Quantile Regression: An Education Policy Research Tool

Edward B. Reeves
Morehead State University

Jesse Lowe
University of Oregon

QUANTILE REGRESSION:
AN EDUCATION POLICY RESEARCH TOOL

EDWARD B. REEVES
MOREHEAD STATE UNIVERSITY

and

JESSE LOWE
UNIVERSITY OF OREGON

ABSTRACT

Ordinary least squares regression is often used in education policy research. Unfortunately, OLS regression coefficients may mislead. In OLS models, the coefficients express the conditional mean relations among the variables. Still, what if this estimate of central tendency in the conditional distribution fails to convey important information about the distribution? Quantile regression is a statistical technique that allows variation in the conditional distribution to be examined. Therefore, it can be used to check the validity and applicable range of OLS coefficients. Following the method of Koenker and Hallock (2001), we compare OLS and quantile regression results, examining variables related to eighth-grade math achievement (NELS:88 database, N = 20,763). The independent variables include school, student and family characteristics that are staples of education policy research. Our graphically-displayed findings reveal striking disparities between the OLS and quantile regression coefficients. We propose that policymakers could be led astray by the simple aggregations achieved by OLS regression to choose "one size fits all" interventions. Moreover, our empirical results suggest that the received wisdom of education policy research may have to be revised because of the new findings that the quantile regression approach will produce.

Francis Galton in a famous passage defending the "charms of statistics" against its many detractors, chided his statistical colleagues

[who] limited their inquiries to Averages, and do not seem to revel in more comprehensive views. Their souls seem as dull to the charm of variety as that of a native of one of our flat English counties, whose retrospect of Switzerland was that, if the mountains could be thrown into its lakes, two nuisances would be got rid of at once (Galton 1889:62).

It is the fundamental task of statistics to bring order out of the diversity—at times the apparent chaos—of scientific observation. And this task is often very effectively accomplished by exploring how *averages* of certain variables depend on the values of the "conditioning" variables. The method of least squares, which pervades statistics, is admirably suited for this purpose. And yet, like Galton, one may question whether the exclusive

focus on conditional mean relations among variables ignores some "charm of variety" in matters statistical (Koenker 2005:xiii).

In this paper,[1] we will argue that when it comes to education policy research, more than "charm of variety" is at stake when the evidence of conditional mean relations among variables is allowed exclusively to guide policymaking. The very success of education policy may hinge upon policymakers acquiring fine-grained knowledge of the varying responses to policy that are likely to occur in the target population. Our plan is to discuss and give an extended empirical example of a statistical procedure that will allow education policy researchers to foresee heterogeneous consequences of proposed policy initiatives. The procedure is quantile regression. The developer and foremost proponent of this procedure is the economist-statistician Roger Koenker, whom we quoted above.

Our presentation is divided into four parts. First, we discuss how the design of policy often relies upon information that reflects only the central tendencies of the target population. Such information is often acquired by using multiple regression analysis, or similar statistical techniques. We also consider the advantages for purposes of policy design of obtaining estimates of the effects of policy variables not on the mean of the response variable but also across entire distribution of the response. Second, we introduce quantile regression as a procedure that allows researchers to make more nuanced analyses of potential policy outcomes. Third, we offer an extended example of the use of quantile regression for analyzing the covariates of eighth grade mathematics achievement. Fourth, we conclude with a discussion of how education policy research would benefit from the enhanced information that comes from supplementing traditional multivariate analysis methods with quantile regression.

POLICY DESIGN AND IMPLEMENTATION, STATISTICAL METHODS, AND THEIR UNINTENDED CONSEQUENCES

Applying Arthur Stinchcombe's (1959) well-known distinction, it may be said that policymaking is craft administered production, while policy implementation is bureaucratically administered production. Herein lies a dilemma: The bureaucratic implementation of a policy often prefigures its design. That is to say, policymakers often envision their task to be that of devising a one-size-fits-all solution for a public

---

[1]This paper was first presented at the annual meeting of the Rural Sociological Society in Louisville, KY, on August 11, 2006.

need, justifying this based on its cost efficiency and reliability of implementation in a bureaucratic setting.

Education policy offers many examples of this dilemma. School busing policy, implemented throughout the United States during the late 1960's to eliminate segregation failed to consider how parents in population centers of varying size might react to the prospect of their children being compulsorily bused. Investigating this matter with national data, Coleman, Kelly, and Moore (1975) found that the response to busing led to the re-segregation of schools in many parts of the country, the opposite of what policymakers had intended. Closer to the present time, the comprehensive No Child Left Behind legislation has been shown to reap unintended consequences. In devising a comprehensive policy for improving educational achievement nationwide, Congress did not foresee how the States or local school districts would react to this new law. Although most States and school districts have complied with the letter of the law, a number have found ways to subvert its intended goals (Chubb 2005; Linn, Baker, and Betebrenner 2002). Many problems of policymaking and implementation, of which these examples are illustrative, stem from insufficient knowledge about the heterogeneity of response that exists within the population that the policy is intended to influence.

Clearly, policymaking is an art, not wholly a science, and it is highly politicized. In the education arena, policymakers must mobilize recognition of a need, such as that of increasing student achievement in mathematics. This means mobilizing diverse groups of stakeholders and constituents—educational administrators and teachers, parents, business leaders, voters, and taxpayers—to perceive the need as important and worthy of the investment of public resources. Already, the magnitude of this task should be evident. Such diverse groups of stakeholders and constituents will have disparate interests, which are not easily harmonized or brought into line to support the policy goal with a single voice. Efforts at persuasion may have to be launched through public forums and the media. Much political capital may be spent to gain the support of opinion leaders and recalcitrant groups. Rarely will this process yield unitary agreement in a straightforward manner.

Nevertheless, let us assume that a critical mass has been reached to support the policy goal. Policymakers are still in for a difficult task because they must fashion a practical means of attaining the goal. Often this will involve eliciting the testimony of experts regarding how best to do it. Among these experts will be policy researchers, whose task is to provide sound information about the characteristics of the population that the policy will influence. If the policy goal is to increase math achievement among eighth graders, for example, policy

researchers will be expected to provide information about the characteristics of eighth-grade students, including the personal and contextual factors that constrain (or facilitate) their ability to learn math (McDonald et al. 2006). In their deliberations, policymakers will be focused on devising ways to alleviate the constraints and take advantage of incentives. The rub is that their knowledge of how the incentives and constraints affect students' learning may be inadequate. We suggest that an important reason for this inadequacy is the manner in which policy researchers go about producing the information.

Multiple regression analysis has been the mainstay of education policy research since the 1960s. During the ensuing decades, there have been noteworthy innovations in analytic techniques so that ordinary least squares (OLS) regression analysis has been supplemented and occasionally replaced by more sophisticated approaches, such as structural equation modeling and hierarchical linear modeling. Nevertheless, these advanced approaches share an essential characteristic with OLS regression: they estimate conditional mean relations between the dependent variable and the predictors. Provided certain assumptions are met, the mean effect is the single best overall estimate of the relationship between a constraint and its influence on student learning, but as often happens the devil is in the details. The estimation of the mean conditional effect is no guarantee that the effect is uniform across the distribution of the student achievement measure. If it is not uniform, policymaking that is based on thinking that students will respond similarly can go seriously awry.

Unfortunately, policymakers are probably predisposed to accept the evidence of the mean effects because it makes the work of devising and implementing policy easier and less expensive. The fact that policy must be implemented through a bureaucratically administered process compels policymakers to seek cost-efficient methods. It is not that bureaucratic agencies cannot implement complex policy. Yet to do so will usually require growing new bureaucratic structures to break down the complexity into manageable components (cf. Simon 1997; Stinchcombe 1990). Policymakers will naturally want to limit the growth of bureaucratic structures for the sake of cost containment. Therefore, without strong evidence of inherent complexity that must be attended to if the policy is going to be successful, a simple policy design will exert great appeal and information that supports the presumption of simplicity in how policy targets will respond will be believed (Wilson 1989).

Policy researchers wish to please their clients, the policymakers, so they provide information on incentives, constraints, and other relevant conditions that is relatively straightforward and not too complex. This may explain why policy

researchers seldom question their own methods and continue to produce summary statistical results that stress mean effects, rather than more thorough, and complex, results about the distribution of the effects. Since policy development is fraught with political pitfalls and uncertainties in its implementation (Lipsky 1980), no one ever thinks to question the adequacy of the policy research methods and how the results are summarized, except perhaps for other researchers. What ensues from the critique of other researchers can only be described as bouts of internecine dispute that are quite opaque to the other stakeholders, including policymakers and implementers.

Multiple regression (as well as structural equation modeling and hierarchical linear modeling) is the common procedure used to obtain a summary of the relationship between a response variable y and a set of covariates x. Ordinary least squares regression captures how the mean of y changes with unit increases in each x. But sometimes—and we suspect more often than not in policy research—the estimation of mean effects is not a sufficient knowledge base for successful policy design. In education policy research it can be important to know, not just the mean effects of the x's, but also how the effects vary across the distribution of y. Suppose, for instance, that the effect of gender differences on math achievement is not constant across the achievement distribution and so it is not well represented by the mean of the effect. In other words, suppose gender differences change with math achievement. An example of this might be that, compared with boys, girls show a significant disparity in math achievement test scores at the high end of the distribution but almost no disparity at the low end. If policymakers wish to elevate math achievement overall, such a finding could prompt them to design interventions that focus on both boys and girls in the low-achievement range of test scores but more on girls instead of on boys in the high-achievement range. To facilitate the latter objective could require the creation of a policy intervention that, first, identifies potentially high performing girls and, second, creates incentives to encourage them to enroll in more rigorous math courses and to put forth greater effort in these courses. Useful policy knowledge such as this is attainable with quantile regression.

QUANTILE REGRESSION: A DIFFERENT APPROACH

The need to examine the distribution of a regression model in a more thoroughgoing manner was proposed by Mosteller and Tukey in 1977:

180   SOUTHERN RURAL SOCIOLOGY

What the regression curve does is find a grand summary for the averages
of the distribution corresponding to the set of x's. We could go further and
compute several different regression curves corresponding to the various
percentage points of the distributions and thus get a more complete picture
of the set. Ordinarily this is not done, and so regression often gives a rather
incomplete picture. Just as the mean gives an incomplete picture of a single
distribution, so the regression curve gives a correspondingly incomplete
picture for a set of distributions (quoted in Koenker and Hallock 2001:20).

A year later, an estimation model that implemented this suggestion was devised by
Koenker and Basset (1978). Instead of using a model that minimizes the sum of
squared deviations about the mean, Koenker and Basset recommend examining
percentiles, i.e., quantiles, by using a method that minimizes the sum of absolute
deviations about the τth percentile cutpoint on the distribution of the response
variables. In other words, instead of estimating the conditioning effect of X on the
mean of Y (as with OLS regression and similar methods), they proposed estimating
the conditioning effect of X on Y at various cutpoints along the distribution of Y.
Employing the latter method, one can observe the conditioning effect of X at
selected quantiles of Y.

Perhaps the most commonly used example of a quantile is the median m, defined
as the point within any given distribution where half the values are greater than m
and the remaining values are equal to or less than m. The use of median m is far
more reliable when dealing with data that is non-normal in its distribution, such as
income.

In an OLS regression model, the effect of a conditioning variable on the mean
of y is obtainable by minimizing $\sum_{i=1}^{n} [y_i - (\beta_0 + \ldots x_i\beta_i)]^2 = \sum_{i=1}^{n} [y_i - y]^2$, which

yields the conditional expectation $E[Y|X=x]$. In a quantile regression model, the
locational effect of a conditioning variable is derived by minimizing

$$\sum_{t \in \{t:} \theta |y_i - (\beta_0 + \ldots + x_i\beta_i)| + \sum (1-\theta)|y_i - (\beta_0 + \ldots + x_i\beta_i)| \text{ where } \theta \text{ represents}$$

the θth conditional quantile of y given x. Thus, the median conditional expectation
of $E[Y|X=x]$ occurs where half the cases are less than or equal to y at θ and the
remaining cases are greater than y, where $\theta = 0.5$. By extension, the conditional
expectation at the 25th quantile occurs where one-quarter of the cases are less than

or equal to *y* at θ and three-quarters of the cases are greater than *y* at θ, where $\theta =$ 0.25. To obtain a more thorough picture of variation in the conditional relation, therefore, a series of quantile effects may be estimated for various cutpoints across the distribution—including the median and select cutpoints on either side of it. (For in depth explanations of quantile regression, see Hao and Naiman 2007; Koenker 2005; Koenker and Hallock 2001; Eide and Showalter 1998; Yu, Lu, and Stander 2003).

Quantile regression is a little-used method for education policy research (though for some exceptions, see Eide and Showalter 1988; Fertig 2003; Schneeweis and Winter-Ebmer 2005). The techniques of quantile regression have found their most extensive use to date in economics, not surprisingly (Koenker and Hallock 2001; Martins and Pereira 2004). The method is currently receiving intense attention from statisticians and has gained broad acceptance in biomedical research (Yu et al. 2003). It is also making inroads into environmental science and ecological studies (Cade and Noon 2003).

AN EMPIRICAL EXAMPLE

Education policy researchers are often asked to estimate the effects of individual student characteristics and the contextual effects of families and schools on student achievement. No valid inferences about an intervention can be made without understanding these influences (McDonald et al. 2006). Consequently a vast research literature, much of it done by sociologists, has developed that details the effects of student status characteristics and context variables on student academic performance (for an overview, see Hallinan 2000; Riordan 2004). Our purpose here is to compare the results of OLS and quantile regression methods using a large sample of data that includes variables denoting math achievement, private versus public schools, school location, student demographic characteristics such as gender and race, and variables denoting family resources. These are variables used in hundreds of studies, and so their effects are generally thought to be well understood. The question that we will address in our empirical example is: Do the mean effects of individual student, family, and school characteristics on student achievement adequately account for the relationships being analyzed, or does quantile regression provides additional information that could be useful for policy design and implementation?

Sample

Quantile regression analysis is more robust when performed on a large sample. The sample that we use for this study consists of questionnaire responses from students, their parents, and school officials collected during the base year of the National Education Longitudinal Survey of 1988 (NELS:88). NELS:88 is a national probability sample of eighth grade students. The survey was commissioned by the National Center for Education Statistics (NCES) to gauge student achievement based on social, demographic, and locational characteristics of students, parents, teachers, schools, and school board administrations. To observe the influences of, and differences between, racial and ethnic factors that affect educational outcomes better both Hispanic and Asian/Pacific Islanders were oversampled (for a more thorough discussion of NELS:88, see Schneider 1993). Although NELS:88 was designed to facilitate the study of student achievement and factors that influence it longitudinally by following the eighth grade cohort for a period of 12 years, in the present study we focus on the student cohort in their eighth-grade year only. The entire eighth-grade cohort making up the NELS:88 sample numbered about 25,000 students. Because we excluded cases with missing values, our study sample is restricted to 20,763 cases.

Data and Variables

In preparing our analysis we examined the math achievement test scores of the eighth-grade students as these relate to variables commonly used in education policy research. To measure students' academic achievements, their mathematics knowledge was assessed via standardized tests given as part of the NELS:88 protocols. Independent variables considered in our study include race/ethnicity, gender, school location, students' postsecondary education plans, parents' highest education level, whether or not the student lives in an intact nuclear family, and family income.

To observe racial/ethnic influences on math achievement, five classifications were used; White (the reference category), African American, Hispanic, Asian, and Native American. Regarding gender, female performance in math was compared with male performance. School location refers to whether the school attended by the student is located in an urban, rural, or suburban area.[2] In our analyses, the

---

[2]The classification of the location of the student's school reflects its status at the time of the 1980 decennial census. 'Urban' means central city; 'suburban' means the area surrounding a central city within a county constituting the MSA (Metropolitan Statistical Area); 'rural' means outside MSA.

academic achievement of students attending urban and rural schools is contrasted with those attending suburban schools. Another predictor variable is postsecondary education plans. This is a measure of student educational aspirations and ranges from "will not finish high school" to "will obtain a doctorate degree." Parents' highest education level is also a ranked categorical variable with an array of categories that is similar to postsecondary education plans. Intact nuclear family indicates whether both of students' biological parents are present within the household; if both parents are present, the nuclear family is intact. Lastly, family income is measured on an interval scale that ranges from no income to $350,000 annually. Here, the original NELS:88 categories of family income were manipulated by recoding the income categories to the midpoint of their respective ranges. For the quantile and OLS regression comparison, family income was manipulated further using a natural log transformation. To facilitate the log transformation, one dollar of income was substituted for "no income."

Table 1 displays the descriptive statistics for all of the variables that we included in our OLS and quantile regression comparisons. In our sample, the standardized math test score averaged 50.9 with a standard deviation of 10.3. Twenty percent of the eighth graders making up the sample attended a private school. Thirty percent attended an urban school and nearly an equal proportion attended a rural school (29 percent). Half the sample is female, and nearly 70 percent is White. African Americans and Hispanics each make up 12 percent of the sample of eighth graders, while Asians and Native Americans comprise 6 percent and 1 percent respectively. On average, student education plans indicate that respondents intend to pursue a four-year degree (although nothing can be said about whether a degree will be obtained, of course). Further, 65 percent of students reside with an intact nuclear family. Regarding yearly family income, the mean income is more than $44,000 (in 1988) but the distribution is characterized by a large positive skew. Parents' highest education level is slightly less than student's postsecondary education plans; on average, parents have completed high school and may have some college, whereas most students plan to obtain a college degree.

RESULTS

In this section, we compare the OLS and quantile regression estimates of the effects of commonly used social and demographic variables on mathematics achievement. In both methods, all independent variables have been mean centered so that the intercept represents the achievement test score when all independent variables are at their mean values. We compare the OLS and quantile regression

Table 1. DESCRIPTIVE STATISTICS

VARIABLE	MEAN	SD	MIN.	MAX.
Standardized math test score	50.91	10.31	33.9	77.2
Private school	0.20	0.40	0	1
Urban	0.30	0.46	0	1
Rural	0.29	0.45	0	1
Gender (female)	0.50	0.50	0	1
African American	0.12	0.32	0	1
Hispanic	0.12	0.32	0	1
Asian	0.06	0.24	0	1
Native American	0.01	0.10	0	1
Postsecondary education plans	4.63	1.28	1	6
Intact nuclear family	0.65	0.48	0	1
Family income ($)	44,302.44	51,565.04	0	350,000
Parents' highest education level	3.13	1.26	1	6

N=20,763 (national probability sample of 8th grade students drawn from NELS:88).

findings in two formats. First, we show and comment on the actual coefficients that were estimated using OLS and quantile regression methods. Second, we resort to graphical displays of the results that permit easy assessment of how adequately each OLS coefficient captures the quantile distribution of the conditional math achievement test score.

OLS and Quantile Regression Coefficients Compared

Table 2 displays the coefficients obtained from the OLS and quantile regression models. Coefficients obtained via quantile regression are pegged for the 5th, 25th, 50th, 75th, and 95th quantiles. The OLS results reveal that all but one covariate has significant effects on the mathematics achievement test score, the one exception being rural school location. Moreover, the significance level of most of the OLS coefficients is uniformly high (p < .001). The direction and magnitude of the OLS coefficients is what one would expect to find after reading the research literature on factors that influence student achievement: Private school students often score

Table 2.   METRIC OLS AND QUANTILE REGRESSION COEFFICIENTS FOR COVARIATES OF 8TH GRADE MATH ACHIEVEMENT (N = 20,763)

VARIABLE	OLS COEF.	Q5	Q25	Q50	Q75	Q95
Intercept	50.90***	38.39***	44.43***	50.34***	56.81***	65.65***
Private school	1.32***	.95***	2.00***	1.83***	1.26***	.16
Urban	-.31*	.01	-.20	-.42*	-.70**	-.03
Rural	.04	.32*	.37*	.01	-.18	-.38
Gender (female) .	-.81***	-.01	-.34*	-.92***	-1.27***	-1.45***
African American	-5.83***	-1.69***	-4.36***	-6.16***	-7.31***	-6.75***
Hispanic	-3.14***	-1.14***	-2.37***	-3.13***	-3.74***	-3.97***
Asian	1.78***	.37	1.05*	2.99***	2.58***	1.39***
Native American	-4.99***	-2.18***	-4.22***	-4.90***	-6.06***	-6.64***
Postsecondary education plans ..	2.18***	.65***	1.54***	2.20***	2.69***	2.74***
Intact nuclear family	.56***	.34**	.37*	.60***	.69**	.47
Family income (ln)	.77***	.27***	.54***	.69***	.74***	.90***
Parents' highest education level ...	1.89***	.75***	1.60***	2.15***	2.23***	1.63***
R^2	.33	.06	.15	.21	.22	.17

NOTE: *p<.05, **p<.01, ***.001; The covariates were mean centered prior to analysis; pseudo-R^2 coefficients for the quantile regression results.

higher on achievement tests than public school students. Urban school students perform less well than their suburban counterparts. The finding of no significant difference in math achievement between rural and suburban schools would surprise some educational researchers perhaps (see Roscigno and Crowley 2001), but other research supports a view that rural school students are not inferior when family background is controlled (Fan and Chen 1999), as it is in our study. Girls performed less well on the math achievement test than boys. African American, Hispanic, and

Native American students performed considerably less well than White students, while Asian students performed somewhat better. No surprises here. Postsecondary education plans are positively related to math achievement. Finally, the family background variables—intact nuclear family, family income, and parents' highest education level—are each positively related to math achievement, just as previous research has shown.

Turning to the conditional quantile coefficients shown in Table 2, we note immediately that the sizes of the coefficients vary (sometimes quite substantially) at different locations within the distribution. The significance level also changes for some variables at different locations. Commonly, the pattern of asymmetry deviates sharply from the estimates obtained with OLS. First, however, note that there is hardly any skewing of the dependent variable (intercept). This is indicated by the very close approximation of the OLS estimated mean coefficient and the median quantile coefficient of 50.3. Let us turn to the effects of the predictor variables.

School Factors. School characteristics are among those variables that change in both coefficient size and statistical significance. First, note that the OLS coefficient for private school is highly significant ($p < .001$) and is shown to positive affect achievement. When observing the coefficient throughout the distribution of achievement scores though, the influence changes at every measured quantile. At the 5th quantile, the influence of private schools has less of an effect on academic achievement that is nearly 0.4 less than the OLS result. At the 25th and 50th quantiles, the effect is *larger* than predicted by the OLS model. At the upper end of the distribution, the private school effect loses both magnitude and statistical significance.

Results for both the urban and rural dummy variables further illustrate changes in the coefficients and significance throughout the distribution of standardized math test scores. The OLS results suggest that students in urban schools perform less well than their suburban counterparts. The quantile regression results show that the significant negative effect of urban location occurs between the 50th and 75th quantiles; the effect is insignificant elsewhere on the distribution. Certainly, we would not want to draw a policy conclusion from this finding alone, but it points forcefully to a need for additional investigation.

Students in rural schools are shown to perform no differently than those in suburban schools by the OLS model; however, at the 5th and 25th quantiles, rural students perform significantly *better* than suburban students. Clearly, the OLS model does not capture this conditional quantile effect. Smaller class sizes coupled with less reliance on tracking in rural schools could be hypothesized to account for

these asymmetric results. From these examples, OLS evidently estimates of the effects of school characteristics on students' academic performance miss important differences that are easily detected in the quantile model.

Student Factors. Student characteristic variables follow the same pattern as school characteristic variables in that the coefficient differs depending on the point within the distribution where it is estimated. First, female student performance on the standardized math achievement test is shown to be significantly less than their fellow male students by the OLS regression model—and this difference is highly significant. Still, this difference increases monotonically in the conditional distribution, from being a nil difference at the 5th quantile to a difference of -1.45 at the 95th quantile. The OLS result only captures the central tendency of this monotonic result.

Student variables measuring the effect of race/ethnic background on mathematics achievement somewhat parallel the gender effect. As the location within the distribution changes from low to high, the absolute value of the effect also often increases. First, the OLS coefficient for African American suggests that, on average, African American students score nearly six points less than White students. Yet the difference between White and African-American academic achievement in math is only 1.7 points at the lower end of the distribution. This difference grows to 7.3 points at the 75th quantile and remains high (6.8 points) at the 95th quantile. From the median to the upper end of the distribution, the African-American performance deficit is larger than what the OLS model predicts. Again, we might not want to draw strong policy inferences from these results alone. Nevertheless, it would appear that the shortfall in African American achievement in math is three or four times more acute at the higher end of the performance distribution than at the lower end. The distributional pattern of conditional quantiles is similar for Hispanic students, although the largest disparity with White students is at the 95th quantile, and the relationship between quantile and coefficient size is monotonic.

Asian students outperform White students by nearly two points according to the OLS model. In comparison, the quantile regression results show that differences in standardized mathematic test scores are not significant for the left tail of the distribution; the difference is most significant in the upper half of the distribution. At the median, Asian students score three points higher than White students. Above the median, the difference in test scores decreases somewhat but remains statistically significant. Regarding Native American students, here too, the relationship between quantile and the coefficient is monotonic (as with Hispanic

188   SOUTHERN RURAL SOCIOLOGY

students). Hence, the OLS model is not giving an accurate picture of the magnitude of effect throughout the distribution. Here, the difference between Native American and White test scores is smallest at the 5th quantile (2.2 points) and is largest at the 95th quantile (6.6 points).

The effect of postsecondary education plans is very significant (p < .001) across the distribution of test scores. However, it also displays a remarkable pattern of monotonic increase. As the location within the distribution changes from low to high, so to does the independent variable's magnitude of effect. The only point in the distribution where the OLS coefficient closely corresponds to the conditional quantile coefficient is at the median. OLS overestimates the effect of postsecondary education plans by 1.5 points at the lower end of the distribution and underestimates it by 0.5 points at the upper end. This difference between modeled results may not seem large enough to warrant our interest; but when we compare the test score of a student that plans to attain a postgraduate degree to the score of a student not wanting to finish high school, the predicted difference is large.

Parental and Family Factors. Parental and family variables have effects that conform to patterns already seen in other variables. These variables all have positive effects that are typically very significant throughout the distribution of the test scores. Usually, the OLS coefficient does not capture these disparate location effects. Furthermore, there are tendencies toward monotonic relationships between these regression coefficients and their locations within the test score distribution.

Regarding intact nuclear family, the OLS result predicts that students who live with both biological parents score 0.6 higher than students in other family arrangements. The quantile regression results show that the intact family advantage varies between 0.3 at the 5th quantile to 0.7 at the 75th quantile. At the 95th quantile, the coefficient is 0.5 and it is not statistically significant. Thus, in the lower end of the distribution, the effect of this measure of family structure is not as large as the OLS model suggests while at 75th quantile the effect is greater than the predicted mean effect though not by a wide margin.

The effects of family income and parents' highest education level on the standardized math achievement score are each highly significant throughout the conditional distribution. The monotonic growth of family income's effect as the quantile increases is clearly marked by the results shown in Table 2. The range of effects varies from 0.3 at the lower end of the distribution to 0.9 at the upper end. The OLS estimate of 0.8 hardly seems to do justice to this variation. The quantile regression results for parents' highest education level are also characterized by wide deviation from the OLS estimate of 1.9. The effect of parents' highest education

ranges from a low of 0.8 at the 5th quantile to a high of 2.2 at the 75th quantile, before declining to 1.6 at the 95th quantile.

The bottom panel in Table 2 shows the goodness-of-fit statistics for the various regression models. The R^2 of 0.33 indicated for the OLS model compares favorably with much of the research that has used large national samples to detect multivariate influences on student achievement. The goodness-of-fit statistics for the quantile regression models are pseudo-R^2 coefficients (based on change in the deviance statistic). Consequently, they cannot be directly compared with the OLS goodness-of-fit statistic, although they can be compared amongst themselves. When this is done, we see that the most robust models are located at the median and 75th quantiles. Goodness-of-fit declines on either side of this region within the distribution. Since more of the variance is explained at the 50th and 75th quantiles, this suggests that these commonly used social and demographic variables are optimal predictors of achievement in the third quartile of the distribution for these data. The very low pseudo R-square of 0.06 at the 5th quantile raises a warning flag about the goodness-of-fit of the coefficients for the left tail of the conditional distribution.

Graphical Results Showing the Adequacy of the OLS Estimates

During the preceding section we compared the OLS and quantile regression coefficients. We repeatedly called attention to the deviations seen in the asymmetric quantile coefficients, and we suggested that these appeared to depart—very substantially in some cases—from the OLS mean-effect estimates. We are now prepared to deal with this issue more forthrightly. Figure 1 shows a series of plots in which the OLS mean effect and its 90 percent confidence interval is compared with the regression quantile effects and their 90 percent confidence intervals. The graphs show the distribution of the quantile estimates between the intervals of .05 and .95 in increments of .05. From these plots we can assess more clearly how well the OLS estimates represent the conditional distribution of math achievement.

The graphs in Figure 1 sometimes illustrate striking differences between OLS and quantile results, but not always. For instance, the effect of private school does not deviate significantly from the OLS estimate across the conditional distribution except at the 90th and 95th quantiles, where the quantile regression estimates are significantly less than the OLS estimate. At no point in the distribution of math achievement do the effects of urban and rural location deviate significantly from the mean effect estimated by OLS. Such results indicate that the mean effects estimated

Figure 1.   COMPARISON OF THE RESULTS OF OLS AND QUANTILE REGRESSION: COVARIATES OF 8TH GRADE MATH ACHIEVEMENT.
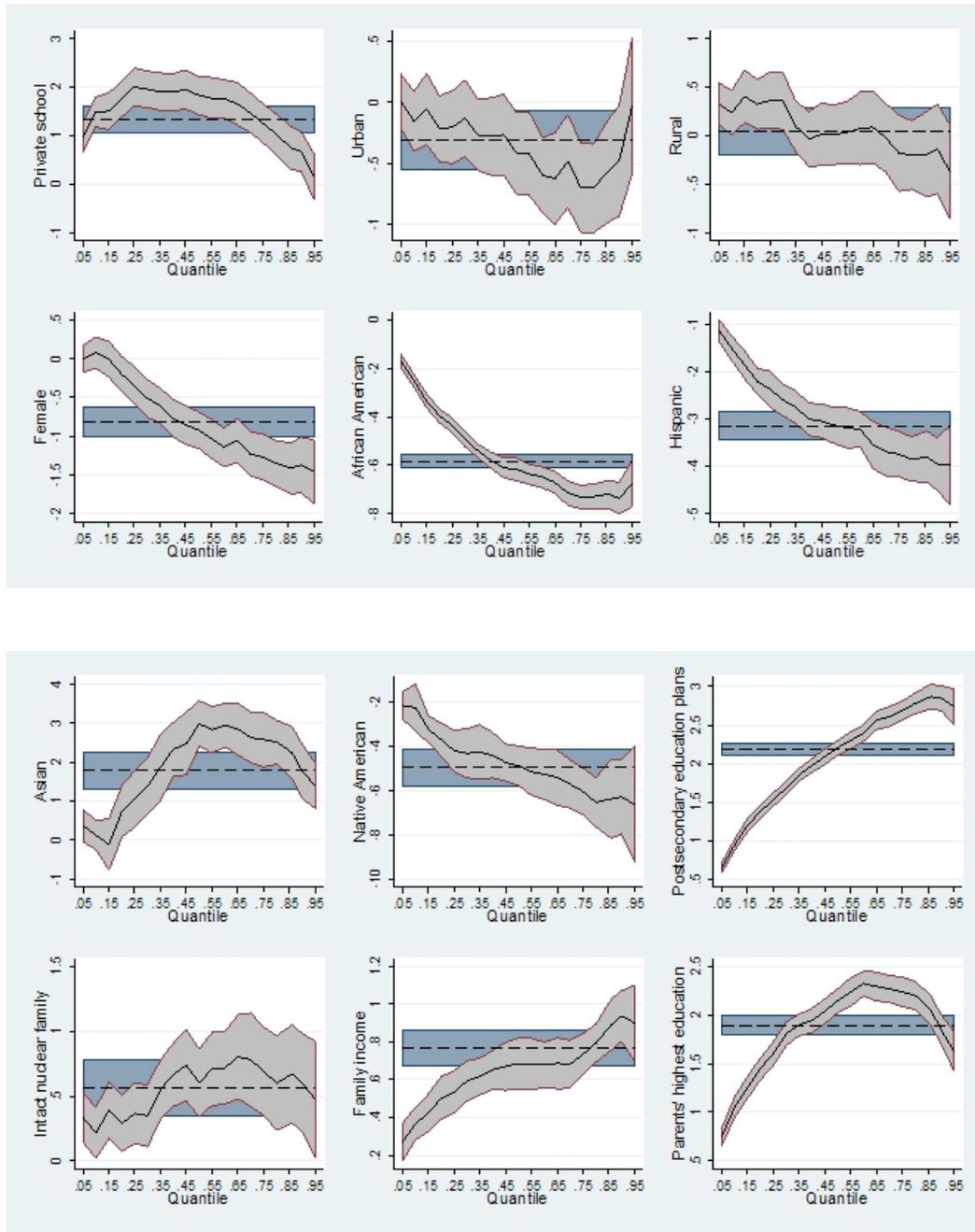
NOTE: 90% confidence bands for OLS and quantile regression coefficients.

by OLS regression are consistent throughout the distribution of math achievement for these school characteristic variables.

All student factors show a marked difference between the OLS and quantile regression results at some points along the conditional distribution. For example, female students are shown to perform better than the OLS estimate in the lower quartile of the distribution. The CIs provided by each approach overlap across the middle of the conditional distribution, suggesting that in these locations there is a negligible difference between the OLS and quantile regression results. Girls often score less than the OLS estimate in the upper quartile of the math achievement distribution.

The quantile results for the ethnicity and race of students deviate in several instances from the OLS estimates. The deviations are especially noteworthy for African American students. The CIs overlap between the 35th and 60th quantiles and at the 90th quantile and above. This suggests that the OLS mean effect represents the math achievement of African Americans for about one-third of the conditional distribution. In the lower end of the distribution, below the 35th quantile, African Americans score significantly higher than the OLS estimate indicates, although they still perform less well than comparable White students. Between the 65th and 90th quantiles, African Americans perform significantly lower than the OLS model indicates. The graph comparing regression results for Hispanic students shows a significant deviation at the lower end of the distribution. In the lowest quartile, Hispanic students do not perform as poorly compared with Whites as the OLS model shows. For Asian students, math achievement is greater than that of White students across nearly the entire distribution of math test scores. Further, the OLS mean effect represents this result well, except the lower 20 percent of the distribution where Asians perform little better than their White counterparts. For Native Americans the findings are essentially reversed. Native American students perform less well than White students across the entire distribution, and the OLS estimate represents this finding with good consistency, except in the lower 15 percent where the disparity in achievement test results between Native American and White students is less pronounced.

Of all model-based deviations shown in Figure 1, postsecondary education plans exhibits the highest degree of difference. The OLS regression model accurately captures the effect of postsecondary education plans in the middle of the distribution, between the 40th and 60th quantiles only. Below this range, the OLS model significantly overestimates the positive effect of postsecondary education plans. Above this range, OLS significantly underestimates this effect. The effect of

192   SOUTHERN RURAL SOCIOLOGY

postsecondary education plans increases monotonically with location in the distribution. This striking result may be due to endogeneity of the relationship between this covariate and math achievement.

The three graphs at the bottom of Figure 1 display the OSL and quantile regression effects of parental and family characteristics. When the student resides in an intact nuclear family, the effect of math achievement is often positive throughout the distribution of achievement test scores. Furthermore, the wide, overlapping CI's show that for this covariate the OLS model yields an estimate that well represents the distribution of conditional effects. The same cannot be said for the estimated effects of the final two covariates. The OLS model represents the positive effect of family income in the upper 75 percent of the distribution. However, in the lower quartile of the distribution, the OLS coefficient significantly overestimates this effect. Thus, the quantile results suggest family income is not as influential on academic achievement as predicted by the OLS regression model among lower achieving students. The comparison of the results of OLS and quantile regression are even more striking with respect to the effect of parents' highest education level. As seen previously for postsecondary education plans, the narrow CI bands associated with the estimates of the effect of parents' education result in a strong conclusion that the OLS-estimated mean coefficient does not represent the distribution of the quantile coefficients. In the lower quartile of the distribution, the OLS model significantly overestimates the positive effect. Between the 50th and 85th quantiles it underestimates this effect. The OLS and quantile coefficients converge in two locations—between the 25th and 50th quantiles and above the 85th quantile. In more than half the conditional distribution, therefore, OLS misrepresents the results. It is also noteworthy that the quantile estimates of parents' education increase monotonically from the lowest location in the distribution through the 60th quantile. Beyond that point of the distribution, however, the quantile regression estimates display a decreasing trend.

These graphical comparisons of OLS and quantile regression results reveal a marked deviation between the two methods at some points in the distribution for nine out of the twelve covariates. Moreover, there are also relationships, such as African American, postsecondary education plans, and parents' highest education level, where the conditional quantile and OLS confidence intervals seldom overlap. Since so many effects are either overestimated or underestimated at some point using the OLS approach, examining the influence of covariates on academic achievement with quantile regression methods in addition to OLS and similar

conditional mean estimation methods seems only prudent for education policy researchers.

A pattern that most quantile-modeled effects demonstrate is the tendency for their coefficients to approach zero in the left tail of the conditional distribution. The effects of these covariates decrease as the quantile decreases. Lastly, the influence of several covariates increases monotonically in a positive or negative direction with the increase in the quantile. To us, these patterns cast doubts on much of the received wisdom of education policy research. It can no longer be assumed that conditional mean relations are optimal summaries of the constraints experienced by educational target groups. Hence, we argue for the advisability of including quantile regression in the tool kit of researchers seeking to investigate influences on educational achievement.

DISCUSSION

1. Chapter 23 of James Coleman's *Foundations of Social Theory*, is concerned with social policy research (Coleman 1990). As Coleman sees it, social policy research arose from a broad and apparently irreversible structural change in American society that first appeared in the 1920s and 30s. This change was marked by the rapid growth in the number and size of corporate actors (see also, Coleman 1982), including the federal government.

In the 1920s, 1930s, and 1940s, the United States underwent a change in the structure of social interaction, from a structure that was almost entirely local and personal to one that had a large component of national and impersonal interactions. World War II provided a stimulus to this change, by inducing an even higher rate of migration from rural to urban areas and, more generally, by inducing persons to move away from the locality and even the region of their birth. All this helped bring about a change which became evident in the 1960s: change in the structure of responsibility from private and local to public and national (Coleman 1990:620).

Social policy research, including research that dealt with education policy, was an outgrowth of this change. Increasingly, federal government branches and agencies needed information about particular populations that could be influenced by policy to fulfill their obligations to care for the general welfare of the nation. Research was necessary to learn how a government program should be designed, or modified, whether it should be continued, or if it was working in the manner that

its designers intended. Because social policy was enacted at the national level while it was implemented at the local level, the old, indirect, informal methods of gaining opinions about these matters were no longer effective (Coleman 1990:623).

Social policy research, including education policy research, nearly always involves the study of the behavior of natural persons directed toward corporate actors or in response to the actions of corporate actors (Coleman 1990:626). According to Coleman, governmental corporate actors are particularly disposed to sponsoring policy research when their legitimacy has been threatened and they are uncertain how to proceed. Coleman goes on to suggest that overcoming the "error of simple aggregation" is vital for policy research (Coleman 1990:646), an error that is often brought on by researchers' use of OLS regression and similar methods. Policymakers want to know the societal, or aggregate, effects of their policy. This is typically supplied by collecting information on the individual responses of a representative sample of policy recipients and then aggregating the responses using statistical methods that involve the calculation of conditional mean functions. Coleman believes that this approach produces distorted information and proposes several ways to avoid the error of simple aggregation. He does not seem to have been aware that quantile regression might also be useful in this regard.

Quantile regression is one among several methods that may inform policymakers and policy implementers about the varying ways in which incentives and constraints influence the responses of policy recipients. Furthermore, quantile regression is especially well suited for assessing the aggregation assumptions built into policy research.

2. One of the foremost issues in education policy today is how effectively to "scale up" exemplary interventions. Exemplary interventions are teaching strategies and/or materials (including technologies) demonstrated effectively through randomized experimental trials.

> Individual scale-up studies investigate whether a given intervention leads to improvements in clearly defined outcomes among a particular population of students. Each produces an essentially dichotomous answer—either the intervention does or does not lead to improvement in a given set of circumstances (McDonald et al. 2006:16).

Because experiments on student achievement are beset with ethical, administrative, and fiscal difficulties, most random experimental trials in education will probably be performed using limited samples of students. This poses a major

difficulty. The sampling frame of the experimental and control groups should be guided by a broad understanding of the larger population of students who may eventually benefit from promising interventions discovered by experimentation. As is well known, the results obtained from experiments have strong internal validity, but their external validity is often weak or unknown.[3] This is precisely the "scaling up" problem. McDonald et al. (2006) broadly discuss methodological issues that must be considered when "scaling up" exemplary interventions and argue for the kinds of research that sociologists have traditionally performed on the characteristics of student populations and on the factors that facilitate and constrain their academic achievement.

Our empirical example shows a clear role for quantile regression to supplement the findings of conditional mean estimation approaches, which have informed nearly all sociological research on education issues. Quantile regression will help researchers handle the "scaling up" problem in two respects. First, studies of the conditional quantile effects of school, student, and family characteristics can make a contribution to better understanding student performance across the distribution of achievement. Such knowledge can guide the experimental researcher to select more efficient sampling frames. For instance, if the disparity in math performance between boys and girls is largely a phenomenon of the upper end of the achievement distribution, then experimental research designed to create an exemplary intervention to augment girls' performance should focus on the appropriate target group. Second, quantile regression may help researchers identify appropriate target groups when "scaling up" exemplary interventions. This approach has the potential to greatly enhance the external validity of experimental research aimed at finding exemplary interventions that will enhance educational achievement in diverse student populations and schools.

3. Presently, the complex error structures found in hierarchical linear modeling and similar methods pose insurmountable difficulties for the estimation of regression quantiles (Koenker and Hallock 2001). This is disappointing because multilevel methods have become the methodology of choice in much education policy research. HLM is one way in which Coleman's concern about the error of simple aggregation has been tackled with new statistical methods. This merely

---

[3]By internal validity we mean that the experiment is powerful in detecting the effect of the intervention *in the experimental group*. By external validity we mean that, because of the restricted sample which experimentation usually requires, one is unable to draw strong conclusions about successfully replicating the experimental results in different locations and with different populations (Shadish, Cook, and Campbell 2002:38 and elsewhere).

serves to reinforce our position that quantile regression will supplement, but not replace, existing conditional mean estimation methods, such as OLS and HLM.

4. Although the present study has focused on the distributed main effects of predictors of math achievement, nothing prevents the exploration of interaction effects with quantile regression. Exploring issues such as how impoverished rural students fare in comparison with impoverished urban students (Alspaugh 1992; Reeves 2003; Reeves and Bylund 2005; Roscigno and Crowley 2001) or how minority students perform in the public schools compared with their counterparts in private schools (Bryk, Lee, and Holland 1993; Coleman and Hoffer 1987; Hoffer, Greeley, and Coleman 1982) could produce results that will unsettle current knowledge of these issues. Reanalysis of earlier findings with quantile regression may suggest novel avenues for empirical exploration and hypothesis testing and could have broad consequences for policy formation and implementation.

5. At times, education policy research seems to be engaged in diagnosing empirical relationships without any theoretical underpinnings. However, as Sørensen and Morgan (2000) argue, such a view impoverishes policy research and yields results that can be criticized because they are ad hoc and the product of unexamined assumptions. Sørensen and Morgan argue convincingly that the choice of method and theory must intersect. We believe that the ability of quantile regression to test the adequacy of conditional mean estimates in the manner that we showed in Figure 1 is a potential boon to the elaboration of powerful, well-supported theories regarding student achievement.

6. We performed the same analysis on the NELS:88 sample with science achievement test scores, instead of math achievement scores. The results of this second analysis were very close to what has been presented here, although the OLS and quantile regression differences were not quite so pronounced. We attribute this to the fact that science knowledge is more diffuse than mathematics, which is highly structured and cumulative. Because of this difference, the measurement of science achievement entails greater error.

CONCLUSIONS

Two objectives of this paper have been to provide a general description of quantile regression and to suggest how this statistical procedure may play a large future role in education policy research. We used empirical evidence of school, student, and family influences on eighth-grade math achievement and compared the results of OLS and quantile estimations to make our case. Our findings raise more questions than we have space to answer here, but we believe that we have provided

ample reason to revisit many findings of education policy research to ascertain how the received wisdom might be altered by greater attention to the "charm of variety" gleaned from using quantile regression. At the very least, we hope that our analysis casts doubt on policy researchers' devotion to the estimation of conditional mean relations among variables.

Matt Riddley, the science writer, when interviewed on Book TV (broadcast November 25, 2006), remarked that advancement in a science often depends upon the development of a new technology.[4] Could quantile regression play a fundamental role in furthering advancements in education policy research? This prospect is not out of the question. The hallmark of science, according to Mr. Riddley, is to see the world in a fresh way and to ask new questions. Quantile regression does seem to hold out the promise of seeing student achievement with a different perspective and could provoke new questions about the conditions that promote achievement.

REFERENCES

Alspaugh, John W. 1992. "Socioeconomic Measures and Achievement: Urban vs. Rural." *Rural Educator* 13:2–7.

Bryk, Anthony, Valerie Lee, and Peter Holland. 1993. *Catholic Schools and the Common Good.* Cambridge, MA: Harvard University Press.

Cade, Bryan S. and Barry R. Noon. 2003. "A Gentle Introduction to Quantile Regression for Ecologists." *Frontiers in Ecology and the Environment* 1:2–20.

Chubb, John E. 2005. "Saving No Child Left Behind." Pp. 1–34 in *Within Our Reach: How America Can Educate Every Child,* edited by J.E. Chubb. Lanham, MD: Rowman & Littlefield.

Coleman, James S. 1982. *The Asymmetrical Society.* Syracuse, NY: Syracuse University Press.

_____. 1990. The Relationship of Sociology to Social Action in the New Social Structure. Pp. 610–49 in *Foundations of Social Theory.* Cambridge, MA: Harvard University Press.

Coleman, James S., Sara D. Kelly, and John A. Moore. 1975. *Trends in School Segregation, 1968-73* (Urban Institute Paper no. 722-03-01). Washington, DC: The Urban Institute.

---

[4]His specific example was the role played by X-ray crystallography in the discovery of the double helix structure of DNA.

Coleman, James S. and Thomas Hoffer. 1987. *Public and Private High School.* New York: Basic Books.

Eide, Eric and Mark H. Showalter, 1998. "The Effect of School Quality on Student Performance: A Quantile Regression Approach." *Economics Letters* 58:345–50.

Fan, Xitao and Michael J. Chen. 1999. "Academic Achievement of Rural School Students: A Multiyear Comparison with Their Peers in Suburban and Urban Schools." *Journal of Research in Rural Education* 15(1):31–46.

Fertig, Michael. 2003. *Who's to Blame? The Determinants of German Students' Achievement in the PISA 2000 Study.* Discussion Paper no. 739. Bonn, Germany: Forschungsinstitut zur Zukunft der Arbeit (Institute for the Study of Labor).

Galton, Francis. 1889. *Natural Inheritance.* London: Macmillan.

Hallinan, Maureen T. (ed.). 2000. *Handbook of the Sociology of Education.* New York: Kluwer Academic/Plenum Publishers.

Hao, Lingxin and Daniel Q. Naiman. 2007. *Quantile Regression* (Quantitative Applications in the Social Sciences, no. 149). Thousand Oaks, CA: Sage.

Hoffer, Thomas, Andrew M. Greeley, and James S. Coleman. 1985. "Achievement Growth in Public and Catholic Schools." *Sociology of Education* 58:74–97.

Koenker, Roger. 2005. *Quantile Regression.* New York: Cambridge University Press.

Koenker, Roger and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica* 46:33–50.

Koenker, Roger and Kevin F. Hallock. 2001. "Quantile Regression: An Introduction." *Journal of Economic Perspectives.* Retrieved December 12, 2005 (http://www.economics.ox.ac. uk/members/s.bandyapadhya/ KoenkerHallock_QR_JEP2001.pdf).

Linn, Robert L., Eva L. Baker, and Damian W. Betebrenner. 2002. "Accountability Systems: Implications of the Requirements of the No Child Left Behind Act of 2001." *Educational Researcher* 31(6):3–16.

Lipsky, Michael. 1980. *Street-Level Bureaucracy.* New York: Russell Sage Foundation.

Martins, Pedro S. and Pedro T. Pereira. 2004. "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Counties." *Labour Economics* 11:355–71.

McDonald, Sarah-Kathryn, Venessa A. Keesler, Nils J. Kauffman, and Barbara Schneider. 2006. "Scaling-Up Exemplary Interventions." *Educational Researcher* 35(3):15–24.

Reeves, Edward B. 2003. "Disentangling the Effects of Nonmetro Location and Student Poverty on School Performance/Improvement: Implications for

Equitable Excellence in Kentucky Public Schools." *Journal of Research on Rural Education* 18(1):17–30.

Reeves, Edward B. and Robert A. Bylund. 2005. "Are Rural Schools Inferior to Urban Schools? A Multilevel Analysis of School Accountability Trends in Kentucky." *Rural Sociology* 70(3):360–86.

Riordan, Cornelius. 2004. *Equality and Achievement*, 2nd ed. Upper Saddle River, NJ: Pearson/Prentice Hall.

Rocigno, Vincent J. and Martha L. Crowley. 2001. "Rurality, Institutional Disadvantage, and Achievement/Attainment." *Rural Sociology* 66(2):268–92.

Schneeweis, Nicole E. and Rudolf Winter-Ebmer. 2005. *Peer Effects in Austrian Schools* (Economic Series, 170). Vienna, Austria: Institute for Advanced Studies.

Schneider, Barbara. 1993. "Parents, Their Children, and Schools: An Introduction." Pp. 1–12 in *Parents, Their Children, and Schools*, edited by B. Schneider and J. S. Coleman. Boulder, CO: Westview Press.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin.

Simon, Herbert A. 1997. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*, 4th ed. New York: The Free Press.

Sørensen, Aage B. and Stephen L. Morgan. 2000. "School Effects: Theoretical and Methodological Issues." Pp. 137–60 in *Handbook of the Sociology of Education*, edited by M.T. Hallinan. New York: Kluwer Academic/Plenum Publishers.

Stinchcombe, Arthur. L. 1959. "Bureaucratic and Craft Administration of Production." *Administrative Science Quarterly* 4:168–87.

———. 1990. *Information and Organization.* Berkeley: University of California Press.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It.* New York: Basic Books.

Yu, Keming, Zudi Lu, and Julian Stander. 2003. "Quantile Regression: Applications and Current Research Areas." *The Statistician* 52(3):331–50.