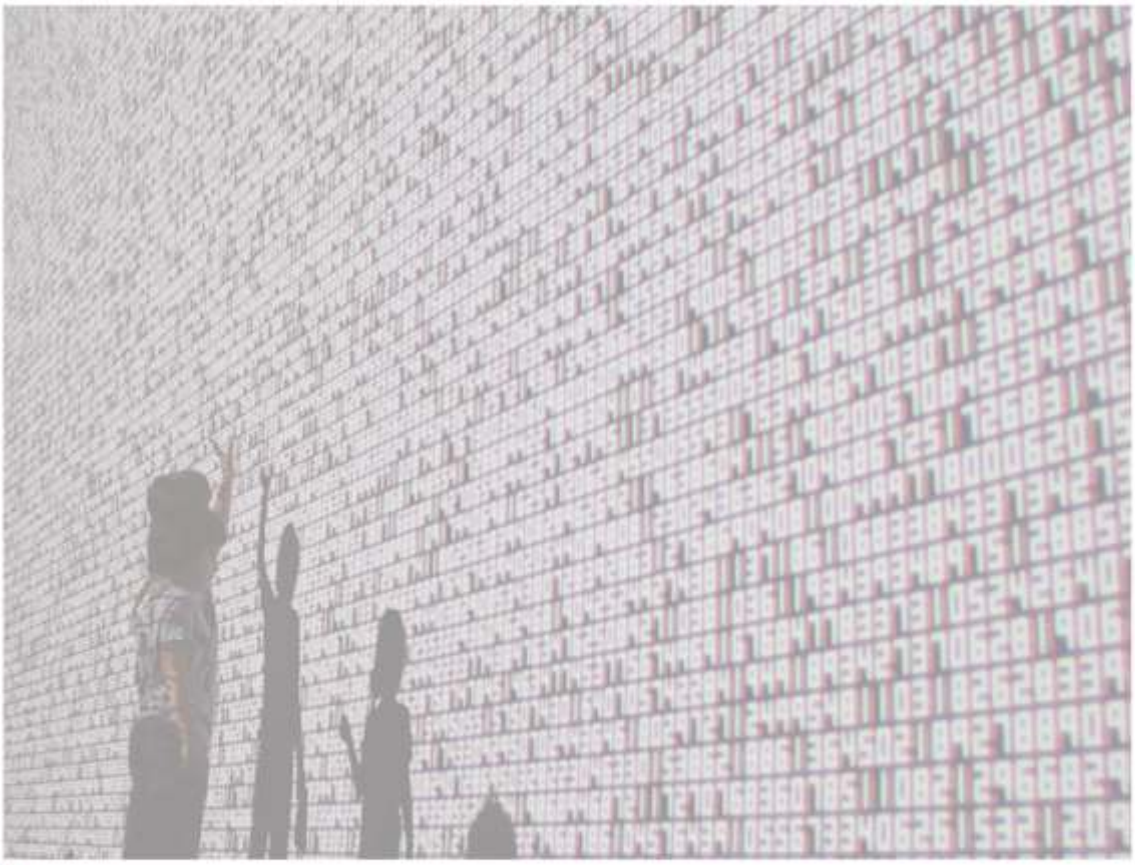
Campaign for Quality Education

# Academic achievement, school performance and spatial heterogeneity



OPEN SOCIETY FOUNDATIONS

# Academic achievement, school performance and spatial heterogeneity

Dr. Muhammad Farooq Naseer
Assistant Professor
Lahore University of Management Sciences

Society for the Advancement of Education

The Society for Advancement of Education (SAHE) is a nongovernmental organization established in 1982 by a group of concerned citizens and academics. It builds on the belief that educational justice entails not just access to school, but to quality education, for all children in Pakistan. SAHE works through an extensive network, the Campaign for Quality Education (CQE), to conduct collaborative research and evidence-based advocacy on key issues to influence educational reform. It has sought such evidence in the realm of data related to school inputs and student outcomes, budgetary analysis, public sector reform and privatization, teacher professional development, language as well as citizenship education.

The research paper has been produced with the support of Open Society Foundations (OSF) and SAHE. Interpretations in the research paper are those of Dr. Muhammad Farooq Naseer and do not necessarily reflect the views of OSF and SAHE.

Copyright © 2013 Society for Advancement of Education (SAHE)

The use of any material in this publication is to be acknowledged.

Published by:
Society for Advancement of Education
65-C Garden Block, New Garden Town, Lahore, Pakistan
www.sahe.org.pk & www.cqe.net.pk

Printed: Balance Publicity
Cover design & layout: Mazhar Saleem

# Foreword

Reliable and valid evidence on access as well as quality dimensions of public education is crucial for determining the direction of education policy and making wise allocation decisions. Yet, amidst the plethora of problems inherent in the management of the education sector in Pakistan, a pressing issue remains the inadequate use of evidence to support policy level decisions and monitoring implementation of policies. This inadequacy does not necessarily imply the absence of data. In fact, all provinces are currently supporting an Education Management Information System (EMIS). However it means that the existing data is not used to inform policy and, as importantly, the quality of data can only improve through the routine use of such data.

Recognizing the opportunity provided by the existence of large school input as well as student outcome data sets in Punjab (the latter being unique to the province), CQE took the initiative to bring the two sets together in order to better understand the correlation between various school inputs and student achievement. CQE reviewed the indicators available in these databases and joined the sets in order to conduct the requisite analysis aimed at school improvement. CQE further explored the situation on the ground by visiting high and low performing schools in district Jhelum, in partnership with the International Growth Center (IGC).

This research paper is based on the findings of this endeavor carried out in collaboration with Development Policy Research Center (DPRC) at LUMS. By combining the datasets we identified learning gaps and derived school-level and regional indicators of learning achievement in Punjab; secondly, we correlated school performance with school inputs and facilities and lastly, examined the spatial heterogeneity in exam scores as a prelude to "mapping" educational achievement.

We would like to gratefully acknowledge the cooperation extended by the PMIU and Punjab Examination Commission (PEC) in making this important venture feasible. We would also like to thank the Open Society Foundation (OSF) for generously supporting this initiative.

Abbas Rashid
Convener, CQE
Executive Director, SAHE

# Contents

# List of Tables and Figures

# Introduction

Improving the quality of education in Pakistan from its current abysmal state is an important policy imperative. Based on results by testing children in a random sample of households, a study by Das et.al. (2006) concluded that children in grade 3 had barely mastered the curriculum for grade 1. Importantly, much of this difference in child learning could be explained by variation in school characteristics rather than individual, household or community level attributes. In particular, children in private schools were found to be learning better than their counterparts in public schools. While there is considerable evidence for the superior learning outcomes from private over public schooling (for instance; Alderman et. al. 2001,& Das et. al. 2006), it is clear that the quality of learning in both types of schools needs significant improvement for Pakistan to reap a demographic dividend.

The education literature is less conclusive on the most effective means of improving the quality of education in schools. For instance, the reasons for improved private school performance in Pakistan do not seem to lie in the academic or professional qualification, salaries or experience of the teaching staff. What seems to matter most is the difference in teachers' incentives between both the types of schools. While the private schools are disciplined by market competition and the profit motive, there is no universally applicable measure of school performance and teacher accountability in the public sector schools.

Given this situation, the Punjab Examination Commission (PEC) was formed in 2006 to conduct standardized exams for grades 5 and 8, and children enrolled in public schools throughout Punjab were required to take these mandatory PEC exams. Before the inception of PEC, exams for these levels were designed in an ad-hoc fashion and marked locally at the district level. Hence these earlier exams varied in quality and curricular relevance and were incomparable across districts.[1]

PEC provides an educational assessment system designed to allow comparison of the educational achievement of schools in Punjab as well as to inform various teacher training and curriculum development initiatives. The PEC website lists the following potential uses for its exam data: "To help determine, i) School wise and gender wise performance of students, ii) comparison of results area wise and public vs. private institutions, iii) areas needing concentration of the administrators, iv) areas wherein teachers require training, v) short-comings/weak areas of content/text-books, vi) the quality of education in schools of the Punjab."

It should be noted that the establishment of PEC is an initiative of the Punjab Education Sector Reform Programme (PESRP) which started in 2003, with financial support from the World Bank. One of the goals of the PESRP program was to introduce provincial wide reforms to improve quality, access, and governance of education. Initially a funding ofapprox $ 600 million was provided from 2003 to 2009. Subsequently, another $300 million credit was approved for extending the program for another three years from 2009-2012. PESRP also established a Programme Monitoring and Implementation Unit (PMIU) outside the provincial Department of Education, tasked with maintaining an adequate financial management system for tracking funds, expenditures and outputs down to the district level.

The PMIU is targeting and accessing data from 62,000 public schools across Punjab for which the Education Management Information system (EMIS) has been revamped and absorbed by the Program. The PMIU conducts its own annual school census as well as collects periodic teacher surveys. Furthermore, PMIU performs ongoing monitoring through district-based field-monitoring teams made up of Monitoring & Evaluation Assistants (MEAs).This monitoring data is collected monthly, for which the PMIU developed some core indicators to measure various school inputs and their progress over time. These monthly monitoring reports are designed to give information and feedback to the districts on the status of school infrastructure and teacher attendance.

[1]For a detailed review of testing conducted by the Punjab Examination Commission, see SAHE (2011).

The objective of this report is to utilize existing government databases such as PEC, EMIS and aggregated data from PMIU Monitoring Reports to generate evidence for education policy formulation and evaluation in Punjab. This is in contrast to other recent education studies on Pakistan such as Annual Status of Education Report (ASER) and Learning and Educational Achievements in Punjab Schools (LEAPS), which have based their insights on conducting large household surveys. Combining and analyzing these disparate education system data sets is challenging, however if institutionalized, this exercise can yield a regular flow of information for the concerned departments in key areas of education reform.

We combine data from these multiple sources to: i) compute school-level and regional indicators of learning achievement in Punjab in order to identify potential learning gaps; ii) correlate school performance with school inputs and an index of school facilities and iii) overlay the spatial heterogeneity in exam scores with other indicators to construct a "learning map" of Punjab.

# Literature Review

The literature regarding the impact of school resources on the quality of education is varied and extensive. Several school inputs have been shown to positively correlate with the quality of education such as having a playground, and spending on school facilities as in a proper hygienic source of food and sanitation, and the number and quality of teaching staff.[2]

While most of these studies produce suggestive evidence on the relation between schooling inputs and learning, attributing causal impact is challenging. The foremost problem facing the majority of these studies is the endogeneity of school inputs - there are many unmeasured and unobserved influences by individual, family and neighborhood factors that conflate the estimates of school resource effects. Another issue is the specification to be used in a particular analysis, since student test scores are affected by both historic and contemporary factors. After adopting various ways to address these two major concerns (among others), most studies conclude that there is no strong or consistent relationship between changes in school resources and student performance. However this does not imply that schools don't make a difference, but merely that the most commonly measured resources do not capture the quality differences in instruction that matters most for producing knowledge (Hanushek, 1997).[3]

Rivkin, Hanushek and Kain(2005) use a rich longitudinal data set on student achievement in the State of Texas to control for fixed effects to account for student, school-by-grade, as well as school-by-year effects. The remaining differences in academic results between grades and cohorts are associated with disparities in school characteristics or teachers. As Loeb and McEwan (2010) describe "they find that a one standard deviation increase in average teacher quality for a grade raises average student achievement in the grade by at least 0.11 standard deviations of the total test score distribution in mathematics and 0.095 standard deviations in reading".[4] The study concludes that as opposed to more restrictive certification standards, effective hiring, promotion and firing practices can improve teacher quality (Rivkin, Hanushek and Kain, 2005).

Meanwhile, research on the impact of ability grouping in schools has yielded mixed results. Arguments for ability grouping focus on effectiveness and the claim that it allows teachers to adapt their instruction according to the ability of their students. Arguments against it assert that it disadvantages lower track students, who receive a poorer quality of instruction, face lower expectations and no behavioral models (Slavin, 1990). A synthesis of the best available evidence conducted by Slavin shows that the effects of ability grouping or tracking (as opposed to no ability grouping) are essentially zero (1990). However this does not include the kind of ability grouping, the number of groups students are assigned to and the duration of study.

An interesting study on the impac t of school inputs by Eide and Showalter (1998) investigates how their effect differs across 'quantiles' in the distribution of math test scores. Using variables such as pupil–teacher ratio, per pupil expenditures at the district level, the fraction of teachers with an advanced degree, enrolment at the school, the length of school year, as well as controls for a number of family, race and neighborhood variables, they find that performance at the top of the math score distribution improves with a lengthened school year, while performance at the bottom stays unaffected. In other words, with things constant smarter students benefit from a longer school year end more teaching time, as compared with those struggling to learn at the bottom of the learning curve.

Moreover, there is a sizable literature on the factors affecting schooling outcomes in Pakistan and many of its conclusions are in line with the aforementioned. The LEAPS Project, a five-year project on Learning and Educational Achievements in Pakistan, finds that higher teacher

[2]More formally, the schooling input variable affects the level of human capital inculcated to the students. The greater the quality of student input, the higher the human capital accumulated.

[3]Hanushek(1997) concludes that centralized and uniform increases in resources are unlikely to have any positive effects, but effective utilization of additional resources along with appropriate incentivization and reward policies for students and teachers could improve learning outcomes.

[4]This also implies that improving teacher quality is a more cost effective way to improve student outcomes than reductions in class size

qualification does not increase student test scores significantly, but increased teacher effort (quantified by teacher attendance rates) results in a remarkable increase (Andrabi et al, 2007). Apart from teachers, no single input can be consistently linked to test scores; learning outcomes are found to be similar across poor and rich households, and more literate and less literate villages. Instead the variation is at the school level, with a large portion of it coming from differences across public and private schools (Andrabi et al, 2007).

According to ASER report, private schools generally have a higher student (90% vs. 85%) and teacher (90% vs. 87%) attendance than government schools. Compared to government schools, a greater proportion of private schools have useable toilet and water facilities, a boundary wall and a playground. A higher percentage of students in private schools are attending out-of-school paid tuition (25.3% compared to 9.7% in public schools) and the average number of rooms for conducting classes is also greater in private schools (3.6 vs. 2.9 rooms, at the primary level) (ASER 2010).

National assessment findings of 2006 presents another related result whereby it concludes that students who have never been punished by their teachers tend to do better, as do students whose performance is rewarded. The forms of assessment identified as being positively correlated with students' achievement in all subjects are written assessments and the assessment of homework (as opposed to verbal assessment, assessment of class performance etc).

On the impact of teacher quality, several studies have attempted to break down teacher quality into its component variables, with one strand of literature focusing on the effects of on-the-job training. The main problem facing such studies is that teacher training is often determined by schools and teachers themselves, resulting in a selection bias in most empirical results.

To overcome this problem, Angrist and Lavy(1998),using a sample of some Jerusalem elementary schools that received funding earmarked for teachers' in-service training, show that after controlling for initial differences in test scores, pupils enrolled in schools where teachers receive in-service training perform better than those enrolled in schools where they do not.

On the other hand, Jacob and Lefgren (2004) reach different conclusions. Using data from Chicago public schools,[5] the authors use an instrumental variables approach to model the effect of training while controlling for race, gender, socioeconomic background and student ability and find no significant impact of teacher training on student achievement. Hence, it may be concluded from the above that the quality and context of a training program has a significant effect on its outcome.

[5]They use data from Chicago public schools where, in 1996, schools where less than 15 percent of students performed at or above the national mean in standardized reading tests were put on probation.

# Data

For the purpose of this report, we analyze data from PEC on individual exam scores for the class 5 and 8 examinations held in 2008 and 2009. In addition, we use information from the EMIS annual school census for the years 2008 and 2009. This dataset is available for government schools (only) and includes information on various school attributes, such as the school type, gender, shift (morning or evening), availability of libraries, laboratories and playgrounds, as well as provision of basic facilities (such as electricity, drinking water and toilets), school council, number of classrooms and basic construction details of the school's building etc. We also had access to district-level data from the PMIU Monitoring Reports for the study period as well as the District Census Report from the Population Census 1998.

In order to conduct school-level analysis, we combine data from PEC and EMIS using unique school identifiers in the two data sets (EMISCODE). There were problems in matching schools based on EMISCODEs, especially for the year 2010 due to which we decided to restrict our analysis to 2008 and 2009 exam years only. Since PEC exams are mandatory for public schools only, the exam data exists only for those private schools which opted for the exam. This implies an important caveat for any public versus private school comparison of scores presented below, as the private schools taking the exam may not be representative of all private schools. Starting with student-level exam score data, we use the average total score obtained by students in each school as a measure of school achievement in this report.

The EMIS data contains information on 49,983 schools from all over Punjab. The location of schools is identified by the administrative unit, including District, Tehsil and Markaz (sub-Tehsil), in which the school is located. There are, on average, 1,445 public schools in one district of Punjab with the minimum number of 770 schools in the Gujrat district. Also, about 90% of the schools in the sample are classified as rural schools and the data has almost equal number of schools for boys and girls.

It may be pertinent to note that most of the public schools were established before 1980. Furthermore, based on the 2008 EMIS data, majority of the schools in the sample had a facility of clean drinking water, electricity, toilets, sewerage system and completely built school boundary wall. On the other hand, majority of the schools did not have a libraryor a playground.

As far as the teachers' academic qualification is concerned, half of the teachers in about 50% of the schools possess graduate degrees, while about 25% schools have a graduate among every 2 out of their 3 teachers. Similarly, the mean experience of the teachers in education sector is 17.2 years, with 10% of the teachers having experience of over 24 years. Majority of the schools in the data set are primary schools, and most primary schools have 2-3 teachers to teach the six classes from katchi to class 5. The student to teacher ratio on average is 56, while about 10% of the schools in our sample have a student to teacher ratio of over 100. All the key variables are summarized in Table 1.

[6]About half of the schools were established before 1972, while about three-fourth of the schools in our sample were established before 1984.
[7]About 86.69% have facility of clean drinking water.
[8]About 59.17% have electricity connection available.
[9]About 76.52% have useable toilets.
[10]About 56.81% have proper sewerage system.
[11]About 79.02% have completely built school wall
[12]About 66.51% do not have a library.
[13]About 54.55% do not have a playground.
[14]About 77% of the schools have class 5 as the highest class in the school.

Table 1: Summary Statistics

| Variable | Mean | Standard Dev. | Quartiles | | | Range |
|---|---|---|---|---|---|---|
| | | | 10% | 50% | 90% | |
| Subjects | | | | | | |
| English | 42.79 | 16.11 | 22.6 | 41.1 | 65.3 | [4,96] |
| Urdu | 50.49 | 13.04 | 33.8 | 50.0 | 68.0 | [8,98] |
| Mathematics | 39.61 | 13.00 | 24.2 | 37.8 | 58.0 | [4,94] |
| Science | 43.26 | 13.44 | 26.8 | 41.9 | 62.0 | [8,97] |
| Social Studies | 39.36 | 12.23 | 24.7 | 37.8 | 56.3 | [8,99] |
| Islamiat | 56.15 | 11.50 | 41.0 | 56.8 | 70.4 | [4,96] |
| Total | 271.66 | 65.41 | 191.3 | 266.0 | 362.1 | [44,528] |
| Inputs | | | | | | |
| Teacher | | | | | | |
| Experience of teachers in the education sector | 17.23 | 6.3 | 7.0 | 18.0 | 24.6 | [1,41] |
| Proportion of teachers with graduate degree | 0.42 | 0.4 | 0.0 | 0.5 | 1.0 | [0,1] |
| Proportion of teachers with intermediate degree | 0.16 | 0.2 | 0.0 | 0.0 | 0.5 | [0,1] |
| Student-Teacher ratio | 55.67 | 38.9 | 18 | 48 | 100 | [0,658] |
| Material Inputs | | | | | | |
| Index of School Material Inputs | -0.06 | 1.02 | -1.57 | 0.22 | 1.16 | [-2.9,1.9] |
| Other | | | | | | |
| School Enrolment | 15.92 | 16.07 | 3 | 11 | 33 | [0,443] |

# Methodology

Firstly, we examined the variation in the performance of schools by district. In addition, we compared the variation of test scores among schools located within the same administrative unit (the "within" variance), to variation in mean test scores across administrative units (the "between" variance): District, Tehsil and Markaz. For this purpose, individual subject test scores as well as total scores were used. To check for robustness of these findings, the within- and between- variance was computed for several different study groups: for class 5 and 8 in the year 2008-2010, and again after dropping outlier schools.

Next we examined learning gaps by comparing the difference in average test scores across different school and student types. However, a simple comparison of average scores has the drawback that it does not control for other variables that might also affect school performance. For instance, when comparing rural and urban schools in a simple bivariate analysis, we do not take into account other factors such as the average teacher experience, which might differ significantly across urban and rural schools. Thus, any rural-urban gap in school performance could actually be due to this third confounding factor. To account for this problem in bivariate analysis, we estimate a multiple regression including these other confounding factors as control variables while comparing urban and rural schools.[15]

## Index for School Material Inputs

The effect of school resources on students' performance was measured with an index for school material input, constructed using factor analysis. The purpose of factor analysis is to discover underlying patterns in the variation among a set of variables. As a statistical method, it seeks to discover if the observed variables can be explained largely or entirely in terms of a much smaller set of latent ('unobserved') variables called factors.

In the present context, we employed factor analysis for the purpose of constructing the school resources index. We used school-level variables, such as availability of clean drinking water, electricity, toilets, sewerage system, good building condition, completely built school boundary wall, library, playground, and the age of school[16] to form an index capturing school resources[17]. Table 1 shows summary statistics for the factor index, which is then used in the analysis.

[16] A dummy variable for age of school was made with 1 for schools established after 1926, and 0 otherwise

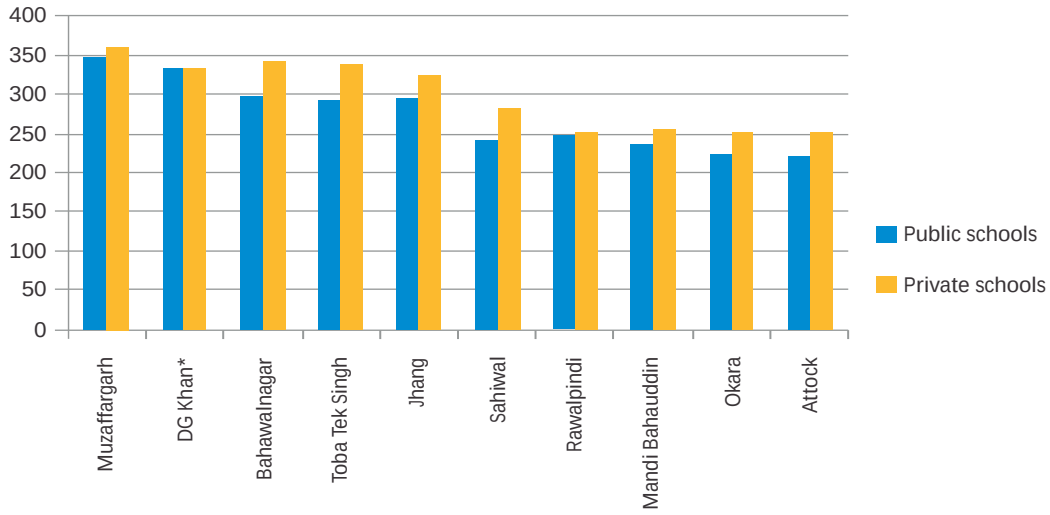
[17] Since all of these were dummy (or binary) variables, we use their polycholric correlations to construct the factor score.

# Results

## Geographic Variation

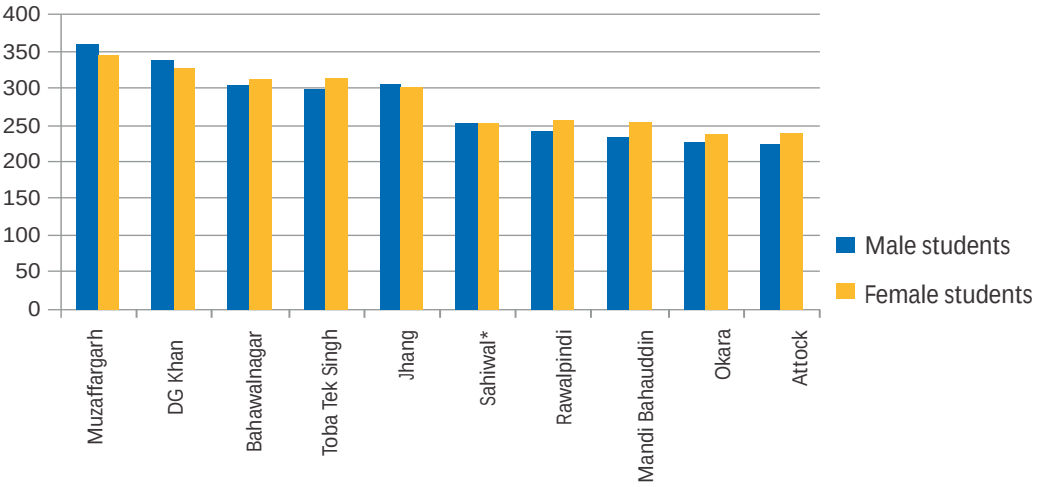
In terms of the average learning achievement at the district level, Muzaffargarh district has the highest overall score followed by DG Khan, Bahawalnagar, Toba Tek Singh and Jhang. In order to see the relative performance clearly, we represent the top 5 and bottom 5 districts in the Figure 1 and Figure 2 below.

Figure 1: Average School Score in Selected Districts by School type




*denotes that the difference is not significant at 5% level between the two groups (public school vs. private school) in this district

Figure 2: Average School Score in Selected Districts by gender




*denotes that the difference is not significant at 5% level between the two groups (male students vs. female students) in this district

Figure 1 shows the difference between public and private schools in these districts, while the other figure shows the difference between the test scores of boys and girls. There is a significant difference in performance of private and public schools in all of the top 5 districts except DG Khan. Further, the performance of boys and girls in all these districts is also significantly different.

On the other hand, district Attock is at the bottom of score distribution with average overall marks less than 250. Attock is followed by Okara, MandiBahauddin, Rawalpindi and Sahiwal. In these districts, also, there is significant difference in the performance of private and public

schools. Moreover, except for Sahiwal, all of the bottom 5 districts have a significant difference in the performance of boys and girls. We see a consistent trend of private schools performing better than the public schools in all these districts. The gender differences reveal that theboys are doing better than girls in Muzaffargarh, DG Khan and Jhang, while in other districts girls are doing better than boys.[18]

## Regional Analysis

In the regional analysis, we divide Punjab into four main regions, namely North, Centre, South and West. FiguresA1 and A2, in the appendix, show the district development index and the literacy rate in each district. The District Development Ranking Index plotted was computed after comparing the district in question with all other districts of Pakistan (based on the Population Census 1998 data as reported in the District Census Reports. District ranking = -1*Ranking Index) while the literacy rate in Figure A2 is measured as the proportion of literate population according to the Population Census, 1998. These figures suggest that the North e.g. Attock, Rawalpindi, Jhelum, etc and Centre e.g. Lahore, Faisalabad, Toba Tek Singh, etc. tend to be more developed and more educated, and some of the districts in these regions have literacy rates of more than 70%. On the contrary, Figures3 and 4below reveal that these two regions perform worse than the South e.g. Multan, Bahawalpur, Rahimyar Khan, etc. and West e.g. Bhakkar, Layyah, Dera Ghazi Khan, etc.

Figure 3: Comparison of different regions by school type

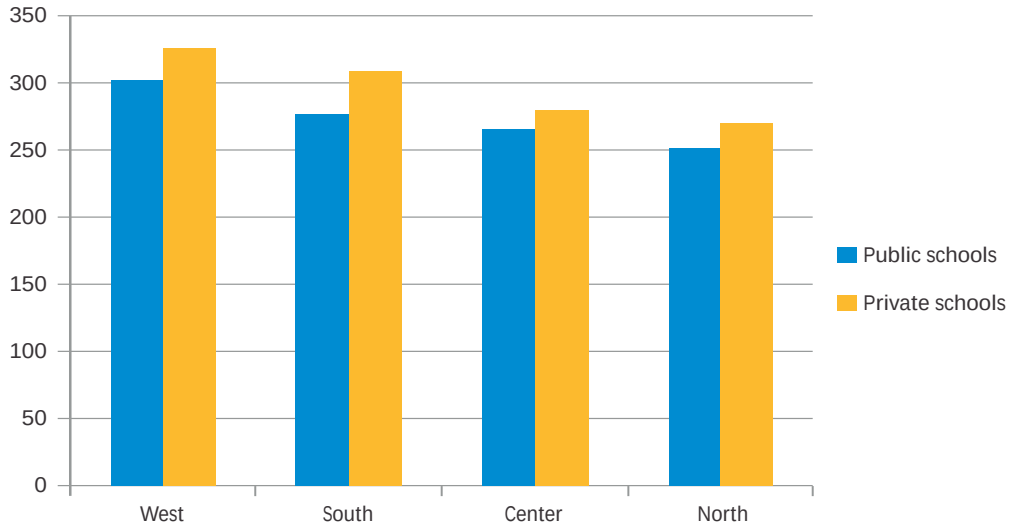
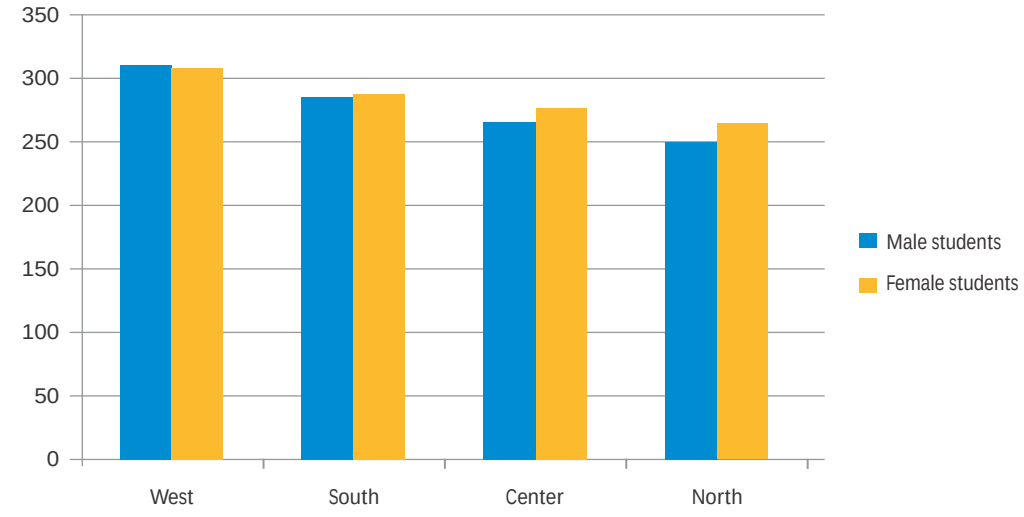



Figure 4: Comparison of different regions by gender




[18]The above analysis is done for class 5 scores in PEC 2009 exams but can be replicated for other levels and years to rank districts on PEC scores.

This is a bit unexpected. Even though, districts in West and South regions have literacy rates not exceeding 41.75%, the West region has an average overall score of more than 300, followed by the South, Centre and North. The Northern region, with an average overall score of around 250, is performing the worst.

## Within and between variation

The main finding of this report is that the key learning gaps in public sector schools are not regional, rather these exist across schools located within a region or administrative unit. This is because the regional gaps are small in comparison to the variation within a particular region or administrative unit. This overturns the existing policy focus on overcoming regional test score gaps.

It is important to outline the difference between the "between" variation and "within" variation before we proceed. The between-variation tells us how different one school is from another (signified by the "slope" of a regression line where the x-axis has the school EMISCODE). The within-variation then tells us the variance of residuals above and below the line for a given EMISCODE. We found that the two are equally great and if anything, the within-variation is larger.

Comparing variance at Markaz, Tehsil and district level for each subject, we noted that the within Markaz variation is lesser than the within Tehsil variation, which in turn is lesser than the within district variation. However the variation within Markaz is still large. In other words, the Markaz and Tehsil are almost equally heterogeneous as the districts. On the other hand, we noted that the between Markaz variation is more than the between Tehsil variation, which in turn is more than the between district variation. Hence, we see a steady increase in the "between" variation and decrease in "within" variation as we move from district to Tehsil to Markaz. The below mentioned Table 2 summarizes these findings.

Table 2: Between and Within Variance of Test scores by Districts, Thesis and Markaz

| | District | | Tehsil | | Markaz | |
|---|---|---|---|---|---|---|
| | between | within | between | within | between | within |
| English | 6.32 | 14.88 | 7.05 | 14.54 | 8.77 | 13.90 |
| Urdu | 4.84 | 12.14 | 5.56 | 11.92 | 6.76 | 11.51 |
| Mathematics | 4.39 | 12.22 | 5.26 | 11.99 | 6.58 | 11.56 |
| Science | 5.60 | 12.20 | 6.63 | 11.97 | 7.43 | 11.53 |
| Total | 26.49 | 59.86 | 31.33 | 58.39 | 36.63 | 55.86 |

For example, the mean total score is 271.6. The standard deviation of total scores within district is 59.86, while the standard deviation of total scores within Tehsil is 58.39 and the standard deviation of total scores within Markaz, 55.86. The standard deviation of total scores between districts is 26.49, while the standard deviation of total scores between Tehsil is 31.33, and the standard deviation of total scores between Markaz is 36.63.

Table 2 also contains within- and between- variation for four subjects i.e. English, Urdu, Mathematics and Science, and the subject scores show similar patterns as the Total PEC score. Finally, as a robustness check on our findings, we check for within- and between- variations across districts, Thesis and Markaz for class 5 in 2008 and class 8 in years 2008 and 2009, and find similar patterns in all the cases. Further, if we drop the high performing and low performing outliers from our sample, i.e. dropping bottom 1% and top 1% performing schools, the variance declines but the overall qualitative conclusion regarding the variance decomposition within districts and schools stays unchanged.

To understand this result: imagine you pick a typical (or 'average') school, say school A, in a given district of Punjab. Next you randomly select another school within the same district (school B), as well as an average school in some other random district of Punjab (school C). Which of the last two schools is likely to be of similar quality to school A? The foregoing analysis tells us that school C is expected to have greater similarity in test scores to school A than school B, even though school B is likely to be located in closer physical proximity.

Figure 5: Comparison of between and within variation at the district level

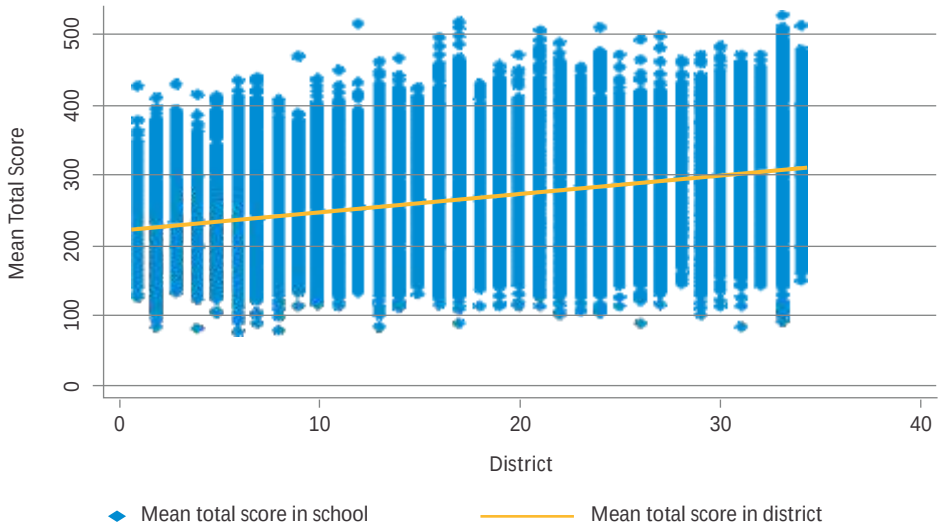



Figure 5 presents the same observation graphically, where a data point corresponds to the average total score obtained by a school and the x-axis value denotes the unique ID given to a district. The smooth upward sloping line connects the mean test scores obtained by the average school in each district. As can be seen here, the variation in test scores across schools within a district (shown by the 'height' of the vertical scatter for a given district),is much larger as compared to the change in mean scores going from left to right (i.e. the difference between the highest and lowest point on the line). Even the highest scoring district (right-most on the graph) has schools that are as bad or worse, than some of the worst performing schools in other less scoring districts. This suggests a potential for improvement in school quality by improved resource management that tries to bridge the learning gaps across schools by focusing attention at the local level.

This is an important finding as in addition to the across-district variation in scores, the high within-Markaz variation and high within-Tehsil variation reveals that there are few Marakiz or Tehsils that have consistently superior education outcomes as measured in terms of uniformly high test scores i.e. higher mean and lower standard deviation. Hence, we conclude that the performance varies a lot within Punjab, which draws attention to further research on the reasons for school-level learning gaps and perhaps improved resource allocation within the province. In the section below, we attempt to further examine average school level learning outcomes using multiple regression analysis.

## Regression Results

Using the set of material inputs to compute a factor index and controlling for other variables that might affect a school's performance, we used a multiple regression to analyze the correlates of test score differences among schools. This is better than a simple comparison of average test scores in different categories of schools and students, as it allows us to compare average test scores while keeping other related factors fixed. The regression analysis is unlikely to reveal cause-effect relationships due to the likely presence of unmeasured, and hence uncontrolled sources of variation in test scores. So the evidence we present in this section should be taken as suggestive of the existence of potential underlying relations among selected factors.

The set of explanatory variables includes, other than urban dummy, gender-of-school dummy, school-level dummy, the proportion of teachers with graduate degrees, proportion of teachers with intermediated degrees, student-teacher ratio, average experience of a school's teachers in the education sector, index for school material inputs, median age of exam-taking students and overall enrolment. These variables can be considered as control variables in examining whether there are significant differences between schools based on different gender, level or location. Table 3 presents the results.

Table 3: Regression Results

| Dependent Variable | Mean Total Score of School | |
|---|---|---|
| Urban | -0.38 | 1.06 |
| Student-to-Teacher Ratio | -0.10*** | 0.01 |
| School Level: High School | -0.10 | 3.83 |
| School Level: Middle School | 7.29* | 3.82 |
| School Level: Primary School | -3.25 | 3.86 |
| School Level: Madrasah | 4.37 | 5.45 |
| Proportion of Teachers with Graduate degree | 8.00*** | 1.30 |
| Proportion of Teachers with Intermediate degree | 4.81*** | 1.27 |
| Experience of Teachers in Education Sector | 0.29*** | 0.07 |
| Enrolment | 0.20*** | 0.02 |
| Index of School Resources | 0.51* | 0.31 |
| Median Age of Students taking PEC Exam | -4.57*** | 0.32 |
| Gender: Male | -5.15*** | 0.59 |
| District Fixed Effects | Yes | |
| Number of Observations | 48,188 | |
| F (46, 48141) | 242.90 | |
| P-value | 0.00 | |
| Adjusted $R^2$ | 0.1788 | |

Notes: Standard errors are robust to heteroskedasticity
Significance levels: * for 10%, ** for 5% and *** for 1%

The results reveal that once the set of control variables are taken into account, the location of the school as rural or urban makes no difference. Thus, we can claim that on average there is no significant difference in the performance of schools in the urban locality as compared to those located in rural areas, once the differential school resources are taken into account.

Similarly, there is no significant difference in performance of Higher Secondary Schools compared with Primary Schools, High Schools and Madrassah Schools, once we take into account additional school inputs. On the other hand, on average, the Middle School students tend to obtain approx 7.3 marks higher than those from the Higher Secondary Schools. Comparing results across schools for different genders, the regression analysis revealed that schools catering to female students on average tend to score 5.2 marks higher than the male schools, keeping all else fixed.

Further, we analyzed various other factors that might affect school performance on PEC exams. The student-to-teacher ratio has a significant effect on the performance of the students. On average, schools with one additional student per teacher tend to score about 0.10 marks lower in examinations, as compared to other schools. Similarly, the proportion of graduate teachers in the school and the proportion of teachers with intermediate degrees, positively contribute to the school's performance. The coefficient suggests that a school with no graduate teacher will be expected to have on average, 8 marks lower than a school with all the graduate teachers, while everything else is the same.[19] Similarly, a school with no teacher holding an intermediate degree, will on average tend to have 4.8 marks lower than a school with everything else the same, but its teachers holding intermediate degrees. Moreover, the average experience of teachers in the education sector also positively (and significantly) affects the school's performance. Other explanatory variables fixed, a school with an additional year of teachers' combined experience will tend to get 0.3 marks higher in the average PEC score.

The index of school material inputs generated using the set of binary variables mentioned in the

[19]The relation does not suggest causality; it merely suggests that as the proportion of graduate teachers increase, the average school PEC score tends to go up as well, without claiming that the former causes the latter change.

earlier part of the paper, positively affects school performance. A school with higher material inputs i.e. a school having completely built boundary wall, library, playground etc, will have a higher index value, and thus will tend to produce better results than a school with lower material inputs. An increase in the index value of one unit is associated with a 0.51 marks improvement in the average test score of the school.

Another variable considered in our analysis is the total school enrolment. Our regression analysis reveals that school enrolment positively and significantly affects the school's performance. Specifically, we would expect an increase in school's enrolment by one student to increase the average school scores by 0.20 marks, which is sizeable. However, this result should be interpreted with caution and the dangers in interpreting regression results too literally should be considered. The result simply sheds light on the fact that larger schools tend to have better results on performance than smaller schools (which could be due to a selection effect, in the sense that it is the “better quality” schools that attract more children and end up having higher enrolment). It should be noted that the result does not imply that enrolment in every school should be increased to achieve better results.

The median age of the students taking the examination is inversely and significantly related to the school's performance. The regression analysis suggests that a school with a higher median age of one year, will tend to have on average 4.6 marks lower than the other school with everything same, but the median age of one year less. This result seems less surprising when we consider that schools with a higher median age may have more proportion of repeaters in grade, or may have more proportion of students who give lesser importance to education, and more importance to other things such as earning money.

Overall, the multiple regression model above explains about 18% of the variation in the schools' performance. The remaining variation may arise due to student-specific or other idiosyncratic factors, such as with PEC exam administration, instead of school level variables and inputs.

# Conclusion

In conclusion, one primary finding from this report is that there is substantial variation in learning outcomes across public sector schools in Punjab. The more developed regions within the province, Centre and North Punjab, do not yield stronger academic performance than the less developed West and South Punjab. Moreover, there is tremendous heterogeneity within the administrative units, i.e. districts, tehsils and marakiz ,in student performance on PEC exams and, by implication, the quality of education being imparted. It is observed that the public system of education comprises of schools with varying quality levels.

The school inputs explain a sizable portion of this variation across schools. The qualification and experience of teaching staff are associated with higher test scores, and the student-teacher ratio is negatively related. Similarly, school facilities' index (based on the building structure, availability of water, sanitation and boundary wall etc.), is also positively correlated with the average score on the PEC exam. However, we found that the test scores of rural and urban schools are statistically neither different nor is there evidence of a large and consistent gap across different school levels, i.e. primary, middle, high and higher secondary schools. While these correlations are suggestive, further research is required for more conclusive evidence on the causal impact of various school inputs on the learning outcomes. However based on this analysis, one can draw the following policy conclusion.

Given that the heterogeneity in school quality within a district dwarfs any variation in “average” school quality across districts, as far as learning outcomes are concerned, there is little value in taking a district as the ultimate unit of policy analysis and implementation. The education reform effort in Punjab would instead do well to devise a school-focused (and possibly demand-focused) approach.

# References


Alderman, H., P. Orazem, & E. Paterno. (2001). "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan." Journal of Human Resources 36(Spring 2001): 304-326.

Amjad, S., J. Anderson, and A. Sandhu.n.d. Monitoring and Evaluation Proposal for PESRP (Government of Punjab, Pakistan). Mimeo.

Angrist, Joshua D. and Lavy, Victor. "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools." National Bureau of Economic Research. 1998. 30 May 2011 <http://www.nber.org/papers/w6781>

"Annual Status of Education Report (Rural) 2010".South Asian Forum for Education Development Jan. 2011. 20 May 2011 <http://www.safedafed.org/aser/home.html>

Das, J., P. Pandey, & T. Zajonc. (2006). "Learning Levels and Gaps in Pakistan." World Bank Policy Research Working Paper 4067.

Eide, Eric. and Showalter, Mark H. "The Effect of School Quality on Student Performance: A Quantile Regression Approach." Economics Letters 58 (1998): 345 – 350

Goddard, Yvonne L., Roger D. Goddard, and Megan TschannenMoran. 2007. A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. Teachers College Record, 109 (4): 877–896.

Hanushek, Eric A. "Assessing the Effects of School Resources on Student Performance: An Update." Educational Evaluation and Policy Analysis 19.2 (1997): 141 - 164

Jackson, C. Kirabo, and Elias Bruegmann. 2009. "Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers." Cambridge, Mass.: National Bureau of Economic Research, Working Paper No. 15202, August

Jacob, Brian A. and Lefgren, Lars. "The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago." The Journal of Human Resources 39.1 (2004): 50 – 79

Loeb, Susanna, & McEwan, Patrick J. "Education Reforms."*In Targeting Investments in Children: Fighting Poverty When Resources are Limited* edited by Phillip B. Levine & David J. Zimmerman,145-178.University of Chicago Press, 2010.

Ministry of Education, Government of Pakistan, (2009). *National education policy, 2009*

"National Assessment Findings 2006." National Education Assessment System, Government of Pakistan 2006. 20 May 2011 <http://www.neas.gov.pk/Document%20Center.html>

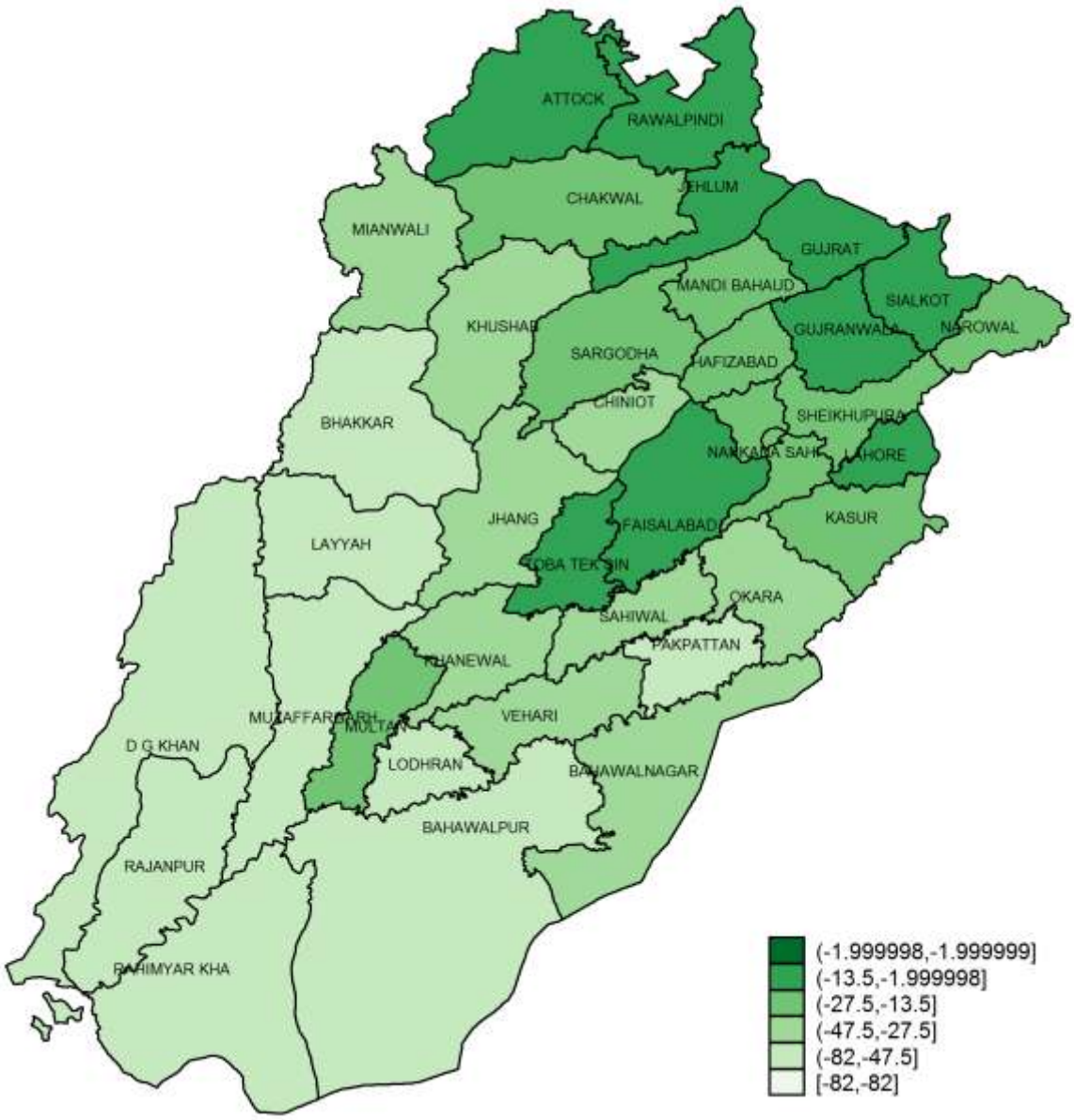
Rivkin, Steven G., Hanushek, Eric A. and Kain, John F. "Teachers, Schools, and Academic Achievement." Econometrica73.2 (2005): 417 – 458

Slavin, Robert E. "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." Review of Educational Research 60.3 (1990): 471-499

Andrabi, T., J. Das, A. Khwaja and T. Vishwanath.2007. "Learning and Educational Achievementsin Punjab Schools (LEAPS):Insights to inform the education policy debate." 20 May 2011 <http://www.leapsproject.org/site/>

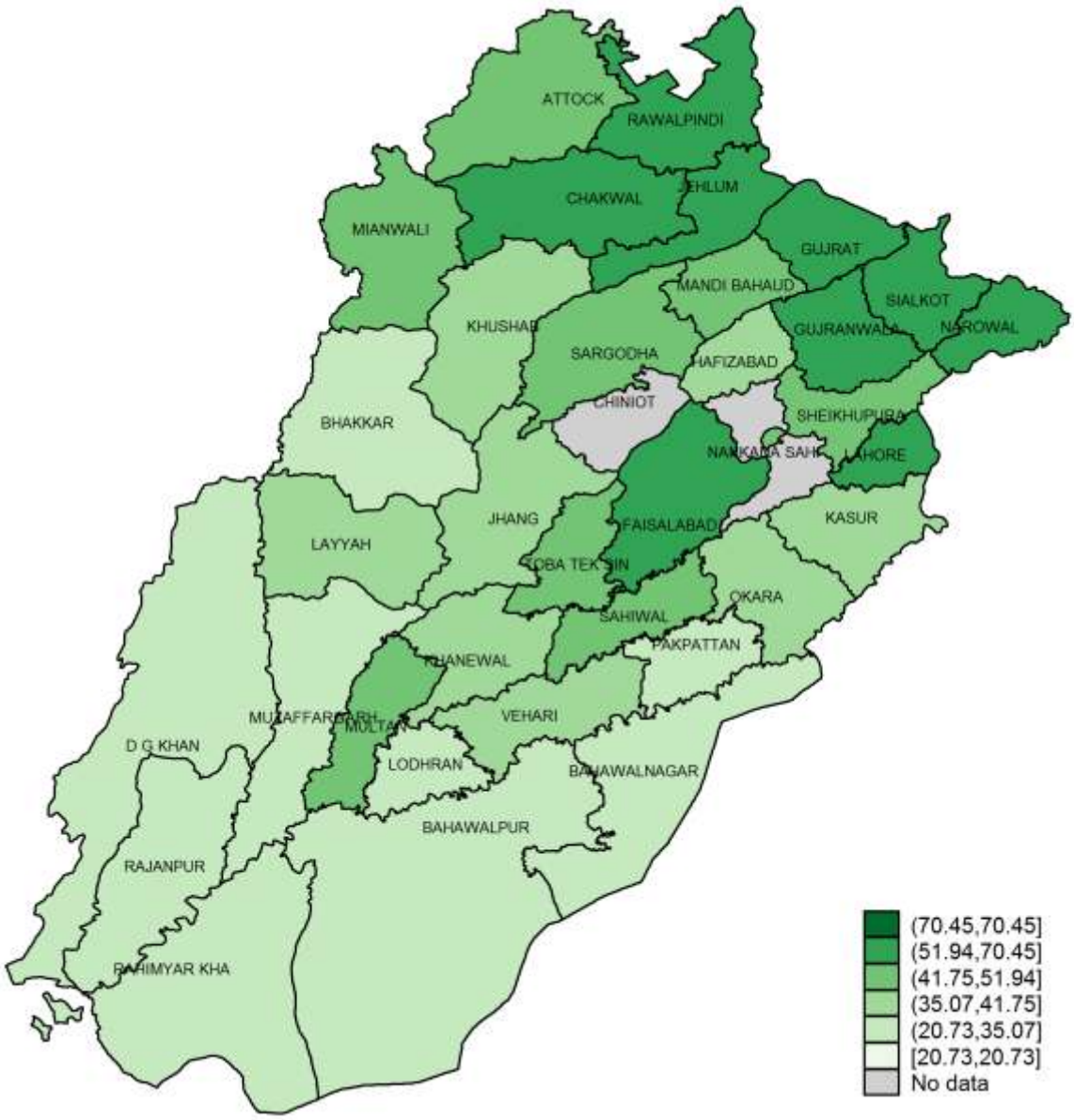
# Appendix

Figure A1: District Development Ranking



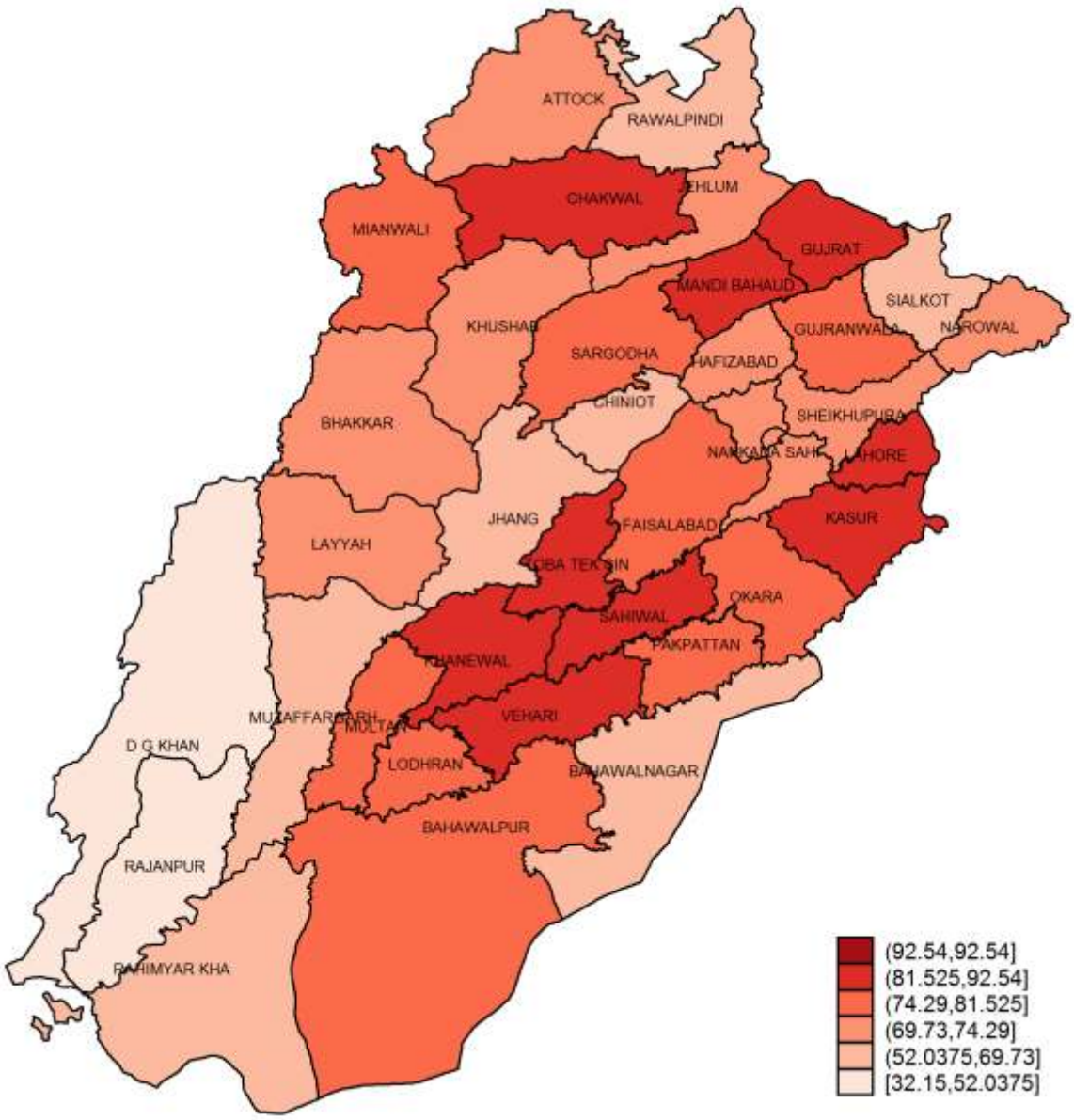

Source: District Census Report 1998

Figure A2: Literacy rate across Punjab



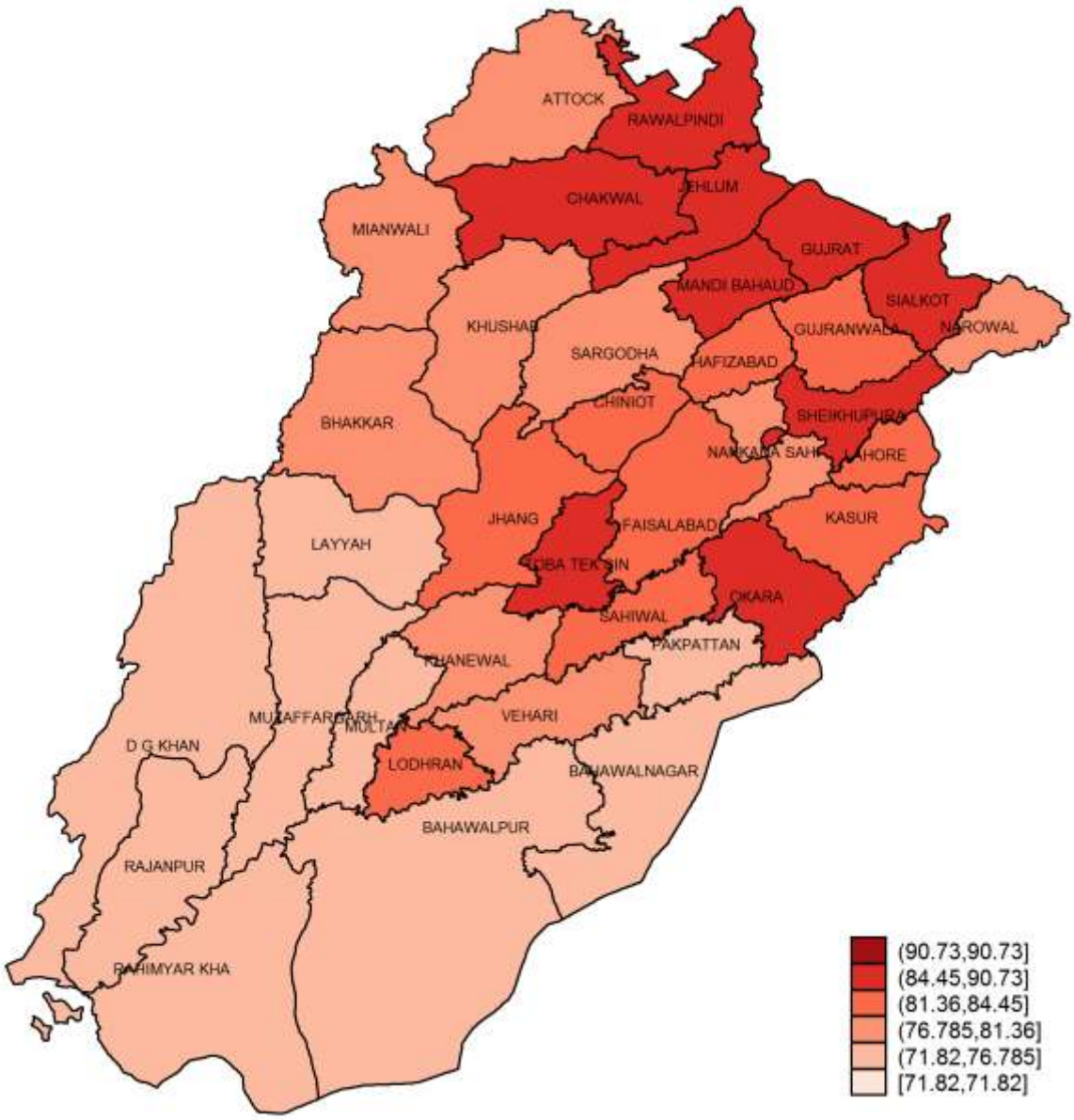

Source: District Census Report 1998

Figure A3: Percentage of Functional Facilities



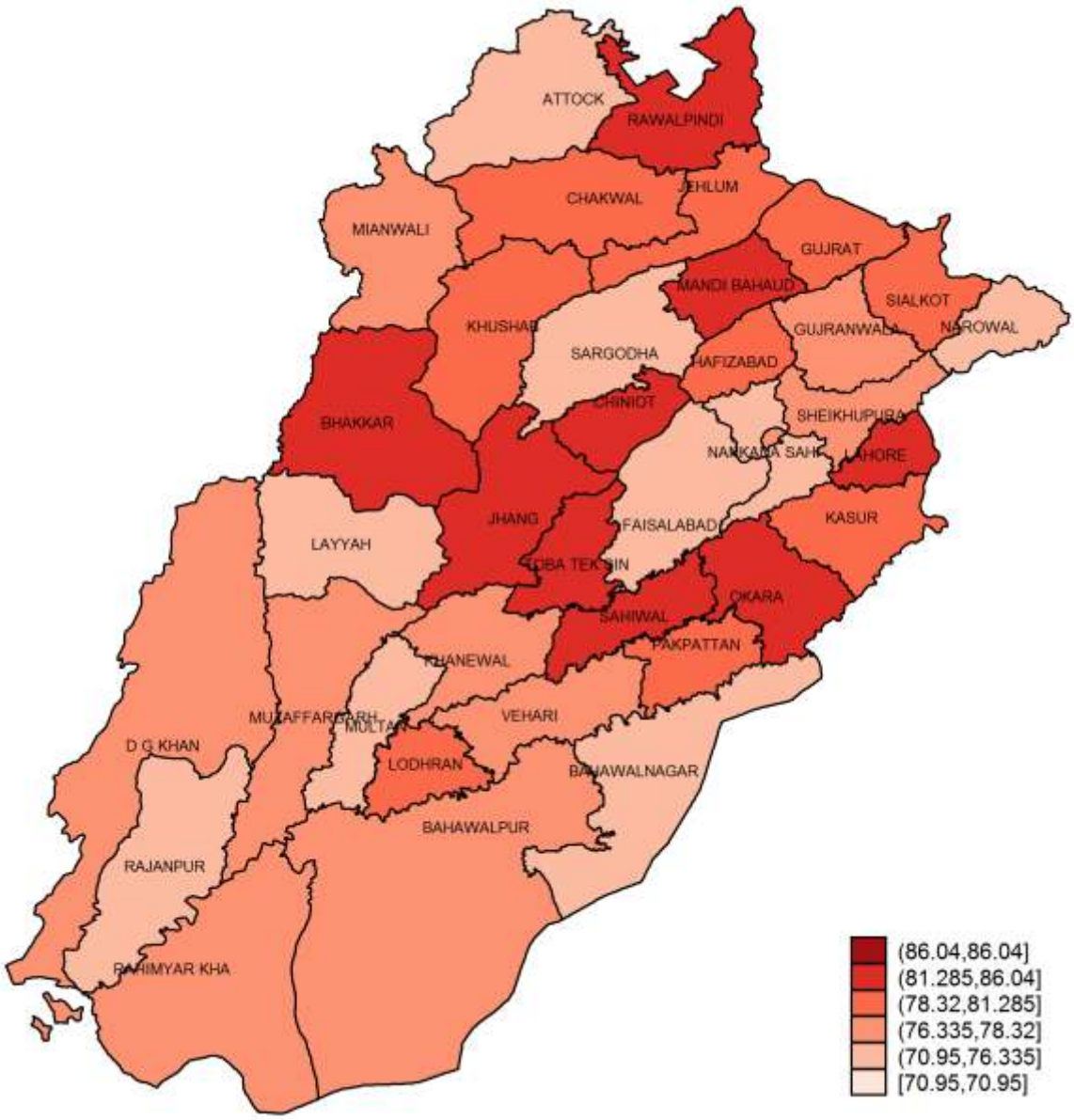

Source: PMIU Monitoring Data, 2009

Figure A4: Student Presence Rate in District (%)



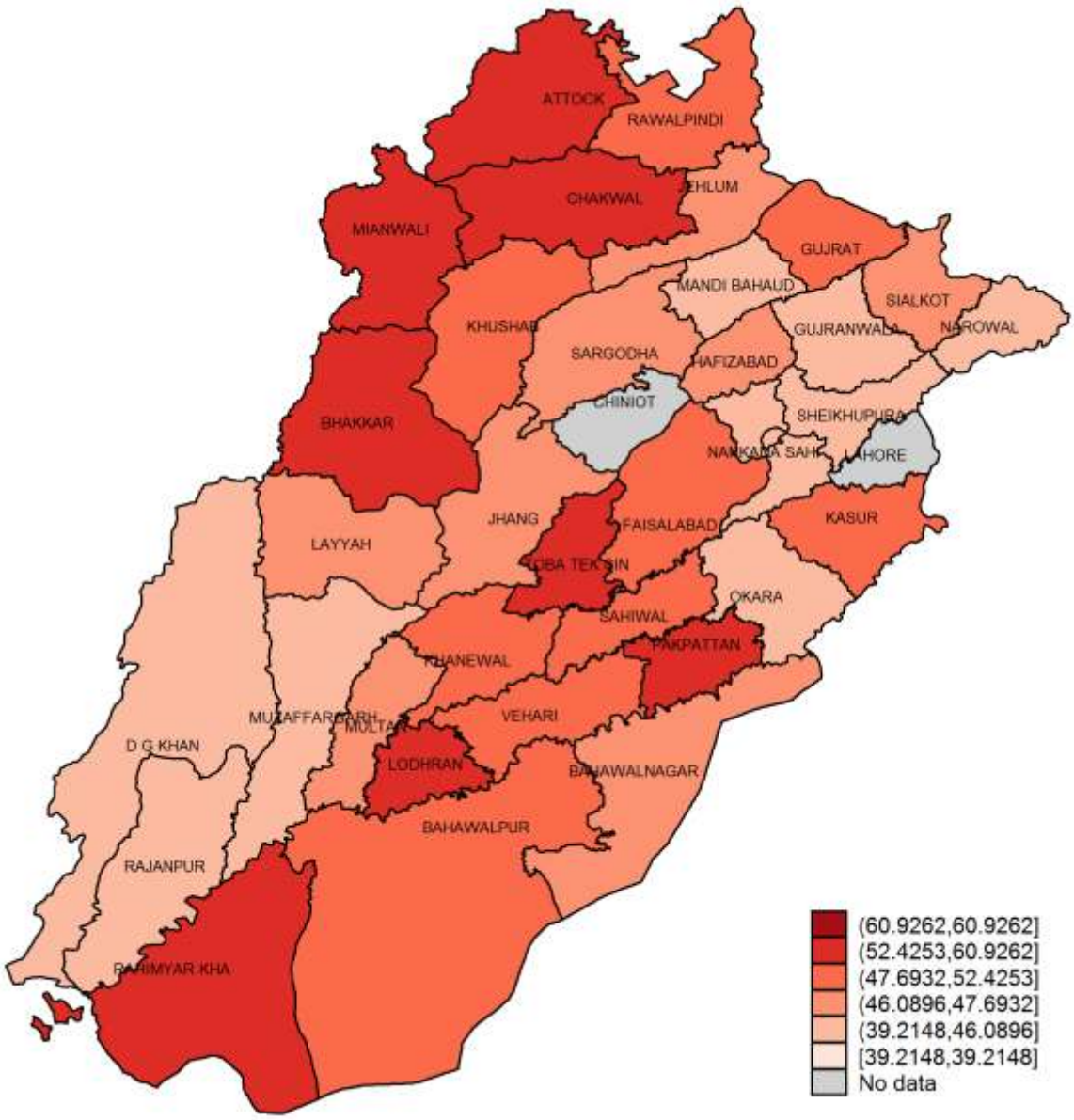

Source: PMIU Monitoring Data, 2009

Figure A5: Teacher Presence Rate in District (%)



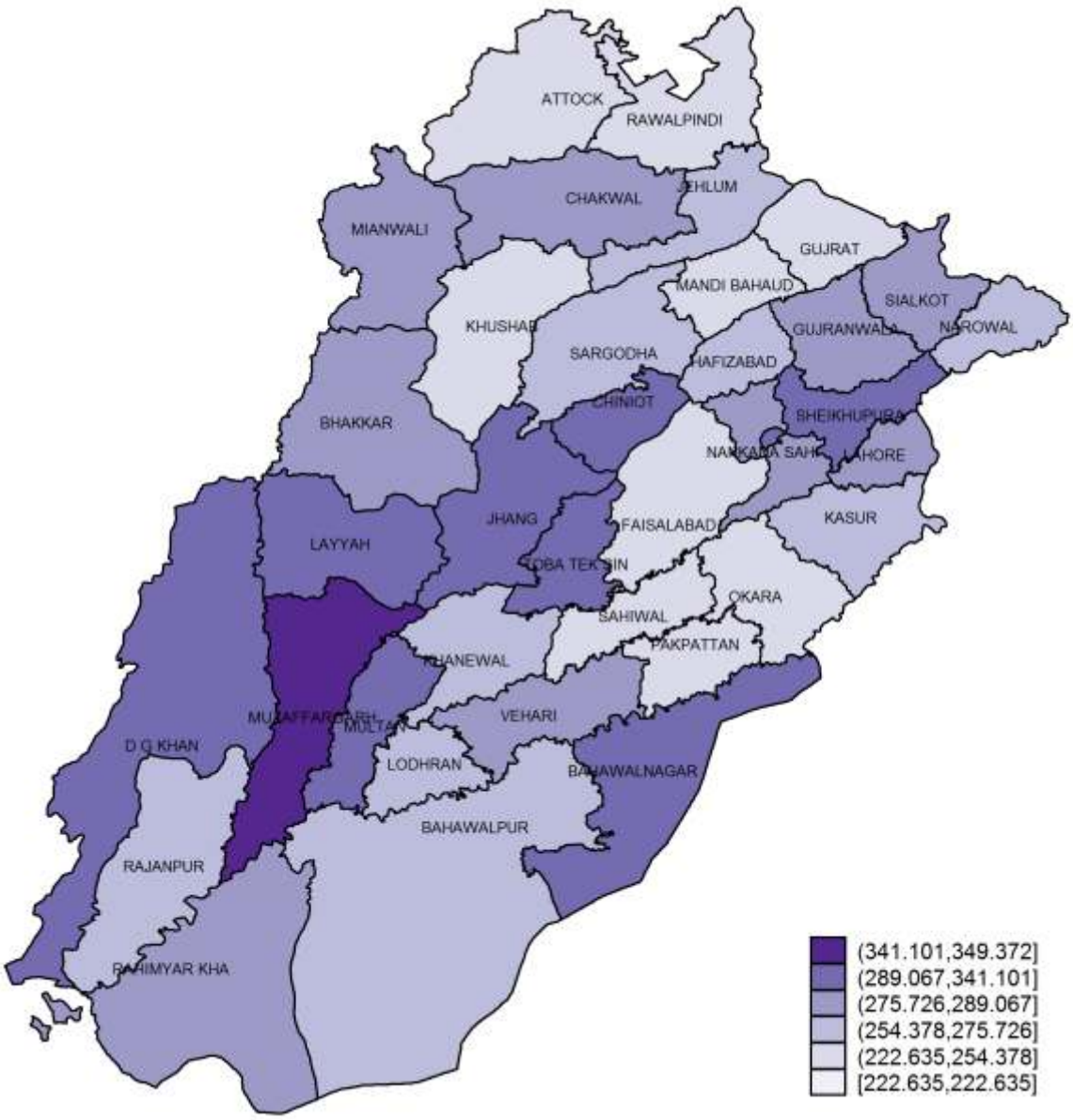

Source: PMIU Monitoring Data, 2009

Figure A6: Percentage of Teachers with Bachelors' Degrees in District



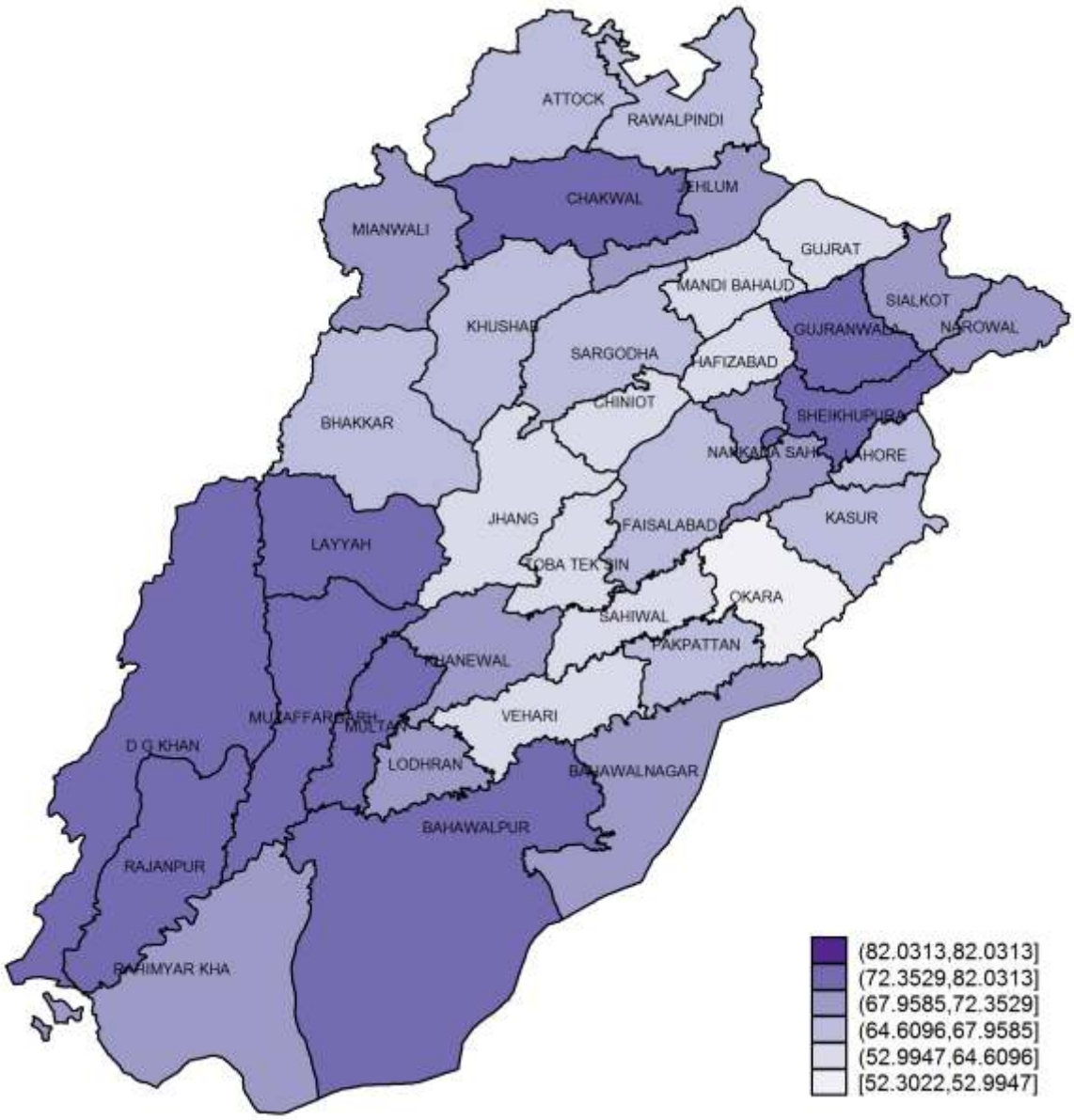

Source: SMIS Data, 2008

Figure A7: Average Total Score in District




Source: PEC Exams Data, 2009

Figure A8: Variance of Total Score in District




Source: PEC Exams Data, 2009

Society for the
Advancement of Education

65-C Garden Block, New Garden Town, Lahore, Pakistan
92-42-35868115-16
www.sahe.org.pk & www.cqe.net.pk