Regional Correlates of Educational Performance In West Virginia Middle and High Schools:

A Study of Demographics, Parental Factors, Teacher Education and School Characteristics

A Monograph Prepared for

Senator Robert Plymale, Chair

Senate Education Committee
The West Virginia Legislature

December 2003

Michael J. Hicks & Viktoriya Rusalkina

Center for Business and Economic Research
Marshall University
One John Marshall Drive
Huntington, West Virginia 25755

CONTENTS

1. INTRODUCTION

This monograph empirically evaluates the link between regional and school characteristics and student performance in West Virginia. This analysis examines all of West Virginia's middle and high schools in 2000-2001 with particular care to evaluate and identify factors that were both within and outside the control of policymakers. This approach led to a holistic analysis of the links between performance and inputs, a traditional focus of economic analysis of education.

While examining several issues, we particularly focus on teacher quality at the school level. While it may seem intuitive that good teachers improve student performance, there is not, as yet, conclusive evidence that measures of teacher quality outweigh other factors in determining student performance. Further, we do not know what, at the aggregate level, are good measures of teacher quality. To paraphrase Justice Potter Stewart, we all may know good teaching when we see it; but this is insufficient guidance for policymakers at any level.

We proceed as follows; we review the existing research on the modeling methods economists use to answer these types of questions. That is followed by a discussion of the data employed in the model, econometric and statistical considerations and the estimation results. For each set of results we provide analysis of the meaning and magnitude of the estimates. We proceed with recommendations for follow on research and end with a summary and conclusion. We include an appendix with summary statistics and bibliography.

2. EXISTING RESEARCH

Typically, school quality is associated with students' educational achievement, which is measured by the results of standardized tests. But, what exactly contributes to the higher test results? Is it higher expenditures, providing pupils with better books, new advanced technology, and nice spacious buildings? Is it higher paid teachers who are more satisfied with their jobs and, subsequently, more productive? Is it family and community characteristics? Or is it is a combination of all these factors? In this study, we are trying to answer at least one of these questions using standard economic analysis.

In order to answer these questions we construct formal models that may be empirically evaluated. As a guide we use information obtained from earlier studies of the factors that link education inputs and outcomes.

The existing economic research in the area of education and school quality does not provide unambiguous answers to the question of which variables (inputs) are significantly related to school quality. The debate concerning the relationship between various inputs and educational outputs continues among prominent economists. For example, Hanushek, the author of several studies on school quality that have enjoyed considerable influence on future research, concludes that, "There is no strong or consistent relationship between school inputs and student performance" (Hanushek, Krueger and Price, 2002:7).

Kruger comes to a different conclusion regarding this relationship. In fact, he finds that smaller classes and higher expenditures per student have positive effect on student attainment. Clearly, the argument between economists continues today, implying that future research on this topic is needed. [1]

There are many different types of models used to estimate educational attainment. This study will concentrate on the production function approach, which evaluates the relationships between various inputs (student-teacher ratio, teacher salary, class size, etc.) and outputs (ACT scores, dropout rates). These are reviewed below.

Levin (2001) uses a quintile regression analysis to estimate the relationship between class size and peer effects on educational achievement. He refers to a longitudinal survey that provides information on Dutch students in the 2nd, 3rd, 6th, and 8th grades in 1994-1995 (Levin, 2001). This survey was based on questionnaires that were distributed to teachers and parents of 800 primary schools, 400 of which are used as "nationally representative" of "regular schools" (Levin, 2001: 228).

[1] Interestingly, these findings were compiled jointly by Kreuger and Hanushek in a collective works volume outlining the source of the disagreement between their research findings.

The results of the questionnaires contain useful information on different variables at the individual as well as class and school levels. Some of the most important variables include: math and language scores in absolute as well as percentile form, class size, number of students with similar IQs, school enrollment, teacher experience and teacher gender (Levin, 2001:244). The number of observations varies from 4,090 to 4,909, depending on the different levels of analysis (i.e. grade level, individual level, etc.)

Levin does not find strong evidence that class size has a direct effect on learning and enhancing student performance. He emphasizes, "class size alone is not the sole determinant of scholastic achievement" (Levin, 2001:235). He believes that the studies that found the existence of the relationship between the class size and educational achievement had in their regressions, "some unobserved factor that proved to be positively correlated with both class size and achievement" (Levin, 2001:235). In addition, the author looks at another factor often neglected in the literature: the existence of "peer effect" which is based on the idea that students learn not only from the teacher, but also from their classmates by observing and working with them. Levin also noted "teachers may target their instruction to groups of children with similar competence levels" (Levin, 2001:236). Therefore, the teacher's time is distributed according to the size of the group and "members of larger groups should receive more instruction time" (Levin, 2001:236). He also concludes that it is likely that peer effect explains, at least partially, the relationship between the class size and educational attainment. In fact, Levin concludes that,"rather than class size reduction, the results provide support for an alternative policy of ability groupings as more viable means to boost scholastic achievement." (Levin, 2001:242)

In analyzing education in the developing world, Hanushek evaluates whether is it better to have more low-quality schools or fewer, but higher-quality schools? Hanushek develops his work on the basis of 96 studies on student performance in developing countries. He also evaluates major studies (400 separate studies) on student performance in the United States and then compares the findings from developing countries to those from the U.S. In fact, he finds significant similarities in results between the two.

Five major variables are used in the study as educational inputs: teacher education, teacher experience, teacher salary, teacher-pupil ratio, and expenditure per pupil (Hanushek, 1995). School outputs in most of the studies reviewed in Hanushek's work, were measured by student scores on standardized achievement tests. Other quantitative measures such as school attendance rates and dropout rates were included (Hanushek, 1995: 228).

The findings of the study support the idea that increasing inputs does not necessarily lead to better student performance. Hanushek does not find statistically significant evidence that students in smaller classes perform better than students in larger classes in either developing countries or in the United States.

The findings for teacher experience are also similar:

> "Although …16 out of 46 studies display significant positive benefits from more teaching experience (the analogous figure for the United States is 29 percent), the majority of the studies-- 28 out of 46—found this input statistically insignificant." (Hanushek, 1995: 230).

Teacher education provides different results for the developing countries and for the United States. While it is an important variable in the developing countries as 35 out of 63 studies supported the idea that better educated teachers positively affect student performance, teacher education is "the least important of all inputs" in the United States studies (Hanushek, 1995: 230).

Another variable used in the study, teacher salaries, does not have a direct affect on student performance in neither type of studies. Similar results are received when evaluating the expenditure per student. In both the studies from developing countries and in the studies from the United States, increasing the expenditure per student is not correlated with an increase in student performance.

Although Hanushek points out that due to data unavailability and, in some cases, data unreliability the results could be distorted, he still emphasizes the idea that increasing inputs does not essentially lead to an increase in student performance and hence, in school quality. He does not find strong evidence that there are "clear and systematic relationships between key inputs and student performance" (Hanushek, 1995:232). However, he says that it does not mean that such relationships do not exist. He simply demonstrates the existing inefficiency of schools. He points out, "Resources are being spent in unproductive ways—ways that do not contribute to improving student performance" (Hanushek, 1995: 243).

Hanushek suggests that one of the ways to improve the existing school inefficiency is to develop appropriate performance incentives. He also emphasizes the idea that it is important to concentrate on "good schools" instead of expanding low-quality schools. Hanushek's idea is that low-quality schools will eventually lead to the waste of resources and, although they provide access to larger number of students, they "may actually be a self-defeating strategy" (Hanushek, 1995: 243). These comments are primarily in a context of limited school access in developing nations.

Several economists disagree with Hanushek's conclusions arguing that his methodology is not appropriate and that his estimates did not control for any background variables. One of the opponents is another well-known economist, Alan

Krueger. In fact, Hanushek and Kruger published a book “The Class Size Debate” in which “two eminent economists debate the merits of smaller class sizes and the research methods used to measure the efficacy of this education reform measure” (Krueger, Hanushek, and Rice, 2002).

Kruger has studied the effect of class size on student performance for a long time. The well-known Tennessee’s Project STAR (the Student Teacher Achievement Ratio study) is an example of several studies he has produced. Harvard statistician Frederic Mosteller (1995) mentions that Project STAR “is one of the most important educational investigations ever carried out and illustrates the kind and magnitude of research needed in the field of education to strengthen schools” (Hanushek, Kruger, 2002:10). Project STAR was an experiment in which each participating school authorized by the legislature was randomly assigned to regular classes, regular classes with full-time aide, or small classes (Bracey, 2000). The results have shown that students studying in small classes were the best academic performers. Kruger finds that “the internal rate of return from reducing class size from 22 to 15 students is around 6 percent”(Krueger, 2000: 28). The experiment greatly influenced policymakers not only in Tennessee, where project STAR was completed, but also in other states such as California.

In “The Class Size Debate”, Krueger re-analyses Hanushek’s literature review and explains his reasons for disagreement with Hanushek’s findings. Krueger argues “Hanushek took more estimates from the studies that had negative, statistically significant results. Sampling bias resulting from smaller subsamples cannot explain this” (Hanushek and Krueger, 2002:17). After reanalyzing Hanushek’s work, Kruger finds “results that point in the opposite direction of his findings: all three alternatives find that smaller sizes are positively related to performance” (Hanushek and Krueger, 2002:18).

Kruger also reviews Hanushek’s findings on expenditures per student and, again, finds that “various weighting schemes indicate that greater expenditures are associated with greater student achievement” (Hanushek and Krueger, 2002:21). He concludes that:

> “Since Hanushek’s results are produced by implicitly weighting the studies by the number of “separate” estimates they present, it seems likely that the opposite conclusion is more accurate: unless one weights the studies of school resources in peculiar ways, the average study tends to find that more resources are associated with greater student achievement” (Hanushek and Krueger, 2002:21).

Driscoll et al. (2003) address the issue of school quality from a slightly different perspective. They examine the impact of school district as well as school and class sizes on student academic performance. Driscoll et al. use school data from California schools because this state has numerous schools varying in sizes, quality, and student body.

Driscoll et al. use 1999 school level data, which was gathered by the California Department of Education. They evaluate 5525 schools in 755 districts in California. The advantage of this study is that it examines size effects at three levels: district, school, and class. The authors also include population density as a regressor because district size and density are correlated (Driscoll et al., 2003). They separately estimate regressions for elementary, middle and high schools. Among major variables included in the analyses are: district size, school size, class size, median household income, and population density. Driscoll et al. use production function approach with the school level standardized test scores as the dependent variable in the regressions (Driscoll et al., 2003:196).

The findings of the study show that "district size has a negative effect on student performance, as measured by standardized scores"[2] (Driscoll et al., 2003: 199). The school size also had a significant negative effect on student performance at elementary school level, but no significant effect on the middle school and high school levels. Similarly, class size was negatively correlated with academic attainment only on elementary school level, and not on the secondary level (Driscoll et al., 2003:199).

Average household income positively affected student attainment and was statistically significant for all three types of schools. Driscoll et al. find that "an increase in the median household income of $10,000 would be associated with an 18-point increase in the API [3] score for elementary school" (Driscoll et al., 2003:200).

Jacques and Brorsen disagree with Hanushek's conclusion that expenditures per student do not affect school quality and student performance. They argue that it is essential to allocate resources to the most productive areas. For example, Jacques and Brorsen note that schools have limited resources and they point out that while an increase in instructional expenditures may positively affect student performance, such an increase in other area (for example, school administration) may, in fact, negatively affect academic achievement because the resources will be taken away from the more productive area to the less productive one.

The authors use the model that utilizes "school district averages of achievement scores on standardized tests as the dependent variable and eleven expense categories by school district as independent variables" (Jacques and Brorsen, 2002: 998). In addition, several socioeconomic factors such as family background were

[2] The coefficient for district size was negative and statistically significant at 1% error level for both elementary and middle school, but it was statistically insignificant for high school regression.

[3] Academic Performance Index—a weighted average of Stanford 9 test scores

included in the model. Unlike other studies that utilize ordinary least squares (OLS), maximum likelihood estimation (MLE) is used in this research "in order to avoid "heteroscedastic disturbances" (Jacques and Brorsen, 2002: 998). The 6,602 observations are used in the analysis. The equation utilized in the study is described below:

Equation 1

$$Y = Sb + Xf + Gb + Ey + u + e$$

where Y accounts for a vector of average test scores for each school district/grade/test combination; X for the socio-economic effects matrix, which includes the percentage of students in special education, the percentage of students receiving free or reduced-price lunches, and four levels of educational attainment of the parents; E represents a matrix of eleven expense variables for each school district, each as a per unit expenditure; u is a random school/grade effect, and *e* is a heteroscedastic error vector (Jacques and Brorsen, 2002: 998).

The authors use required data from the Oklahoma Department of Education (1996), and then they test results from the Criterion Referenced Test (CRT) and the Iowa Test of Basic Skills (ITBS) for the year 1994-1995. Jacques and Brorsen also utilize school district census data from 1990 to obtain parental information, especially the education level of the parents. The study takes into consideration a percentage of students who took the test (since some schools eliminate special needs children from taking the test), and the school size, which is measured by the average daily membership (Jacques and Brorsen, 2002: 999).

The 11 expenditure categories used in the model include the following: instructional expenditures, instructional staff support services, student support services, student administration, general administration and business activities, student transportation services, operations, maintenance, child nutrition, and community services, facilities acquisition and construction, "other outlays" (debt service, clearing account, etc.), scholarships, student aid, and staff awards, and "repayments" (Jacques and Brorsen, 2002: 999).

The results of the study show that instructional expenditures, student support, and student transportation are statistically significant variables while the remaining variables are statistically insignificant. One conclusion from this research is that policymakers should be advised to concentrate on the most productive areas and increase expenditures in those specific areas to avoid misallocation of the resources. It is important to note that the results strongly suggest that "money spent on instruction leads to a small increase in student performance", while money spent on teachers, teacher supplies, and training have a more significant effect on the average test score (Jacques and Brorsen, 2002: 1002).

The authors of this study (Dewey, et. al., 2000) attempt to compare their findings with those of Hanushek who says that school inputs do not matter. Dewey et al. conduct a "meta re-analysis of education production function literature" (Dewey et al., 2000:30). In this "re-analysis" they examine 127 regressions from 46 different studies. The authors evaluate the effect of different variables on the output, which is measured but the results of the standardized test. The variables (inputs) include teacher experience, teacher education, teacher test score, ranking of teacher's college, teacher salary, teacher per pupil, expenditure per pupil, and school size among others (Dewey et al., 2000:30).

The studies that examine the effect of school resources are initially divided into two categories: "good" and "bad" regressions. "Bad" regressions are those that fail to consider parental inputs, income, and socioeconomic status. Almost three-quarters of all the studies are labeled as "bad". [4]

The 414 coefficients derived from all 127 regressions are evaluated using a one-tail test. The results demonstrate that 37 percent of all the coefficients were positive at a 5 percent significance level (Dewey et al., 2000:31), which failed to reject that school inputs are effective, a conclusion different from that in Hanushek's studies.

The authors employ the technique from meta-analysis and find that:

> "teacher education, teacher experience, teacher salary, other teacher characteristics, teachers per pupil, and expenditure per pupil each have a significantly positive impact on test scores in the set of studies reviewed. Only the hypothesis that students learn more in larger schools is unconfirmed" (Dewey et al., 2000:42).

Moreover, Dewey et al. conduct their own empirical study utilizing OLS method. They use data from Project TALENT from 1960 as well as state data for 1987-1992. The results of regression analysis demonstrate that each of the inputs used, to some degree, affects achievement. Not surprisingly, when income is added to the regression, "most school input measures become less significant". This is explained with the stronger correlation between school inputs and income measure (Dewey et al., 2000: 42).

[4] We explain this more fully later. In this context, the 'bad' regressions are this known to suffer from omitted variable bias, a type of statistical error that distorts the magnitude and statistical significance of the parameter estimates.

In conclusion, the authors partially agree with Hanushek that simply increasing school spending does not inevitably increase academic achievement. However, they find that school inputs do matter and that the most important issue is to allocate school resources correctly and if necessary, to increase spending in the most productive areas.

The study by Andrews et. al. [2002] takes a close look at school consolidation and the authors attempt to come to a consensus on how school and district size affects costs and student performance (Andrews et al., 2000). The extensive literature review of the existing studies on economies of scale in education is also included in this paper.

Andrews et al. examine 15 cost function studies and 12 production function studies to answer the following questions: do school size and school district size matter and is consolidation generally an effective policy? They conclude that,

> "…moderation of in district and school size may provide the most efficient combination. Under some conditions, consolidation of very small rural districts may save money, as long as schools are kept moderate size and transportation times remain reasonable" (Andrews et al, 2002:256).

Cost functions used in the research, for the most part, lead to a conclusion that there is an opportunity to save significant administrative and instructional costs when moving from a small district with 500 or less students to a larger district with 2,000-4,000 students (Andrews et al, 2002). Andrews et al. note that per student costs may also continue to decline until the enrollment reaches approximately 6,000 students. That is the point where diseconomies of scale start working (Andrews et al, 2002).

Since the studies using cost function do not consider the opportunity costs of increased travel time for students and parents in the case of consolidation, the optimal enrollment, according to the authors, is in fact lower than the studies suggest. This leads to the recommendation that any school district considering consolidation should determine total travel times. If those times are too high, then it could be concluded that any potential cost savings due to consolidation will result not from savings in teacher salaries or maintenance and capital costs from consolidation of school buildings, but from cutting the administrative expenditures and support staff and services (Andrews et al, 2002:255).

Production function analysis shows that large schools in many cases negatively affect student performance, especially if those schools have a sizable number of disadvantaged students. However, the authors warn that many of the studies do not consider a nonlinear relationship between enrollment and student performance. Andrews et al. also find that "decreasing returns to size may begin to emerge for high schools above 1000 students and elementary schools above 600 students" (Andrews et al., 2002:255).

Rubenstein, et. al. [1999] examines four different types of measuring school efficiency: Adjusted Performance Measures[5] (APMs), the production function approach (which are both tested in the process of empirical estimation of efficiency), cost-function based efficiency measures, and Data Envelopment Analysis (DEA).

The biggest advantage of APMs is that they are easy to measure and explain. The typical equation used to estimate APMs is the following:

Equation 2

$$TS_{sdt} = g_0 + g_1 Z_{sdt} + g_2 X_{sdt} + q$$

where TS_{sdt} represents the output of school s, in district d, in the year t; Z_{sdt} is a vector of uncontrollable factors, X_{sdt} is a vector of resource variables and q is an error term (Rubenstein et al., 1999:267).

The authors, however, argue that these advantages of simplicity and data availability are offset by unreliability of the results. Rubenstein et al. point out that in many cases APM simply demonstrate how one school is different from the other, but they really do not show the level of the school efficiency (Rubenstein et al, 1999:268).

Production function approach appears to be more reliable than APMs because it takes into consideration much more factors and involves more microeconomic theory (Rubenstein et al, 1999). Production functions provide the opportunity to not only compare one school to another, but also to find out which inputs affect outputs the most. The typical production function equation is the following:

Equation 3

$$TS_{sdt} = f(TS_{sdt-1}, ST_{sdt}, P_{sdt}, SC_{sdt}, DT_{dt}, T, D, S, e)$$

where TS_{sdt} is the output of school s in district d at time t, TS_{sdt-1} is the same output one period ago, ST is a vector of student characteristics, P is a vector of peer characteristics, SC is a vector of school inputs, DT is a vector of district characteristics, T is a vector of time dummies, D is a vector of district dummies, S is a vector of school dummies, and e is an error term (Rubenstein et. al, 1999:268).

[5] Adjusted Performance Measures are output measures that are regression-adjusted to control for mitigating characteristics of the school (Rubenstein et al., 1999)

Rubenstein et al. test both the APM method and production function method to measure the efficiency of elementary public schools in Georgia. In their analysis they use a three-year panel of school data. The inputs include total school enrollment, pupil-teacher ration, teachers with a master's degree or higher. To account for family income, the authors include the percentage of students eligible for free and reduced price lunches. They also include the percentage of African American students in their analysis. The outputs in the equations are the percentage of 5th grade students in a school scoring above the national median on the Iowa Test of Basic Skills (ITBS) in reading[6] (Rubenstein et al, 1999**:** 268).

When using the APM equation, "the APMs explain over 50 percent of the variation in the 5th grade test scores. All variables show significant relationship with performance" (Rubenstein, et al., 1999:269). The APM method shows that most of the schools have consistent values over the years, which implies that the APM method can be valuable in evaluating the efficiency of groups of schools (efficient schools and inefficient schools) rather than that of the individual school.

The R-squared in the production function approach demonstrates that "the additional resource variables included in the cross-sectional production functions do not improve the fit of the model significantly. Neither resource variable has a significant relationship with student performance in any year, while the coefficients on the variables measuring school characteristics are of identical sign and similar magnitude to those found in APMs" (Rubenstein et al., 1999:270).

On the basis of their analysis, Rubenstein et al. conclude that all four methods described in the paper are useful in measuring school efficiency. However, APMs and production function appear to be used more often due to the fact that they are more comprehensible and easier to implement. The authors suggest that since all four models have their advantages as well as drawbacks, future research is needed to find a model that eliminates the disadvantages of each approach described and as much bias as possible.

Chakraborty et al. employ the stochastic and non-stochastic production function approach to evaluate the technical efficiency level for 40 school districts in Utah during the academic year of 1992-1993. The stochastic approach utilizes the production function method, from which the authors derive the following measure of technical efficiency:

[6] ITBS is given to all students in 3rd and 5th grade (Rubenstein et al., 1999:268.)

Equation 4

$$\frac{\text{actual output}}{\text{potential output}} = \frac{g_i}{Y_i} = \frac{a_0 e^{-u} \prod_{j=1}^{k} x_j^{a_j} e^v}{a_0 \prod_{j=1}^{k} x_j^{a_j} e^v}$$

where a_0 accounts for a parameter common to all districts, u—for the degree of technical inefficiency that varies across school districts, e^{-u} — for the component of inefficiency, x_j are exogenous inputs, and v is the stochastic disturbance term (Chakraborty, 2001:893).

In the nonstochastic approach, the Data Envelopment Analysis (DEA) is utilized. The DEA is a type of linear programming technique. In this model, school efficiency "is measured by the reciprocal of the output distance function, which is obtained by maximizing ? subject to the restriction imposed by the assumptions of input and output disposability and returns to scale" (Chakraborty et al, 2001: 895). The DEA approach is directed towards maximizing outputs (efficiency) while minimizing or keeping the same level of inputs.

The inputs used in the analysis are various school and nonschool characteristics. The school variables include: student-teacher ratio, percentage of teachers with an advanced degree, and percentage of teachers with over 15 years of experience. These are controlled variables. The nonschool characteristics are uncontrolled variables and consist of the following: the percentage of students who qualify for Aid to Families with Dependent Children (AFDC) subsidized lunch, percentage of district population having completed high school, and net assessed value per student (Chakraborty et al, 2001: 896).

The results of the production function approach indicate that most schools in the school districts of Utah are technically efficient and that one of the major determinants of student performance is the level of parental education (Chakraborty et al, 2001). The model also supports the findings of other studies that lower student-teacher ratios are correlated with higher educational attainment of students. However, no significant evidence was found proving that there is a relationship between school district size and school efficiency (Chakraborty et al, 2001).

The DEA approach provides parallel results. Similar to the production function approach, the outcomes of the DEA model suggest that socioeconomic and environmental factors have the most significant effect on students' educational attainment. Since both methods provide analogous findings, it may be concluded that they both are sufficient for measuring school quality and that economists may choose either of them for future research.

The methods described in the study are valuable because they not only allow measuring technical efficiency of schools, but they also provide an opportunity to determine the inefficient areas to be corrected. For example, knowing that a certain school's quality is highly dependent on the controllable variables (student-teacher ratios, teacher salaries, expenditures per student, etc.) allows policy makers to concentrate on correcting those specific characteristics that need immediate attention.

In their study, Imazeki and Reschovsky discuss school quality in Texas. Specifically, they examine the relationship between school finance and students' educational attainment as well as evaluate several policy implications for the educational system in Texas.

The cost function approach is used in this study to determine the minimum amount of resources each district must spend in order to provide its students with adequate education (Imazeki and Reschovsky, 1998). The authors utilize the following equation:

Equation 5

$$E_{it} = h(S_{it}, P_{it}, Z_{it,} F_{it}, e_{it}, u_{it})$$

where E_{it} represents per pupil expenditures, P_{it} – for student, family, and neighborhood characteristics; e_{it} – for a vector of unobserved characteristics of the school district[7], and u_{it} -for a random error term.

The study employs data from 1995-1996 for Texas K-12 school districts. The school output is represented by the student performance on standardized test scores, such as TAAS (Texas Assessment of Academic Skills[8]), SAT and ACT exams. Imazeki and Reschovsky consider teacher salaries as a measure of input prices. They "treat the teacher salary index as endogenous when estimating the cost function" (Imazeki and Reschovsky, 1998: 277).

The cost function is estimated utilizing two-stage least squares, "with the school output variables and the teacher salary index treated as endogenous" (Imazeki and Reschovsky, 1998: 277). The results demonstrate that resources are currently spent ineffectively and that school districts can spend less money to achieve a certain level

[7] One of such unobserved factors is "inefficiency": the extent to which district spending exceeds the amount necessary to obtain its chosen level of output (Imazeki and Reschovsky, 277).
[8] TAAS is a standardized test which is required for all students in grades 3 through 8 and in grade 10, and it tests reading and math skills (Imazeki and Reschovsky).

of educational advancement. The authors also find that "spending is negatively, though not significantly, correlated with the percentage of students who are in high school" (Imazeki and Reschovsky, 1998: 278). The analysis shows that school expenditures and costs vary greatly from one district to another. The authors specifically find that "the district with the lowest costs could achieve an average level of student achievement by spending about one-fifth as much per pupil as the district with average costs" (Imazeki and Reschovsky, 1009:279).

The study particularly emphasizes that there is a strong link between costs and educational quality. It also provides valuable information for policymakers by determining minimum costs to provide adequate education to students in each of the school districts studied. However, the authors point out that simply providing enough financial resources will not necessarily lead to achieving higher educational quality. They propose using strict accountability standards and financial incentives in order to control the appropriate utilization of the resources, which would lead to achieving a high quality educational system.

The literature cited above represents the most influential work in this field, and provides a few important points. First, there is general agreement that family, community and poverty all play a role in educational outcome. The magnitudes of these impacts vary modestly, with most researchers finding that, at the very least, the sizes of these effects are much larger than other inputs in education.

Researchers are divided on other factors. Notably, there is considerable disagreement (and much ambiguity in research findings) regarding a link between teacher quality, school quality and educational outcome. The differences manifest themselves in studies that vary by technique, region and time. Also, specification problems (with omitted variables bias presenting the greatest concern) seem to present considerable concern to researchers hoping for a strong conclusion.

It is within this background that we explore the link between teacher quality and educational outcome. In so doing we will attempt to extend the findings from earlier studies into a more detailed understanding of the current link between teacher education and educational outcome in West Virginia.

3. AN APPLICATION TO WEST VIRGINIA

Employing the information gleaned from these studies we attempt to isolate a particular factor – teacher quality – in our analysis of school performance in West

Virginia. As the literature review presented above should make clear, isolating a single factor contributing to educational outcome is possible, but requires complex methods of estimation. Simple correlations (and sometimes more complex models) risk omitted variable bias. {This phenomenon is what we referred to earlier as 'bad' regressions}. Frankly, any study that offers conclusions on such complex issues as these through simple correlation analysis is, at best, not useful.[9]

Thus, while our goal here is to measure teacher quality, to do so we must also include estimates of other critical variables in order to minimize omitted variable bias. While we can never fully exclude this possibility we can be comfortable that we have made every possible effort to do so. This began with an extensive data collection effort.

One area we will not explore in this analysis is technical efficiency. That is, we will not estimate whether or not schools are combining inputs in the most efficient manner. This is an important question that goes to the heart of public investment in education. We save this analysis for later research.

The data we collected for this project included all publicly available data from each West Virginia middle and high school. The data collected were from 1997 through 2001 and were available from the West Virginia Department of Education. These data included, but are not limited to all available test scores, attendance and enrollment data for each of the years. Data on number of teachers, administrators and other staff members and their average salaries were collected. The numbers of teachers who met certain categories of educational achievement, information regarding advanced placement and enrollment in languages, mathematics, science, and social studies were also available and were collected. The Department of Education also made available information regarding the number of students taking the SAT 9 tests under standard conditions and those that missed the examination. School construction dates (from which we calculated ages) were also provided by the Department of Education. All of these data permitted us to make additional variables through averaging and three year changes to the levels.

We made several assumptions about each of these data elements that are central to our analysis. We use teacher education as a proxy for teacher quality, other useful measures being largely unavailable. In addition to the number or percentage of teachers in each reported instructional category (e.g. BA+15 or MA degree) we were able to compute a mean number of post secondary years of education for each school.

[9] An example of omitted variable bias would perhaps better illustrate this dilemma. We know that women in the United States earn roughly 70 percent as much as men. However, addition of more variables such as education, average job tenure, employment breaks and age reduces this gap to less than 5 percent. While gender bias may still cause several of these outcome differences, an appropriate specification of the model better isolates the causative factors.

We use test scores of differing types to measure school quality. We also did this with attendance, though it is clear that the direction of causation may occur in either direction. Similarly, as we estimated the duration a teacher had served we recognized that good schools might be magnets for better teachers, so the direction of causation is reversed in our analysis. In cross sectional analysis it is not possible to establish this direction of causation or endogeneity.[10] It remains a theoretical, not empirical issue for which clear determination is not forthcoming at present.

We matched these data on individual schools with demographic information for the surrounding region. Here we used the local zip codes in which the school was located to proxy school district demographics. Matching zip codes to school districts in a consistent fashion proved too costly an approach. Demographic variables included all Census data for 2000. Here again, it is not always clear in which direction causality flows. For example, median house value is often employed as a measure of wealth, but a number of studies have found (to no one's surprise) that school quality affects home prices.

In some instances we also used county or binary variables for such things as rural/urban dichotomy and county population density. We were also able to combine or scale variables to create such variables as proportion of college graduates in a region. As with any statistical study, the application of proxy variables and assumptions suggests that careful interpretation of the results is warranted. We will endeavor to make these interpretations clear in later sections. There are well over 175 variables available for analysis. All received some level of review (and happily, many were rejected early in this process). We will not present an exhaustive discussion of each variable not used in the final results presented later. We will discuss those not used due to an absence of statistical significance which itself may have important policy implications. The individual variables and summary statistics are available in the appendix to this report.

4. MODELING EDUCATION INPUTS IN WEST VIRGINIA

A number of options are available to the economist modeling the connection between various institutional and individual factors that influence educational outcome. The vast economic literature on this issue was reviewed in Rusalkina and Hicks [2002] and in the preceding section. Here we briefly discuss the options available along with our decision to choose the method we describe below. While there are theory driven reasons to select each approach, any selection is also heavily predicated on data availability.

[10] This issue has spawned considerable research into both technical and theoretical issues arising from this problem (see Ericsson and Irons, 1994 and Pearl, 2000).

The human capital approach views educational decisions as a function of rational consumers. In this approach the decision on quality, quantity and type of education achieved is the result of a number of factors influencing individuals. This method obviously benefits from its ability to analyze individual decisions. However, data limitations make this technique limited to small samples. This approach is perhaps most appropriate when the available data is at the individual level, it is less appropriate when examining regions.

A reduced form model is a flexible approach that imposes no restrictions on the available theory. This loose approach is often employed when data is aggregated to the school district, county or state level. The reduced for model is especially useful when trying to answer broader questions involving economic growth or migration with regional education as an aggregate input to these decisions. Our data is sufficiently disaggregated that a more sophisticated method is available.

As mentioned earlier the production function approach permits testing various 'inputs' of education on 'outputs' and is thus very appropriate when attempting to measure the influence of a particular 'input' on 'outputs.' The production function approach in education is not typically characterized by the use of specific functional form as often occurs in other industries. This is often a drawback to modeling when an issue, such as scale or scope economies, is estimated. As noted earlier, we will not be estimating scale or scope economies directly since these are issues of technical efficiency (though scale benefits in production will be estimated).[11] This largely removes concern for development of specific non-linear functional form in this model.

We chose not to use the Data Envelope Analysis since the application of a non-stochastic model would not answer this question of teacher quality directly. Also, we have not yet employed (but did calculate) adjusted performance measures (APM's) as part of this study. We discuss them in detail in later sections.

The modeling of these data that are currently being performed is a production function approach at the individual school and grade level. The production function approach is one of three main modeling methods employed by economists to measure the relationship between school inputs and outputs. It is the most appropriate method of modeling data on aggregate school performance data. The model functionally relates school outputs (or performance) based upon inputs and control variables (such as teacher quality, regional demographics, etc.). We express this as:

[11] To be clear, whether or not economies of scale or economies of scope exist is a different question than whether or not scale influences outcomes. The former set of questions must be framed within a cost analysis (since duality theory is not applicable in this type of public good setting). A number of authors in the education literature have employed these terms incorrectly.

Equation 6

$$Y = f(X, M, C)$$

where **Y** is a matrix of school quality measures (such as test scores), **X** a vector of teacher quality measures (education levels), **M** a matrix of other school variables (such as school size and age), and **C** a matrix of control variables (such as per capita income within the school district).

This approach is flexible, comprehensive and suggested by an understanding of the extensive research on educational performance. However, selecting the appropriate specification among many alternatives is an econometric issue of some magnitude.

5. Issues in Econometric Analysis

We test this model using several multivariate statistical techniques including ordinary, weighted and non-linear least squares estimates, instrumental variable, principal components and simple correlation analysis. This analysis is extensive, since we have at least 21 school performance indicators and over 170 explanatory variables from which to estimate the impact of teacher quality on school performance. This potentially results in $21*(170^{169})$ total possible specifications making selection of the most appropriate model challenging. Even calculating the number of possible combinations is not computationally feasible in most settings, and so calls for some selection criterion to generate useful results.

In order to best represent the correlation between various explanatory variables and school performance we employ several test statistics that point to the best fitting model. This approach involves testing and comparing each equation against all possible variations. The chief method for selecting the most appropriate model is the use of the Akaike Information Criterion (AIC). The AIC balances the variance of the estimated equation with a penalty for over use of variables (absence of parsimony). This test statistic is widely used in advanced time series and large data set estimations. The AIC takes the form:

Equation 7

$$AIC = -2\left(\frac{l}{t}\right) + 2\left(\frac{k}{t}\right)$$

where there are *k* parameters and *t* observations estimating the log likelihood function *l*.[12]

The process also includes other test statistics that measure particular elements of goodness of fit or correct for common problems in multivariate estimation. Chief among these are the significance tests (F-statistic) and the Durbin-Watson (D-W) statistic. The Durbin-Watson statistic is a serial correlation test that, in this instance, is a generally understood measure of model selection where serial autocorrelation is not an issue (as with this data).[13]

Also, all variances are treated by White's [1980] heteroscedasticity invariant variance-covariance matrix. This is recommended by the observation of severe heteroscedasticity among some studies mentioned in the literature review. This matrix takes the form:

Equation 8

$$M = \frac{t}{t-k}\left(X'X\right)^{-1}\left(\sum_{i=1}^{T} u_i^2 x_i x_i'\right)\left(X'X\right)^{-1}$$

where *t* are the number of observations on *k* regressors and *u* is the ordinary least squares residual term.

The chief concern facing this analysis is the clear likelihood of multicollinearity among the regressors. The existence of this condition will certainly prove fatal to a number of specification options. For example, it is clearly impossible to regress four regional income variables without generating this problem to such a degree that the results are meaningless.[14]

Multicollinearity can be handled by a variety of means. First we can eliminate variables that are collinear through simply choosing that variable which minimizes the AIC. This is how we intend to pick the best model, which by definition would mitigate (if not formally minimize) collinearity.[15] This is a preferred approach for

---

[12] $l(b) = \sum_{i=0}^{n} y_i \log\left(1 - F\left(-x_i'b\right) + \left(1 - y_i\right)\log F\left(-x_i'b\right)\right)$ establishes the log likelihood function which is estimated using a second derivative iterative algorithm (Bernt, Hall, Hall and Hausman algorithm).

[13] The Durb in Watson statistic is calculated as $u_t = ru_{t-1} + e_i$ which should, under ideal conditions roughly be equal to two.

[14] Multicollinearity exists when two or more regressors are linear functions of each other. In this condition, the variance between estimators is low (asymptotically approaching zero in some cases). When computing least squares estimators we use these between variable variances, which when very low may make the estimate unsolvable.

[15] Minimizing multicollinearity involves deriving first and second order conditions of eigenvectors, which are here not computationally feasible. Mitigating the ever present problem of collinearity is sufficient for our purposes.

demographic variables, but is not a solution for our measures of teacher quality. Clearly, levels and hours of schooling in a particular school are likely to be collinear, so that selecting some alternative is necessary. One method is the principal components method. A principal components regression is a statistical device that allows variables that are linear combinations to present a subset of relationships that describe the underlying variability in the data. This method was employed in the model selection with less success than the actual average number of hours of teacher education created from the underlying data.

These criterion combined with correction for typical concerns of ordinary least squares provides a basis for analyzing the data collected on schools.

6. ESTIMATION RESULTS

For our modeling efforts we faced considerable specification choices that reduced to a single model applied across all outcome measures. All of these variables (or some linear combination of them) have been reviewed in the preceding literature. This model appears as:

Equation 9

$$\begin{aligned} Outcome = f(&Teacher\ \text{Education, Percent of Families in Poverty, Percent of Adults with College Degree,}\\ &\text{Age of School, Enrollment , Average Class Size, Average Attendance, Drop Out Rate,}\\ &\text{Population Density, stochastic error, intercept)} \end{aligned}$$

The outcome measures we employ include SAT 9 test scores for grades 7 through 11, PSAT tests for 10th and 11th Grade Students, ACT test scores, enrollment in English, Foreign Languages, Math, Science and Social Studies and advanced placement examinations for grades 10 through 12.

Table 1, Summary Statistics of Selected Independent Variables

Variable	Mean	Median	STD Dev
TEACHERED	2078.516	2122.367	438.0475
PERFAMINPOVERTY	0.25682	0.242733	0.139196
PERCOLLGRAD	0.131407	0.122172	0.081689
AGEOFSCHOOL	43.41608	41	25.02711
ENROLL	543.7975	465.1667	326.5225
AVGCLASSSIZE	19.17002	19.5	3.13397
AVGATTEND	93.33403	93.53333	5.614497
DROPOUTPERCENT	1.733102	0.733333	1.935583
POP/SQUAREMILES	32.64	11.52	65.58

Table 2, Summary Statistics of Selected Outcome Measures

	Mean	Median	Std. Dev.
APTT 10TH	0.093716	0	0.258673
APTT 11TH	2.767213	0.5	4.212661
APTT 12TH	4.628962	2.75	6.036362
SAT9 GRADE10	58.85366	58.66667	5.796079
SAT9 GRADE11	59.80759	59.66667	5.255582
SAT9 GRADE6	63.98855	63.33333	5.845672
SAT9 GRADE7	60.31532	59.66667	6.231638
SAT9 GRADE8	61.50541	61	6.045268
SAT9 GRADE9	58.69065	59.33333	7.51227
SAT	1038.21	1038.333	44.63131
ACT	55.75738	55.90833	11.1063
ACTCOMP	19.72369	20	2.040371

As previously mentioned, this model specification is the result of eliminating available variables in each of our three categories through the use of simple correlation measures (competing variables in the same category were selected by the highest simple correlation). So, for example, when faced with median household income and per capita income in the same region as a proxy for regional economic conditions we chose the variable or variables that exhibited statistically significant correlation with outcome measures. The second step was in choosing the combination of these variables that minimized the AIC. The results appear in the following tables. Analysis of results follows.

TABLE 3, REGRESSION RESULTS FOR SAT 9 TESTING

	7th Grade	8th Grade	9th Grade	10th Grade	11th Grade
C	-74.1577	-6.88178	-90.21067**	-16.1401	17.97182
TEACHERED	-0.00235	-0.00072	0.000788	0.003181***	0.002452**
PERFAMINPOVERTY	2.866106	6.145378	-7.5867	-12.40336**	-9.54453**
PERCOLLGRAD	24.04691**	16.30438	18.93412*	23.94337***	23.67902***
AGEOFSCHOOL	-0.01925	-0.00357	-0.01502	0.002089	0.002338
ENROLL	-0.00474	-0.00407	0.002418	0.000747	0.001878
AVGCLASSSIZE	0.047106	0.252234	-0.589524**	-0.24358	-0.394027**
AVGATTEND	1.463939**	0.677199	1.690819***	0.790491**	0.466266
DROPOUTPERCENT	-1.27229	-1.08899	-0.738763***	-0.533839*	-0.535813*
POP/SQUAREMILES	0.014972	0.012843*	0.012672**	0.005919[a]	0.006195
R-squared	0.325888	0.219623	0.372292	0.511019	0.507367
Adjusted R-squared	0.202071	0.076289	0.316906	0.471726	0.467781
S.E. of regression	5.76433	5.976264	6.310056	4.226905	3.848594
Sum squared resid	1628.147	1750.071	4061.314	2001.073	1658.908
Log likelihood	-181.588	-183.719	-360.004	-343.753	-332.314
Durbin-Watson stat	1.868552	1.940471	1.799698	2.356341	2.018412
Mean dependent var	59.17514	60.72599	57.95387	58.83333	59.79508
S.D. dependent var	6.453074	6.218157	7.634716	5.815577	5.275417
Akaike info criterion	6.494521	6.566734	6.607211	5.799229	5.611705
Schwarz criterion	6.846646	6.918859	6.849935	6.029067	5.841543
F-statistic	2.632019	1.532244	6.721766	13.00531	12.81665
Prob(F-statistic)	0.014376	0.163251	0	0	0

*denotes statistical significance to the .10 level, ** statistical significance to the .05 level and *** statistically significance to the .01,[a]statistically significant to the .15 level employing asymptotic t-statistics. All standard errors are treated with White's [1980] matrix.

In these sets of results, between 21 and 51 percent of the variation in SAT 9 test scores are explained by the variables we present above. The better performing models offer considerable explanatory power for cross sectional analysis. However, the regressions for 7th and 8th grade scores suggest little value in interpretation, with the 8th Grade regression not possessing statistical significance (F statistic not significant).

Importantly, the teacher education enjoyed a positive and statistically significant impact on SAT 9 scores for 10th and 11th grade students. While the magnitude of the impact is not large, it is important to realize that this is a rough proxy for teacher quality. This is consistent with the findings of both Hanushek and Kreuger. For this variable we used both aggregate years of education and mean years of education. Both results were provided almost identical results. From this evidence alone it is clear that a link between teacher quality and educational outcomes is important, though it calls for considerable additional study.

Not surprisingly, and consistent with all other studies we have observed, measures of income and education play the dominant role in overall explanations of educational outcome. Here, the percent of families in poverty explains a considerable proportion of educational outcome. Also, the percentage of adults in the zip code in which the school is located possessing a college degree explains much of the variation in test scores. This finding is consistent with virtually all earlier research. The direction of these impacts is as expected, and are illustrative of the persistence effect felt by school performance in regions that suffer poor educational achievement.

School size plays a small, positive role in higher test scores among high school students. The effect is modest (and linear in follow up tests) and without significant interaction with other variables. An additional 250 to 330 students to a high school is associated with a one point increase in SAT 9 scores among 10^{th} and 11^{th} grade students. Importantly, this study does not estimate scale economies so the efficiency of larger schools cannot be determined without costs data. It should be again noted that several other studies have attempted to measure scale economies, with mixed success. Without a better understanding of efficiency gains associated with scale in schools, no policy recommendations are supportable. However, it is equally clear that among middle and high school students, larger schools are not adversely impacting SAT 9 test scores. This is consistent with other economic studies that find only very large school districts suffering ill effects of size on educational outcome.

Class size was statistically significant and negative in this analysis. For high school students, a one-pupil reduction in the average class size resulted in a half a point increase in the average SAT 9 test score. These findings strongly support the work of Krueger. There was no apparent interaction between class size and teacher education, thus no evidence exists that teacher education and class size are either substitutes or complements.

Average attendance rates are positively correlated with SAT 9 scores, though the relationship is not large, and not subject to clear policy recommendations beyond the obvious, keeping children in school improves performance, even when measured at the aggregate levels. Also, we cannot establish the direction of causation among these variables.

Lower drop-out rates were also strongly correlated with SAT 9 test scores, but as with attendance, the magnitude of the impact is not large. The primary culprit in constraining the usefulness of adjusting these variables is that there is little variation in these variables across schools, and the drop out and rates of absence are not large. This does not mean they cannot be improved, only that the impact of a percentage point change in either variable will have a limited impact on SAT 9 scores. Notably, a one-percentage point improvement in either represents a fairly substantial change in the total.

Population density also affects SAT 9 scores. Population density is a continuous (not dichotomous) representation of rurality. So, schools in more rural areas are associated with lower SAT 9 scores in middle schools. This affect is not large, nor is its cause clear. A number of factors that are correlated with rural areas may lead to this result, though we have attempted to correct for these through our demographic data. It may be also that some unmeasured variable such as duration of school bus rides (to pick a popular topic) generates this result.

These results suggest that there are several factors at issue in determining educational outcomes. Indeed, the appropriate measure we use for outcomes should be evaluated for robustness. At issue is whether several proxies for educational outcome are similarly correlated with the inputs we employ. If they are, we can feel more certain in interpreting the results of our estimation. To this end we estimate the impacts of ACT component testing. See Table 4 below.

These results offer much of the same support for teacher education as a correlate of educational outcomes as do the earlier results. Here the ACT composite results are less well explained than the proportion of total students taking the exam. For the proportion of students taking the test, this model explains more than a third of variation between schools. As with the earlier estimates, teacher education, the percent of adults with college degrees in the zip code in which the school is located, average class size, and percentage of students dropping out the previous year all enjoy strong statistical significance. Poverty, rurality and enrollment are not clearly significant across the board. The latter does for composite scores. This contrasts with the earlier results, though the magnitude of the differences is not large.

Table 4, Regression Results for ACT Testing

	% ACT Takers	ACT Composite
C	17.45026	23.56969*
TEACHERED	0.007626***	-0.000493
PERFAMINPOVERTY	-11.32774	-2.041008*
PERCOLLGRAD	29.37365**	1.56048
AGEOFSCHOOL	-0.028572	-0.008635
ENROLL	0.002551	0.000752*
AVGCLASSSIZE	-0.948661***	0.223576
AVGATTEND	0.510277	-0.076096
DROPOUTPERCENT	-2.630301***	0.077861
POP/SQUAREMILES	-0.002202	0.001535
R-squared	0.40032	0.274299
Adjusted R-squared	0.351255	0.214378
S.E. of regression	8.967562	1.822494
Sum squared resid	8845.888	362.0419
Log likelihood	-428.2856	-235.0556
Durbin-Watson stat	1.937067	1.861153
Mean dependent var	55.90583	19.71485
S.D. dependent var	11.13365	2.056172
Akaike info criterion	7.30476	4.118581
Schwarz criterion	7.537051	4.35212
F-statistic	8.15901	4.577724
Prob(F-statistic)	0	0.000039

*denotes statistical significance to the .10 level, ** statistical significance to the .05 level and *** statistically significance to the .01, [a]statistically significant to the .15 level employing asymptotic t-statistics. All standard errors are treated with White's [1980] matrix.

Overall, these second set of results supports the findings in the first set of results. There are still additional data to explore, the following table illustrates the results from the combined SAT scores.

Table 5, SAT Results

	Coefficient
C	102.2105
TEACHERED	-0.00204
PERFAMINPOVERTY	-28.53001***
PERCOLLGRAD	63.21019***
AGEOFSCHOOL	0.05767
ENROLL	0.007201***
AVGCLASSSIZE	0.032002
AVGATTEND	-0.99239
DROPOUTPERCENT	-0.65109
POP/SQUAREMILES	0.052419***
R-squared	0.602967
Adjusted R-squared	0.570483
S.E. of regression	8.333429
Sum squared resid	7639.063
Log likelihood	-419.485
Durbin-Watson stat	2.192583
Mean dependent var	12.54889
S.D. dependent var	12.7155
Akaike info criterion	7.158082
Schwarz criterion	7.390373
F-statistic	18.56171
Prob(F-statistic)	0

*denotes statistical significance to the .10 level, ** statistical significance to the .05 level and *** statistically significance to the .01, [a]statistically significant to the .15 level employing asymptotic t-statistics. All standard errors are treated with White's [1980] matrix.

These results also support the earlier findings, albeit with some notable exceptions. Most importantly, for the purposes of this research, teacher education no longer enjoys statistical significance. Regional poverty, percentage of college graduation and rurality appears to have strong impacts on SAT scores.

The findings with respect to enrollment in foreign languages, math, science, English and social studies all provide similar findings. So, too does the PSAT testing. In each of these cases there is some variation in the size and significance of the impacts. Generally the same variables matter: educational achievement of adults in the region, poverty rates, class size, size of school, rurality and, most importantly for our purposes teacher education. To place the impacts in relative size we should point out that for our best estimates, we can account for only a little more than half the variation in test scores. This is better than most of the other studies reviewed above, but it is clear that there's much more research needed in support of policy.

Similarly, several other studies note that the 'within school' variation in test scores is greater than the 'between school' variation. Again, there is clearly much to be learned. For the variables that do explain school level differences in test scores we believe relative magnitudes are important to illustrate. We cannot directly compare each variable since in our specification they have different units of measure. But, to see how test scores vary with changes in explanatory variables we calculate the impacts and illustrate them below:

Table 6, Magnitude of Variable Effects (at the Margin) on 11th Grade

A CHANGE IN EACH VARIABLE		EFFECT ON SAT-9 SCORE
One 4 hour class increase per teacher	***is correlated with:***	1 percentage point increase in test scores
One percentage point decrease in the rate of families in poverty	***is correlated with:***	9 percentage point increase in test scores
One percentage point increase in the proportion of parents with a college degree	***is correlated with:***	23 percentage point increase in test scores
One fewer student per class	***is correlated with:***	0.39 percentage point increase in test scores
One percentage point decrease in the dropout rate	***is correlated with:***	0.5 percentage point increase in test scores

Importantly, these impacts are not the direct correlations, but are controlled for other variables. These are impacts at the margin, not the average. Notably, the size of these variables should offer some pause for policymakers. For example, increasing the proportion of parents with a college degree by one percentage point is, in some districts, a 20 percent change.

This section has outlined findings from these models (and some we do not illustrate for brevity). But what we have found is not significant in the modeling process is also potentially of some importance. These other findings are reviewed below.

7. OTHER FINDINGS

There are considerable issues not thus far addressed in this analysis. These may be characterized in three areas: answers we did not find because we did not ask; answers to questions we do not know; and variables we found do not impact education and didn't make it into our analysis. We will address them in reverse order.

We found that none of the three-year changes in inputs explained any significant issue in educational output. The main reason is that three-year changes in these variables are largely non-existent. That is to say, that while there have been changes in the actual values, they are not statistically significantly different over the three-year period at the school level. This means that any measured improvement in the inputs that affected West Virginia school quality during the study period is largely a chimera. The same can be said for aggregate test scores, which do not show statistically significant improvement over the study period. The absence of findings in these areas is perhaps one of the leading results of this study.

Like Hanushek (1995) we found no correlation between teacher experience and educational outcomes. Similarly, teacher salaries were not correlated with improved educational outcomes. This differed from other studies and may partly be explained by a highly centralized compensation system in West Virginia. Simply, pay and outcomes are not designed to be connected in the State funding formula, and they are not.

We also could find no strong relationship between measures of staff experience and anything measuring quantity of administrators to school performance. In our quest to employ every possible variable we even tested the impact of teacher gender on school performance. We posited the possibility that gender may be correlated with some other unobserved variable that influences the school environment (such as prior military service in male principals). While that possibility may exist, there is no statistically significant relationship in these data between principal gender and educational outcomes.

We do not yet know which schools deliver educational services most efficiently. We know which schools combine inputs most effectively (that is get the greatest impact from the inputs and situations they have been given) but we have not yet assigned a cost to this finding. That is a wholly additional study.

Also, a good many variables that may be good proxies for regional income or demographics are not significantly correlated in this study. Much to our surprise the final specification of the model appears very similar to the more extensive studies reviewed in the early part of this monograph. This is helpful for two reasons. Firstly, these results provide considerable support for our estimation. Secondly, these specifications provide strong support for a Bayesian approach to modeling education. In the Bayesian approach we would be modeling teacher quality with strong expectations regarding the specification of the model and the parameter estimates. Had we estimated these models in Bayesian setting the prior selection of a positive relationship between teacher quality and outcomes would have been supported. This is strong support for continued research into this area and is also strong support for the findings of this study.

We found modest correlations between the density of private schools and public school performance. This is potentially important because it supports the main argument for private schools (they stimulate regional competition in school performance) and refutes the main argument against private schools (they are cherry picking students). However, these results are very tentative.

Also importantly, we found some puzzling relationships between the way SAT-9 tests are administered and the scores of these tests. We do not understand the results. The proportion of students taking the SAT-9 test under non-standard conditions ranges, at the school level from between 4 and 25 percent. This degree of variation is, as its most charitable characterization, puzzling. There is little to explain this degree of variation beyond program failure. At the county level, high rates of non-standard test taking result in higher individual school test scores. The link between non-standard testing and scores at the school level is more complex. For example both high and very levels of non-standard test taking are correlated with better educational outcomes (though this result is blurred by small sample sizes). Overall, these statistical relationships should not occur if the program is effectively administered.

Keeping in mind that there should be, in actuality, little difference between schools in the proportion of students eligible for non-standard testing there are a number of explanations for a relationship between test scores and the proportion of students taking the tests in a non-standard setting. In the extreme, out-migration may have generated real differences in the demographic characteristics of a school that would generate high levels of students appropriately provided non-standard testing environments. Schools with fewer resources may not be able to identify students effectively to permit them to participate in the program. Parents may feel compelled to direct students into, or away from the program in different regions, thus providing higher variance in the data. However, the demographic homogeneity in the State combined with a school funding formula designed to mitigate differences in school funding suggests that these reasons are not likely causes of these correlations.

More likely explanations for the high variability in non-standard testing rates and the relationships between rates and test scores are that the data are manipulated to provide either increased financial resources to individual schools (due to programmatic issues) or simply to manipulate scores. This is an important additional avenue of research.

Finally, there are two important issues we did not explore in this study. The first, is the aforementioned estimate of technical efficiency in schools. This study permits us to rank schools based upon their performance after correcting for different variables beyond their control. This is the Adjusted Performance Measure mentioned earlier (but without the problems mentioned by Rubenstein). By this we mean that through the process outlined in this study we can rank schools on how well they are doing given the demographics, poverty, education, rurality and other factors beyond their control. We have not included these results at this juncture since we propose to use these findings in a later double-blind analysis of the top and bottom performing schools.

However, simply ranking schools or establishing the effectiveness of the different variables does not establish whether or not schools are efficient. More complex modeling is needed to perform this analysis. For example, we find that teacher education and larger schools are both correlated with higher test scores among certain groups. This is important, but it does not tell us if this would be an efficient use of resources or what schools are combining resources most efficiently.

8. AVENUES FOR ADDITIONAL RESEARCH

The findings presented in this monograph represent an unusually high level of explanation for variation in school quality. When compared to the most extensive studies of educational outcomes the quality and quantity of data employed, the robustness and explanatory power of the models and the inferences derived from the estimates are all substantial. However, there is considerably more analysis needed to fully understand the issue of educational inputs and performance. We believe the following areas of research should be pursued.

An extension of this modeling process to elementary schools is warranted. Education of younger children differs in important ways so the variables that explain educational outcomes of younger children may change.

The geographic extension of this research to several other states, with the same level of analysis is warranted. This may help us determine what policies at the state level effect educational outcomes.

Other projects are currently ongoing to identify school and teacher level measures of quality. We believe that combining the data from these studies with the variables we have identified may provide useful policy changes. For example, knowing what correlates of teacher quality correlate to education would provide a useful measure of the quality of this variable in explaining teacher quality. Similarly, estimating interaction effects between variables would be highly helpful. For example, we find additional teacher education is correlated with better educational outcomes. We think it important to evaluate whether or not this relationship holds across all academic majors.

We believe that a significant potential study of schools can be performed in a quasi-experimental setting. This method would be different, but much less resource intensive than the Tennessee STAR experiments of the past decade that is widely believed to be among the most credible studies in the field. We propose to rank schools by an adjusted performance measure (APM) and dispatch field evaluators to the highest and lowest performing schools. This experiment would enjoy a double blind nature in that absent the analysis presented in this paper, it is impossible to know which schools are high and low performers. So, examiners can approach schools knowing what questions not to ask (since we already know the answers) and delve more deeply into each school's operation.

Also, we believe that a robust estimate of the technical efficiency of schools in West Virginia be performed. For example, knowing how much spending on instruction increases performance, and whether the benefits are offset by some other choice (say administrative costs) is important for individual schools, districts and the state. Knowing how efficiently schools allocate resources may provide important insight into the overall educational finance in the State.

9. SUMMARY AND CONCLUSIONS

Existing research provides mixed results on the link between inputs to education such as teacher quality and educational outcomes such as test scores. This debate continues among prominent economists. At issue are both the link between the inputs and outputs as well as whether or not any of the measured improvements result in a more efficient use of resources than some other intervention. This is not a trivial discussion, for it goes to the heart of the issue of public support of secondary education. Nor is it a normative discussion. What economists are centrally attempting to understand is whether or not deployment of an additional $100 in funding to public schools will raise school performance more than some alternative use of the resources such as a tax cut that reduces poverty. We have not yet attempted to answer these questions.

As with other research performed on this subject we found that poverty and proxies for family education (the proportion of adults with college degrees) provided the strongest explanations for variation in school test scores.

We found that bigger schools tend to lead to modestly better educational outcomes in some settings, as do more urban counties, lower drop out rates and higher attendance. None of these findings were unexpected. Nor are any of these impacts very large relative to family education and poverty.

Most importantly, teacher quality, even when measured roughly, is correlated with higher test scores in West Virginia high schools. This result holds when correcting for other school inputs, demographic and regional economic variables. Though the magnitude of the effect was not large (which may be a function of its rough measurement) it is nevertheless an important result.

10. REFERENCES

Altonji, Joseph G. 1998. "Effects of Personal and School Characteristics on Estimates of the Return to Education," *Economic Perspectives* 22, pp.65-80.

Andrews, Matthew; William Duncombe and John Yinger.2002. "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?" *Economics of Education Review* 21, pp. 245-263.

Angrist, Joshua D. and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106, pp. 979-1014.

Arabsheibani, Reza G. and Lamine Manfor. 2001. "Non-Linearities in Returns to Education in Libya," *Education Economics* 9, pp.139-144.

Arias, Omar; Kevin F. Hallock and Walter Sosa-Escudero. 2001. "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data," *Empirical Economics* 26, pp.7-40.

Ashenfelter, Orley and Alan Krueger. 1994. " Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review* 84, pp.1157-1174.

Ashenfelter, Orley; Colm Harmon and Hessel Oosterbeek. 1999. "A Review of Estimated of the Schooling/Earnings Relationship, with Tests for Publication Bias," Industrial Relations Section working paper, Princeton: Princeton University Press, pp.425-449.

Black, Sandra E. 1998. "Measuring the Value of Better Schools," *Economic Policy Review* 4, pp. 87-94.

Boissiere, M; J.B. Knight and R.H. Sabot. 1985. "Earnings, Schooling, and Cognitive Skills," *The American Economic Review* 75, pp.1016-1030.

Bradbury, Katharine L. 2002. "Education and Wages in the 1980s and 1990s: Are All Groups Moving Up Together?" *New England Economic Review*, pp.19-46.

Card, David. 1998. "The Casual Effect of Education on Earnings," The Center for Labor Economics working paper 2.

Card, David and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *The Journal of Political Economy* 100, pp.1-40.

Casse, Daniel and Bruno V. Manno. 1998. "The Cost and Price of College and the Value of Higher Education," *Academic Questions* 11, pp. 38-55.

Chakraborty, Kalyan; Basudeb Biswas and W. Cris Lewis. 2001. " Measurement of Technical Efficiency in Public Education: A Stochastic and Nonstochastic Production Function Approach," *Southern Economic Journal* 67, pp. 889-905.

Census: More Degrees Equal Bigger Bucks. CNN.com. 07/18/02. fyi.cnn.com/2002/fyi/teachers.ednews/07/18/degree.dollars.ap/index.html Access Date: 07/18/02.

Chu, Hong-Yih. 2000. "The Impacts of Educational Expansion and Schooling Inequality on Income Distribution," *Quarterly Journal of Business & Economics* 39, pp. 39-50.

Consumer Price Index. Federal Reserve Bank of Minneapolis. http://minneapolisfed.org/research/data/us/calc/hist1913.cfm Access Date: 09/14/02.

Cosca, Theresa. 1998. "Earnings of College Graduates in 1996," *Occupational Outlook Quarterly* 42, pp. 20-29.

Datcher, Linda. 1982. "Effects of Community and Family Background on Achievement," The *Review of Economics and Statistics* 64, pp.32-41.

Dewey, James; Thomas A. Husted and Lawrence W. Kenny. 2000. "The Ineffectiveness of School Inputs: A Product of Misspecification?" *Economics of Education Review* 19, pp.27-45.

Dinopoulos, Elias and Paul Segerstrom. 1999. "A Schumpeterian Model of Protection and Relative Wages," *The American Economic Review* 89, pp. 450-472.

Dothan, Uri and Joseph Williams. 1981. "Education as an Option," *The Journal of Business* 54, pp.117-139.

Driscoll, Donna; Dennis Halcoussis and Shirly Svorny. 2003. "School District Size and Student Performance," *Economics of Education Review* 22, pp. 193-202.

Eide, Eric R; Mark H. Showalter and David P. Sims. 2002. "The Effects of Secondary School Quality on the Distribution of Earnings," *Contemporary Economic Policy* 20, pp. 160-170.

Ericsson, Neil R and John S. Irons (ed). 1994. Testing Exogeneity. Oxford University Press, Oxford. 1994.

Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes," *Journal of Economic Literature* XL, pp. 436-482.

Grosskopf, Shawna; Kathy Hayes, Lori L. Taylor and William L. Weber. 1998. "Allocative Inefficiency and School Competition," *Proceedings of the 91 Annual Conference on Taxation, National Tax Association,* pp.282-290.

Haider, Steven J. 2001. "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics* 19, pp.799-837.

Hanushek, Eric A. 1995. "Interpreting Recent Research on Schooling in Developing Countries," *The World Bank Research Observer* 10, pp. 227.

Hanushek, Eric A.; Alan B. Krueger and Jennifer King Rice. 2002. *The Class Size Debate,* Washington, D.C: Economic Policy Institute.

Harmon, Colm and Ian Walker. 1995. "Estimates of the Economic Return to Schooling for the United Kingdom," *American Economic Review* 85, pp. 1278-1287.

Haveman, Robert; Barbara Wolfe and James Spaulding. 1991. "Childhood Events and Circumstances Influencing High School Completion," *Demography* 28, pp. 133-157.

Haveman, Robert and Barbara Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature* 33, pp. 1829-1878.

Hecker, Daniel E. 1992. "College Graduates: Do We Have Too Many or Too Few? College Earnings and Why They Vary," *Occupational Outlook Quarterly* 36, pp. 13.

Heckman, James. 1998. "What Should Be Our Human Capital Investment Policy?" *Fiscal Studies* 19, pp.103-119.

Hedges, Larry V. and Rob Greenwald. 1996. Have Times Changed? The Relation between School Resources and Student Performance. Brookings Institution Press.

Hoxby, Caroline M. and Bridget Terry Long. 1999. "Explaining Rising Income and Wage Inequality Among the College-Educated," Harvard University, NBER working paper.

Imazeki, Jennifer and Andrew Reschovsky. 1998. "Measuring the Costs of Providing an Adequate Public Education in Texas," *Proceedings of the 91Annual Conference on Taxation, National Tax Association*, pp.275-281.

Jaeger, David A. and Marianne E. Page. 1996. "Degree Matter: New Evidence on Sheepskin Effects in the Returns to Education," *The Review of Economics and Statistics* 78, pp.733-740.

Jacques, Charles and Wade B. Brorsen. 2002. "Relationship Between Types of School District Expenditures and Student Performance," *Applied Economics Letters* 9, pp. 997-1003.

Jolliffe, Dean. 2002. "Whose Education Matters in the Determination of *Household Income? Evidence from a Developing Country," Economic* Research Service, U.S. Department of Agriculture, working paper.

Juhn, Chinhui; Kevin M. Murphy and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill," *The Journal of Political Economy* 101, pp.410-442.

Kane, Thomas J. and Cecilia Elena Rouse. 1995. "Labor-Market Returns to Two- and Four-Year College," *American Economic Review* 85, pp.600-615.

Kremer, Michael R. 1995. "Research on Schooling: What We Know and What We Don't – A Comment on Hanushek," *The World Bank Research Observer* 10, pp. 247.

Lassibille, Gerard and Lucia Navarro Gomez. 1998. "The Evolution of Returns to Education in Spain 1980-1991," *Education Economics* 6, pp. 3-10.

Levin, Jesse. 2001. "For Whom the Reductions Count: A Quintile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement," *Empirical Economic* 26, pp. 221-246.

Mallier, Tony and Timothy Rogers. 1995. "Measuring Value Added in Higher Education: A Proposal," *Education Economics* 3, pp. 119-133.

Marlow, Michael L. 2000. "Spending, School Structure, and Public Education Quality. Evidence from California," *Economics of Education Review* 19, pp. 89-106.

Mariani, Matthew. 1999. "High-Earning Workers Who Don't Have a Bachelor' Degree," *Occupational outlook Quarterly* 43, pp.8-15.

McHaffie, Patrick H. 1998. "Contingency in the Local Provision of Public Education," *Growth and Change* 29, pp. 196-215.

Miller, Paul and Charles Mulvey. 1995. "What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings," *American Economic Review* 85, pp. 586-600.

Moll, Peter G. 1998. "Primary Schooling, Cognitive Skills and Wages in South Africa," *Economica* 65, pp.263-284.

Newell, Andrew and Barry Reilly. 1999. "Rates of Return to Educational Qualifications in the Transitional Economies," *Education Economics* 7, pp. 67-85.

Odden, Allan. 1990. " School Funding Changes in the 1980s," *Educational Policy* 4, pp. 33-48.

Pearl, Judea. 2000. Causality: Models, Reasoning and Inference. Cambridge University Press, Cambridge. 2000

Pritchett, Lant. 2001. "Where Has All the Education Gone?" *World Bank Economic Review* 15, pp.367-391.

Pryor, Frederic L. and David Schaffer. 1997. "Wages and the University Educated: A Paradox Resolved," *Monthly Labor Review* 120, pp.3-15.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update," *World Development* 22, pp. 1325-1344.

Psacharopoulos, George and Eduardo Velez. 1996." Returns to Education During Economic Boom and Recession: Mexico 1984, 1989, and 1992," *Education Economics* 4, pp. 219-231.

Psacharopoulos, George. 1995. "The Profitability of Investment in Education." Working Paper.

Rosenzweig, Mark R. 1995. "Why Are There Returns to Schooling?" *The American Economic Review* 85, pp.153-158.

Rubenstien, Ross, Amy Ellen Schwartz and Leanna Stiefel.2000. "Conceptual and Empirical Issues in the Measurement of School Efficiency" Proceedings of the National Tax Association, v 91. pp. 267-274.

Rupert, Peter and Mark E. Schweitzer. 1996. "Earnings. Education, and Experience," *Economic Review* 32, pp. 2-13.

Selected Interest Rates. Federal Reserve Statistical Release http://www.federalreserve.gov/releases/h15/data.htm. Access Date: 09/12/02.

The Data Page. University of New York. http://pages.stern.nyu.edu/~adamodar/New_Home_Page/data.html Access Date: 09/10/02.

Toh, Mun Heng and Chai Shing Wong. 1999. "Rates of Return to Education in Singapore," *Education Economics* 7, pp.235-253.

Vila, Luis E. 2000. "The Non-Monetary Benefits of Education," *European Journal of Education* 35, pp. 21-33.

Weinberger, Catherine J. 1998. "Race and Gender Wage Gaps in the Market for Recent College Graduates," *Industrial Relations* 37, pp. 67-84.

White, Halbert (1980) "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–838.

Wilson, Kathryn. 2001. "The Determinants of Educational Attainment: Modeling and Estimating the Human Capital Model and Education Production Functions," *Southern Economic Journal* 67, pp. 518-551.

Wolter, Stefan C. 2000. "Wage Expectations: A Comparison of Swiss and US Students," *Kyklos* 53, p

Appendix: Data, Summary Statistics, and Additional Estimation Results

As mentioned in the text, the data on school outcomes was collected from the West Virginia Department of Education, West Virginia Report Cards, for various years. Most of these data are self explanatory and can be found at the Department of Education website: **http://wvde.state.wv.us/data/report_cards/**.
Data on the number of nonstandard testing in the SAT 9 was provided by the WVEIS Director, Marshall Patton.

Data on demographic and economic information was obtained from the Department of Census, various tables. We matched the reported school zip code to the zip code census information for the 2000 Census. As previously mentioned, aligning school districts with Federally available demographic and economic information is an enormous task, that would likely yield little or no improvement in data quality.

Summary statistics for selected variables are contained in the following tables.

Table A-1: Demographic Data

	Mean	Median	Std. Dev.
POP	9718.293	5621	11593.82
POPCHANGE	-359.6493	-319.5	4470.145
POVERTYFAMWITHCH	246.4216	155	284.7881
HSGRAD	2494.944	1463	2870.332
MARRIEDWCHILD	804.5192	482	911.7686
MEDAGE	39.49512	39.7	2.962784
MEDGROSSRENT	371.8163	372	80.47546
MEDHHINC	28873.77	28051	7193.665
SOMECOLLEGE	1485.533	731	1923.347
SQUAREMILES	476.3993	423	210.5559
TOTALADULTS	6614.338	3794	7834.679
TOTALADULTWCOLL	463.4007	155	720.5107
UNMARRIEDPARTNHH	181.1359	83	252.4871
CHILDLT5	584.3763	311	718.1433
CHILD6TO11	664.7003	373	778.6519
CHILD15TO17	345.5087	206	402.0033
CHILD12TO14	335.8955	197	393.7643

Table A-2 Selected 3-Year Changes

	Mean	Median	Std. Dev.
DSAT9GRADE7	1.464158	1	9.434279
DSAT9GRADE8	1.073874	0	9.602357
DSAT9GRADE9	-1.09048	0.5	15.26598
DSAT9GRADE10	2.122951	2	7.162319
DSAT9GRADE11	2.786885	2.5	4.875694
DSAT	-1.34262	-0.9	7.285148
DPSAT10TH	-1.81721	0	6.55166
DPSAT11TH	-1.63689	-2.55	12.402
DAPTT10TH	0.078689	0	0.472418
DAPTT11TH	0.598361	0	4.654279
DAPTT12TH	0.596721	0	5.844697

Table A-3 Selected Inputs

	Mean	Median	Std. Dev.
AVG CLASS SIZE	19.17002	19.5	3.13397
AVG PRINC SALARY MI	56819.97	57493.87	5100.459
AVG PRINSALARY HIG	61443.78	61441.4	5350.859
AVG SALARY HIGH SCH	37378.27	37065.07	1543.115
AVG SALARY MIDDLE S	36905.96	37016.92	1563.546
ENROLL ENG LANG	95.21934	98.56667	16.31715
ENROLL FOR LANG	25.45751	24.95	14.21794
ENROLL MATH	85.19897	84.73333	14.02039
ENROLL SCIENCE	81.85082	82.65	17.26867
ENROLL SOCIAL STUD	86.94568	89.26667	17.56052
FTE HIGH SCHOOL	108.2515	78	89.61389
FTE MIDDLE SCHOOL	96.60063	70.68	81.58229
FTE PRINC HIGH SCHO	3.571181	3	2.748957
FTE PRINCIPAL MIDD	3.954545	3	2.941143

We performed regression analysis on several independent variables mentioned, but not illustrated in the text. The following tables illustrate these findings.

Table A-4, Results of Enrollment in English Language

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.939645	2.944113	-0.319161	0.7502
TEACHERED	0.001352	0.002211	0.611482	0.5421
PERFAMINPOVERTY	-15.39785	15.48059	-0.994655	0.3219
PERCOLLGRAD	1.833978	4.639311	0.395313	0.6933
AGEOFSCHOOL	0.056633	0.046405	1.220419	0.2247
ENROLL	-0.005867	0.003912	-1.499508	0.1364
AVGCLASSSIZE	0.119566	0.146136	0.818181	0.4149
AVGATTEND	0.975436	0.066069	14.76380	0.0000
DROPOUTPERCENT	2.069221	1.513188	1.367458	0.1741
POP/SQUAREMILES	0.007753	0.009763	0.794036	0.4288
R-squared	0.545415	Mean dependent var		96.46823
Adjusted R-squared	0.510743	S.D. dependent var		12.46045
S.E. of regression	8.715700	Akaike info criterion		7.243034
Sum squared resid	8963.685	Schwarz criterion		7.465848
Log likelihood	-453.5541	F-statistic		15.73080
Durbin-Watson stat	0.258851	Prob(F-statistic)		0.000000

Table A-5, Results of Enrollment in Foreign Language

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-20.53829	2.988344	-6.872800	0.0000
TEACHERED	0.002535	0.001904	1.331033	0.1857
PERFAMINPOVERTY	-22.41758	5.528427	-4.054966	0.0001
PERCOLLGRAD	38.01375	10.40400	3.653763	0.0004
AGEOFSCHOOL	0.031487	0.031878	0.987740	0.3253
ENROLL	0.004521	0.002675	1.689713	0.0937
AVGCLASSSIZE	-0.376631	0.349613	-1.077280	0.2836
AVGATTEND	0.471844	0.068471	6.891129	0.0000
DROPOUTPERCENT	0.183793	0.650926	0.282356	0.7782
POP/SQUAREMILES	0.015267	0.010742	1.421232	0.1579
R-squared	0.460042	Mean dependent var		26.13268
Adjusted R-squared	0.418859	S.D. dependent var		11.17009
S.E. of regression	8.515254	Akaike info criterion		7.196500
Sum squared resid	8556.127	Schwarz criterion		7.419315
Log likelihood	-450.5760	F-statistic		11.17061
Durbin-Watson stat	1.262012	Prob(F-statistic)		0.000000

Table A-6, Enrollment in Mathematics

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.678059	2.450953	-0.276651	0.7825
TEACHERED	-0.001334	0.001447	-0.922439	0.3582
PERFAMINPOVERTY	0.566711	4.523877	0.125271	0.9005
PERCOLLGRAD	2.560279	7.722737	0.331525	0.7408
AGEOFSCHOOL	-0.028771	0.024740	-1.162926	0.2472
ENROLL	0.002036	0.001833	1.110617	0.2690
AVGCLASSSIZE	-0.198806	0.241830	-0.822091	0.4127
AVGATTEND	0.989488	0.041413	23.89312	0.0000
DROPOUTPERCENT	-0.746473	0.483623	-1.543503	0.1254
POP/SQUAREMILES	0.023236	0.006521	3.563009	0.0005

R-squared	0.657025	Mean dependent var	83.45729
Adjusted R-squared	0.630866	S.D. dependent var	9.799879
S.E. of regression	5.954050	Akaike info criterion	6.480925
Sum squared resid	4183.184	Schwarz criterion	6.703740
Log likelihood	-404.7792	F-statistic	25.11654
Durbin-Watson stat	1.529209	Prob(F-statistic)	0.000000

Table A-7, Enrollment in Science

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.902748	3.430713	-0.263137	0.7929
TEACHERED	0.002001	0.002399	0.833925	0.4060
PERFAMINPOVERTY	-12.90848	14.05160	-0.918649	0.3602
PERCOLLGRAD	6.673177	7.706241	0.865944	0.3883
AGEOFSCHOOL	0.029082	0.045968	0.632652	0.5282
ENROLL	-0.006051	0.003793	-1.595217	0.1133
AVGCLASSSIZE	0.333637	0.230359	1.448332	0.1502
AVGATTEND	0.788452	0.068876	11.44735	0.0000
DROPOUTPERCENT	0.652575	1.370163	0.476275	0.6348
POP/SQUAREMILES	0.032253	0.013240	2.435966	0.0163

R-squared	0.424261	Mean dependent var	80.27734
Adjusted R-squared	0.380349	S.D. dependent var	11.93302
S.E. of regression	9.393431	Akaike info criterion	7.392803
Sum squared resid	10411.91	Schwarz criterion	7.615617
Log likelihood	-463.1394	F-statistic	9.661546
Durbin-Watson stat	0.624031	Prob(F-statistic)	0.000000

Table A-8, Enrollment in Social Studies

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.880906	5.133878	-0.171587	0.8641
TEACHERED	0.002370	0.002591	0.914826	0.3621
PERFAMINPOVERTY	-14.80476	15.60759	-0.948562	0.3448
PERCOLLGRAD	5.537102	9.280801	0.596619	0.5519
AGEOFSCHOOL	0.084222	0.049574	1.698922	0.0920
ENROLL	-0.002251	0.004085	-0.550949	0.5827
AVGCLASSSIZE	0.397670	0.254795	1.560748	0.1213
AVGATTEND	0.798679	0.081879	9.754446	0.0000
DROPOUTPERCENT	0.803951	1.453745	0.553021	0.5813
POP/SQUAREMILES	-0.006707	0.012741	-0.526395	0.5996

R-squared	0.405153	Mean dependent var	86.63568
Adjusted R-squared	0.359783	S.D. dependent var	12.92134
S.E. of regression	10.33882	Akaike info criterion	7.584594
Sum squared resid	12613.17	Schwarz criterion	7.807409
Log likelihood	-475.4140	F-statistic	8.930040
Durbin-Watson stat	0.754617	Prob(F-statistic)	0.000000

Table A-9, Advanced Placement Test Average, 10th Grade

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.444876	1.657366	0.268424	0.7889
TEACHERED	3.94E-06	4.90E-05	0.080420	0.9360
PERFAMINPOVERTY	-0.059191	0.117351	-0.504395	0.6150
PERCOLLGRAD	1.288947	0.396734	3.248892	0.0015
AGEOFSCHOOL	0.000941	0.000737	1.277399	0.2042
ENROLL	1.02E-06	8.34E-05	0.012191	0.9903
AVGCLASSSIZE	7.95E-05	0.008617	0.009230	0.9927
AVGATTEND	-0.006129	0.016849	-0.363742	0.7167
DROPOUTPERCENT	0.012730	0.016827	0.756506	0.4510
POP/SQUAREMILES	-0.000624	0.000281	-2.218581	0.0286

R-squared	0.153624	Mean dependent var	0.095278
Adjusted R-squared	0.084375	S.D. dependent var	0.260549
S.E. of regression	0.249315	Akaike info criterion	0.139459
Sum squared resid	6.837397	Schwarz criterion	0.371750
Log likelihood	1.632458	F-statistic	2.218425
Durbin-Watson stat	1.976002	Prob(F-statistic)	0.025900

Table A-10, Advanced Placement Test Average, 11th Grade

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.662937	29.08353	0.022794	0.9819
TEACHERED	0.001531	0.000857	1.787264	0.0766
PERFAMINPOVERTY	1.958040	2.895918	0.676138	0.5004
PERCOLLGRAD	17.85260	4.155834	4.295793	0.0000
AGEOFSCHOOL	0.011436	0.017935	0.637640	0.5250
ENROLL	0.002156	0.001620	1.330493	0.1861
AVGCLASSSIZE	-0.077529	0.117806	-0.658112	0.5118
AVGATTEND	-0.049451	0.300834	-0.164379	0.8697
DROPOUTPERCENT	0.039273	0.300089	0.130870	0.8961
POP/SQUAREMILES	0.004156	0.009454	0.439568	0.6611

R-squared	0.275983	Mean dependent var	2.813333
Adjusted R-squared	0.216745	S.D. dependent var	4.232486
S.E. of regression	3.745821	Akaike info criterion	5.558814
Sum squared resid	1543.429	Schwarz criterion	5.791105
Log likelihood	-323.5288	F-statistic	4.658900
Durbin-Watson stat	1.955368	Prob(F-statistic)	0.000031

Table A-11, Advanced Placement Test Average, 11th Grade

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.49186	40.22040	0.260859	0.7947
TEACHERED	0.001066	0.001766	0.603788	0.5472
PERFAMINPOVERTY	10.18688	6.061628	1.680552	0.0957
PERCOLLGRAD	33.04025	10.40157	3.176469	0.0019
AGEOFSCHOOL	-0.003821	0.027996	-0.136469	0.8917
ENROLL	0.000532	0.002823	0.188470	0.8509
AVGCLASSSIZE	0.011529	0.199915	0.057667	0.9541
AVGATTEND	-0.170708	0.414622	-0.411719	0.6813
DROPOUTPERCENT	0.185260	0.453603	0.408418	0.6838
POP/SQUAREMILES	-0.001887	0.009321	-0.202392	0.8400

R-squared	0.198612	Mean dependent var	4.681111
Adjusted R-squared	0.133044	S.D. dependent var	6.070005
S.E. of regression	5.651809	Akaike info criterion	6.381484
Sum squared resid	3513.724	Schwarz criterion	6.613775
Log likelihood	-372.8890	F-statistic	3.029103
Durbin-Watson stat	1.344195	Prob(F-statistic)	0.002852

Table A-12, PSAT 10th Grade

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	148.8246	55.09671	2.701152	0.0080
TEACHERED	-0.001028	0.002045	-0.502781	0.6161
PERFAMINPOVERTY	-15.75672	5.646753	-2.790404	0.0062
PERCOLLGRAD	46.36999	15.09807	3.071252	0.0027
AGEOFSCHOOL	-0.009775	0.035058	-0.278810	0.7809
ENROLL	0.002018	0.002618	0.770758	0.4425
AVGCLASSSIZE	-0.022001	0.319074	-0.068952	0.9452
AVGATTEND	-1.474680	0.582495	-2.531663	0.0128
DROPOUTPERCENT	-1.042591	0.608859	-1.712369	0.0896
POP/SQUAREMILES	0.004168	0.012111	0.344172	0.7314

R-squared	0.329501	Mean dependent var	7.540278
Adjusted R-squared	0.274642	S.D. dependent var	8.617461
S.E. of regression	7.339313	Akaike info criterion	6.904023
Sum squared resid	5925.207	Schwarz criterion	7.136314
Log likelihood	-404.2414	F-statistic	6.006332
Durbin-Watson stat	1.748299	Prob(F-statistic)	0.000001

Table A-13, PSAT 11th Grade

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-70.47136	95.28781	-0.739563	0.4611
TEACHERED	-0.000748	0.002421	-0.309129	0.7578
PERFAMINPOVERTY	-23.80169	6.293446	-3.781981	0.0003
PERCOLLGRAD	36.76743	19.72576	1.863929	0.0650
AGEOFSCHOOL	-0.018930	0.049984	-0.378726	0.7056
ENROLL	-0.006226	0.003813	-1.632984	0.1053
AVGCLASSSIZE	-0.293687	0.472290	-0.621837	0.5353
AVGATTEND	1.122037	1.021877	1.098016	0.2746
DROPOUTPERCENT	0.335079	0.739714	0.452985	0.6515
POP/SQUAREMILES	0.029214	0.013738	2.126552	0.0357

R-squared	0.216169	Mean dependent var	23.17319
Adjusted R-squared	0.152037	S.D. dependent var	11.02121
S.E. of regression	10.14887	Akaike info criterion	7.552257
Sum squared resid	11329.96	Schwarz criterion	7.784548
Log likelihood	-443.1354	F-statistic	3.370707
Durbin-Watson stat	2.122560	Prob(F-statistic)	0.001104