



ARTICLE



<https://doi.org/10.1057/s41599-025-06080-4>

OPEN

The trade-off between enrolment expansion, school resources, and learning outcomes in Punjab's public schools

Saba Arshad^{1,2✉} & Elizabeth Washbrook¹

Driven by global commitments like the Sustainable Development Goals (SDGs), developing countries have rapidly expanded primary and secondary school enrolment. Yet, this focus on access has often come at the expense of learning. In Punjab, education reforms since 2004 have significantly increased enrolment in public schools (grades 1-8). However, the potential trade-off between enrolment expansion and learning outcomes has yet to be systematically studied. This study assesses the link between enrolment expansion and learning outcomes and examines how school resources mediate this relationship. It uses a school-level dataset on enrolment, test scores, and school resources in primary and middle public schools in Punjab, with longitudinal data from 2008 to 2018. The analysis follows a three-step approach: first, pooled OLS regression is used to examine the relationship between enrolment expansion and learning outcomes; second, a change-in-change analysis assesses how changes in enrolment over time affect changes in learning outcomes; and third, a mediation analysis explores whether school resources can account for the relationship between enrolment expansion and learning outcomes. OLS estimates show a weak positive link between enrolment and pass rates, suggesting larger schools may perform slightly better. However, change-in-change analysis reveals that faster enrolment growth is associated with declines in pass rates—by 3.6 and 5.8% points at the primary and middle levels. Negative effects on pass rates were not replicated for average test scores, indicating that those most affected were low-achieving students near the pass/fail borderline. These trends are consistent across boys' and girls' schools. Mediator analysis shows school resources do not explain the trade-off. While some negative effects on pass rates exist, the limited impact on test scores overall suggests Punjab has managed to expand access without compromising learning for the majority of students, although learning of the lowest achieving students may have been negatively impacted.

¹School of education, University of Bristol, Bristol, UK. ²Nuffield Department of Primary Health Care, University of Oxford, Oxford, UK. ✉email: prhc0953@ox.ac.uk

Introduction

It is now well-recognised that education plays an important role in a country's economic and social development. In recognition of this, widening access to primary and secondary education has been the goal of many developing countries (Glewwe and Muralidharan, 2016). Universal and equitable access to education has been promoted by various international frameworks, such as Education for All (EFA), the Millennium Development Goals (MDGs), and the Sustainable Development Goals (SDGs), which recognise education as a fundamental human right (Sifuna, 2007). For example, Goal 4 of the SDGs (2016–2030) commits countries to improving access to quality education for all, emphasising inclusivity and lifelong learning opportunities.

As a result of these international commitments, the world has witnessed a rapid expansion of primary and secondary education in developing countries. However, this expansion has often placed significant strain on the financial and human resources of these countries in certain contexts (Taylor and Spaul, 2013; Valente, 2019). Crouch and Vinjevold (2006) argued that many developing countries have over-emphasised access at the expense of learning outcomes. Recent evidence from India also shows that policies that promote school enrolment may negatively affect test scores because students in these schools are not learning as much as their counterparts in more developed countries (Shah and Steinberg, 2019). Thus, this imbalance in policy preferences creates tension between expansion in enrolment and learning outcomes.

The key contribution of this study is to provide the first comprehensive correlational analysis of the association between enrolment expansion and learning outcomes in Punjab. It employs a school-level dataset on enrolment, test scores and school resources in primary and middle public schools in Punjab, from 2008 to 2018. This is the first large-scale study to use province-wide school-level examination data to measure learning outcomes. Even though the province of Punjab has conducted annual examinations of students in grades 5 and 8 (these are terminal grades for primary and middle schools, respectively) since 2008, these test scores have rarely been used in research on learning outcomes in Punjab (Barrera-Osorio and Ganimian, 2016).

There are at least three reasons why rapid expansion in enrolment might influence observed test scores. First, because of expansion, the composition of the student population changes, and the level of prior achievement, on average, drops, a phenomenon often known as the compositional effect (Bold et al., 2014). Rapid enrolment expansion may have instigated the self-selection of lower-ability students or enabled less-prepared children to attend school (Zuze and Leibbrandt, 2011). A second possible reason is peer effects, whereby the changes in student composition may also lead to adverse peer effects, which exert downward pressure on average test scores (Bold et al., 2014). The third reason is rooted in how rapid expansion in enrolments may affect the distribution of school resources. For example, when the number of students in school increases, the student-teacher and student-classroom ratios may also rise if inputs do not increase simultaneously. This could lead to children getting less individual attention and lowering their test scores (Chimombo, 2005).

The compositional and peer effects are related to student's background characteristics that are not within the education system's immediate control. The distribution of school resources is the only mechanism influencing changes in the school's test scores that education service providers can control (Lucas and Mibiti, 2014). It is not inevitable that expansion will be linked with lower test scores, particularly given the supply-side investments made under Punjab's Public Education Sector Reform

Programme (PESRP). Since 2004, under the PESRP, the Government of Punjab has substantially improved the basic facilities in public schools. For example, the changes in basic school facilities (access to drinking water, boundary walls, toilets, and electricity) are remarkable, as virtually 99% of public schools had these facilities by 2012. Hence, this study conducted a mediation analysis, using school inputs as a mediator variable in the association between enrolment expansion and learning outcomes (PESRP, 2021). This mediation analysis enables the researcher to explore which school inputs, when controlled, strengthen or weaken the measured association between enrolment and learning outcomes. Specifically, it considers the role of changes in class size, changes in student-teacher ratios, changes in teachers' qualifications, and changes in basic school facilities (drinking water, boundary walls, toilets, and electricity).

It is pertinent to mention that it is beyond the scope of the school-level data used in this study to explain the overall association between enrolment expansion and learning outcomes, as this would require access to complete student-level data with household information. While this study does not claim to provide causal estimates, the availability of school-level panel data like the Punjab Annual Public-School Census (PAPSC) and Punjab Examination Commission (PEC) enables us to provide informative, descriptive analyses. Furthermore, this study will also add a methodological contribution to the literature by using two different regression models, a simple pooled OLS model and the change-in-change estimator, to measure the trade-off between expansion and learning outcomes. As the name suggests, in a pooled OLS model, multiple cross-sections are pooled or combined for the regression model. The pooled OLS model results compare school variations without distinguishing between- and within-school differences (Wooldridge, 2012). In comparison, the change-in-change model exploits within-school variation while controlling for between-school differences. The change-in-change estimates effectively compare the same school at different points instead of comparing different schools to each other (Lucas and Mbiti, 2014).

The remainder of the study is organised as follows. Next section provides background on Punjab's educational reform programme. Subsequent section reviews the literature on the effect of rapid enrolment expansion on school resources and learning outcomes and presents research questions. The subsequent section summarises the modelling framework and describes the datasets. This is followed by a presentation of the empirical results. The final section concludes the study.

Background and context of the study

To meet the targets of access, quality, and gender inequality in education set by international obligations (EFA, MDGs, and SDGs), many developing countries (such as India, Bangladesh, Kenya, Tanzania, Malawi, and Uganda) have launched education sector reform programmes (World Bank, 2012). Pakistan is no exception; Punjab, its largest province constituting 53% of the total population, has also started a large-scale education reform programme to improve access, learning outcomes and reduce gender inequality in education (World Bank, 2015). Since 2004, the Government of Punjab has implemented an extensive education reform programme known as Punjab Education Sector Reform Programme (PESRP), with the financial and technical assistance of development partners (the World Bank and UK's Department for International Development/FCDO). The desired strategic objectives of PESRP are to improve access, learning outcomes and gender equality in education (World Bank, 2017).

The PESRP consists of multiple phases (three phases so far), each targeting specific areas of improvement within the education sector. The first phase of PESRP (PESRP-I), which spanned the years from 2004 to 2011, primarily focused on foundational reforms in the public schools of Punjab. Critical interventions introduced under PESRP-I included abolition of tuition fees, distribution of free textbooks, recruitment of more teachers, building new classrooms, and provision of basic facilities, including electricity, boundary walls, toilets, and drinking water (World Bank, 2012). Furthermore, a gender-targeted conditional cash transfer (CCT) programme was introduced, known as the Female School Stipend Programme (FSSP) to reduce gender inequality in education in 16 historically disadvantaged districts of Punjab. Additionally, a robust monitoring system was instituted to gather school-level administrative data pertinent to monitoring the implementation of PESRP initiatives and the performance of public schools, alongside standardised testing mechanisms to track student learning outcomes (Barber, 2012).

Building on the efforts of phase I, the Government of Punjab introduced its second phase, PESRP-II, which spanned from 2012 to 2018. In addition to building on the interventions introduced in its first phase, PESRP-II also implemented additional measures to improve access, promote gender equality and improve learning outcomes in education (World Bank, 2012). Under PESRP-II, the Government of Punjab emphasised the role of teachers in improving access, quality, and gender equality in education. The key initiatives introduced under the PESRP-II included strengthening the recruitment process and hiring more qualified teachers, implementing teacher training programmes for the continuous professional development of teachers, and emphasising the public-private partnership programme to expand access to education (World Bank, 2012). Collectively, the interventions introduced under PESRP-I and II aimed to improve access to quality education for all, particularly for marginalised communities such as females and children from less advantaged backgrounds (Ansari, 2021; Halai and Durrani, 2020).

With a focus on consolidating previous gains and addressing remaining challenges in the education sector, the Government of Punjab is currently implementing the third phase of PESRP (PESRP-III). This phase emphasises the importance of quality assurance mechanisms, including textbook reforms and Early Childhood Education (ECE) programmes.

The existing evidence on the performance of PESRP in Punjab primarily highlights an increase in enrolment and participation rates in public schools across the province (Barber, 2012; World Bank, 2012; Abdullah and Akhtar, 2019). Global evidence suggests that there may be a trade-off between enrolment expansion and learning outcomes (see literature review section below). However, given the substantial investments in school resources (teachers, classrooms, and basic facilities) under PESRP I and II, it is possible that the Government of Punjab may have mitigated this trade-off. Therefore, this study examines the complex relationship between enrolment expansion and learning outcomes, while exploring the mediating role of these costly supply-side interventions in both expanding educational access and improving learning outcomes.

Literature review. Past studies from Punjab have reported significant improvements in enrolment in public schools following the introduction of the PESRP (Habib, 2013; Aslam et al., 2019). However, no large-scale study has yet explored whether enrolment expansion is negatively linked to learning outcomes. This gap in the literature motivates us to draw on global studies to better understand the potential trade-off between enrolment growth and learning outcomes. This issue is further examined in

the following section, where we review the relevant literature on the relationship between enrolment expansion and educational quality, as well as the role of school resources in mitigating this trade-off. However, before delving into global insights, this section first provides a concise discussion of the dual theoretical framework adopted and its role in shaping the research objectives of this chapter.

Theoretical framework. Changes in enrolment and in learning outcomes are generally assessed through two key theoretical lenses: the Human Rights-Based Approach (HRBA) and Human Capital Theory (HCT). HRBA emphasises education as a fundamental human right, asserting that every child should have access to quality education regardless of socioeconomic status (UNESCO, 2016). HCT, on the other hand, views education as an investment in human capital, highlighting its role in enhancing individual productivity and contributing to economic growth (Hanushek et al., 2007; World Bank, 2018).

This study adopts a dual theoretical approach, integrating both HRBA and HCT to provide a comprehensive analysis of education reforms. While HRBA underscores the moral and legal obligation of governments to provide equitable access to education, HCT stresses that education should lead to skill development and economic returns. However, as Unterhalter (2023) warns, prioritising access without ensuring learning risks undermining the very purpose of education. Similarly, Alcott and Rose (2015) argue that governments must not only enrol children in schools but also ensure they are actively learning.

From both the HRBA and HCT perspectives, equitable learning environments and strong learning outcomes are crucial for realising the full benefits of education. While HRBA focuses on education as a right that fosters personal and societal development, HCT highlights how learning outcomes shape workforce readiness and economic progress. Thus, reforms should be designed to balance access and learning rather than trading one off against the other (UNESCO, 2016).

By adopting this dual framework, this study examines whether the Punjab Education Sector Reform Programme has successfully achieved this balance. Understanding how the programme aligns with both theoretical perspectives is essential to evaluating its impact on educational equity and economic development in Punjab.

The association between enrolment expansion and learning outcomes. The evidence from Kenya and Tanzania of a great relevance to this study, as both countries have implemented reform programmes (free primary education programmes) to enhance access to schools. Several studies have found evidence that enhancing access to schools has negative, but limited, effects on learning outcomes. Such as studies reported that enhancing access to schools through Free Primary Education programs in Kenya and Tanzania did not adversely affect learning outcomes. Bold et al. (2011) were the first to examine a trade-off between enrolment and learning outcomes after implementing free primary education programmes (FPE¹) in Kenya. They used average scores from the nationwide primary school exit examination dataset, known as the Kenya Certificate of Primary Education (KCPE), to measure the learning outcomes between 1998 and 2005. Using a fixed effect model, Bold et al. (2011) found a negative effect of increased enrolment on test scores. However, this negative effect of enrolment expansion was relatively small. The exam average score declines by 3.7 points for each additional thousand students enrolled in the district. Further, as Bold et al. (2011) pointed out, their estimation results could be biased due to district-level idiosyncratic shocks (such as positive changes in

income at the district level) that directly influence exam performance.

Building on the work of Bold et al. (2011), Lucas and Mbiti (2014) also examined the impact of FPE (Free Primary Education) on learning indicators in Kenya. Lucas and Mbiti (2014) used a change-in-change strategy to overcome the district-level biases. Lucas and Mbiti (2014) used school-level enrolment data from education management information system (EMIS) from the Ministry of Education and average KCPE scores to measure learning outcomes from 2000 to 2007. They concluded that after the introduction of FPE, school access substantially increased, and this expansion in enrolment led to a small negative effect (-0.05 percent points) on average KCPE test scores.

Similarly, Valente (2019), also found a small decrease in test scores in Tanzania after implementing FPE. Valente (2019) used regional school-level enrolment data and Southern and Eastern Africa Consortium for Monitoring Educational Quality (SAC-MEQ) data for the years 2000 and 2007 to examine the impact of enrolment growth on test scores. The impact estimates of expansion in enrolment of test scores showed a decrease in the reading (math) score of -0.15 and (-0.16) of a standard deviation with an increase in enrolment by one standard deviation. Valente (2019) concluded that given the size of the enrolment growth and the magnitude of test score drops, FPE has statistically significant negative effects on learning outcomes, but they were not large.

The relationship between expansion in enrolment and school resources. To understand patterns of changes in enrolment and test scores, I turn to the aspects of schooling that might be associated with expansion in enrolment. If enrolment is increasing and school capacity is not, this might crowd classrooms and increase student-teacher ratios. Previous studies have argued that reform programmes like PESRP that significantly increased student enrolment could have some adverse consequences regarding school inputs, such as increasing the student-teacher ratio (STR) and class size (school-classroom ratio or SCR) and reducing the qualifications per teacher due to labour supply issues (see McMillen, 2004). For example, Sifuna (2007) discussed how rapid growth in enrolment in Kenya and Tanzania has also increased class sizes and caused a shortage of teachers. He reported that teachers in Kenyan and Tanzanian primary schools must manage large classes (60–80 students per teacher) after the introduction of FPE, which makes it difficult for teachers to give the individual attention to students they deserve, leading to lower test scores.

Valente (2019) also reported that the rapid increase in enrolment in Tanzania brought some adverse consequences, such as a substantial increase in student-teacher ratios (an STR increase of 6.9% in primary schools). Further, Valente (2019) pointed out that with rapid expansion the supply of qualified and well-trained teachers is limited in the short term, as it takes time to train teachers, so schools must resort to hiring less qualified teachers when the more qualified ones become scarce.

In contrast, Shah and Steinberg (2019) reported that with an increase in enrolment after the passage of the RTE (Right to Education) programme in India, the STR declined (to less than 30 in most schools). They therefore discuss an example of where an increase in enrolment brings improvements in infrastructure, such as gender-specific toilets and libraries. However, they found that these increase in enrolment has a negative association with the learning outcomes, auguring that merely school infrastructure and enrolment is not enough to improve learning outcomes.

The PESRP has also implemented many supply-side interventions focusing on increasing school capacities, such as (1)

building new classrooms, (2) provision of basic facilities, (3) hiring new teachers, (4) provision of free textbooks, and (5) introducing teacher training programmes (Abdullah and Akhtar, 2019). The association between school access and some school resources (such as basic facilities) is often perceived as a two-way link. These supply-side interventions may increase student enrolment and vice versa. Gillani (2021), using annual status of educational report (ASER) data for 2013–2018, reported a positive and significant association between toilet facilities and increased enrolment. Gillani (2021) used pooled estimation to assess this association between toilet facility and enrolment and found an average increase of 12.8 students in public schools of Punjab which have toilet facilities compared to a school that did not have a toilet facility.

The other link is that policymakers may act differently and prioritise providing these basic facilities to the schools that have increased enrolment. In other words, higher enrolment increases the chance for a school to get funds from the authorities to ensure access to basic facilities. School-level fund allocation formulas are often based on the total enrolments (PESRP, 2021). Valente (2019) also confirmed that enrolment expansion is positively associated with physical facilities in schools because the enrolment growth upsurges the capitation grant for the school's non-salary budget. However, little is known about how changes in school enrolment have affected school resources in Punjab.

The association between school inputs and test scores. Inputs that can affect student learning outcomes include those at the child-level (such as students' own cognitive and other abilities); the household-level (socioeconomic status, parental income, parental qualifications, family size, etc.); and the school-level (such as the number of teachers, number of classrooms, teachers' qualifications and experiences, teaching methods and practices, availability of basic schools physical facilities such as electricity, toilets, libraries, playgrounds, and computer labs) (Barrett et al., 2019). There is ample evidence of the impact of child-level and school-level inputs on educational attainment and student performance (see Glewwe and Muralidharan, 2016). The literature on the causal effects of child-level inputs on learning is vast and beyond the scope of this study. This section discusses studies on the school-level inputs measured in this study and their associations with test scores.

Some researchers have maintained that school resources have played a positive role in improving test scores. For example, Glewwe and Kremer's (2005) systematic review across developing countries found that schools with lower STR (number of students per teacher) achieved higher test scores. Similarly, Duflo et al. (2012), using a randomised control trial study, found that simply reducing STR in the government school in Kenya combined with interventions to reduce class size (number of students per classroom) positively impacted student performance. Valente (2019) included STR and teacher qualifications as covariates in the difference in difference analysis to understand the role of these school-level inputs on test scores. She found that lower STR and higher teacher qualification positively correlate with student performance in reading and math scores.

One past study also found a positive association between school physical inputs and learning outcomes in Pakistan. Arshad and Tayyab (2019) evaluate the effect of school physical facilities on student achievement using survey data for only one district in Punjab and test scores of the grade 8 PEC examination scores. Arshad and Tayyab (2019) found that all the physical facilities in the school (ventilation, plants, playgrounds, first aid medical box, LCD/LED) improved (by 15.4%) students' test scores. In particular, the availability of electricity has a strong positive

association with both school enrolment and learning outcomes. Arshad and Tayyab (2019) discussed how summer in Punjab is one of the longest seasons, and often the temperature is too high (around 40°). Children may find it difficult to attend school without electricity to run fans. Further, schools without electricity would not have computer labs, which could negatively affect learning outcomes. However, their study was based on a small sample of only 100 schools in one out of 36 districts in Punjab.

However, more recent evidence from natural experiments and randomised evaluations suggests that public provision of school inputs does not always lead to improved quality of education. For example, Glewwe and Muralidharan (2016), using randomised controlled trials in Kenya, suggest no effect of provision of flipcharts or textbooks on test scores. Similarly, Adukia (2017) studied the impact of the school's basic facilities (availability of toilets) on students' learning outcomes and found no significant impact of the construction of toilets on students' test scores in India. Likewise, Banerjee et al. (2013) argue that learning is not about coming to school, having a smaller class size or lower teacher-student ratio, or having toilets in school; it is about whether we are serious about learning. Teaching practices and teachers' motivation are more important indicators (Alcott et al., 2020). Despite considerable concern about the effect of expensive supply-side interventions on learning outcomes, little evidence is available from Punjab.

Research contribution. Based on the existing global knowledge, the effect of enrolment expansion and school resources on learning outcomes is unclear. The Government of Punjab has taken significant steps to address the dual challenge of enrolment expansion and maintaining school learning outcomes. Efforts such as hiring more teachers to maintain an appropriate student-teacher ratio and constructing additional classrooms to ensure adequate student-classroom ratios have been prioritised (Barber, 2012). Nonetheless, no past studies from Punjab have examined whether these resources have kept pace with the expansion in enrolment. Additionally, measures such as providing electricity for adequate lighting and ventilation play a crucial role in improving learning environments and attracting more students (Arshad et al., 2018). Therefore, the PESRP, with its multiple intervention mechanisms, could potentially have mitigated the trade-off between enrolment and learning outcomes. Despite its significance, however, no large-scale study from Punjab has explored the association between enrolment expansion and the critical role of school resources within this context. Therefore, this is the first study to systematically examine the relationship between enrolment expansion and learning outcomes in Punjab while assessing whether school resources have mitigated potential trade-offs. By analysing the role of interventions such as teacher recruitment, classroom expansion, and infrastructure improvements, this study provides new insights into whether the Government of Punjab's efforts under PESRP have successfully balanced increased access to education with maintaining learning outcomes.

Methodologically, assessing the causal links between enrolment and test scores is problematic because it requires extensive child- and household-level data. Even with access to complete school and child information (including the child's prior level of achievement), in the absence of random allocation of students to schools, confounding effects from selection bias cannot be ruled out (Cook et al., 2002). While this research study does not use a quasi-experimental design or claim to provide causal estimates, the availability of panel datasets (as explained in the introduction section) enables us to provide informative, descriptive analyses. This descriptive study will identify whether there is

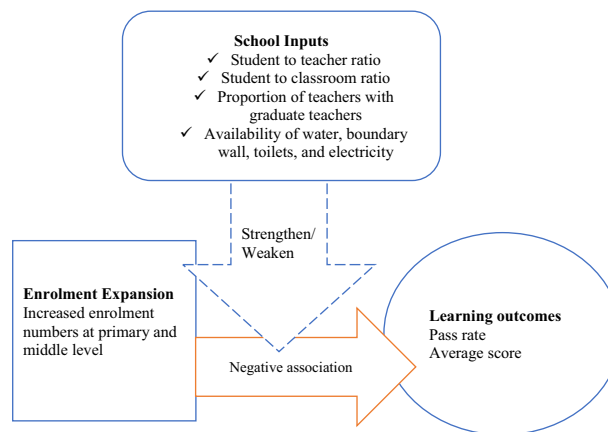


Fig. 1 Modelling framework. This framework is a tailored version of Scheeren's (2019) school input-process-outcome model.

any inverse relationship between enrolment expansion and learning outcomes. It is not inevitable that expansion will be linked with lower test scores, particularly given the supply-side investments made under the PESRP reform programme. Further, on the suggestion of Valente (2019), this study also examined the mediating role of school inputs in the trade-off between school enrolment and learning outcomes. The availability of data on various school-level inputs allows us to investigate whether controlling for school resources weakens or strengthens the measured association between enrolment and learning outcomes. Even without establishing causality, this mediation analysis will extend the existing literature on trade-offs between enrolment expansion, supply of school facilities and learning outcomes.

Research questions.

- Was expansion in enrolments between 2008 and 2018 in Punjab associated with poorer results in exit exams at the school level at (a) the end of primary schooling and (b) the end of middle (lower secondary) schooling?
- To what extent did any association reflect within-school changes, as opposed to time-invariant between-school differences in school size and performance?
- What was the role of changes in measured school-level inputs in accounting for any association between enrolment expansion and student achievement at the school level?

Modelling approach

This section commences with an elaboration on the modelling approach employed to examine the association between enrolment expansion, school resources, and learning outcomes. It then provides a brief overview of the datasets. Subsequently, it describes the computation of key variables of interest outlined in the model.

The study hypothesises that the Government of Punjab has mitigated the negative impact of enrolment expansion on learning outcomes, measured by average scores and pass rates in terminal exams by simultaneously increasing school resources. In this framework, enrolment at both primary and middle levels are the independent variables, learning outcomes serve as the dependent variable, while school resources act as a mediator, influencing how enrolment expansion translates into learning outcomes.

Figure 1 presents modelling framework used in this study, which is based on mediation model. The core idea behind this

mediation model is to evaluate whether school resources have kept pace with enrolment growth. Specifically, the study examines whether interventions such as teacher recruitment, and infrastructure improvements have effectively reduced the potential trade-offs between increased access to education and the quality of learning outcomes.

Figure 1 underscores the interdependent relationship between these components. Without adequate resources, a surge in enrolment may lead to overcrowded classrooms, overburdened teachers, potentially resulting in lower academic achievement. Conversely, well-managed school resources can mitigate the negative consequences of rapid enrolment growth, ensuring that learning outcomes in schools are maintained despite increased demand.

The following section briefly discusses the data sources, and computation process of independent, dependent and mediator variables for this study, as well as the assumed relationship between them.

Datasets. This study relied on two longitudinal school-level² datasets to evaluate the trade-off between enrolment expansion and learning outcomes in public schools in Punjab, between 2008 and 2018. The first secondary dataset source is the Programme Monitoring and Implementation Unit (PMIU), which is responsible for the monitoring and evaluation of public schools in Punjab. The PMIU has administered the PAPSC since 2003. The PAPSC dataset contains information on school-level enrolment, inputs (such as teachers, classrooms, and basic school facilities) from 2003 to 2018.

The second data source on learning outcomes comes from the PEC, which conducts grade 5 and grade 8 exams, representing the terminal grades for primary and middle levels, respectively. The dataset of grade 5 and grade 8 test scores is available from 2008 to 2018. As the dependent variable (test scores) series only starts from 2008, we used data from that year onward. It is important to note that the PMIU continues to collect the PAPSC data annually, and the PEC still produce examination data every year. However, access was only available up to 2018 for this study. After July 2018, a newly elected government in Punjab introduced changes to data-sharing regulations and protocols. Despite regular requests in 2019, it was not possible to gain access to any further years of data. Due to this data unavailability, the focus of this study is limited to the first two phases of the Punjab Education Sector Reform Programme (PESRP), spanning 2003 to 2018.

Key variables of interest. The key independent variable used in this study is school enrolment measured by the logarithm of total enrolment (absolute numbers) in grades 1–5 (for primary schools) and the logarithm value of the sum of enrolment in grades 6–8 (for middle schools), (i.e., $\ln \text{enrol}_{i,t}$ for school i at the time t). Following Egalite Kisida (2016), the log specification is used because it allows for a non-linear effect of enrolment on test scores. Also, this logged transformation of the independent variable is used in regression models for interpretability. Data reveals that most of the schools in Punjab have multi-grade teaching, and the school inputs are also categorised at the whole school level, so the whole primary/middle section enrolment rather than grade 5/8 enrolment were used as the independent variables.

Dependent variables. This study used two dependent variables to measure learning outcomes. The first was a continuous variable measured by the pass rate. Candidates who obtained 33% marks were considered to pass. If a student failed to achieve 33% in total in the PEC examination, they usually must repeat grade 5 or

grade 8 accordingly. Any student who failed the PEC exam must repeat the same grade for the following year. The second was the school's average test score, a continuous variable measured by the total average score of the school per year.

A school's pass rate tells the ratio of students who pass each year by the 33% threshold. A high pass rate suggests that most students meet the minimum requirements and progress through their education. In contrast, the school's average score will provide supplemental information on learning outcomes, as a higher score indicates that students pass the exam with better scores/ are performing well. Thus, both pass rate and average score were used in this study to comprehensively understand the school's learning outcomes.

Mediator variables. As discussed earlier, enrolment expansion can affect school resources in two ways. This expansion may stretch existing resources. On the other hand, the expansion may attract more funding and grants and improve the school's facilities. Based on the expected sign of association between expansion in enrolment and school-level inputs, two separate sets of mediators are used in this study.

The first set of mediators is expected to be negatively associated with enrolment: if the supply-side interventions do not pace up with the expansion, the student-teacher ratio, the student-classroom ratio, and the supply of qualified teachers might be stretched. The first group of mediators included the student-teacher ratio (STR), student-classroom ratio (SCR), and teacher qualifications (proportion of teachers with a college degree in a school). The mediators STR and SCR are calculated using the total number of students divided by the total number of teachers/classrooms. The school-level inputs are calculated at the whole school level and not the class, as teachers, classrooms and other facilities are available to the whole school and are not allocated to individual grades. Teacher qualification is a proportion variable computed using the sum of teachers with a college degree (BA, MA, MPhil, or PhD) divided by the total number of teachers posted in that school.

Considering the efforts of PESRP-I and II to improve school facilities, enrolment expansion is likely to be positively associated with a second set of school facilities. This second set of mediator variables consists of the school's basic facilities (toilet, boundary wall, electricity, and drinking water). Binary variables have a value of 1 if the school has drinking water, toilets, electricity, and a boundary wall or zero otherwise.

Estimation strategy. The main question of this study is whether there is an inverse relationship between enrolment expansion and test scores of grades 5 and 8 terminal examinations in the public schools of Punjab. The estimation strategy consists of three parts. First, following the suggestions of Kuziemko (2006), the analysis begins with the classical cross-sectional model. The nature of PEC and PAPSC data provides an obvious choice of pooling the observations from all years and analysing them with pooled ordinary least squares (OLS). Thus, the OLS estimator is used to grasp the long-term relationships between access and learning outcome indicators. These pooled cross-section models will tell us about the overall association between enrolments and test scores. District-level cluster-robust standard errors are used to produce efficient and consistent parameter estimates. However, as Wooldridge (2012) highlighted that OLS estimates ignore within-school variations, a change-in-change model is used to isolate the within-school changes in this trade-off between enrolment expansion and learning outcomes. Lastly, to identify the mediation role of school inputs, two sets of school facilities indicators are added in the OLS and change-in-change models.

The trade-off between enrolment and learning outcomes. A simple linear regression model is used to answer the question: Is there a trade-off between access and learning outcomes?

$$Y_{it} = \alpha + B_1X_{it} + B_2year_t + \mu_i + \varepsilon_{it} \quad (1)$$

Here, i is the index for the individual school, and t is the index for the time. Y_{it} is the dependent variable (i.e., school-level pass rate or average score), α represents the intercept for the model, i.e., the predicted value of Y_{it} when the predictors are zero values. X_{it} is the logged transformation of the independent variable (enrolment in absolute numbers). B_1 captures the magnitude of the effect of an increase in enrolment, that is, the effect of a one unit increases in X_{it} . As X_{it} is equal to logged enrolment, the coefficient value can be divided by 10 for interpretation purposes. $B_1/10$ is, therefore, the expected change in Y_{it} for an increase in X_{it} of 0.1, approximately equivalent to a 10% increase in school enrolment. For instance, presume the coefficient value of log-enrolment is -0.20 . The results will be interpreted as showing that a 10% increase in enrolment in a school was associated with a decrease in the average school pass rate of approximately 0.02 (0.20/10) percentage points. The expected sign for this coefficient is negative.

Following Wooldridge's (2012) suggestion, year dummies are added to the model to account for an overall time trend in test scores, which commonly affects schools, for example, annual changes in the format or difficulty of the exam. B_2 captures the time trend in the Y_{it} . The error term is split into two units, μ_i and ε_{it} . The error terms represent measurement error in the dependent variable and all unknown explanatory influences on school-level learning outcomes that have not been controlled in the model (omitted variables such as family SES, parents' involvement, community participation, school management system, and peer effects). μ_i denotes the time-invariant errors and ε_{it} consist of the time-variant measurement error portions.

As an additional step, some time-invariant school characteristics in the POLS estimations that may be correlated with school enrolment were added to the model. In Eq. 2, S_i represents the vector of school-level variables: school location (urban vs. rural), school level (in terms of grades served), school gender, and district dummies.

$$Y_{it} = \alpha + B_1X_{it} + B_2year_t + B_3S_i + \mu_i + \varepsilon_{it} \quad (2)$$

This vector of school-level fixed characteristics will adjust for bias in B_1 that results if these characteristics are systematically associated with enrolment and test scores. For example, if schools with enrolment for all grades 1–10 both perform better than, and tend to have more students than, stand-alone primary schools (grades 1–5), then the estimate of B_1 from Eq. 1 (which omits this variable) will lead to an upwards-biased estimate of the effect of enrolment on the outcome. B_3 coefficients capture the effect of these time-invariant school-level variables on the pass rate and average scores. However, any unmeasured school-level characteristics will remain in the error term ($\mu_i + \varepsilon_{it}$) and if these are correlated with enrolment, they will continue to bias the estimate of B_1 .

Change-in-change model. Linear change-in-change models were used to gauge the association between enrolment expansion and learning outcomes in the second stage. The change-in-change approach has each unit (school in this case) as its own control, eliminating any influence of school time-invariant confounding

variables (Lucas & Mibity, 2014).

$$Y_{it} - Y_{it-1} = (\alpha - \alpha) + B_1(X_{it} - X_{it-1}) + B_2year_t + B_3(S_i - S_i) + B_3(\mu_i - \mu_i) + (\varepsilon_{it} - \varepsilon_{it-1}) \quad (3)$$

The conventional change in change estimation described above in Eq. 3 implicitly compares the same school at specifically adjacent points in time instead of comparing different schools to each other. The change-in-change model isolates within-school variation over time to estimate the relationship between changes in enrolment expansion and test scores. As year dummies are used to control for the time trend for Y_{it} , the time dummies were not differenced.

Equation 3 can be rewritten as

$$\Delta Y_{it} = B_1 \Delta X_{it} + B_2 year_t + \Delta \varepsilon_{it} \quad (4)$$

Equation 4 regresses the outcome ΔY (change in pass rate or average score in school i between year $t-1$ and year t). ΔX_{it} is now the predictor variable and presents the growth rate in student numbers (changes in logged enrolment). $B_1/10$ now has the interpretation of the expected change in the dependent variable for a 10-percentage point increase in the growth rate of enrolment, for example, a difference in the growth rate from 0.05 (5%) between year $t-1$ and t to 0.15 (15%). This change-in-change model differences out the time-invariant characteristics of schools that are fixed over time. The fixed characteristics of schools, such as location, remain the same in both periods. Thus, the changes in S_i and μ_i are equal to zero.

Enrolment expansion and school inputs. As suggested by the literature, the expansion in enrolment might stretch school resources. Several school-level inputs (STR, SCR, teacher qualifications, availability of basic facilities) are regressed on school enrolment using a simple linear regression POLS and change-in-change model.

$$Z_{it} = \alpha + B_1X_{it} + \mu_i + \varepsilon_{it} \quad (5)$$

Here Z_{it} is the school-level resources and X_{it} is logged school enrolments. B_1 captures the association between expansion in enrolment and school inputs and is expected to be positive. For instance, if the B_1 associated with presence of electricity is positive, this indicates larger schools are more likely to have access to electricity.

$$\Delta Z_{it} = B_1 \Delta X_{it} + \Delta \varepsilon_{it} \quad (6)$$

In Eq. 6, the key dependent variable is the change in resources, and the independent variable is the change in logged enrolment. Here, the B_1 associated with presence of electricity captures whether the schools that expanded between $t-1$ and t are more likely to have acquired electricity during that period.

Mediation analysis. Further, on the suggestion of Valente (2019), this study used school resources as a potential mediator between school expansion and school quality. Equation 7 below shows a POLS model with two sets of potential mediators (Z_{it} and C_{it}), and Eq. 8 shows the equivalent change-in-change model.

$$Y_{it} = \alpha + B_1X_{it} + B_2year_t + B_3S_i + B_4Z_{it} + B_5C_{it} + \varepsilon_{it} \quad (7)$$

$$\Delta Y_{it} = B_1 \Delta X_{it} + B_2 year_t + B_3 \Delta Z_{it} + B_4 \Delta C_{it} + \Delta \varepsilon_{it} \quad (8)$$

Here Z_{it} is the vector of potential mediator variables hypothesised to be negatively linked with enrolments (e.g., STR); C_{it} is the vector of potential mediator variables hypothesised to be positively linked with enrolments (e.g., provision of electricity). As mentioned earlier, through this mediation analysis, this study aimed to identify whether controlling for school inputs

strengthens or weakens the estimated trade-off between enrolment and test scores. If the magnitude of the B_1 coefficient is reduced when the mediators are added, this indicates that school inputs have partially accounted for the total association between enrolment and test scores estimated in Eqs. 2 and 4.

Sample construction. Supplementary Table S1 online in the presents the full sample's descriptive statistics (mean, standard deviation, minimum and maximum values). The PAPSC data has 528,873 school-year observations for the primary level and 152,176 for the middle level for 2008–2018. Of these, only 394,063 have matching PEC (test score) data at the primary level and 69,859 at the middle level. It is to highlight that participation in the PEC was not mandatory, allowing individual schools the choice to take part. This voluntary nature meant that the data collected did not encompass all schools, preventing a complete match with the school census information. However, supplementary analyses conducted using a logit model revealed no significant systematic differences between the schools that chose to participate and those that did not.

As discussed above, the key dependent variable, the pass rate for grades 5 and 8, comes from the PEC dataset. Using the unique ID code (EMIS code), the PAPSC and PEC data were merged to create one longitudinal cross-sectional dataset. To implement the change-in-change estimator, at least two observations per school from two consecutive time points are needed. Therefore, sample size used in this study is further restricted to schools with PEC exam observation and enrolment data from two consecutive points. This restricted the sample to 294,656 observations at the primary level and 33,680 at the middle level.

The PAPSC data had some outliers in the dataset, such as a maximum STR of 396. Cook's distance was used to identify the outliers in this study. Cases are considered significant influences on the regression estimates if the values of Cook's distance are over $4/n$. There were, in total, 16,141 cases (outliers) for the primary level dataset and 1829 middle school cases that could influence the change-in-change regression estimations for this sample (see Supplementary Tables S2 and S3 online). Change-in-change regression analyses with and without the identified outliers were compared and estimates were found to be similar but slightly attenuated when the influential cases were included (see Supplementary Table S4 online). As exclusion of outliers does not have any significant impact on the results, influential cases were dropped from the sample. The analytical sample for this study is therefore restricted to schools that have matching values of test scores and enrolment indicators for two consecutive years without the influential cases. The analytical sample consists of 278,515 primary-level observations and 31,851 middle-level observations from 2009 to 2018.

Table 1 presents the summary statistics of the restricted sample. On average, a school has a pass rate of 78% in grade 5, and 83% in grade 8. The average primary school score in grade 5 exams is 287 out of 500, and the average score for middle schools in grade 8 is 357 out of 500. The average enrolment in a year in a primary school is 99 students, and the average enrolment at the middle level (grades 6–8) is 213 students. The average STR and SCR at primary (middle) level schools in Punjab are 27 (30) and 26 (42), respectively. Table 1 suggests that most public schools (almost 95%) have basic facilities such as toilets, boundary walls, drinking water, and electricity. On average, 54 (71) percent of primary (middle) schoolteachers have college degrees.

Empirical results

This section explores the association between enrolment expansion and learning outcomes first through POLS estimations, and

then through the change-in-change model. The POLS model results show no significant association between logged enrolments and either the pass rate for grade 5 or average scores for grade 8. Small but significant positive effects were found for average scores for grade 5 and pass rates for grade 8 (see Supplementary Table S5, columns 1 and 4). A 10% increase in enrolment is associated with an increase in the grade 8 pass rate of 0.06 percentage points (which is tiny when the average pass rate is 83.3%). A 10% increase in logged enrolment is also associated with an increase of 0.6 average score points for grade 5 (again, a minuscule effect when considering that scores are measured out of 500, with a mean of 287 score points and a standard deviation of 60, as shown in Table 1).

After adding year dummies and school-level cross-sectional covariates (school location, school gender, school level, and district dummies), the POLS estimation changes very little (see Supplementary Table S5, columns 3 and 6, in the online appendix).

The association between enrolment and learning outcomes remains positive for most outcomes and strengthens slightly for grade 8 outcomes, but the effects of a 10-percentage increase in logged enrolment remain under 0.1 percentage points on the pass rate and under 1 score point at grade 8. A significant adverse effect emerges only for the grade 5 pass rate, but the effect is again tiny – for a 10-percentage increase in enrolment, the pass rate decreases by 0.06 percentage points at the primary level (in the context of a mean pass rate of 78.2%). The addition of time-invariant school-level controls suggests that school size is not associated systematically with other years, districts, or observed school characteristics that predict exam performance, as these do not appear to be significant sources of bias.

Overall, the POLS model shows that an increase in enrolment is positively associated with the learning outcomes (apart from the grade 5 pass rate). The interpretation in this context might be that higher enrolment can lead to better learning outcomes, possibly because of increased access to education, better allocation of resources, or greater emphasis on education quality due to increased demand. However, this association is nuanced, and while the data may show a positive correlation, the magnitudes of the POLS coefficients imply that the effects are likely small and might be biased as the model did not control for all unobserved time-invariant school characteristics. Schools vary in characteristics that do not change over time, such as their location, community support, or school culture. POLS assume these characteristics are constant or do not influence the number of enrolments, but this assumption may not hold in reality. Ignoring such factors can lead to misleading conclusions about the actual effect of enrolment changes.

The change-in-change regression estimates shown in Table 2 below report a significant negative association between enrolment growth rate and exam pass rate at primary and middle levels. A 10-percentage point faster growth between year $t-1$ and year t is associated with a reduction of $(-36.97/10)$ 3.7 percentage points in the pass rate for grade 5 (primary level) in year t , and a 5.5 percentage points reduction at the middle level. This negative association suggests that while higher enrolment may signal demand for education and potentially increase investment, institutions may struggle to absorb larger student populations effectively. Moreover, these are quantitatively meaningful reductions when compared with the average pass rates (78.22 and 83.29 for grades 5 and 8 respectively) in Table 1.

In contrast, in terms of average scores, a small but significant positive association was found between enrolment growth and average scores at both grade 5 and grade 8. A 10%-point faster growth in enrolment in a school between $t-1$ and t is associated with an increase in grade 5 and 8 average scores in year t of 0.6

Table 1 Summary statistics for the analytical sample.

Variable	Observation	Mean	Std. dev.	Min	Max
Primary level- grades 1-5					
Pass rate grade 5	278,515	78.22	23.38	0	100
Average score grade 5	278,515	287.40	60.43	3	444
Enrolment	278,515	99	78.77	2	2495
Logged enrolment	278,515	4.37	0.67	0.69	7.99
School gender (0 male,1 female)	278,515	0.51	0.50	0	1
School Location (0 urban,1 rural)	278,515	0.90	0.30	0	1
Primary grades only	278,515	0.86	0.35	0	1
Primary with middle level grades	278,515	0.09	0.29	0	1
Primary with middle and high level grades	278,515	0.05	0.21	0	1
Primary with middle, high, and higher secondary grades	278,515	0.00	0.06	0	1
Student-teacher ratio (STR)	278,515	27.21	15.70	0.75	396
Student-classroom ratio (SCR)	278,515	30.66	18.96	0.56	316
Proportion of teachers with a college degree	278,515	55.00	19.49	0.56	342
Availability of boundary wall	278,515	0.95	0.38	0	1
Availability of drinking water	278,515	0.94	0.24	0	1
Availability of electricity	278,515	0.97	0.16	0	1
Availability of toilet	278,515	0.96	0.20	0	1
Middle school (grades 6-8)					
Pass rate grade 8	31,851	83.29	18.97	0	100
Average score grade 8	31,851	357.68	87.10	10	492
Enrolment	31,851	213.38	225.08	3	1978
Logged enrolment	31,851	4.93	0.93	1.39	7.59
School gender (0 male,1 female)	31,851	0.47	0.50	0	1
School Location (0 urban,1 rural)	31,851	0.76	0.43	0	1
Middle grades only	31,851	0.01	0.08	0	1
Middle with primary grades	31,851	0.40	0.49	0	1
Middle with primary and high school grades	31,851	0.57	0.49	0	1
Middle with primary, high, and higher secondary grades	31,851	0.02	0.16	0	1
Student-teacher ratio (STR)	31,851	26.36	10.86	1	294
Student-classroom ratio (SCR)	31,851	42.48	22.00	0.625	369
Proportion of teachers with a college degree	31,851	0.71	0.31	0	1
Availability of boundary wall	31,851	0.97	0.16	0	1
Availability of drinking water	31,851	0.99	0.08	0	1
Availability of electricity	31,851	0.98	0.14	0	1
Availability of toilet	31,851	0.99	0.10	0	1

Author's calculations: Source PAPSC and PEC 2009-2018.

Table 2 Change-in-change regression model estimates of the effect of enrolment numbers on learning outcomes.

	Change in pass rate	Change in average score
Primary level grade 1-5		
Change in logged enrolment	-36.973*** (1.007)	6.26** (1.819)
r2	0.142	0.090
N	278,515	278,515
Middle-level grades 6-8		
Change in logged enrolment	-54.557*** (1.334)	12.190*** (2.621)
R2	0.225	0.616
N	31,815	31,815

Year dummies are included in each model. Robust standard errors are reported in brackets, ***p < 0.01, **p < 0.05.

and 1.2 score points respectively. Given that the average scores are approximately 287 and 358, the magnitude of the effect is very small. In terms of contextualising these results, non-trivial negative effects of enrolment growth were found on pass rates but not on average scores that summarise the entire achievement distribution. This suggests that it is students near the pass/fail

boundary (a mark of 33%) that performed more poorly in schools that had experienced rapid enrolment expansion. Students higher up the achievement distribution do not seem to have been impacted in the same way and even benefited from growing enrolments, such that their gains outweighed the losses at the bottom of the distribution, leading to slightly higher test scores on average. For the majority, therefore, the positive association between enrolment growth and scores could be interpreted as an indication that increased access to education could lead to slightly better learning outcomes, possibly due to factors such as resource allocation, increased investment in schools, or the dynamic nature of growing educational environments. Declining pass rates suggest that weakest students were less likely to meet the minimum standard in schools that grew rapidly. This could reflect diversion of resources from those most in need of support, but it could also be a compositional effect, in which disproportionately lower-achieving students were drawn into school by the reforms.

The results of this study are consistent with existing research, including Valente (2019) and Bold et al. (2011), which also found that while increased enrolment is associated with small reductions in educational outcomes, the effects are modest. This suggests that the relationship between enrolment and academic performance is not necessarily linear or large in magnitude, reinforcing the idea that other systemic factors such as teacher quality, curriculum design, socioeconomic background, and parental

Table 3 Change in change regression estimations between enrolment and school inputs.

	Change in STR	Change in SCR	Change in the proportion of teachers with a college degree	Change in drinking water	Change in electricity	Change in toilet	Change in boundary wall
Primary level- Grade 1-5							
Change in logged enrolment	16.222*** (0.875)	22.457*** (0.946)	-0.02 (0.015)	-0.004 (0.002)	0.021*** (0.004)	-0.005 (0.003)	-0.001 (0.002)
r ²	0.072	0.195	0	0	0	0	0
N	278,515	278,515	278,515	278,515	278,515	278,515	278,515
Middle level- Grade 6-8							
Change in logged enrolment	9.136*** (0.448)	15.379*** (0.805)	0.060** (0.02)	-0.003 (0.004)	0.012* (0.005)	0 (0.007)	-0.011* (0.005)
r ²	0.031	0.032	0	0	0	0	0
N	31,851	31,851	31,851	31,851	31,851	31,851	31,851

Robust standard errors are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.10.

involvement, could explain why the changes in scores due to enrolment growth are relatively minor.

The effect of expansion in enrolment on school resources.

Using the simple POLS approach, all the school inputs are positively associated with enrolment (see Supplementary Table S6 online). Larger schools have much higher STR and SCR than smaller ones, which may negatively affect enrolment growth and learning outcomes. Unexpectedly, higher enrolments were also associated with increased teacher qualifications. However, by simply looking at POLS estimations, it is unclear whether this reflects responses to enrolment expansion or other long-standing characteristics. At the same time, enrolments are positively associated with access to basic facilities at the primary and middle levels. As discussed earlier, this could be due to reverse causation. Perhaps schools with more facilities attract more students, or policymakers spend more on basic facilities for schools with higher enrolment. The change-in-change models help clarify this by looking at whether enrolment changes coincide with input changes.

The change-in-change estimations in Table 3 below show that enrolment growth between year *t-1* and year *t* is associated with significantly higher STR and SCR at primary and middle levels. The coefficients in columns 2 and 3 in Table 3 below show that a 10-percentage point growth rate in enrolments is associated with, on average, 1.6 (16.22/10) more students per teacher and 2.2 more students per classroom at the primary level, and 0.9 more students per teacher and 1.5 per classroom at the middle level. Regarding teachers' qualifications, there is no significant association between enrolment growth and changes in teachers' qualifications at the primary level. Though the estimates show a positive significant association between enrolment growth and teacher qualification changes at the middle level, the magnitude of this association is small. A 10-percentage point enrolment growth rate is associated with an increase of the proportion of teachers with a college degree by less than 1 percentage point (0.006), which is negligible, considering the average proportion of graduate teachers in a school is 71%. Nevertheless, it is possible that highly educated teachers are more attracted to faster-growing schools.

Table 3 further shows that enrolment growth is not significantly positively associated with the likelihood of availability of the school physical facilities at both primary and middle levels, with the exception electricity. With a 10-percentage point growth rate, the probability of having electricity increased by 0.2 percentage points at the primary level and 0.1 percentage point at the middle level. This could have helped to mitigate the adverse effects of faster enrolment growth, but the effects are minimal. To sum up, changes in basic facilities are not important potential mediators in the relationship between enrolment expansion and learning outcomes. In this context, changes in STRs and SCRs are more promising indicators for understanding why enrolment growth was associated with declining pass rates, but whether they matter for student learning remained unanswered.

School inputs as mediators. The results in Table 3 are consistent with literature showing the enrolment expansion is associated with increasing student-teacher ratios (STRs) and class sizes (SCRs) (Bold et al., 2011; Glewwe and Kremer, 2006; Valente, 2019) and provision of school infrastructure (Hayat, 2017). However, most inputs are not significantly associated with exam outcomes in cross-sectional POLS models (see Supplementary Table S7 online). As expected, STR significantly negatively affects the pass rate at the primary level. With a one-unit increase in STR (i.e., one additional student per teacher), the school pass rate on

Table 4 Change in change regression estimation with the school-level inputs as mediators.

	Change in the pass rate.		Change in the average score.	
Primary level- Grades 1-5				
Change in logged enrollment	-36.973*** (1.007)	-37.452*** (0.947)	-37.457*** (0.946)	6.122** (1.785)
Change in STR		0.004 (0.006)	0.004 (0.006)	0.022 (0.019)
Change in SCR		0.012 (0.007)	0.012 (0.007)	0.034 (0.021)
Change in teacher qualifications		0.835* (0.348)	0.838* (0.348)	-0.390 (0.868)
Change in the availability of drinking water.			0.540* (0.243)	-1.275 (1.316)
Change in electricity		0.412 (0.205)	0.418* (0.206)	0.190 (0.773)
Change in toilet		-0.079 (0.250)	-0.061 (0.253)	-0.845 (1.224)
Change in boundary wall		-0.341 (0.298)	-0.328 (0.299)	-1.813 (0.941)
r2	0.142	0.140	0.140	0.290
N	278,515	278,515	278,515	278,515
Middle level- Grades 6-8				
Change in logged enrollment	-54.557*** (1.334)	-56.845*** (1.378)	-56.823*** (1.380)	12.008*** (2.372)
Change in STR		0.152*** (0.020)	0.152*** (0.020)	0.184** (0.054)
Change in SCR		0.060*** (0.009)	0.061*** (0.010)	0.048 (0.024)
Change in teacher qualifications		0.283 (1.861)	0.275 (1.862)	-2.156 (4.917)
Change in the availability of drinking water		-0.854 (0.826)	-0.744 (0.846)	0.209 (3.378)
Change in electricity		-1.416* (0.596)	-1.272* (0.552)	-2.352 (2.717)
Change in toilet		1.276* (0.566)	1.155 (0.582)	1.661 (3.449)
Change in boundary wall		0.611 (0.549)	0.927 (0.523)	-0.975 (3.533)
r2	0.225	0.229	0.230	0.628
N	31,815	31,851	31,815	31,851

Year dummies included in all models. Robust standard errors are reported in brackets, ****p* < 0.01, ***p* < 0.05, **p* < 0.1.

average decreases by 0.056 percentage points, and the average score decreases by 0.010 points at the primary level. Unexpectedly, there is a positive association between SCR and average scores at both primary and middle levels. The primary (middle) level SCR coefficient suggests that one additional student per class increases the average score by 0.9 (0.7) points. The estimated coefficients are statistically significant, but the effect size is tiny. Also, unexpectedly, teacher qualifications are associated with poorer middle school outcomes.

These unexpected positive associations could be due to several factors. Schools with higher SCR might have more resources or effective teaching methods that help mitigate the challenges of larger class sizes (Abdullah and Saeed, 2019). Additionally, these schools could foster a more motivated student body or employ strategies like group-based learning that enhance performance despite larger classes (Ellis, 2024). It's also possible that this relationship reflects contextual factors, such as socioeconomic conditions or school-specific characteristics, that weren't fully captured in the data (Bold et al., 2011). Change-in-change estimates can help throw light on the role of unobserved time-invariant school-level factors such as these.

Regarding the set of school infrastructure variables, the availability of electricity is positively associated with the primary pass rates. This suggests that access to electricity may improve the learning environment, enabling better resources, such as lighting for after-school study or the use of technology in the classroom, which could contribute to higher academic performance (Hayat, 2017). However, the availability of a boundary wall is associated with worse primary average scores. This may reflect underlying issues such as overcrowding or a less conducive learning environment. Alternatively, it could suggest that schools with boundary walls are in areas where the resources are not effectively used or distributed, leading to poorer academic outcomes despite the physical infrastructure (Alcott and Rose, 2015).

Mediation analysis. Table 4 below shows the mediation analysis for the change-in-change model. Overall, changes in most school inputs are not significantly associated with simultaneous changes in learning outcomes in the same year. As expected, an increase in teachers' qualifications (more college-educated teachers) is associated with an improvement in the primary pass rate (0.10 increase in proportion with a degree associated with 0.08 percentage point improvement in pass rate). This is consistent with previous research suggesting that more qualified teachers improve student learning outcomes. Higher education levels in teachers may contribute to better teaching methods, deeper subject knowledge, and more effective classroom management, all of which can positively influence student performance (Valente, 2019).

Similarly, the availability of drinking water in a school is associated with a 0.54 percentage point improvement in the primary pass rate, and the supply of electricity is associated with a 0.42 percentage point improvement in the primary pass rate. For middle-level schools, having toilets shows a positive association, i.e., the introduction of toilets in year t brings a change of 1.2 percentage points in the pass rate (significant at a 10% level). These findings are consistent with the previous work of Hayat (2017) and highlight the crucial role of basic amenities in supporting student well-being and academic performance.

However, changes in the STR are positively associated with the changes in the pass rate and average score at the middle level. Also unexpectedly, electricity has a negative association: introducing electricity in a school at the middle level is associated with a decline in the pass rate of 1.3 percentage points in the same year. A possible reason is that introducing electricity may coincide with

school expansions or infrastructural changes that temporarily disrupt learning. Another possibility is that students may be distracted by access to electronic devices, leading to lower academic focus.

Regarding the role of inputs as a mediator in the relationship between enrolment and test scores, I have hypothesised that the introduction of the first set of key inputs (STR, SCR, and teacher qualifications) in Eq. 8 would increase the coefficient on enrolment growth variable, because a 'negative pathway' from enrolments to outcomes has now been statistically controlled. This hypothesis was not supported for any outcome, with coefficients decreasing by varying magnitudes rather than increasing.

On the other hand, the inclusion of the second set of mediators (basic school facilities) was hypothesised to reduce the enrolment growth coefficients, by statistically controlling for a 'positive pathway'. As discussed earlier, expanding enrolment could increase access to basic facilities in the school and protect against any adverse effects of expansion on test scores. Columns 3 and 7) of Table 4 added the set of changes in the binary school infrastructure variables, and the comparison with the estimates listed in columns 1 and 5 of Table 4 provides little support for this hypothesis. The estimated effect of enrolment growth on learning outcomes drops slightly for only two of the four outcomes (the grade 5 pass rate and grade 8 average scores), and for these only by a negligible amount. So there is little evidence that enrolment expansion brought forth infrastructure improvements that had measurable effects on learning outcomes.

Overall, controlling for all inputs (columns 4 and 8) does little to account for the observed associations between enrolment growth and learning outcomes, either positive or negative. The exception is at the middle level, where the positive effect of enrolment growth on average scores was partly accounted for by the fact that STRs increased when growth was faster and (unexpectedly), higher STRs predicted higher average scores.

To sum up, while improvements in infrastructure, teacher qualifications, and student-teacher ratios are often expected to influence academic performance, their effects in this analysis were frequently not statistically significant and occasionally contrary to the direction predicted. This suggests that short-term changes in these inputs may not immediately translate into measurable improvements in pass rates or average scores.

Gender disaggregated mediation analysis. To assess whether the association between enrolment expansion and learning outcomes varies differently for school gender, the change-in-change mediation analysis was disaggregated for male and female schools. Since girls' schools were consistently supported under the Punjab Education Sector Reform Program, it was expected that they would experience distinct effects compared to boys' schools (Barber, 2012). Although the direction of the impact was not entirely clear, it is likely influenced by gender-specific factors. For instance, the Female Stipend Support Program (FSSP) may have encouraged higher female enrolment, potentially leading to overcrowding and resource constraints in girls' schools (World Bank, 2012). On the other hand, boys' schools, which were not affected by the absence of female teachers, may have had a relatively better teacher-student ratio and access to resources. These dynamics suggest that gender may play a role in how school expansion affects learning outcome.

Table 5 shows the gender-wise breakdown of the change-in-change regression model. The gender-segregated analysis shows that overall, the association between enrolment expansion and learning outcomes is similar between male and female schools. Both male and female schools show a significant inverse

Table 5 Gender-wise- change in the change regression model.

Variables	Female		Male	
	Pass rate	Average score	Pass rate	Average score
Primary level grades 1-5				
Change in logged enrollment	-38.967*** (1.32)	5.985* (2.539)	-37.370*** (1.632)	5.705** (1.657)
Change in STR	0.005 (0.009)	0.046 (0.029)	-0.011 (0.008)	-0.004 (0.026)
Change in SCR	0.012 (0.010)	0.025 (0.028)	0.007 (0.008)	0.038 (0.023)
Change in teacher qualifications	1.213** (0.435)	-0.0952 (1.104)	0.106 (0.543)	0.026 (1.006)
Change in the availability of drinking water	0.926 (0.504)	-2.427 (1.551)	0.586 (0.393)	-0.129 (1.454)
Change in electricity	0.186 (0.422)	0.885 (1.064)	0.732** (0.263)	-0.214 (0.853)
Change in toilet	0.158 (0.430)	0.129 (1.784)	-0.052 (0.270)	-0.474 (1.208)
Change in boundary wall	0.248 (0.404)	-3.822 (21.96)	0.059 (0.339)	-0.644 (1.82)
R2	0.13	0.274	0.152	0.309
N	140,863	140,863	137,652	137,652
Middle-level grades 6-8				
Change in logged enrollment	-58.391*** (1.439)	5.790 (3.076)	-56.122*** (2.106)	14.530*** (3.743)
Change in STR	0.075** (0.021)	0.017 (0.060)	0.095*** (0.029)	0.188* (0.073)
Change in SCR	0.054** (0.015)	0.066** (0.031)	0.049*** (0.012)	0.046 (0.038)
Change in teacher qualifications	-4.913 (2.769)	-10.801 (5.409)	-4.266 (2.288)	-2.361 (8.893)
Change in the availability of drinking water	-0.770 (1.291)	-5.576 (6.649)	0.635 (0.539)	4.339 (4.430)
Change in electricity	-1.414 (0.871)	-0.999 (4.846)	1.034 (1.165)	-1.550 (4.541)
Change in toilet	3.997** (0.994)	1.818 (6.914)	0.759 (0.760)	1.871 (0.335)
Change in boundary wall	0.594 (1.032)	-1.835 (1.909)	0.642 (0.539)	1.782 (3.613)
R2	0.208	0.639	0.254	0.625
N	14,857	14,857	16,958	16,958

***Notes to the table- robust standard errors are reported in brackets. ***p < 0.01, **p < 0.05, *p < 0.10.

relationship between enrolment growth and changes in pass rates, and positive associations were found between enrolment growth and average scores. For example, a 10% growth in enrolment in female primary (middle) schools is associated with a 3.9 (6.0) percent point decrease in pass rates. Considering the vast expansion of enrolment for females in public schools of Punjab due to PESRP-I and II, the similarity in trends suggests that neither male nor female learning outcomes were disproportionately harmed. The evidence suggests that uniform policy interventions under the Punjab Education Sector Reform Program influenced both types of schools in a comparable manner. While female schools have additional support mechanisms like the Female Stipend Support Program, their impact on learning outcomes might be offset by other factors, such as overcrowding or resource constraints.

Like results for the combined sample, most of the inputs are not significantly associated with the changes in learning outcomes for male and female primary-level schools. The gender-disaggregated analysis also confirms that school-level inputs do not mediate the association between enrolment expansion and test scores, as the enrolment growth coefficients with and without school input controls vary little. In other words, the set of school-level inputs do not explain either the negative or positive effects of expansions in enrolment on learning outcomes. Additional child-level and household information that enable the computation of compositional and peer effects are needed to identify the mechanisms underlying the estimated associations.

Furthermore, the fact that both male and female schools exhibit the same pattern—a significant inverse relationship between enrolment growth and changes in pass rates, alongside a positive association between enrolment growth and scores averaged over the distribution as a whole—suggests that gender-specific factors are not be the primary drivers of these trends. This could indicate that broader structural issues, such as resource allocation, teaching quality, or school management practices, affect both types of schools similarly. Further analysis is needed to explore whether these trends hold across different school types and geographic locations, and whether gender-sensitive policies have any differential long-term effects on student performance.

Conclusion

The aim of this study was to assess whether the government of Punjab has successfully mitigated the potential trade-off between enrolment expansion and learning outcomes because, along with the expansion of enrolment, the Punjab Education Sector Reform Programme (PESRP) also enhanced school infrastructure and recruited more qualified teachers (Barber, 2012). We hypothesised that, unlike other reform programmes globally, PESRP would not experience the negative consequences typically associated with rapid expansion.

This paper adds to the existing literature in several ways. Firstly, this study is among the first to highlight the association between enrolment and test scores in public schools in Punjab. This involved the creation of a large-scale school-level panel dataset using ten annual rounds of PAPSC administrative data and the PEC terminal examination (grades 5 and 8) data. The school-level observations used for this study are novel and have not been used before for this analysis. Further, analysing school resources has helped us understand to what extent school input dilution can explain any observed association between enrolment and test scores and how much is left 'unexplained.'

The estimations of this study are not designed to provide causal estimates and so should be interpreted cautiously. Even previous studies that attempted to assess the causal effect of enrolment expansion using fixed effect estimates with extensive child and

household-level data have concluded that the subject matter involves many complexities, and casual estimates should be interpreted cautiously. (Shah and Stenberg, 2019). However, the school-level panel dataset allows us to add a methodological contribution. More central to the current study is that the bulk of previous research on this topic has been cross-sectional. Therefore, this research also contributed to the literature methodologically by using pooled cross-sectional and change-in-change estimates. Unlike a simple POLS model, which is unable to isolate within-school changes, the change-in-change model effectively compares the same school at different points instead of comparing different schools to each other (Kuziemko, 2006).

The simple POLS estimation shows a small positive (except grade 5 after controlling for the year dummies) associations between learning outcomes and enrolment, (i.e., larger schools tend to have better performance). When these advantages are controlled by isolating within-school changes (using a change-in-change-model), an adverse effect of enrolment growth emerges. The change-in-change model shows that a 10% growth rate in enrolment is negatively associated with the pass rate at both primary and middle levels (by 3.6 and 5.8 percentage points respectively). However, enrolment growth is weakly positively associated with the changes in average scores. This pattern of association remains similar for both male and female schools. The implication is that students near the pass/fail boundary performed more poorly in schools that had experienced rapid enrolment expansion. This could reflect a compositional effect, in which the students drawn into participation by the reform programme were relatively weaker academically than the current students. However, it is also possible that resources and support for the lowest achieving students were adversely affected by increasing student numbers. In contrast, higher achieving students were less affected or even benefited from being in larger schools and classrooms, such that test scores, on average, increased slightly in schools where enrolments rose.

The results of this study are broadly consistent with previous work in this field (Lucas and Mbiti, 2014; Valente, 2019). These studies reported that enrolment expansion has negative, but minor, effects on learning outcomes. Given that the PESRP has improved enrolment substantially and opened the door of education for disadvantaged students, and particularly females, across the province, this study concludes with cautious optimism that the PESRP remains successful in increasing access to primary and middle education without worsening the school's learning outcomes for the majority of students. When focusing on the average student, under PESRP-I and II, the Government of Punjab has been successful in achieving the balance between enrolment expansion and learning outcomes. However, for the lowest achieving students, expansion may have had a cost in reducing the likelihood of passing the exit exam and progressing to the next schooling stage.

Further, the mediation analysis shows that rapid expansion was associated with the dilution of resources. With enrolment expansion in any year, student-teacher and student-classroom ratios significantly increased at both primary and middle levels. However, enrolment growth was not associated with a decline in average teacher qualifications. This could be linked to the fact that since 2012, the Punjab government has been actively recruiting qualified teachers. Further, as discussed earlier, because the Government of Punjab was continually trying to improve school capacities through the PESRP, the provision of school facilities (electricity, boundary walls, drinking water, and toilets) is positively associated with the expansion in enrolment.

However, the mediator analysis also revealed that school-level resources did not explain the associations between expansion and learning outcomes in the case of public schools in Punjab. The

effects of most of the school resources on learning outcomes were not statistically significant and/or of opposite in sign to what was predicted, such as higher STR and SCR being associated with positive changes in learning outcomes, and provision of electricity and boundary walls being negatively linked with the learning outcomes. These unexpected estimates point to the presence of remaining omitted variables bias from time-varying confounders.

Indeed, a key limitation of this study was that only a fraction of potential influences on learning outcomes were measured in the dataset. A future study with extended school, child and household level indicators is recommended to explore additional mechanisms linking enrolment expansion to school performance. Importantly, a study incorporating child- or family-level could help us understand whether the decline in pass rates associated with enrolment expansion was due to changes in learning conditions or in the composition of students. Many other school-level factors may also be important. For example, Glewwe and Muralidharan (2016), argue that teacher absenteeism has a more meaningful impact on learning outcomes than the simple student-teacher ratio. Also, looking at the distribution of learning outcomes across different socioeconomic groups might reveal differences in the consequences of rapid expansion across the population. Further, for future studies, it is recommended that the study be replicated in other provinces of Pakistan, where schools have experienced similar expansions in enrolment, to validate or dispute the study results. A future study incorporating more recent data would further enhance our understanding of the ongoing trends and provide valuable insights into the long-term impact of the Punjab Education Sector Reform Programme.

Main conclusion is that with the rapid expansion in enrolment in public schools of Punjab after the introduction of PESRP, there were no substantial adverse effects on learning outcomes for the majority of students. The largely successful achievement of a balance between enrolment expansion and learning outcomes by the Government of Punjab positions them as a compelling case study for similar reforms facing challenges in this regard. However, the finding of small, but non-trivial, negative effects on pass rates merits further consideration of how the lowest-achieving group of students were impacted.

In terms of school inputs, enrolment expansion stretched schools' capacity in terms of STR and SCR. However, this stretch in resources is not significantly associated with variation in test scores, meaning it does not appear to have driven the drop-in pass rates. Therefore, in Punjab, where resources are limited, our findings suggest that a strategy of strengthening existing teacher resources could be more effective in enhancing learning outcomes than simply increasing the number of staff. This could be achieved by better equipping teachers with the necessary tools and providing them with ongoing learning opportunities. For example, giving teachers regular access to workshops, modern teaching materials, and technology could significantly enhance their teaching effectiveness, ultimately improving student learning outcomes (Ellis, 2024). The results of this study suggest that whatever future strategies are used to promote student learning, particular attention should be paid to the lowest achieving students – those most at risk of failing to meet the minimum passing standard at the end of an educational stage – in times of expanding enrolment.

Data availability

Part of the dataset supporting this study's findings is publicly available. The remaining data, specifically related to learning outcomes, is available upon request from the Department of School Education, Punjab, Pakistan, upon official request. However, the data analysis is available from the authors upon

reasonable request and with the permission of the School of Education.

Received: 21 October 2024; Accepted: 7 October 2025;

Published online: 20 December 2025

Notes

- 1 The president of Kenya launched the FPE in the year 2003. Through this programme, school fees were completely abolished in all public schools in Kenya.
- 2 Both PEC (Punjab Examination Commission) and PAPSC publicly share data only at the school level; therefore, this study relies exclusively on school-level data. Obtaining individual-level data, such as socioeconomic background, is challenging in Punjab due to the absence of a child-centred census. Additionally, the last population census was conducted in 2017, following a long gap since 1998, making it difficult to access up-to-date demographic information at the student level.

References

- Abdullah NA, Saeed Akhtar M (2019) A document analysis of PESP II: UK aid programme for education in Pakistan. *J Elem Educ* 29(1):47–60
- Adukia A (2017) Sanitation and education. *Am Econ J: Appl Econ* 9(2):23–59. <https://doi.org/10.1257/app.20150083>
- Alcott B, Rose P (2015) Schools and learning in rural India and Pakistan: who goes where, and how much are they learning? *Prospects* 45(3):345–363
- Alcott B, Banerji M, Bhattacharjee S, Nanda M, Ramanujan P (2020) One step forward, two steps back: transitions between home, pre-primary, and primary education in rural India. *Comp J Comp Int Educ* 50(4):482–499
- Ansari A (2021) Evaluating the Effectiveness of Punjab's Public Private Partnership Programmes in Education (Doctoral dissertation)
- Arshad M, Tayyab M (2019) Assessing the effects of school support facilities on Academic Achievement at Punjab Education Foundation Partner Schools. *Eur Online J Nat Soc Sci* 8(2):214–222
- Arshad M, Qamar ZA, Gulzar FH (2018) Effects of Physical Facilities at Public Schools on Students' Achievement in Punjab, Pakistan. *Glob Soc Sci Rev* III(IV):102–113. [https://doi.org/10.31703/gssr.2018\(iii-iv\).07](https://doi.org/10.31703/gssr.2018(iii-iv).07)
- Aslam M, Malik R, Rawal S, Rose P, Vignoles A (2019) Do government schools improve learning for poor students? Evidence from rural Pakistan. *Oxf Rev Educ* 45(6):802–824
- Banerjee A, Glewwe P, Powers S, Wasserman M (2013). Expanding access and increasing student learning in post-primary education in developing countries: A review of the evidence. Abdul Latif Jameel Poverty Action Lab
- Barber (2012) The Good News from Pakistan. *New Statesman*, 137(4884), 23. <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=29979842&site=ehost-live>
- Barrera-Osorio F, Ganimian AJ (2016) The barking dog that bites: test score volatility and school rankings in Punjab, Pakistan. *Int J Educ Dev* 49:31–54
- Barrett P, Treves A, Shmis T, Ambasz D (2019) The impact of school infrastructure on learning: a synthesis of the evidence
- Bold T, Kimenyi M, Mwabu G, Sandefur J (2011) Does abolishing fees reduce school quality? Evidence from Kenya. *Cent Study Afr Econ Working Pap* 4:1–47
- Bold T, Kimenyi M, Mwabu G, Sandefur J, James CD, Hanson K, Mcpake B, Gwatkin D, Hopwood I, Kirunga C, Meesen B, Morris SS, Meessen B, Tibouti A, Laokri S, Weil O, Maxime K (2014) Can free provision reduce demand for public services? Evidence from Kenyan Education. 293–326. <https://doi.org/10.1093/wber/lht038>
- Chimombo JPG (2005) Quantity versus quality in education: case studies in Malawi. *Int Rev Educ* 51(2–3):155–172
- Cook TD, Campbell DT, Shadish W (2002) *Experimental and quasi-experimental designs for generalised causal inference* (1195). Boston, MA: Houghton Mifflin
- Crouch L, Vinjevold P (2006) South Africa: access before qualit, and what should we do now? *Profr Rev de Curriculum y Formón del Profr* 10(1):16–16
- Duflo E, Dupas P, Kremer M (2012) School governance, teacher incentives, and pupil teacher ratios experimental evidence from Kenyan schools. 9–25
- Egalite AJ, Kisida B (2016) School size and student achievement: a longitudinal analysis. *Sch Effect Sch Improv* 27(3):406–417
- Ellis E, Alexander S, Altepeter C, Ferguson A (2024) *Teacher Recruitment and Retention: Leadership Strategies for Equity and Workforce Stability* (Doctoral dissertation, Saint Louis University)
- Gillani A (2021) The association between the presence of sanitation facilities and school enrolment in Pakistan. *World Dev Perspect* 21(July 2020):100289
- Glewwe P, Kremer M (2006) Schools, teachers, and education outcomes in developing countries. *Hand Econ Educ* 2:945–1017
- Glewwe P, Kremer M (2006) Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries. *Handb Econ Educ* 2(06):945–1017. [https://doi.org/10.1016/S1574-0692\(06\)02016-2](https://doi.org/10.1016/S1574-0692(06)02016-2)

- Glewwe P, Muralidharan K (2016) Improving education outcomes in developing countries: evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education* 5:653–743. Elsevier B.V. <https://doi.org/10.1016/B978-0-444-63459-7.00010-5>
- Habib M (2013) Education in Pakistan's Punjab: Outcomes and Interventions. In *Lahore J Econ* 18:21–48
- Halai A, Durrani N (2020) School Education System in Pakistan: Expansion, Access, and Equity. In *Handbook of Education Systems in South Asia*
- Hanushek EA, Jamison DT, Jamison EA, Woessmann L (2007) Education and economic growth: It is not just going to school but learning something while there that matters. *Education* 8(2):62–71
- Hayat FA (2017) The relationship between access to toilets and school. Graduate School of Arts and Sciences of Georgetown University, Washington, DC. https://repository.library.georgetown.edu/bitstream/handle/10822/1044655/Hayat_georgetown_0076M_13782.pdf?sequence
- Kuziemko I (2006) Using shocks to school enrolment to estimate the effect of school size on student achievement. *Econ Educ Rev* 25(1):63–75. <https://doi.org/10.1016/j.econedurev.2004.10.003>
- Lucas AM, Mbiti IM (2014) Effects of school quality on student achievement: discontinuity evidence from Kenya. *Am Econ J: Appl Econ* 6(3):234–263
- Mcmillen BJ (2004) School size, achievement, and achievement gaps. *Educ Policy Anal Arch* 12(58):12
- Punjab Education Sector Reform Programme (2021) Performance Evaluation of the Punjab Education Sector Programme (PESP2). Review of Education Sector Performance: Final. March
- Shah M, Steinberg B (2019) The right to education act: Trends in enrollment, test scores, and school quality. In *AEA Papers and Proceedings* (Vol. 109, pp. 232–238). 2014 Broadway, Suite 305, Nashville, TN 37203: American Economic Association
- Sifuna DN (2007) The challenge of increasing access and improving quality: An analysis of Universal Primary Education interventions in Kenya and Tanzania since the 1970s. *Int Rev Educ* 53(5–6):687–699. <https://doi.org/10.1007/s11159-007-9062-z>
- Taylor S, Spaull N (2013) The effects of rapidly expanding primary school access on effective learning: The case of Southern and Eastern Africa since 2000. Working Pap 41:47–59. <https://doi.org/10.1016/j.ijedudev.2014.12.001>
- UNESCO (2016) Global Education Monitoring Report 2016 [online]. Available at: <http://en.unesco.org/gem-report>
- Unterhalter E (2023) An answer to everything? Four framings of girls' schooling and gender equality in education. *Comp Educ* 59(2):145–168
- Valente C (2019) Primary education expansion and quality of schooling. *Econ Educ Rev* 73. <https://doi.org/10.1016/j.econedurev.2019.101913>
- Wooldridge JM (2012) Pooling cross sections across time: simple panel data methods. *Introd Econ Mod Approach*, 5th, 448–483. <http://www.amazon.com/dp/0324113641>
- World Bank (2012) The World Bank Annual Report 2012. Retrieved from <https://documents.worldbank.org/curated/en/168831468332487486>
- World Bank (2015) Girls' education in the 21st century gender equality, empowerment, and economic growth. Retrieved from <https://siteresources.worldbank.org>
- World Bank (2017) World development report: Learning to realise education's promise
- World Bank (2018) Pakistan Development updates. Retrieved from <https://openknowledge.worldbank.org/handle/10986/30527?show=full>
- Zuze TL, Leibbrandt M (2011) Free education and social inequality in Ugandan primary schools: a step backwards or a step in the right direction? *Int J Educ Dev* 31(2):169–178. <https://doi.org/10.1016/j.ijedudev.2010.06.013>

Acknowledgements

I would like to extend my sincere thanks to the Monitoring and Evaluation Department at the Programme Monitoring & Implementation Unit (PMIU), Punjab. Their support was instrumental in facilitating access to the data and in helping us understand the coding and structure of the original dataset. This research forms part of the author's PhD thesis, which was fully supported by the University of Bristol. No specific funding was received for the preparation or publication of this manuscript.

Author contributions

S.A. led the research design, data collection, analysis, and manuscript writing. L.W. (contributing author) provided supervisory support and made substantial contributions

to the development of the methodology and refinement of the manuscript. Both L.W. and S.A. reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval

This study did not involve any direct interaction with human participants. The analysis was based exclusively on anonymised and aggregated administrative datasets obtained through an email request to the monitoring & evaluation unit, data custodian of the School Education Department (SED), Government of Punjab, Pakistan. A part of this dataset is used in this study is also publicly accessible via the Programme Monitoring & Implementation Unit (PMIU) portal of the School Education Department, Government of Punjab: <https://openpunjab.pesrp.edu.pk/schools/>. According to the Punjab Transparency and Right to Information Act (2013), anonymised public data may be used for research without requiring separate institutional ethics approval. Additional safeguards are outlined in the Prevention of Electronic Crimes Act (PECA, 2016) and the Draft Personal Data Protection Bill (2023), which regulate the lawful use of personal data and permit research use of anonymised data under appropriate safeguards. In the United Kingdom, the study complies with the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018, which allow the use of anonymised data for research purposes. Because no personally identifiable information was accessed and no direct contact with individuals occurred, formal ethics approval was not required. The research was carried out in accordance with the principles of the Declaration of Helsinki, ensuring that the rights, dignity, and privacy of individuals were respected throughout.

Informed Consent

This study did not involve direct participation of human subjects. The data analysed were collected through official administrative procedures of the PEC and the School Education Department (SED), where schools and students provide information as part of mandatory registration and examination processes. Consent for data collection and its use were therefore embedded within these official procedures under the Ministry of Education, Government of Pakistan. The datasets made available by PMIU and PEC are deidentified prior to public release. In accordance with the Personal Data Protection Bill (2023) of Pakistan, anonymised data may be used for research, archiving, or statistical analysis with appropriate safeguards for privacy. Similarly, the UK GDPR and the Data Protection Act 2018 permit the use of anonymised data for research without requiring individual informed consent. Because this study relied exclusively on secondary, anonymised, and publicly accessible datasets, no individual informed consent was required.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-025-06080-4>.

Correspondence and requests for materials should be addressed to Saba Arshad.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025