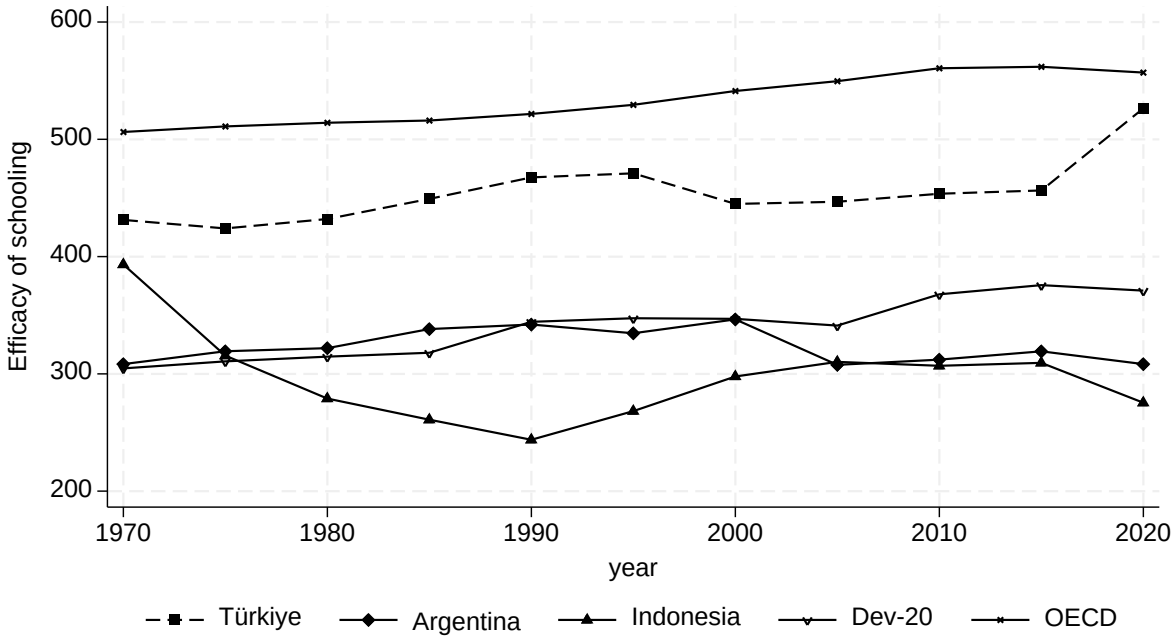
Source: authors, based on Altinok & Diebolt (2024, 2025)

Figure A.2: Quality of schooling across different countries and regions (1970-2020)



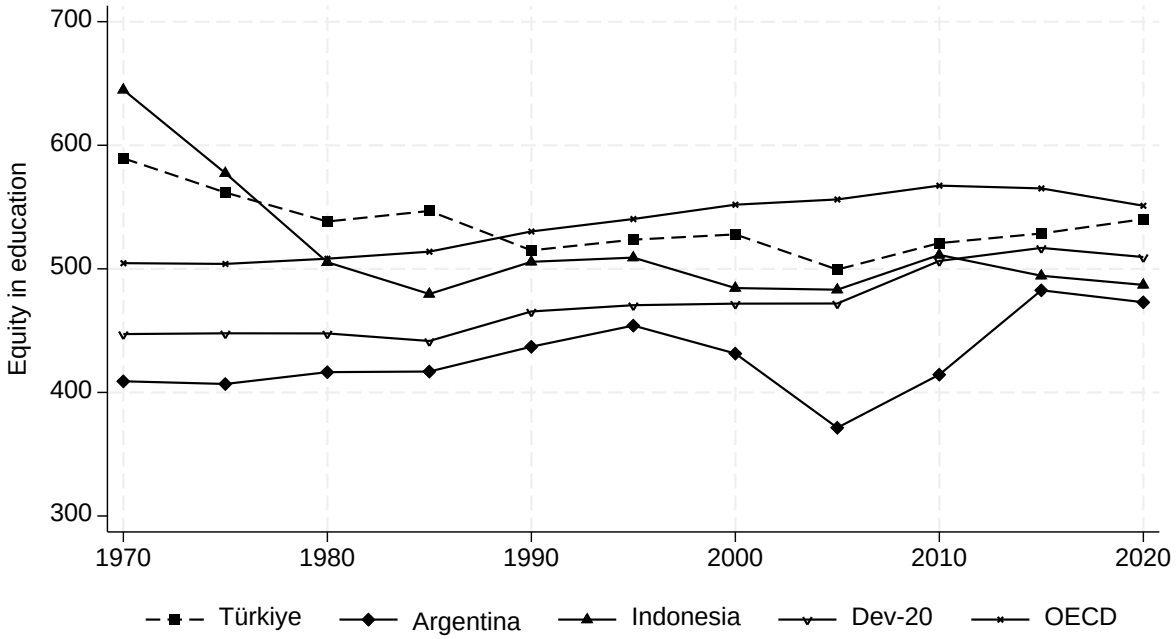
Source: authors, based on Altinok & Diebolt (2024, 2025)

Figure A.3: Efficacy index of schooling across different countries and regions (1970-2020)



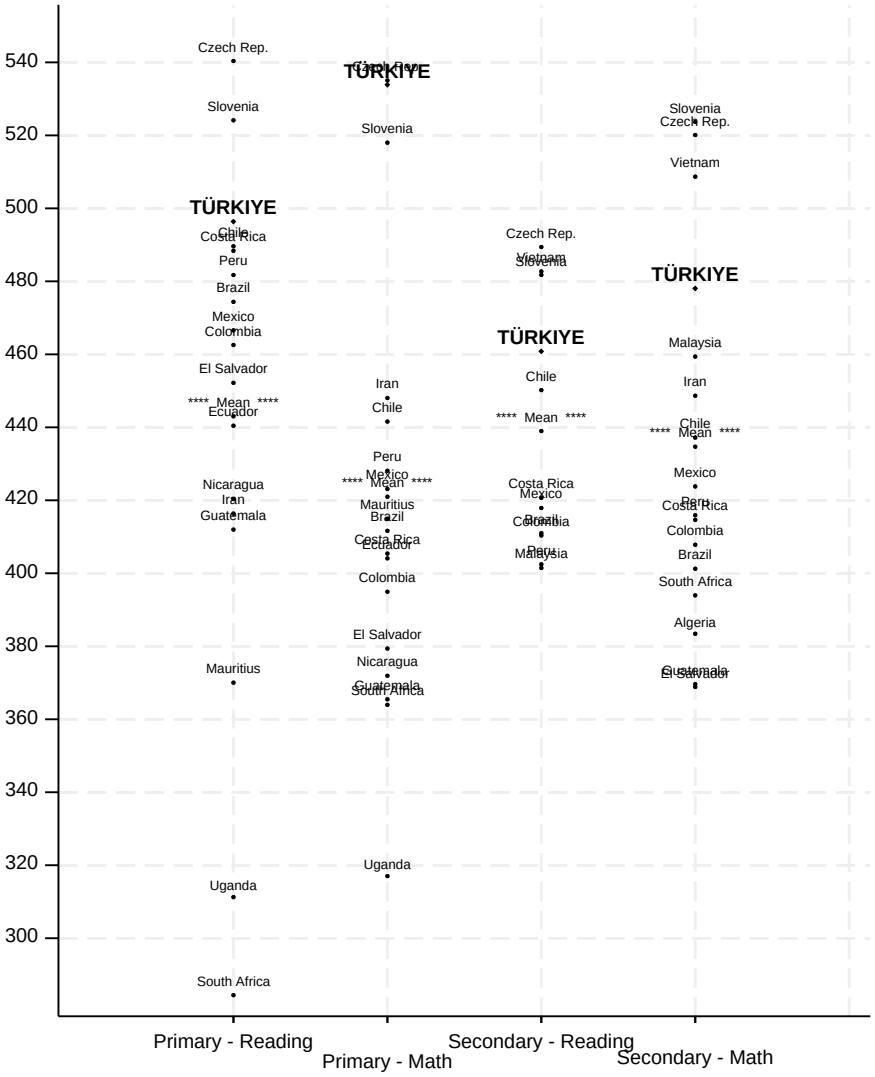
Source: authors, based on Altinok & Diebolt (2024, 2025)

Figure A.4: Equity index of schooling across different countries and regions (1970-2020)



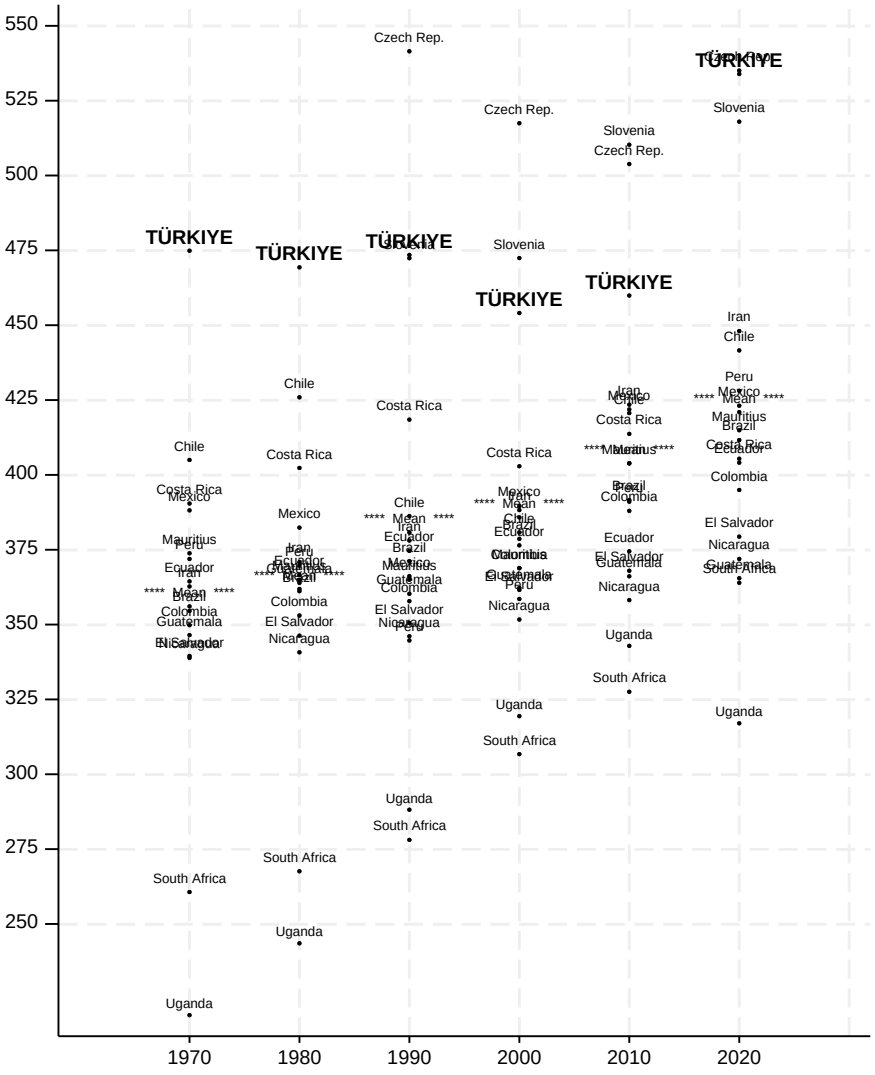
Source: authors, based on Altinok & Diebolt (2024, 2025)

Figure A.5: Performance of Türkiye compared to similar countries in 2020



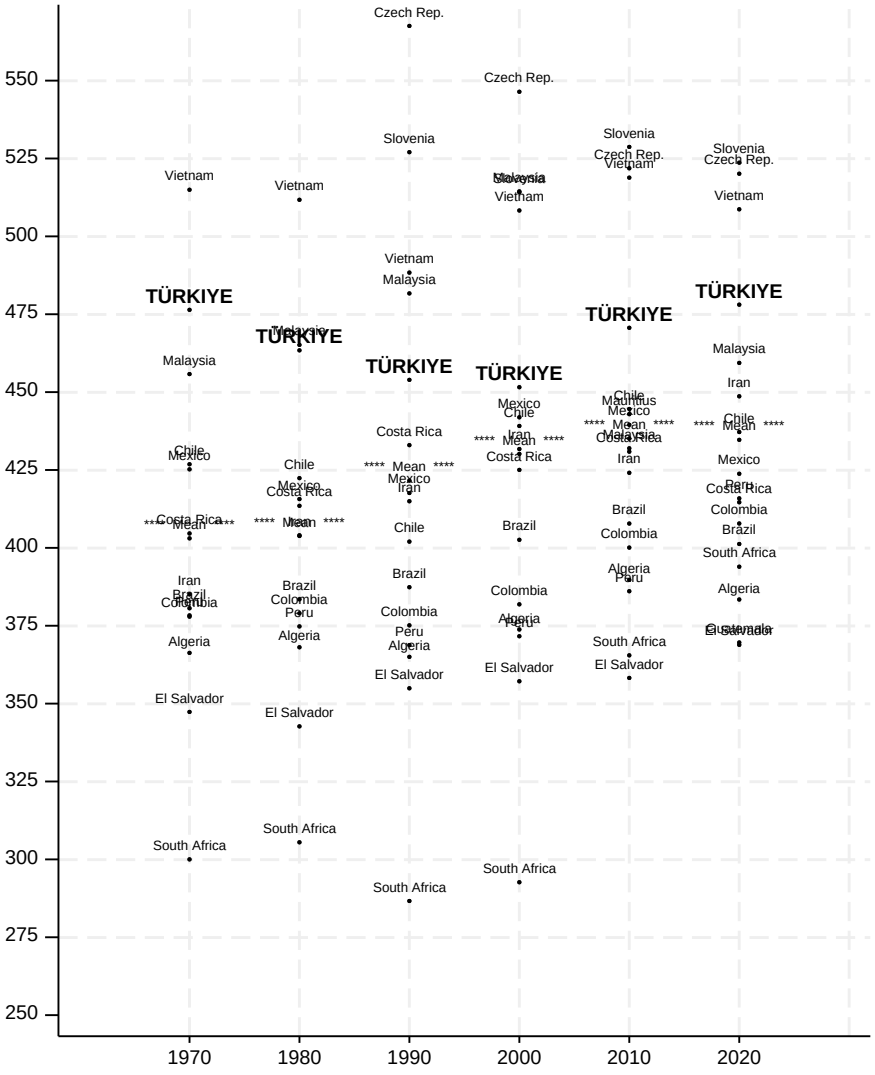
Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.6: Trends on school performance in mathematics, primary education



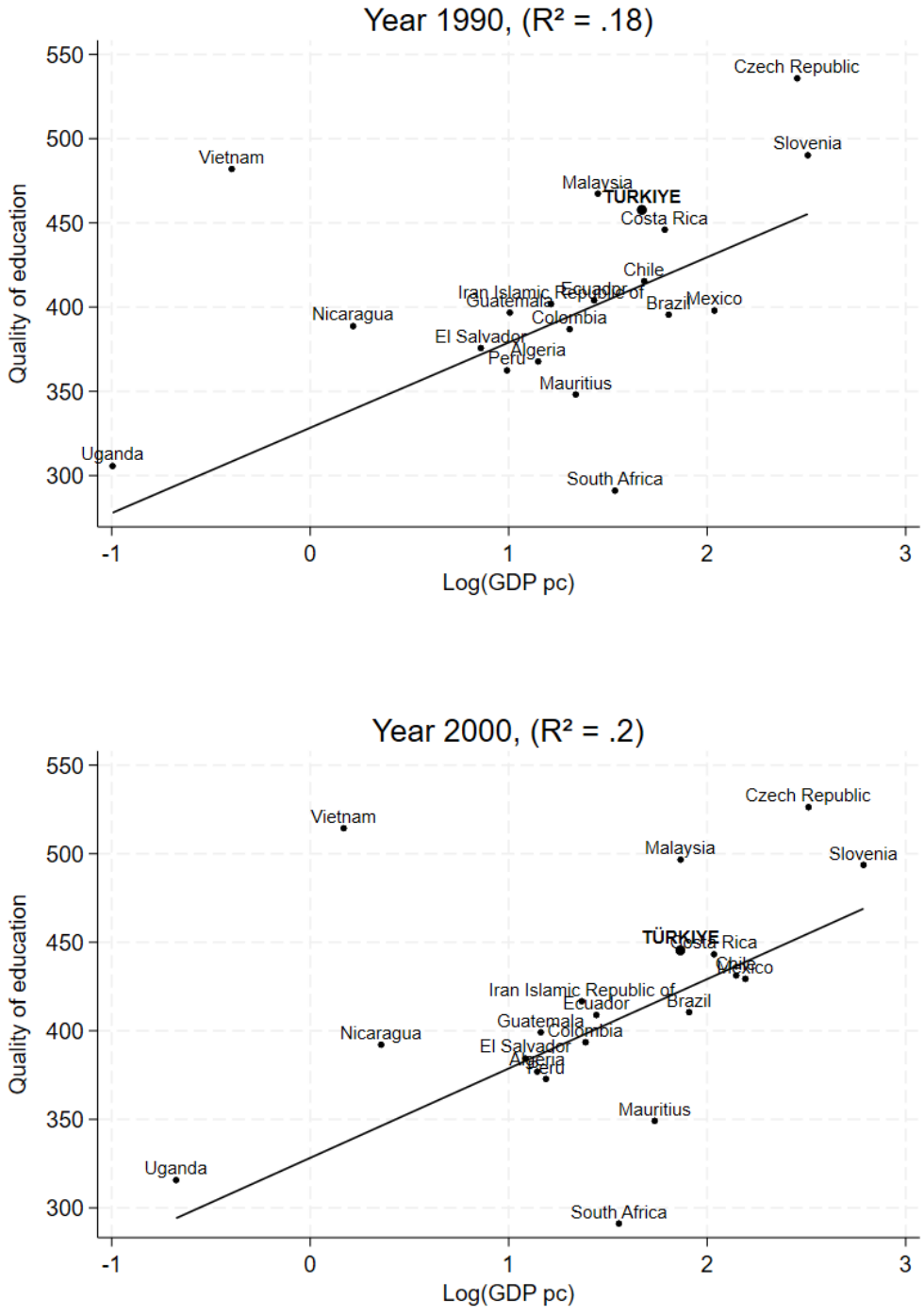
Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.7: Trends on school performance in mathematics, secondary education



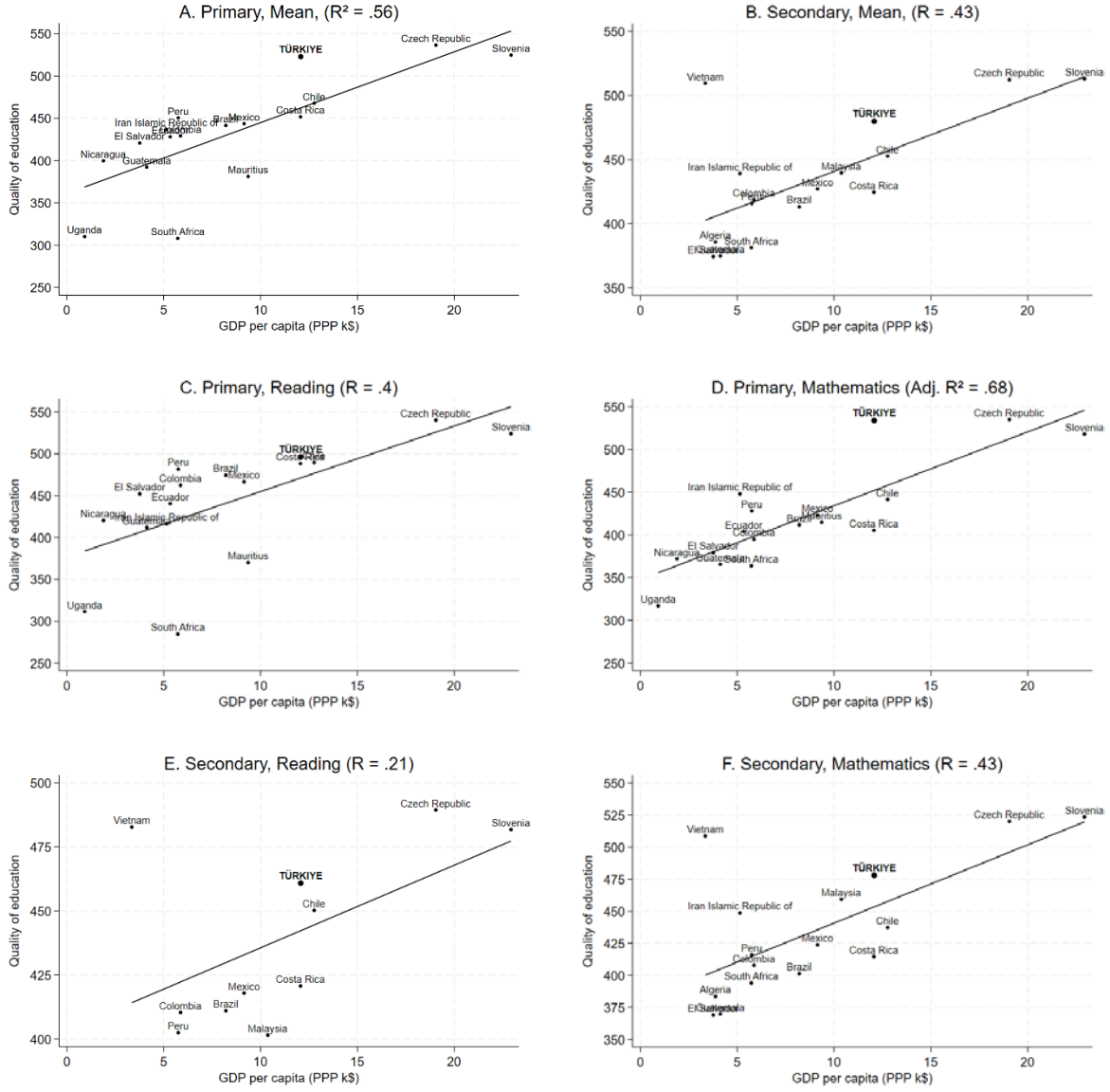
Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.8: Relationship between GDP per capita and quality of education (1990 & 2000)



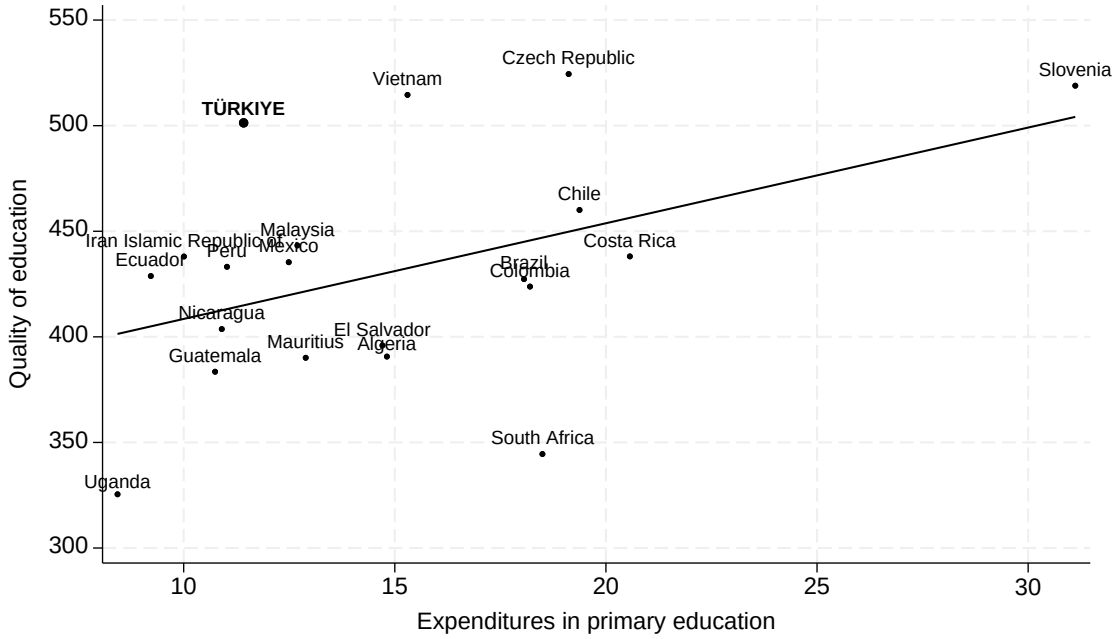
Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.9: Relationship between performance in each skill and level and GDP per capita



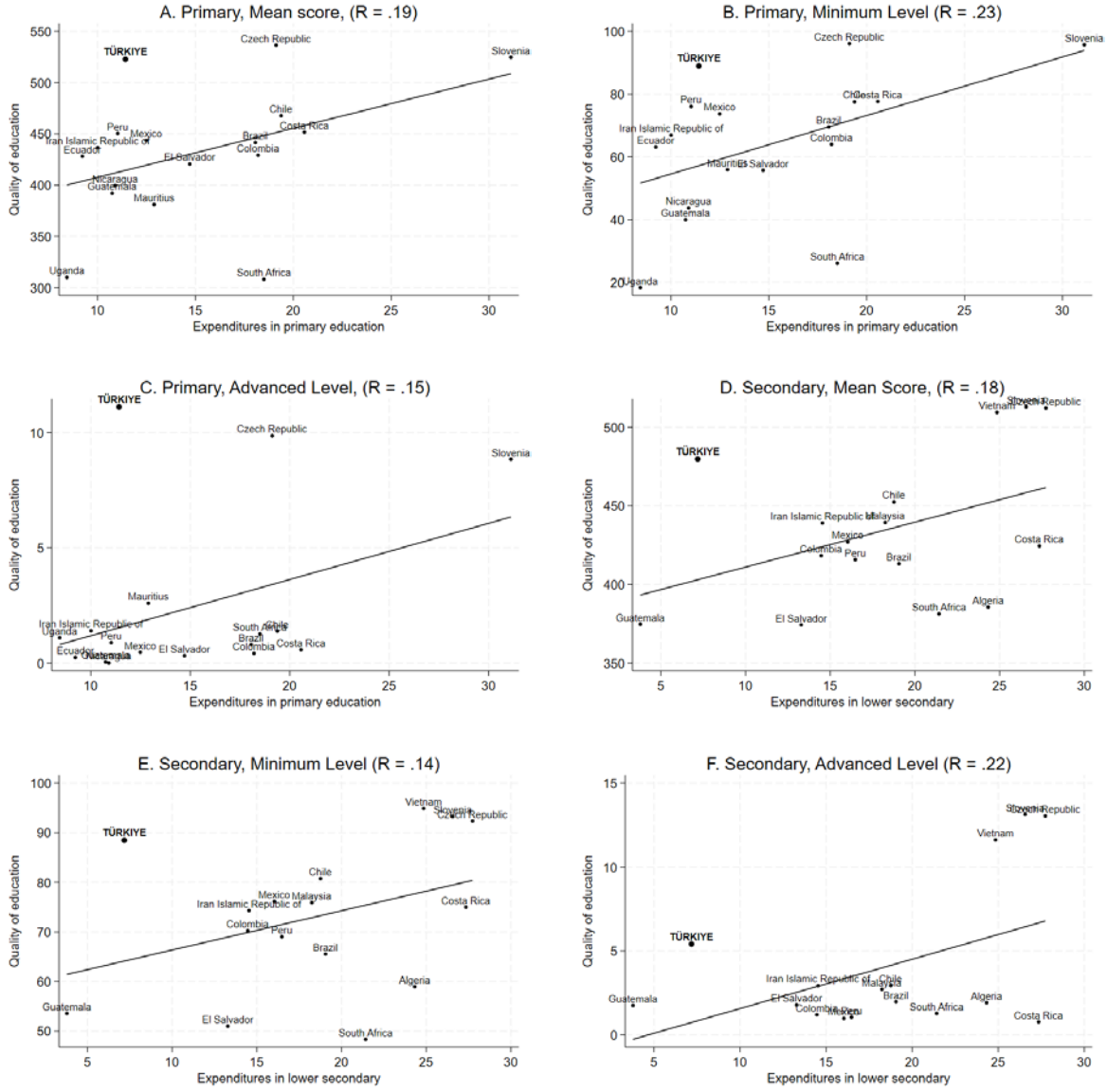
Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.10: Expenditures on education and quality of education in 2020 ($Adj\ R^2 = 0{,}15$)



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Figure A.11: Expenditures in education and quality of education (2020)



Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Table A.1: Schooling level performance and expenditures in education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log(GDP pc)	42.448***	7.003	23.141***	14.922**	11.415	17.604***	64.504***	29.115***	46.769***
	(3.563)	(4.660)	(7.589)	(7.052)	(7.164)	(4.282)	(5.041)	(8.712)	(5.596)
Expenditures on ed	-25.859***	63.702***	-16.151*	-3.345	25.981***	-0.968	-1.294	88.833***	22.285***
	(5.183)	(8.378)	(8.687)	(7.443)	(6.566)	(4.682)	(6.596)	(14.146)	(6.126)
Intercept	316.453***	370.781***	418.567***	405.307***	398.803***	400.304***	179.644***	84.066***	200.381***
	(20.598)	(18.595)	(30.607)	(26.601)	(24.282)	(16.392)	(30.386)	(31.762)	(25.118)
Level	Primary	Secondary	Both	Both	Both	Both	Primary (adj.)	Secondary (adj.)	Both (adj.)
Skill	All	All	Reading	Maths	Sciences	All	All	All	All
Adjusted	No	No	No	No	No	No	Yes	Yes	Yes
R-squared	0.593	0.488	0.382	0.498	0.452	0.420	0.709	0.632	0.643
Observations	502	433	293	337	305	935	502	427	929
Nb. countries	17	17	20	19	20	20	17	17	20
1970-2020									
Ranking	3	5	4	4	5	5	1	2	2
Residuals	-20.5	-40.4	-10.6	-33.3	-18.8	-22.1	64.8	129.6	116.1
Best ranked	Czech Rep.	Czech Rep.	Brazil	Czech Rep.	Vietnam	Czech Rep.	Turkey	Vietnam	Vietnam
2000-2020									
Ranking	1	3	2	3	2	2	1	3	2
Residuals	9.2	-2.9	24.5	2.8	12.2	11.8	49.5	88.2	87.1
Best ranked	Turkey	Vietnam	Vietnam	Vietnam	Vietnam	Vietnam	Turkey	Vietnam	Vietnam

Source: authors' calculations based on (Altinok & Diebolt, 2024, 2025).

Appendix B

This appendix is adapted from Hanushek & Woessmann (2011) and Hanushek et al. (2017a). For more information, see Hanushek & Woessmann (2011).

The projection model is based on four different phases. First, we introduce the reform (Phase A), where the effect of the quality of education is passed on over a 10-year period. In the next phase (Phase B), older workers are replaced by reform workers (i.e. young people affected by the reform and having graduated enter the labor market). This means that the reform only becomes fully effective over a period of 30 years. Third, workers who were only partially affected by the reform are replaced (Phase C). Finally, all workers affected by the reform are replaced, and the effect of the reform comes to an end (Phase D).

Phase A (2020-2030): Introduction of reform

During the first 10 years of the educational reform, the additional growth in GDP per capita due to the reform for year *t* is given by :

$$\Delta_{growth}^{t} = growth\ coeff. \times \Delta Scores \times \frac{1}{Working\ life} \times \frac{t-2020}{10} + \Delta_{growth}^{t-1} \qquad \text{(B.1)}$$

where the *growth coefficient* represents the effect of the one standard-deviation increase in test scores obtained in the growth models of Hanushek & Woessmann (2012) and $\Delta\, Scores$ is the change in educational performance due to the reform analyzed. Each year, only part of the workforce is replaced by new workers who have received a better education (i.e. the one related to the reform). We take this lag into account with $\frac{1}{Working\ life}$, with a working life of 40 years. The $\frac{t-2020}{10}$ term shows that it takes 10 years for the reform to be fully effective.[27]

Phase B (2031-2060): Replacement of older workers by new ones who have benefited from reform

The educational reform is now fully effective, and the academic quality of all students remains stable at this new level. However, given that the working life is assumed to be 40 years, there are still workers with previous levels of schooling who will be replaced by new

[27]In the first version of their projection model Hanushek & Woessmann (2011) assumed that the reform would only be fully effective after 20 years; we have used the latest version of their model, in which this period is reduced to 10 years. However, the results change little if this period is reduced to 20 years.

workers with a higher level of quality. Thus, over the next 30 years, additional growth should follow the following path:

$$\Delta^{t}_{growth} = growth\ coeff. \times \Delta Scores \times \frac{1}{Working\ life} + \Delta^{t-1}_{growth} \quad \text{(B.2)}$$

Phase C (2061-2070): Replacement of workers who have received part of the reform with a higher-quality product

After 40 years, all workers who have not experienced educational reform have been replaced by new workers. At the same time, the cohorts of the first 10 years were only partially affected by the educational reform. During this phase, these workers are replaced by cohorts who have received all the educational reform. The additional growth for the following 10 years is thus:

$$\Delta^{t}_{growth} = growth\ coeff. \times \Delta Scores \times \frac{1}{Working\ life} - \left(\Delta^{t-40}_{growth} - \Delta^{t-41}_{growth}\right) + \Delta^{t-1}_{growth} \quad \text{(B.3)}$$

Phase D (after 2070): Educational reform for all workers

The entire workforce has been exposed to the reform via the renovated education system. Additional economic growth is now at a constant long-term level:

$$\Delta^{t}_{growth} = growth\ coeff. \times \Delta Scores \quad \text{(B.4)}$$

The cumulative effect of the reform can be assessed by calculating various indicators. The total value of the reform can be obtained by calculating the present difference between GDP with and without the reform. We therefore calculate the gain from improving the education system over a period of 80 years, using the discount rate (*d*) which is similar to that used in Hanushek et al. (2017a) and set at 3%:

$$Value\ of\ reform = \sum_{t=2020}^{t=2100}(GDP^{t}_{reform} - GDP^{t}_{without\ reform}) \times (1+d)^{-(t-2020)} \quad \text{(B.5)}$$

Expressed as a percentage of current GDP, the total value of the reform is obtained using the following equation:

$$Value\ of\ reform_{\%\ of\ current\ GDP} = \frac{Value\ of\ reform}{GDP^{2020}_{without\ reform}} \times 100 \quad \text{(B.6)}$$

Other indicators can also be obtained, such as the ratio between the total value of the reform and the discounted GDP obtained over the entire duration of the reform (equation B.7).

$$Value\ of\ reform_{of\ discounted\ future\ GDP} = \frac{Value\ of\ reform}{\sum_{t=2020}^{t=2100} GDP^{t}_{without\ reform} \times (1+d)^{-(t-2020)}} \times 100 \quad \text{(B.7)}$$

Furthermore, we can also calculate how much additional GDP growth the reform could generate for any year, such as 2090, using equation B.8:

$$GDP_{Increase\ in\ 2090\ (\%)} = \frac{GDP_{reform}^{2090} - GDP_{without\ reform}^{2090}}{GDP_{without\ reform}^{2090}} \times 100 \quad \text{(B.8)}$$

References

Alacacı, C., & Erbaş, A. K. (2010). Unpacking the inequality among Turkish schools: Findings from PISA 2006. *International Journal of Educational Development*, *30*(2), 182–192.

Altinok, N., & Diebolt, C. (2024). Cliometrics of Learning-Adjusted Years of Schooling: Evidence from a New Dataset. *Cliometrica*, *18(3)*, 691-764. https://doi.org/https://doi.org/10.1007/s11698-023-00276-x

Altinok, N., & Diebolt, C. (2025). Mapping Educational Inequality Worldwide: A New Dataset (1970-2020). Mimeo, January 2025.

Angrist, N., Djankov, S., Goldberg, P. K., & Patrinos, H. A. (2021). Measuring human capital using global learning data. *Nature*, *592*(7854), 403–408.

Angrist, N., Patrinos, H. A., & Schlotter, M. (2013). An Expansion of a Global Data Set on Educational Quality: A Focus on Achievement in Developing Countries. *The World Bank: Policy Research Working Papers*. https://doi.org/10.1596/1813-9450-6536

Avvisati, F., & Givord, P. (2023). The learning gain over one school year among 15-year-olds: An international comparison based on PISA. *Labour Economics*, *84*, 102365.

Aydın, A., Sarıer, Y., & Uysal, Ş. (2012). Sosyoekonomik ve sosyokültürel değişkenler açısından PISA matematik sonuçlarının karşılaştırılması. *Eğitim ve Bilim*, *37*(164).

Aydın, T., & Çilek, A. (2024). PISA 2022 Sınav Sonuçları Bağlamında Türk Eğitim Sisteminin Genel Görünümü. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *24*(4), 2361–2378.

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.

Boman, B. (2022). Vietnam's exceptional educational achievement: a thematic review of the emerging literature. *Discover Education*, *1*(1), 14. https://doi.org/10.1007/s44217-022-00014-x

Borsch-Supan, A. (2000). A Model Under Siege: A Case Study of the German Retirement Insurance System. *The Economic Journal*, *110*(461), F24–F45. https://doi.org/10.1111/1468-0297.00506

Çelik, Z., & Gür, B. S. (2013). Turkey's education policy during the AK party era (2002-2013). *Insight Turkey*, *15*(4), 151.

Coleman, J. S. (1966). *Equality of Educational Opportunity* (Vol. 1). US Department of Health, Education, and Welfare, Office of Education.

Dinçer, M. A., & Oral, I. (2013). Türkiye'de devlet liselerinde akademik dirençlilik profili. *İstanbul: Eğitim Reformu Girişimi Yayınları*.

Dincer, M. A., & Uysal, G. (2010). The determinants of student achievement in Turkey. *International Journal of Educational Development*, *30*(6), 592–598.

Dolu, A. (2020). SOSYOEKONOMİK FAKTÖRLERİN EĞİTİM PERFORMANSI ÜZERİNE ETKİSİ: PISA 2015 TÜRKİYE ÖRNEĞİ. *Journal of Management and Economics Research*, *18*(2), 41–58.

Filmer, D., Rogers, H., Angrist, N., & Sabarwal, S. (2020). Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education. *Economics of Education Review*, *77*, 101971.

Foshay, A. W., Thorndike, R. L., Hotyat, F., Pidgeon, D. A., & Walker, D. A. (1962). *Educational Achievements of Thirteen-year Olds in Twelve Countries: Results of an International Research Project, 1959-61* (Vol. 4). Unesco.

Gust, S., Hanushek, E. A., & Woessmann, L. (2024). Global universal basic skills: Current deficits and implications for world development. *Journal of Development Economics*, *166*, 103–205.

Hanushek, E. A. (2019). *Addressing Cross-National Generalizability in Educational Impact Evaluation*. National Bureau of Economic Research.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *The American Economic Review*, *90*(5), 1184.

Hanushek, E. A., Ruhose, J., & Woessmann, L. (2017a). Economic Gains from Educational Reform by US States. *Journal of Human Capital*, *11*(4), 447–486.

Hanushek, E. A., Ruhose, J., & Woessmann, L. (2017b). Knowledge Capital and Aggregate Income Differences: Development Accounting for US States. *American Economic Journal: Macroeconomics*, *9*(4), 184–224.

Hanushek, E. A., & Woessmann, L. (2011). How much do educational outcomes matter in OECD countries? *Economic Policy*, *26*(67), 427–491.

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, *17*(4), 267–321. https://doi.org/10.1007/s10887-012-9081-x

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. MIT Press.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(1), 1–47.

Koca, H., Yakar, A., Dev, F., & Şen, G. (2024). TIMSS ve PISA verileri üzerinden Türkiye'nin eğitim performansının gelişiminin analizi. *Ulusal Eğitim Dergisi*, *4*(2), 644–660.

Köseleci, N. (2015). Progress towards good-quality education for all in Turkey: a qualified success? *Comparative Education*, *51*(4), 555–574.

Lim, S. S., Updike, R. L., Kaldjian, A. S., Barber, R. M., Cowling, K., York, H., Friedman, J., Xu, R., Whisnant, J. L., & Taylor, H. J. (2018). Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016. *The Lancet*, *392*(10154), 1217–1234.

MEB. (2018). *Turkey's Education Vision 2023*. https://planipolis.iiep.unesco.org/sites/default/files/ressources/turkey_education_vision_2023.pdf

Mullis, I., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). TIMSS 2019 international results in mathematics and science. In *Retrieved from Boston College, TIMSS & PIRLS International Study Center website: https://timssandpirls. bc. edu/timss2019/international-results*. https://timss2019.org/reports/wp-content/themes/timssandpirls/download-center/TIMSS-2019-International-Results-in-Mathematics-and-Science.pdf

Mullis, I., von Davier, M., Foy, P., Fishbein, B., Reynolds, K., & Wry, E. (2023). *PIRLS 2021 International Results in Reading*. https://doi.org/10.6017/lse.tpisc.tr2103.kb5342

OECD. (2021). *OECD Economic Surveys: Turkey 2021*. OECD Publishing. https://doi.org/doi:https://doi.org/10.1787/2cd09ab1-en

OECD. (2023a). *OECD Economic Surveys: Türkiye 2023*. OCDE. https://doi.org/https://doi.org/10.1787/864ab2ba-en

OECD. (2023b). *PISA 2022 Results (Volume I)*. OECD. https://doi.org/doi:https://doi.org/10.1787/53f23881-en

Özdemir, C. (2015). *Relationship between equity and excellence in education: multilevel analysis of international student assessment data with a focus on turkey*.

Paccagnella, M. (2016). Literacy and Numeracy Proficiency in IALS, ALL and PIAAC. *OECD Education Working Papers*, *142*. https://doi.org/doi:https://doi.org/10.1787/5jlpq7qglx5g-en

Parandekar, S., & Sedmik, E. (2016). *Unraveling a secret: Vietnam's outstanding performance on the PISA test*.

Piketty, T. (2013). Le capital au XXIe siècle. In *Les Livres du nouveau monde*. Seuil.

Polat, M., Karabiyik, T., & Kartal, G. (2024). *TIMSS 2023: National Report Türkiye*.

Sarıer, Y. (2020). TIMSS uygulamalarında Türkiye'nin performansı ve akademik başarıyı yordayan değişkenler. *Temel Eğitim*, *2*(2), 6–27.

Sarıer, Y. (2021). PISA UYGULAMALARINDA TÜRKİYE'NİN PERFORMANSI VE ÖĞRENCİ BAŞARISINI YORDAYAN DEĞİŞKENLER. *Türkiye Sosyal Araştırmalar Dergisi*, *25*(3), 905–926.

Schwerdt, G., & Wiederhold, S. (2019). A Macroeconomic Analysis of Literacy and Economic Performance. *Mimeo*. https://doi.org/http://www.dataangel.ca/docs/A%20Macroeconomic%20Analysis%20of%20Literacy_February2019.pdf

Spaull, N. (2018). Who makes it into PISA? Understanding the impact of PISA sample eligibility using Turkey as a case study (PISA 2003–PISA 2012). *Assessment in Education: Principles, Policy & Practice*, 1–25.

Suna, H. E., & Özer, M. (2021). Türkiye'de sosyoekonomik düzey ve okullar arası başarı farklarının akademik başarı ile ilişkisi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, *12*(1), 54–70.

UNESCO Institute of Statistics. (2019). Costs and Benefits of Different Approaches to Measuring the Learning Proficiency of Students. *UIS Information Paper*, *53*. http://uis.unesco.org/sites/default/files/documents/ip53-costs-benefits-approaches-measuring-proficiency-2019-en.pdf

Uysal, N. K., & Gelbal, S. (2018). Inequality in Education in Turkey in the Scope of PISA and TSI. *Kastamonu Education Journal*, *26*(4), 1187.

Waldow, F. (2009). What PISA Did and Did Not Do: Germany after the 'PISA-shock.' *European Educational Research Journal*, *8*(3), 476–483. https://doi.org/10.2304/eerj.2009.8.3.476

World Bank, T. (2011). *Vietnam: High Quality Education for All by 2020*. World Bank. https://openknowledge.worldbank.org/handle/10986/27450

Wu, M. (2010). Comparing the Similarities and Differences of PISA 2003 and TIMSS. *OECD Education Working Papers*, *32*, 0_1.

Yediyıldız, B., & Ustun, U. (2024). Deciphering socioeconomic inequality in science literacy: Evidence from PISA 2015 in Türkiye. *International Journal of Educational Development*, *111*, 103173.

Yilmaz, G. (2015). Turkish middle income trap and less skilled human capital. *Iktisat Isletme ve Finans*, *30*(346), 9-36.

Résumé long

Lors de la dernière enquête TIMSS 2023 (*Trends in International Mathematics and Science Study*), la Turquie s'est distinguée par une performance remarquable, se classant au 8e rang sur 58 participants avec un score moyen de 553, soit bien au-dessus de la moyenne internationale (Koca et al., 2024). Toutefois, les résultats de l'enquête PISA (*Programme for International Student Assessment*) de l'OCDE ne confirment pas pleinement cette performance. Cette apparente contradiction met en évidence le rôle déterminant de la qualité des données utilisées pour évaluer l'efficacité d'un système éducatif.

Dans cette recherche, nous exploitons un ensemble de données couvrant la qualité des systèmes éducatifs d'un grand nombre de pays. Nous mobilisons à la fois les scores moyens obtenus lors des évaluations internationales, comme PISA, et des mesures complémentaires basées sur des indicateurs plus spécifiques. Ces derniers s'articulent autour de deux dimensions essentielles :

1. La qualité universelle de l'éducation, qui mesure la capacité d'un élève à acquérir des compétences fondamentales en lecture, mathématiques et sciences.
2. Les compétences avancées, indispensables à l'innovation et au développement économique.

En combinant ces deux dimensions, nous proposons un indicateur original de l'efficacité des systèmes éducatifs. Au-delà de l'efficacité, nous intégrons également deux indicateurs complémentaires permettant d'évaluer l'équité des systèmes éducatifs. L'équité est ici définie comme l'écart relatif entre les déciles de performance aux tests internationaux, ajusté par le taux de scolarisation afin de tenir compte de la sélectivité des systèmes éducatifs.

En exploitant les données issues de deux études distinctes, nous testons dans quelle mesure la performance du système éducatif turc se compare à celle d'économies de niveau similaire. Ces données couvrent la plupart des pays de l'OCDE ainsi qu'un large échantillon de pays en développement. L'approche adoptée repose sur une analyse comparative et une méthodologie contrefactuelle. En particulier, la méthode des doubles différences permet d'évaluer si l'évolution de la performance éducative en Turquie se distingue significativement de celle des pays affichant un développement comparable.

Nos résultats confirment en partie la surperformance de la Turquie dans les évaluations des acquis des élèves, notamment depuis le début des années 2000. Si la plupart des pays du groupe témoin ont enregistré des progrès en matière de performance éducative, la Turquie affiche une amélioration nettement plus marquée. Cette tendance est corroborée par nos indicateurs alternatifs de performance : en 2020, 88 % des élèves turcs atteignent le seuil minimal de compétence au secondaire, contre seulement 73 % dans les pays témoins.

L’analyse en doubles différences entre 2000 et 2020 confirme cet écart positif, qui s’établit à 3,6 %.

Cependant, malgré une efficacité éducative relativement élevée, des faiblesses persistent en matière d’équité, c’est-à-dire dans la capacité du système éducatif à offrir des chances de réussite équivalentes à tous les élèves. Notre indicateur d’équité mesure l’écart relatif entre les 10 % des élèves les plus performants et les 10 % les moins performants : plus cet écart est important, plus le système est inégalitaire. Nos résultats montrent que cet indicateur d’équité a diminué au fil des décennies en Turquie, passant de 589 à 540 points, soit une baisse moyenne de 10 points tous les 10 ans. En revanche, l’équité ajustée du taux de scolarisation s’est nettement améliorée, en particulier dans l’enseignement secondaire. Ces évolutions témoignent des efforts de la Turquie en faveur de la scolarisation de masse : le taux d’achèvement du secondaire est ainsi passé de 60 % à 90 % entre 2000 et 2023, notamment sous l’effet des réformes de 2012 (réforme *4+4+4*), qui ont porté la durée de la scolarité obligatoire à 12 ans.

Toutefois, notre analyse contrefactuelle met en évidence que les pays comparables à la Turquie ont enregistré de meilleurs résultats en matière d’équité. Cette observation soulève la question de la répartition de la qualité du capital humain au sein de la population turque. En outre, de fortes inégalités éducatives subsistent entre les provinces (Uysal & Gelbal, 2018) et entre les établissements scolaires (Polat et al., 2024).

Enfin, nous réalisons un exercice de prospective économique afin d’évaluer les gains potentiels d’une réforme éducative axée sur l’amélioration de la qualité de l’éducation. En mobilisant un modèle de prévision (Gust et al., 2024), nous montrons qu’une politique combinant amélioration de la qualité et renforcement de l’équité générerait des bénéfices économiques considérables à l’horizon 2100. En effet, cette approche permettrait une hausse de 111 % du PIB contre 68,1 % pour une réforme ne visant que la qualité du système éducatif. Ces résultats soulignent qu’au-delà de la recherche de meilleurs scores aux évaluations internationales, la Turquie devrait, à l’instar du programme *Vision 2023* (MEB, 2018), faire de l’équité un objectif stratégique majeur, en mettant en place des politiques éducatives garantissant un accès équitable à une éducation de qualité sur l’ensemble de son territoire.

Extended Summary

In the latest TIMSS 2023 survey (*Trends in International Mathematics and Science Study*), Türkiye achieved remarkable results, ranking 8th out of 58 participants with an average score of 553, well above the international average (Koca et al., 2024). However, the results from the OECD’s PISA survey (*Programme for International Student Assessment*) do not fully confirm this strong performance. This apparent contradiction highlights the crucial role of data quality in assessing the effectiveness of an educational system.

In this study, we utilize a comprehensive dataset on the quality of education systems across a large number of countries. We incorporate not only the average scores obtained in international assessments such as PISA but also additional measures based on two key dimensions:

1. The universal quality of education, which assesses students' ability to acquire fundamental skills in reading, mathematics, and science.
2. Advanced competencies, which are essential for fostering innovation and economic development.

By combining these two dimensions, we propose an original indicator to measure the efficiency of education systems. Beyond efficiency, we also introduce two complementary indicators to assess equity in education. Equity is defined here as the relative gap between the highest and lowest deciles of performance in international tests, adjusted by school enrollment rates to account for the selectivity of education systems.

Using data from two distinct studies, we evaluate whether Türkiye's educational performance is superior to that of economies with similar development levels. These data cover most OECD countries as well as a large number of developing nations. Our approach relies on comparative analysis and counterfactual methodology. In particular, the difference-in-differences method allows us to determine whether the evolution of Türkiye's educational performance significantly differs from that of countries with similar levels of economic development.

Our findings partially confirm the exceptional improvement in Türkiye's student achievement scores, particularly since the early 2000s. While most peer countries have experienced progress in educational performance indicators, Türkiye has exhibited a significantly stronger upward trend. This is corroborated by alternative performance thresholds: in 2020, 88% of Turkish students met the minimum proficiency level in secondary education, compared to only 73% in the peer group. The difference-in-differences analysis between 2000 and 2020 confirms this positive trend, with Türkiye showing a relative gain of 3.6%.

However, despite its relatively strong educational efficiency, Türkiye still faces challenges in terms of equity—i.e., its ability to provide equal opportunities for all students. Our equity indicator measures the relative gap between the top 10% and bottom 10% of students: the larger the gap, the less equitable the system. Our results show that this equity index has declined over the past decades in Türkiye, decreasing from 589 to 540 points, an average drop of 10 points per decade. Conversely, enrollment-adjusted equity has significantly improved, particularly in secondary education. These trends reflect Türkiye's efforts to expand mass education, with the secondary school completion rate rising from 60% to 90% between 2000

and 2023. This progress was largely driven by the 2012 education reform (*4+4+4*), which extended compulsory schooling to 12 years.

Nevertheless, our counterfactual analysis reveals that countries comparable to Türkiye have achieved better results in terms of educational equity. This raises concerns about the distribution of human capital quality across the Turkish population. Furthermore, significant educational disparities persist across Turkish provinces (Uysal & Gelbal, 2018) and between schools (Polat et al., 2024).

Finally, we conduct an economic foresight exercise to assess the potential long-term gains of an education reform focused on improving quality. Using a forecasting model (Gust et al., 2024), we demonstrate that a policy combining both quality improvement and enhanced equity would generate substantial economic benefits by 2100. Specifically, such a mixed policy could increase GDP by 111%, compared to a 68.1% increase for a policy focusing solely on quality enhancement. These findings suggest that beyond improving international test scores, Türkiye should, in line with its *Vision 2023* program (MEB, 2018), prioritize policies aimed at fostering greater equity in education across its regions and schools.