Aggregation and the Estimated Effects of School Resources

Eric A. Hanushek, Steven G. Rivkin, Lori L. Taylor

The Review of Economics and Statistics, Volume 78, Issue 4 (Nov., 1996), 611-627.

Stable URL:
http://links.jstor.org/sici?sici=0034-6535%28199611%2978%3A4%3C611%3AAATEEO%3E2.0.CO%3B2-B

AGGREGATION AND THE ESTIMATED EFFECTS OF SCHOOL RESOURCES

Eric A. Hanushek, Steven G. Rivkin, and Lori L. Taylor*

Abstract—This paper helps reconcile the contradictory findings about school resources and school effectiveness by developing the implications of data aggregation. With model misspecification, the theoretical impact of aggregation is generally ambiguous. When important state differences in school policy are omitted, however, aggregation implies clear upward bias of estimated school resource effects. Review of past studies and new empirical analysis provide strong evidence that aggregation inflates the coefficients on school resources. The pattern of results is also inconsistent with an errors-in-variables explanation. These results provide further support to the view that additional expenditures alone are unlikely to improve student outcomes.

I. Introduction

A key element of the policy discussion surrounding schools has been the effect of additional resources on student performance. In simplest terms, if schools effectively turn added resources into higher student achievement, policy makers can concentrate on the appropriate level and distribution of resources and can let the local school districts concentrate on uses of those resources. However, if schools do not effectively turn resources into performance, then education policy that is aimed at either the level or distribution of outcomes becomes much more complicated. Policy makers either must concentrate on picking good approaches and processes—something that school districts themselves may not be able to do—or they must turn to different incentive mechanisms that might alter the way districts spend their available resources.[1]

A large body of research casts doubt on the effectiveness of local school districts at turning added resources into higher student achievement. In comprehensive summaries of empirical evidence, Hanushek (1986, 1989) found that there was no consistent or systematic relationship between achievement and either pupil–teacher ratios, teacher salaries, years of teacher schooling, years of teacher experience or per-student expenditure. The inefficacy of smaller pupil–teacher ratios is particularly noteworthy, given that the widely-held belief that lower pupil–teacher ratios improve educational outcomes has played a prominent role in the crafting of educational policies. Between 1940 and 1990 the average pupil–teacher ratio in the United States public schools has fallen from 28 to less than 16.[2] Though the expansion of special education accounts for a portion of the decline, the vast majority of the change reflects a fall in mainstream class sizes.[3]

The conclusion that higher expenditures (including lower pupil–teacher ratios) have not raised achievement has not been universally accepted. Proponents of meta-analysis[4] point out that weighted averages of the individual estimates of school resource effects are positive, even if many coefficients are very imprecisely estimated or even negative. Others argue that concentration on standardized test scores as the measure of achievement in most research ignores the potential impact of smaller class sizes or higher teacher salaries on other student outcomes.[5] For example, school resources may be linked with future earnings even if they do not affect test scores. Betts (1996) summarizes the evidence in favor of such a linkage, including two widely cited recent studies by Card and Krueger (1992a, 1992b) which find that smaller classes and higher teacher salaries increase the wage premium associated with an additional year of schooling. While Betts casts doubt on the robustness of any relationship between school resources and the subsequent earnings of students, some studies still suggest positive resource effects.

This paper attempts to reconcile the contradictory results in the debate over the effectiveness of additional school resources in raising achievement. We concentrate on two specific and important differences among the various studies of educational performance: the level at which school characteristics are aggregated and the number and type of controls for differences in student backgrounds, opportunities and academic preparation.[6] As described below, many production function studies analyze the relationship between student outcomes and individual school and school district characteristics, while others, including Card and Krueger, use state average school characteristics. There is even greater variation in the variables used to control for differences in socio-economic backgrounds and opportunities.

The probability that studies report a positive and statistically significant relationship between achievement and school resources rises dramatically along with the level of aggregation (see below). This pattern holds when achieve-

Received for publication February 24, 1995. Revision accepted for publication April 8, 1996.

* University of Rochester, Amherst College, and Federal Reserve Bank of Dallas, respectively.

We would like to thank Julian Betts, Jeff Grogger, Finis Welch, Robert Willis, Geoffrey Woglom, and participants at the NSF/*Review of Economics and Statistics* conference on *School Quality and Educational Outcomes* (Harvard University, December 1994) for helpful comments.

[1] A discussion of alternative policy approaches in the absence of clear relationships between resources and policies is found in Hanushek with others (1994).

[2] These calculations include school principals in addition to teachers.

[3] Hanushek and Rivkin (1997) analyze the contribution of special education programs to the pupil–teacher ratio decline. Between 1980 and 1990, the period most often cited for the influence of special education, less than 40% of the fall in the pupil–teacher ratio could possibly be attributed to the expansion of special education.

[4] See Hedges et al. (1994) and the discussion below.

[5] Approximately two-thirds of the production function studies reviewed in Hanushek (1989) use some standardized test measure of outcomes, while the rest consider a range of outcomes including continuation in schooling, dropping out of high school, and subsequent income. The review here includes the Card and Krueger (1992a, 1992b) studies along with other studies meeting the criteria for inclusion. Several earlier analyses of the effects of school resources on earnings (for example, Welch (1966) and Johnson and Stafford (1973)) are not included, however, because they do not include family background differences and thus do not distinguish between resource effects and other influences on performance.

[6] Betts (1995) examines specific aspects of the Card and Krueger methodology in order to understand why their findings contradict much of the previous literature. Heckman, Layne-Farrar, and Todd (1996) examine the sensitivity of the estimates to sample and specification and question key identifying assumptions. Speakman and Welch (1995) also demonstrate the sensitivity of the estimates to specification. A third important difference, highlighted in Betts (1996), is the period in which students attended school. Hoxby (1995) examines the impact of the rise in teacher unionization on the efficiency of additional school resources.

ment is measured by test scores or other outcomes.[7] The interpretation of this is, however, controversial. Most researchers suggest that disaggregated analysis is generally superior, although some argue that aggregation of the relationships may actually have beneficial effects through reduction in measurement error or the bias introduced by the endogeneity of school and residential location choices. Thus a crucial issue is identifying whether aggregation reduces a downward bias present in school-level studies or introduces a positive bias which inflates school resource coefficients.

We present a theoretical model demonstrating that aggregation alters the degree of omitted variables bias, even when the true marginal impacts of included variables are constant across observations. Omitted variables have their clearest effects on estimates when the data are aggregated to the level of the omitted factors (such as when the data are aggregated to the state level and state-level determinants of students' performance are neglected). In this case, aggregation increases the magnitude of omitted variable biases. Given the dominant role of states in school organization, financing, and regulation, it is likely that state-level resources are correlated with a variety of important state influences on school performance. Therefore, aggregation-induced changes in the magnitude of omitted variables bias provide a plausible explanation for the pattern of school resource results.

The hypothesis that aggregation increases estimated effects of school resources by enlarging omitted variable biases is supported by direct empirical analysis. Using the High School and Beyond (HSB) data set, we estimate aggregated and disaggregated relationships between the primary determinants of school expenditures (teacher–pupil ratios and teacher salaries) and two measures of educational outcomes (standardized test scores and years of post-secondary schooling). Consistent with prior research into educational performance, our analysis finds that aggregation inflates the coefficients on school characteristics for both outcome measures. More importantly, the pattern of results is inconsistent with an errors-in-variables explanation. Rather, aggregation appears to exacerbate problems of omitted variables bias and produce incorrectly large school resource coefficients.

II. Prior Evidence on the Effects of School Resources

Since the Coleman Report (Coleman et al (1966)), a substantial number of analyses have been directed at uncovering the effects of different resources on student performance. A previous summary of these results (Hanushek (1989)) found 187 separate estimates of educational production functions, spread across 38 separate publications.[8] Here we expand on that universe by including both more recent studies and studies missed in previous searches. This updating through 1994 yields 377 studies spread across 90 publications.

For the purposes here we focus on two aspects of school resources: teacher–pupil ratios and expenditure per student. These resources are key to much of the policy making in education, are the most important determinants of the growth in schooling costs (see Hanushek and Rivkin (1977)), and enter centrally into recent debates about the estimated effectiveness of resources. Because data are seldom collected with analyses of student performance in mind, however, many less-than-perfect sources of data have been employed in an effort to tease out the effects of the key school resources. These analyses often involve merging data from different sources, and this merging frequently implies that specific resources cannot be directly matched with the students receiving them. Indeed, resource data are often aggregated across groups of students, say, at the school district level, and then related to individual or group performance. But, suspiciously, this aggregation of school resources in the empirical analyses tracks closely the results that are obtained.

Table 1 displays the overall results for teacher–pupil ratios.[9] Conventional policy arguments suggest that the teacher–pupil ratio should be positively related to student performance. The first thing to note from the table is that the combined estimates over the 277 separate investigations give little reason to believe that smaller classes are related to higher student performance.[10] Fifteen percent (41/277) of the estimates are statistically significant and positive, while 13% (37/277) are also statistically significant but negative. Disregarding statistical significance does not change the pic-

[7] Betts (1996) concentrates on studies in which earnings are used to measure achievement and finds in a broader set of studies than reviewed here that this conclusion is quite general. (His larger set includes studies that do not adequately measure family background—which will bias the expenditure coefficients upwards since expenditure tends to be positively related to the educational strength of families.)

[8] For these purposes, a "study" is a separate estimate of an educational production function. An individual publication may include several studies, pertaining to different grade levels or measurement of outcomes. Alternative specifications of the same basic model were not double counted; nor was publication of the same basic results in different sources. The attempt was to record the information from all published studies that included one of the central measures of resources (either real resources of pupil–teacher ratios, teacher experience, teacher education, facilities, or administrator characteristics or monetary resources of expenditures per student or salaries), recorded information about the statistical significance of the estimated relationships, and included some measure of family and nonschool inputs. The 1989 summary missed a few estimates that were available prior to mid-1988, and these have been incorporated in the update here. A complete listing of the included studies along with a more extensive analysis of the results can be found in Hanushek (1996a).

[9] The best way to tabulate the results across different studies has been the subject of lively debate; see Hedges et al. (1994) and Hanushek (1994, 1996b). Under certain circumstances (which do not appear applicable here), the simple tabulations presented below may lead to misleading conclusions. The statistical requirements to do anything beyond these simple counts are, however, unlikely to be met in the studies reviewed.

Technical issues aside, there is agreement on all sides of that debate that resources do not consistently matter regardless of how they are employed. Further, no consensus exists on when or when not resources are likely to matter.

[10] Not all studies contained information on each specific resource. Of the 377 studies looking at any of the identified resources, 277 analyzed either teacher–pupil ratios, pupil–teacher ratios, or class size. (All analyses of pupil–teacher ratios or of class size are put in terms of teacher–pupil ratios by reversing the signs.)

TABLE 1.—PERCENTAGE DISTRIBUTION OF ESTIMATED EFFECT OF TEACHER–PUPIL RATIO ON STUDENT PERFORMANCE (277 estimates)

Level of Aggregation of Resources	Number of Estimates	Statistically Significant		Statistically Insignificant		Insignificant, Unknown Sign
		+	−	+	−	
Total	277	15%	13%	27%	25%	20%
Classroom	77	12	8	18	26	36
School	128	10	17	26	28	19
District	56	21	16	39	20	4
County	5	0	0	40	40	20
State	11	64	0	27	9	0

Note: Rows may not add to 100 because of rounding.

ture of results centered on zero, or no effect. The second thing to note is that the positive estimates (both statistically significant and total) come disproportionately from studies which aggregate the school data to the state or even district level. While relatively few studies employ state level measures (11), almost two-thirds find positive and statistically significant effects of teacher–pupil ratios. At the district level, 21% find positive and statistically significant results and 39% find positive but statistically insignificant effects. When the resources are measured closer to the students, at the classroom or school level, any hint of disproportionate impact of smaller classes goes away.

Table 2 repeats the same consideration of aggregation for the estimated effects of expenditure on performance. For the 163 estimates of school expenditure effects, 87 come from relating spending in individual schools to the performance of the students in those schools while the remaining studies are aggregated to the district level or above. At the school level, 17% find positive and significant effects of spending on performance, while 7% incredibly find a statistically significant *negative* effect of spending. But this picture is again altered when the resources are no longer related to specific students through aggregation to higher levels. At the state level, 64% of the estimates suggest that higher spending is associated with student performance in a statistically significant manner. Only one of the 28 estimates of state-level resources (4%) is negative. Of the significant positive expenditure results in all work to date, 43% come from state-level aggregate estimates and another 29% come from aggregation to district levels.[11] This pattern of results—where findings of positive resource effects come disproportionately from highly aggregated measures—is precisely what is found in our own subsequent analysis of school resource effects.

Some previous speculation about why the production function results appear to differ from the results for earnings have centered on the measurement of outcomes. Table 3 divides the expenditure results into those related to test score measurement of outcomes and all other measures. Two-thirds of the studies consider test score measures with the remainder evaluating school attainment, dropout behavior, earnings, and other measures. This table makes it clear, however, that aggregation produces disproportionate numbers of estimates indicating that expenditure affects outcomes, regardless of how outcomes are measured. Seventy-five percent of the state-level estimates of expenditure on test performance and 60% of the state-level estimates of expenditure on nontest performance are positive and statistically significant, compared to only 19% and 12% of the estimates of classroom and school expenditure for test and nontest measures, respectively.[12]

Whether or not the evidence on school resource effects is consistent with a positive relationship between student outcomes and school resources depends crucially on the appropriateness of more aggregate specifications. If aggregation reduces downward bias present in school level studies, as might follow from measurement errors, the finding of no systematic effects in these more disaggregated studies should be weighted less. On the other hand, if aggregation worsens a specification error bias, the disaggregated study

[11] The separate analysis of resource effects by Hedges et al. (1994) concentrates on expenditure studies without regard to the aggregation of the underlying analysis.

[12] Betts (1996) finds a very similar pattern for wage and school attainment studies.

TABLE 2.—PERCENTAGE DISTRIBUTION OF ESTIMATED EFFECT OF EXPENDITURE PER PUPIL ON STUDENT PERFORMANCE (163 estimates)

Level of Aggregation of Resources	Number of Estimates	Statistically Significant		Statistically Insignificant		Insignificant, Unknown Sign
		+	−	+	−	
Total	163	27%	7%	34%	19%	13%
Classroom	4	0	0	0	0	100
School	83	17	7	35	23	18
District	43	28	9	37	26	0
County	5	0	0	40	20	40
State	28	64	4	32	0	0

Note: Rows may not add to 100 because of rounding.

TABLE 3.—PERCENTAGE DISTRIBUTION OF ESTIMATED EFFECT OF EXPENDITURE PER PUPIL ON STUDENT PERFORMANCE BY OUTCOME MEASURE AND AGGREGATION OF RESOURCE EFFECTS (163 estimates)

Measure of Outcome and Aggregation of Resources	Number of Estimates	Statistically Significant		Statistically Insignificant		Insignificant, Unknown Sign
		+	−	+	−	
A. Test Score Outcomes[a]						
Total	109	25%	9%	28%	21%	17%
Classroom	4	0	0	0	0	100
School	57	19	9	28	21	23
District	38	26	11	37	26	0
County	2	0	0	0	50	50
State	8	75	13	13	0	0
B. Other (Nontest) Outcomes[b]						
Total	54	31%	2%	46%	15%	6%
School	26	12	4	50	27	8
District	5	40	0	40	20	0
County	3	0	0	67	0	33
State	20	60	0	40	0	0

Note: Rows may not add to 100 because of rounding.
[a] All studies measure student performance by some form of standardized test score.
[b] All studies employ some outcome measure (such as income or school attainment) other than a standardized test score.

result of no systematic relationship between outcomes and spending should be accepted as the best measure of the impact of additional resources on student performance.

One specific class of specification errors seems particularly relevant. In the United States, schools are organized and regulated by the separate states, and the states maintain very different policies that directly affect the schools. For example, in 1990–91 state average spending per student varied between \$2,960 and \$8,645, while the local share of this expenditure varied between 0.5% and 86.8% (Hanushek, Rivkin, and Taylor (1996)). These spending patterns were motivated by state funding formulae with very different incentives (Gold et al. (1991)). While consideration of expenditure variation is fairly straightforward to incorporate, other policy attributes, which may be as important, are more difficult to integrate into any analysis. For example, 37 states had some form of teacher competency testing, although the details of these differ dramatically. Similarly, almost all states have course requirements for student graduation, but these range from more or less complete lists of basic courses to a single course in physical education as the only requirement. Other regulations govern teacher certification requirements, teacher tenure, and teacher pay.

Virtually none of the previously identified studies considers how the state structure for schooling affects outcomes or, relevant to this analysis, affects the estimates of resource effects. The previous comparisons by aggregation level of school resources also provide an incomplete picture of the potential impact of unmeasured state variations. While analyses aggregated to the state level necessarily must include data across different states (and regulatory regimes), the analyses at lower levels of aggregation may or may not draw data from multiple states. Importantly, any unmeasured, regulatory factors common to an individual state will not bias the estimated resource parameters if analysis is conducted entirely within an individual state. Of the 377 total studies, 58% use data within individual states and thus would not be affected by state-level omitted variables. But, just looking at studies where resources are measured below the state level (e.g., school or district), 35% of the estimates employ samples from more than one state.

Table 4 separates the studies by whether samples are drawn from individual states or multiple states. For those drawn from multiple states, it also distinguishes between the highest level of aggregation of measured resources (the state level) and other aggregation levels, i.e., whether or not there is within-state variation in the resource measures. The top panel considers teacher–pupil ratios. In the 157 studies of schools within single states, a disproportionate number find negative effects of teacher–pupil ratios. When multiple state samples are considered, however, studies of teacher–pupil ratios tend to find greater numbers of positive effects and greater numbers of statistically significant effects on student performance. For multiple state samples, 18% find positive and statistically significant effects of teacher–pupil ratios with only 8% finding statistically significant negative effects. This contrasts sharply with 12% and 18%, respectively, for results from single state samples. Moreover, while the number of estimates based on state aggregates is small, aggregation further reinforces the disproportionate finding of positive effects.

Similar results emerge from the bottom panel that considers the effects of expenditure per pupil. Results from multistate samples yield disproportionate numbers of positive and statistically significant effects: 35% versus 1% negative and statistically significant. Again, while the difference between multistate and single state samples is seen at all levels of aggregation, aggregation to the state level dramatically reinforces the prevalence of positive versus negative estimated effects. Finally, for both teacher–pupil ratios and expenditures, going to the combined point estimates (ignoring statistical significance) shows the same but stronger patterns:

TABLE 4.—PERCENTAGE DISTRIBUTION OF ESTIMATED EFFECT OF TEACHER–PUPIL RATIO AND EXPENDITURE PER PUPIL BY STATE SAMPLING SCHEME AND AGGREGATION

State Sampling Scheme and Aggregation of Resource Measures	Number of Estimates	Statistically Significant		Statistically Insignificant		Insignificant Unknown Sign
		+	−	+	−	
		A. Teacher–Pupil Ratio				
Total	277	15%	13%	27%	25%	20%
Single state samples[a]	157	12	18	31	31	8
Multiple state samples[b]	120	18	8	21	18	35
With within-state variation[c]	109	14	8	20	19	39
Without within-state variation[d]	11	64	0	27	9	0
		B. Expenditure per Pupil				
Total	163	27	7	34	19	13
Single state samples[a]	89	20	11	30	26	12
Multiple state samples[b]	74	35	1	39	11	14
With within-state variation[c]	46	17	0	43	18	22
Without within-state variation[d]	28	64	4	32	0	0

Notes: Rows may not add to 100 because of rounding.
[a] Estimates from samples drawn within single states.
[b] Estimates from samples drawn across multiple states.
[c] Resource measures at level of classroom, school, district, or county, allowing for variation within each state.
[d] Resource measures aggregated to state level with no variation within each state.

positive resource effects are more likely in multistate samples than in single state samples, and this likelihood is strongest in samples using aggregate state data.

The bias introduced by neglect of state financial and regulatory influences on achievement depends on the correlation between resources and state achievement factors. These results suggest omitted state factors tend to bias resource effects toward being more positive and toward appearing more statistically significant. The bias is strongest when resources are aggregated to the state level. The fact that resources appear more important in multistate samples regardless of the level of aggregation, however, points strongly toward misspecification bias as opposed to errors in variables (see below).

III. Aggregation, Omitted Variable Bias, and Measurement Error

Following the literature on estimating the value added by schools (see, for example, Hanushek (1979), Aitkin and Longford (1986) or Hanushek and Taylor (1990)), we model the relationship between educational attainment and student and family characteristics as

$$A_{ij} = \mathscr{F}(T_{ij}, F_{ij}, C_{ij}, S_{ij}) \qquad (1)$$

where A_{ij} is the level of educational achievement for individual i in school j, T_{ij} is a standardized pretest score, F_{ij} is a vector of individual and family characteristics, C_{ij} is a vector of community variables, and S_{ij} is a vector of school characteristics.

Adequate controls for differences in family background, community environment, and student preparation are needed in order to isolate the effects of school characteristics, because education goes on both inside and outside schools. The performance of any specific student will combine the influences of the school and of the outside environment, particularly the family. Moreover, parents may systematically select school districts through migration in accordance with their preferences (Tiebout (1956)) or otherwise attempt to secure good school resources for their children. In such a case, unmeasured parental inputs could be correlated with measured school resources.

Accounting for pre-existing differences in academic preparation is necessary in order to capture the impact of school factors during a given period. Education occurs over time, so that the achievement, say, of a ninth grader is determined in part by schools (and family) in the ninth grade and in part by these inputs in prior years. Since data on the past history of educational inputs are frequently unavailable, strategies that will isolate the achievement gains that might be related to specific measured inputs such as through estimation of value-added models of achievement are frequently employed. In addition, schools may use past performance in determining input allocations (e.g., lower class size for poorly performing students). Such a possibility reinforces the need to control for past performance in analyzing the effects of school resources.

These general issues in the estimation of educational production functions are set out at the beginning because they pervade the discussion here. Clearly, inadequate controls for academic preparation, family inputs, and the like will bias the estimated effects of school characteristics on achievement. In addition, the failure to account for differences in local and state institutional structures for their schools will also introduce bias. The previously presented stylized facts about past research suggest that the absence of controls for differences in state structures leads to overestimates of the effects of school resources on student outcomes, and that the magnitude of such biases tends to increase along with the level of aggregation. Because virtually no attention has been given to how aggregation of data might interact with such specification biases, we develop a conceptual model that permits examination of this issue.

The level of aggregation can influence the estimated relationship between educational outcomes and specific school characteristics in a number of ways.[13] This section examines aggregation related issues in the simplest form using regression equation (2):

$$A_{ijs} = \alpha_{js} + \beta_{ijs}T_{ijs} + \eta_{ijs}F_{ijs} + \theta_{ijs}C_{js} + \psi_{ijs}S_{js} + \epsilon_{ijs} \qquad (2)$$

where the subscript s indexes state of residence and ϵ_{ijs} is an i.i.d. random error. To fix ideas, assume that S_{js} represents a single measure of school quality, say, per-student educational expenditure.[14]

Most studies of educational production functions implicitly assume that the marginal impact of a change in school expenditure is the same regardless of socio-economic background, academic skill, or even the value of school expenditure: $\psi_{ijs} = \psi$ for all i, j, and s. Under such conditions, aggregation to the district or state level will not alter the estimated relationship between attainment and school expenditure (although the form of the data will generally affect the efficiency of the estimates).[15]

Unfortunately, equal marginal effects for all students is a sufficient condition for perfect aggregation only when the empirical model is correctly specified. In practice, information for certain relevant variables might be unavailable. Such data limitations are frequently most severe when aggregate data are employed, as opposed to more detailed survey or individual record information. If the omitted variables are correlated with school expenditure, then the estimated school expenditure coefficient will be biased regardless of the level of aggregation. The issue taken up here is whether the degree of bias is a function of the level of aggregation of school resource data.[16]

One type of specification error that is sensitive to the level of aggregation is the endogeneity of school selection. Because families that place a greater value on education likely choose to live in better school districts, it is imperative to control for differences among family resources and preferences regarding education in order to isolate the impact of school characteristics on student outcomes. Inadequate information on families may generate inflated estimated school resource effects which confound the true impact of school resources with the influence of unmeasured family characteristics. The magnitude of biases is likely to be higher in school or school district level studies, because families are most likely to choose among schools within a limited geographic area. However, the fact that more aggregate estimates tend to be larger and not smaller than school or district level estimates indicates that the endogeneity of school selection is not driving the observed pattern of results.

A. *Aggregation and Omitted Variables in a Simple Two-Variable Case*

We consider omission of an important community factor using equation (3), a simplified version of equation (2) that ignores all variation in pretest scores and family backgrounds and subsumes the influences of all community and state factors into a single measure of community environment. This omitted factor, C_{js}, might be difficult to measure institutional factors such as state regulations on teacher certification or might be a more abstract element such as the community's tastes for education. By assumption, the marginal impacts of both school expenditure and community environment are identical for all students.

$$A_{js} = \theta C_{js} + \psi S_{js} + \epsilon_{js}. \qquad (3)$$

If there is no information on the relevant aspects of community environment, a regression of academic attainment on school expenditure will produce the following biased estimate of the school expenditure coefficient:

$$\hat{\psi} = \psi + \theta\phi, \qquad (4)$$

where ϕ equals the school expenditure coefficient in an auxiliary regression of community environment on school expenditure. Even if equation (3) satisfies the conditions for perfect aggregation, aggregation will alter the bias in the estimate of ψ through its impact on the coefficient ϕ.

The effects of aggregation on the coefficient ϕ are analyzed assuming that both C and S are driven by a common underlying factor, Λ, which indexes, say, tastes for education or the organizational incentives affecting schools. Λ_{js} is the

[13] See Theil (1971) for a comprehensive discussion of aggregation bias.

[14] We implicitly introduce an assumption that community factors and school resources vary only at the school level. While previous research indicates significant variations in school inputs at the classroom level, this added complexity does not provide any additional insights for the analysis here. Similarly, nothing is gained here by adding an individual-specific component of achievement as has been done in past modeling of schools.

[15] Theil (1971) shows that this assumption is a sufficient condition for perfect aggregation for a linear specification. If this assumption is not valid, the estimate of ψ can no longer be interpreted as the marginal impact of a change in school expenditure, because there is no single marginal impact. In specific instances it can be meaningfully interpreted as an average marginal impact as derived from a random coefficients model (see, for example, Swamy (1970)). However, depending on the distribution of the underlying parameters, changes in student composition or the influences of other variables could alter the coefficient estimate and interpretation. Proponents of hierarchical linear modeling (e.g., Bryk and Raudenbush (1992)) and others argue that the impacts of school characteristics vary by socio-economic background, and Summers and Wolfe (1977) and Ferguson (1991) found significant nonlinearities in the effects of class size on student achievement. Exploratory analysis in our empirical modeling indicated that the estimated impacts of school characteristics on attainment did not vary significantly by either student pretest score or race. Therefore, we will maintain the assumption that the marginal effects of school characteristics are equal for all students throughout the analysis.

[16] This conception of the problem was developed early in Grunfeld and Griliches (1960).

local value of Λ for school j, and $\bar{\Lambda}_s$ is the state average. With a linear specification,

$$C_{js} = \mu(\Lambda_{js} - \bar{\Lambda}_s) + \xi\bar{\Lambda}_s + u_{js}, \tag{5}$$

$$S_{js} = \gamma(\Lambda_{js} - \bar{\Lambda}_s) + \delta\bar{\Lambda}_s + v_{js}. \tag{6}$$

We allow both local and state values of Λ to affect school expenditure, reflecting the influence of both the state and local political processes in determining school revenues and expenditure.[17] We define C_{js}, Λ_{js} and $\bar{\Lambda}_s$ so that μ, ξ, γ and δ are all non-negative, and restrict δ to be greater than γ so that school expenditures do not fall as $\bar{\Lambda}_s$ rises. u_{js} and v_{js} are i.i.d. random errors that are independent of each other and the variable Λ.

We first examine the impact of aggregation when a local community characteristic is omitted. Following that, we turn to the case were a state-level variable is excluded.

Omitted Local Factors: The coefficient ϕ equals the covariance of C and S divided by the variance of S. Using equations (5) and (6), and assuming that only local values of Λ affect the community environment ($\mu = \xi$), this equals

$$\phi = \frac{\xi(\gamma\sigma^2_{\Lambda_w} + \delta\sigma^2_{\Lambda_b})}{\gamma^2\sigma^2_{\Lambda_w} + \delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_w} + \sigma^2_{v_b}} \tag{7}$$

where the variances of Λ and v (σ^2) have been partitioned into within state (subscript w) and between state (subscript b) components.

When the data are aggregated to the state level, equations (5) and (6) are rewritten, with bars indicating state average values, as

$$\bar{C}_s = \xi\bar{\Lambda}_s + \bar{u}_s \tag{8}$$

$$\bar{S}_s = \delta\bar{\Lambda}_s + \bar{v}_s \tag{9}$$

and $\bar{\phi}$, the aggregate auxiliary regression coefficient, equals

$$\bar{\phi} = \frac{\xi\delta\sigma^2_{\Lambda_b}}{\delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_b}}. \tag{10}$$

The difference in the auxiliary regression coefficients, $\phi - \bar{\phi}$, equals

$$\phi - \bar{\phi} = \frac{\xi(\gamma\sigma^2_{\Lambda_w} + \delta\sigma^2_{\Lambda_b})}{\gamma^2\sigma^2_{\Lambda_w} + \delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_w} + \sigma^2_{v_b}} - \frac{\xi\delta\sigma^2_{\Lambda_b}}{\delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_b}}. \tag{11}$$

[17] State mean expenditure levels reflect both local and state expenditures. Therefore, within state variation results both from differences in local values of A and the method by which states allocate money to districts.

Aggregation increases the omitted variable bias when $\phi - \bar{\phi} < 0$, and decreases the bias when $\phi - \bar{\phi} > 0$. Let Γ be the positive term

$$\Gamma = \frac{\xi\delta\gamma\sigma^2_{\Lambda_w}\sigma^2_{\Lambda_b}}{(\gamma^2\sigma^2_{\Lambda_w} + \delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_w} + \sigma^2_{v_b})(\delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_b})}. \tag{12}$$

Dividing equation (11) by Γ does not change the sign of $\phi - \bar{\phi}$:

$$\frac{\phi - \bar{\phi}}{\Gamma} = \delta - \gamma - \frac{\sigma^2_{v_w}}{\gamma\sigma^2_{\Lambda_w}} + \frac{\sigma^2_{v_b}}{\delta\sigma^2_{\Lambda_b}}. \tag{13}$$

The following partial derivatives describe the influences of the variance components and regression coefficients on the sign of $(\phi - \bar{\phi})/\Gamma$ (abbreviated as D):

$$\frac{\partial D}{\partial\sigma^2_{\Lambda_b}} = -\frac{\sigma^2_{v_b}}{\delta(\sigma^2_{\Lambda_b})^2} < 0,$$

$$\frac{\partial D}{\partial\sigma^2_{\Lambda_w}} = \frac{\sigma^2_{v_w}}{\gamma(\sigma^2_{\Lambda_w})^2} > 0,$$

$$\frac{\partial D}{\partial\sigma^2_{v_b}} = \frac{1}{\delta\sigma^2_{\Lambda_b}} > 0,$$

$$\frac{\partial D}{\partial\sigma^2_{v_w}} = -\frac{1}{\gamma\sigma^2_{\Lambda_w}} < 0, \tag{14}$$

$$\frac{\partial D}{\partial\delta} = 1 - \frac{\sigma^2_{v_b}}{\delta^2\sigma^2_{\Lambda_b}}\ ?$$

and

$$\frac{\partial D}{\partial\gamma} = -1 + \frac{\sigma^2_{v_w}}{\gamma^2\sigma^2_{\Lambda_w}}\ ?.$$

Increases in the between-state variance of Λ (the underlying factor determining both attainment and school expenditure) and the within-state variance of v (the error in the auxiliary school expenditure equation) raise the probability that aggregation exacerbates omitted variable bias. On the other hand, increases in the within-state variance of Λ and the between-state variance of v lower the probability that aggregation worsens omitted variable bias. Thus, even in the case of a single omitted variable, the effect of aggregation on the magnitude of the bias is ambiguous in general. It depends on the relative sizes of the variance components of Λ and v.

Omitted State Effects: The ambiguity of the aggregation bias disappears, however, if the omitted community effects are determined at the state level with no within-state variation ($\mu = 0$). This might occur, for example, if overall state regulations and policies directly affected the character of student learning. In such a case, the relevant community factor for both the school and the state level relationships is determined by equation (8). In this situation, the value of the relevant school level coefficient, ϕ, becomes:

$$\phi = \frac{\xi\delta\sigma^2_{\Lambda_b}}{\gamma^2\sigma^2_{\Lambda_w} + \delta^2\sigma^2_{\Lambda_b} + \sigma^2_{v_w} + \sigma^2_{v_b}}. \tag{15}$$

Compare this with the aggregate auxiliary regression coefficient, $\bar{\phi}$, from equation (10). Since the numerators are identical and the denominator in equation (15) is unambiguously larger than that in equation (10), the aggregate estimate of ψ is biased upward from the microlevel estimates. With schools, where many key policies are made at the state level, this latter model structure appears very relevant.

B. *Measurement Error*

To this point, we have assumed that the school characteristics are measured without error. An additional concern about the estimation of school performance models is the possibility of measurement error in the school variable. If data are collected by surveys, their quality may be low when the local respondent is uncertain about the precise values, say, of expenditure or even number of students. On a related issue, if the educational models are aggregated over time instead of employing the basic value-added formulation sketched in equations (1) and (2), year-to-year fluctuations in the data may provide a misleading picture of the relevant historical data.[18] In other words, measurement error is likely to be particularly important as empirical specifications diverge from the ideal described in equation (1).

The effect of measurement error in simple linear models is well-known (see, e.g., Maddala (1977)). If the observed school input $\tilde{S}_{js}$ is

$$\tilde{S}_{js} = S_{js} + \nu_{js} \tag{16}$$

then the estimate of ψ will be inconsistent and biased toward zero, even when the measurement error, ν, is i.i.d. with mean zero. This attenuation can be eliminated through the use of instrumental variables techniques, if it is possible to identify an appropriate instrument that is correlated with the true value of the regressor but uncorrelated with the random measurement error.

The strategy of aggregating within defined groups (e.g., states) and replacing the microdata with the group average is a variant of instrumental variables, in which the group average is used as the instrument.[19] Group averaging reduces bias if it raises the signal to noise ratio. Consequently, random assignment into groups will not solve the measurement error problem even if aggregating eliminates all noise, because the group averages will be uncorrelated with the true values of the individual observations and thus inappropriate instruments. Observations must be grouped on the basis of a characteristic that is correlated with the true underlying values of the regressor.

In the case of school resource estimates, the pattern of past empirical results is broadly consistent with what would be expected if state average resource values provide valid instruments for the mismeasured school level data; that is, aggregation suggests stronger resource effects than found in the disaggregated samples. However, as demonstrated, the presence of omitted variables could also produce the identical pattern of larger state level estimates, if omitted factors are correlated with state average resource levels. When aggregation preserves between group differences in an omitted factor which is itself related to both achievement and school resource levels, the state averages are no longer valid instruments, and the sources of the effects of aggregation are not identified.

C. *Complications*

The previous analyses produce straightforward results. Without measurement error, aggregation has an ambiguous effect on the estimated schooling parameter when local community factors are omitted. If, however, the omitted factor applies at the state level, aggregation of all variables to the state level will unambiguously bias the schooling parameter upward. Finally, with classical errors-in-variables, aggregation will tend to increase the magnitude of the schooling parameter by virtue of reducing the downward bias imparted by measurement error.

All of these derivations were, nonetheless, conducted within a very simplified model. First, none of this analysis generalizes in a clear way to nonlinear models. Second, just a single variable was omitted and only a single behavioral parameter was being estimated. The omitted variable bias becomes much more complicated when multiple included and excluded variables are considered, and the effects of aggregation can no longer in general be ascertained. Third, the analysis of measurement error assumed errors independent of the included variables, but systematic measurement error will not have the clear effects previously derived. Fourth, the measurement error model no longer yields any simple predictions when more than one variable is measured with error (see, e.g., Maddala (1977)). With multiple measurement errors, the coefficients are not necessarily biased

[18] Card and Krueger (1994) highlight fluctuations in capital expenditure or inexplicable annual movements in pupil–teacher ratios as evidence of this sort of error. While annual movements in pupil–teacher ratios or other inputs have little impact on value-added models, they will obviously introduce serious measurement errors in the cumulative inputs.

[19] In a linear model the replacement of the microdata with group averages would produce coefficients with the same expected values but larger standard errors than would the use of the group averages as instruments for the microdata (see Rivkin (1995)).

toward zero, but instead depend on the entire pattern of covariances among the exogenous variables.

The dependence of the theoretical conclusions about aggregation on the specifics of the model and data lead us to examine aggregation empirically within the context of school and state variations in educational performance. The effects of school resources on both cognitive achievement and schooling attainment are considered at both the school level and the state level to investigate aggregation within a consistent model specification and data source.

IV. Data

Data for the empirical analysis come from the High School and Beyond (HSB) longitudinal survey. The survey is administrated by the National Opinion Research Center under contract with the Department of Education. The base year of the survey is 1980, at which time approximately 36 high school sophomores and 36 high school seniors from almost 1,000 high schools were interviewed. Follow-up surveys were completed in 1982, 1984, and 1986, yielding six years of information for individuals retained in all follow-up surveys. The base year survey reports parental schooling and family income levels and contains information on the high schools. Six years of schooling histories are contained in the follow-up surveys. In addition, a battery of standardized mathematics, verbal and science tests are administered along with the base year and first follow-up surveys.

Two separate measures of student achievement (A_{ij}) are employed: composite 12th grade test scores and years of schooling attained. The 12th grade composite test score is a linear combination of mathematics and reading scores, where the weights equal the estimated parameters for the two test scores in a regression of the probability of continuing high school on the test scores and other explanatory variables.[20] The rankings of students by this achievement measure is thus not arbitrary, but related to their academic preparation for continuing school. Similar weights are also produced when samples are divided by race and gender.

The second attainment measure is years of post-secondary schooling for high school graduates. It varies from zero (no post-secondary schooling) to eight (a Ph.D., M.D., etc.).[21] We measure school attainment at roughly age 24 (6 years following high school graduation). Approximately 20% of individuals are still in school, and recent trends suggest many nonstudents will return to school at some point in the future. Thus, we are measuring educational attainment 6 years after high school, not total years of schooling.

The value added framework uses a pretest score as a regressor in order to isolate the contribution of high schools. We condition on the 10th grade composite test score in the analysis of 12th grade test scores and on the 12th grade test score in the analysis of post-secondary schooling. Therefore any impact of high schools on post-secondary schooling is in addition to their effects on cognitive achievement. The vector F_{ij} includes race, gender, parental schooling and family income.[22]

Because they are the primary determinants of educational expenditures, we use teacher–pupil ratios and relative teacher salaries to measure school characteristics.[23] We divide starting teacher salaries by the average earnings of college-educated residents in the community (normalized to 40-week salaries) to derive relative salaries. Scaling salaries by the local wage level controls for any differences in the alternative wage opportunities available to teachers. (The results are not sensitive to the way in which salaries are deflated.)

We experiment with the inclusion and exclusion of a set of community factors that are both likely to be important and frequently excluded from analysis. In the school attainment regressions, C_{ij} includes the percentage of adult community residents who have a college degree, the local unemployment and wage rates for high school graduates, and the resident tuition at state universities. Because the link between high school test scores and the opportunity cost of post-secondary schooling is likely to be quite weak, the labor market and tuition variables are excluded from the test score regressions.

Both local and state aggregates of community factors are analyzed. The percentage of college educated residents captures environmental effects on educational expectations and achievement. Local wages and unemployment rates reflect variations in the opportunity cost of attending college, while the state resident tuition at public universities indexes the monetary cost of college attendance.[24] Each of the community variables was computed over individuals 20 to 49 years of age. These older residents attended high school during an earlier period, and many of them attended schools outside of the community in which they live. This makes it extremely unlikely that the community variables are themselves affected by the quality of schools in the sample.

We do not have measures of relevant institutional influ-

[20] These results come from preliminary analyses of continuation in high school using the sophomore cohort. The weight on the mathematics test score is three times greater than the weight on the reading test score, a result that is consistent with other work (e.g., Bishop (1992)) suggesting that mathematical skill is disproportionately valued.

[21] Students might complete a degree associated with eight years of post-secondary schooling in fewer than eight calendar years.

[22] Because of missing data problems, the family income variable is not directly included in the analyses of 12th grade test scores. Instead, the high school family income distribution for same race students who report family income is added as a school characteristic. Individual parental education is included in the first stage equation.

[23] Data on expenditure per pupil are seldom available at the individual school level, so it is not possible to look at total expenditure here. The empirical analysis, using normalized starting salaries, will not yield precisely the level of classroom expenditure per student but will be highly correlated with real expenditure differences across schools and districts.

[24] The High School and Beyond survey contains very little information on community environment. Therefore we use U.S. Census data to construct measures of local unemployment, wage and college completion rates. The information on university tuition is taken from Peterson's Guide to Four Year Colleges (1982). Rivkin (1991) describes the construction of the community characteristics.

ences such as financing or regulatory conditions for each state. Thus, the community factors analyzed are designed to indicate how explicit misspecification can influence the estimated resource parameters, but the more complete models should not be interpreted as being fully specified. They mirror past analyses of multistate data, and are subject to the same upward biases on resource effects.

We restrict our attention to non-Hispanic Blacks and Whites attending public high schools, and omit all observations in schools that have fewer than five observations with nonmissing data and schools with a pupil–teacher ratio above 40. We also exclude states with only a single high school in the final sample. We use the sophomore cohort to analyze test scores, and the senior cohort to analyze post-secondary schooling. The sophomore cohort sample includes 11,386 observations for students who attended 627 schools in 46 states, and the senior cohort sample has 2,309 observations for students who attended 307 schools in 38 states. Variable means and standard deviations are presented in table A1.

V. Empirical Results

We use a two stage estimation framework to analyze the impact of aggregation on schooling coefficients in order to account for the grouping of students into schools.[25] This approach mirrors the estimation strategy of Card and Krueger (1992a) and allows us to focus on the key aspects of aggregation and school resource effects. Because we used the same procedures in both the school- and state-level analyses, we will describe only the school-level specifications.

In the first stage, we regressed academic attainment on the pretest score, student and family characteristics, and a separate indicator variable for each high school, E_{ij}:

$$A_{ij} = \beta T_{ij} + \eta F_{ij} + \sum_j \alpha_j E_{ij} + \mu_{ij}, \tag{17}$$

where E_{ij} equals 1 if student i attends school j and 0 otherwise and where μ_{ij} is an i.i.d. random error. The stage one regression produces estimates of the average school effects, α_j, controlling for differences in student characteristics.

The second stage regression estimates the following relationship:

$$\alpha_j = \lambda + \theta C_j + \psi S_j + \epsilon_j. \tag{18}$$

However, the actual school intercepts are not observed. Instead, the first stage regressions generate fitted values for the school intercepts. Using the first stage predicted values implies an additional error to the second stage regression because sampling error yields:

$$\hat{\alpha}_j = \alpha_j + \eta_j. \tag{19}$$

[25] See Moulton (1986) for a discussion of the impact of grouping on the precision of estimated effects.

Thus, the second stage becomes a random components model.

$$\hat{\alpha}_j = \lambda + \theta C_j + \psi S_j + \epsilon_j + \eta_j. \tag{20}$$

Because the sampling variances of the predicted values differ across schools, η_j is heteroskedastic. We assume that the variance of η_j is proportional to the stage 1 sampling variance of α_j, which suggests the use of weighted least squares in the second stage regression. But as Hanushek (1974) points out, this implicitly assumes that the other component of the error term, ϵ_j, has a variance that is either proportional to η_j or zero. In order to allow for the more general random effects specification, we use a specialized form of generalized least squares in the second stage. We first estimate equation (18) using ordinary least squares.[26] Next, we regress the square of the residuals on the sampling variance of the school intercepts. Finally, we use the inverse of the predicted square of the residuals from this auxiliary regression as the weight in the GLS estimation of equation (20).

Appendix table A2 presents the coefficient estimates from the first-stage regressions (equation (17)).[27] As expected, both measures of academic attainment are positively related to the early achievement score, parental education and family income. Whether school or state dummy variables are included appears to have little impact on the coefficient estimates. There are also significant between-school attainment differences. The hypotheses that the high school dummy variables do not add to the explanatory power of the regressions can be rejected at any conventional level of significance.[28]

A. *Basic Aggregation Effects*

We turn now to the second stage regressions which contain information about the effects of school resources. Two sets of regressions are computed for each attainment measure. The first set uses estimates of school value added as the dependent variable, while the second set uses estimates of the state value added. At each level, estimates are computed over a series of regression specifications. Some specifications include the previously discussed community variables, while others do not. These particular variables are obvious measures of community environment, particularly in the school attainment regressions, but they are by no means the only or even the most important community or state determinants of academic outcomes. The sensitivity of the school resource estimates to their exclusion does not provide an absolute measure of the impact of specification

[26] Borjas (1987) describes this procedure.

[27] Because the High School and Beyond Survey was not designed to produce representative samples of each states' schools, the estimated state fixed effects do not provide a rank ordering of states according to school quality.

[28] The F-test statistics equal 12.52 (625, 10752 degrees of freedom) for the test score regression and 2.32 (306, 1993 degrees of freedom) for the post-secondary schooling regression.

TABLE 5.—GENERALIZED LEAST SQUARES ESTIMATES OF SCHOOL-SPECIFIC TEST PERFORMANCE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Local Community Factor									
Local college completion rate		1.05	0.60		1.28	0.84		1.05	0.59
		(1.62)	(0.87)		(2.04)	(1.25)		(1.65)	(0.87)
State Community Factor									
State college completion rate			1.94			1.92			1.93
			(1.83)			(1.81)			(1.84)
School Factors									
Teacher–pupil ratio	0.66	0.17	−0.08	0.56	−0.01	−0.26			
	(0.31)	(0.08)	(−0.04)	(0.27)	(−0.01)	(−0.12)			
Teacher salary	−0.82	−0.67	−0.68				−0.82	−0.66	−0.68
	(−1.97)	(−1.55)	(−1.57)				(−1.97)	(−1.55)	(−1.58)

Note: $n = 627$. t statistics are in parentheses. All models include a dummy variable for being in a rural area, a dummy variable for being in the South, and the average income of families in the school.

error, but rather some information on how the error likely varies with changes in the level of aggregation.

Tables 5 and 6 contain the generalized least square regression results for the test score specifications, and tables 7 and 8 contain the results for the educational attainment specifications. Each table reports the results of regressions which contain both the teacher–pupil ratio and teacher salary variables as well as regressions in which the two school characteristics are entered separately.[29]

Consistent with the prior research summarized above, there is no evidence that either increasing the teacher–pupil ratio or raising teacher salaries increases the composite test score in any of the school-level regressions that are reported in table 5. The coefficients for the teacher–pupil ratio are all very small and insignificantly different from zero, and the teacher salary coefficients consistently have negative signs. The level of college completion at the state level is strongly related to gains in student achievement—substantively suggesting that state differences in educational policies are important and analytically suggesting the strong possibility of upward biases through aggregation that ignores such state differences. Local community variations appear less important for student achievement once state community variation is considered. Though the community education level is positively related to 12th grade test scores, there is little evidence that omitting this variable substantially alters the estimated school characteristic effects in these school-level regressions. In other words, bias introduced by the omission of these community characteristics does not appear to be a major problem at the school level of aggregation (although these may not accurately reflect relevant omitted factors).

The state-level regressions reported in table 6 produce very different results. Columns 2 and 4 in table 6 show that the teacher–pupil coefficient estimate is large and closer to conventional statistical significance in specifications that exclude the average community education level. The inclusion of community differences, however, substantially reduces the magnitude of the teacher–pupil coefficient and leaves it far below its standard error. The sensitivity of the teacher–pupil ratio coefficient estimate to the inclusion of the community education level is evidence that aggregation worsens problems caused by the exclusion of relevant variables. As in the school-level regressions, the teacher salary coefficient remains negative regardless of whether the community education level is included.

Results for the educational attainment specifications in tables 7 and 8 provide additional evidence that aggregation exacerbates the problem of omitted variable bias. Table 7 demonstrates that local variations in schooling costs (including opportunity costs) are very important, although the addition of state community factors adds little. More importantly, the table shows that school-level coefficients for the teacher–pupil ratio and teacher salary are statistically insignificant in all specifications. While some people have suggested that school resources have important indirect effects on students through influencing school completion, these

[29] The state average values of the community variables are not the actual state averages but rather the averages computed over the communities in which students attended high school. Thus, we do not artificially introduce measurement error because of the sampling of schools in each state.

TABLE 6.—GENERALIZED LEAST SQUARES ESTIMATES OF STATE-SPECIFIC TEST PERFORMANCE

	(1)	(2)	(3)	(4)	(5)	(6)
Community Factor						
State college completion rate		4.73		4.60		5.07
		(2.18)		(2.01)		(2.57)
School Factors						
Teacher–pupil ratio	9.93	2.95	12.30	5.29		
	(1.34)	(0.39)	(1.65)	(0.66)		
Teacher salary	−2.83	−3.04			−3.28	−3.16
	(−1.82)	(−2.02)			(−2.12)	(−2.15)

Note: $n = 46$. t-statistics are in parentheses.

TABLE 7.—GENERALIZED LEAST SQUARES ESTIMATES OF SCHOOL-SPECIFIC EDUCATIONAL ATTAINMENT MODELS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Local Community Factors									
Local college completion rate		2.11	2.31		2.00	2.29		2.20	2.46
		(3.19)	(3.34)		(3.10)	(3.24)		(3.34)	(3.45)
Local unemployment rate		3.62	4.22		3.79	4.25		3.48	4.18
		(2.50)	(2.54)		(2.70)	(2.59)		(2.40)	(2.51)
Local wage		−0.0030	−0.0008		−0.0032	−0.0016		−0.0031	−0.0009
		(−2.68)	(−0.38)		(−2.92)	(−0.84)		(−2.80)	(−0.42)
State Community Factors									
State College completion rate			−1.50			−1.68			−1.50
			(−1.17)			(−1.33)			(−1.17)
State unemployment rate			−1.83			−1.75			−2.17
			(−0.71)			(−0.68)			(−0.84)
State wage			−0.0019			−0.0012			−0.0019
			(−0.82)			(−0.52)			(−0.80)
In-State public tuition		−0.0001	−0.0001		−0.0001	−0.0001		−0.0001	−0.0001
		(−0.84)	(−0.88)		(−0.97)	(−1.06)		(−0.88)	(−0.91)
School Factors									
Teacher–pupil ratio	2.86	2.31	2.39	3.01	2.44	2.55			
	(1.50)	(1.22)	(1.26)	(1.59)	(1.30)	(1.35)			
Teacher salary	0.47	0.33	0.45				0.51	0.38	0.51
	(1.24)	(0.80)	(1.03)				(1.34)	(0.92)	(1.16)

Note: n = 307. t-statistics are in parentheses

TABLE 8.—GENERALIZED LEAST SQUARES ESTIMATES OF STATE-SPECIFIC EDUCATIONAL ATTAINMENT

	(1)	(2)	(3)	(4)	(5)	(6)
Community Factors						
State college completion rate		4.89		4.89		5.54
		(2.64)		(2.65)		(2.91)
State unemployment rate		8.51		9.40		7.72
		(2.95)		(3.50)		(2.51)
State wage		−0.0041		−0.0049		−0.0063
		(−1.71)		(−2.29)		(−2.36)
In-State public tuition		0.0001		0.0000		0.0000
		(0.47)		(0.25)		(0.24)
School Factors						
Teacher–pupil ratio	11.63	11.66	12.33	11.94		
	(2.50)	(2.53)	(2.59)	(2.61)		
Teacher salary	1.91	0.94			2.21	1.21
	(1.91)	(0.88)			(2.03)	(1.04)

Note: n = 38. t-statistics are in parentheses.

results do not support such a notion. Though the exclusion of the community variables does increase the magnitude and significance of the estimated effects of changes in the teacher–pupil ratio and teacher salaries, the estimates remain statistically insignificant.

Table 8 vividly demonstrates, however, how the apparent picture changes when aggregated to the state level. Aggregation to the state level substantially increases the magnitudes and statistical significance of the estimated relationship between post-secondary schooling and both the teacher–pupil ratio and teacher salary. The teacher–pupil ratio has a statistically significant impact in the expected direction regardless of whether the community variables are included in the regression, and its magnitude is approximately five times that found in the school-specific estimates of table 7. Similarly, the estimated effect of teacher salary is 3–4 times as large in the aggregate as in the school-specific regressions. Moreover, excluding the community variables has a dramatic impact on the teacher salary coefficient: The magnitudes of the coefficient and the t-statistics are roughly twice as high in specifications where the community variables are excluded.

These results are broadly consistent with the notion that aggregation heightens problems of omitted variable bias. The exclusion of the community variables usually had a larger impact on the magnitude of the school characteristic coefficients at the state level as opposed to the school level of aggregation. Because aggregation tended to exacerbate the bias from excluding the observed community factors, it is also likely that aggregation increased the bias due to the omission of other relevant factors. Again, no measures of state education policy except the postsecondary tuition are included even in the expanded models.

There were substantial differences in the sensitivity of the school characteristic estimates to both the omission of the community variables and to aggregation depending upon which measure of academic performance was used. These differences across the two attainment measures indicate that the excluded determinants of test score growth differ from the factors that impact post-secondary schooling. This might be one reason why there is little correlation between the test score growth and post-secondary schooling estimates of school value added. The correlation coefficient between these two measures of school value added equals -0.07. At first glance the lack of correlation suggests that schools choose between a variety of objectives and that schools which are good at raising test scores might not have characteristics which are conducive to post-secondary school attainment. Yet because the estimates of school value added capture both school and nonschool influences, differential impacts of nonschool variables on the different attainment measures might potentially conceal the fact that higher quality schools increase most types of attainment.

B. *Measurement Error*

While the prior results are consistent with an omitted variables bias explanation, they do not conclusively demonstrate that aggregation-aggravated omitted variable bias is the sole problem. Aggregation moved the school resource and community environment coefficients away from zero, as would also be predicted with classical measurement error. Though the five-fold increase in the estimated effect of the pupil–teacher ratio on years of schooling is consistent with a prominent role for measurement error only if the signal to noise ratio in school level data is extremely low, the results do not rule out the possibility that aggregation reduces bias caused by the mismeasurement of school resources. The difficulty of separately identifying the contributions of specification bias and measurement error reduction is that grouping schools by states likely violates the criterion that the grouping variable must be uncorrelated with any relevant omitted factors.

In order to identify better the contributions of specification and measurement error to the observed pattern of estimates, an appropriate instrument must be found that divides schools among groups in a way that preserves differences in the true values of the regressor but does not preserve differences in omitted factors related to both outcomes and school resources.[30] Our approach is to regroup the schools into "pseudo-states," where schools are assigned to groups based on values of a variable that is assumed to be correlated with the true values of the school characteristics but uncorrelated with the measurement error and with the student outcomes except through the regressor. The use of an instrumental variable other than the state breaks the link between school characteristics and unobserved state determinants of student outcomes. We somewhat arbitrarily construct groups equal in number to the actual states in the test score and post-secondary schooling analyses, so as not to confuse the empirical results with possible differences related to sample size.

One plausible set of grouping variables simply uses the ordering of the school resources themselves—i.e., the ordering by the magnitude of the teacher–pupil ratio or the teacher salary. Grouping across teacher–pupil ratios will lead to consistent estimates if schools are correctly categorized on the basis of true (error-free) teacher–pupil ratios or true salaries, and there are no relevant omitted factors correlated with the true values of school resources. Because errors in measurement can lead to misclassification, we divide all schools into the "state" groups and omit two schools at each group boundary, since these schools are the most likely to be wrongly classified.[31] However, Pakes (1982) shows that even the use of trimmed group means will not necessarily eliminate or even reduce the attenuation of measurement error, depending on the distribution of the errors and true variables.

An alternative approach begins by assigning actual states to four artificially created divisions based on a state characteristic which is presumed to be correlated with school resource decisions but not directly correlated either with measurement errors or with student outcomes.[32] Of course, the last condition is difficult to verify and, indeed, is suspect when grouping is based on income or wealth—which might well be correlated with measured factors influencing school performance. This uncertainty over the validity of the instruments leads to a strategy of comparing results across a number of different grouping variables and looking for consistency in the findings. We use three state characteristics as instruments: (1) the state per capita assessed property value; (2) the state poverty rate; and (3) the state percentage of workers unionized.

After states are assigned to divisions on the basis of the selected state characteristics, schools are randomly assigned to "pseudo-state" within each division. This means that schools are no longer divided into groups on the basis of their actual state locations, rather they are now randomly divided into groups (equal in total number to the number of states in the sample) within the constructed divisions. We then use the same two-stage estimation procedure that was used with the actual state groupings to produce point estimates of school resource effects when the school characteristics are aggregated by pseudo state. Because any single random allocation within a "division" may give misleading point estimates, this random allocation of schools to pseudo-states is repeated thirty times for each state characteristic

[30] A similar approach of looking for between-jurisdiction differences has been independently proposed in Heckman, Layne-Farrar, and Todd (this *Review*, this issue). Their discussion provides a public choice rationale for the existence of such differences.

[31] Early proposals for grouping and instrumental variables concentrated on bivariate models and analyzed the trade-off between bias and efficiency from aggregating to two groups (Wald), three with an omitted center category (Bartlett), and multiple groups (Durbin); see Maddala (1977). The approach here is an extension of these. For a general consideration of instrumental variables in cross-sections, see White (1982).

[32] Note that divisions so constructed are not the geographical divisions commonly defined in the census but instead contain states from different geographical areas.

TABLE 9.—ALTERNATIVE GROUPED ESTIMATES (IV) OF TEST PERFORMANCE

	Pseudo-states Grouped by Teacher–Pupil Ratio		Pseudo-states Grouped by Teacher Salary		Random Pseudo-States within Divisions Grouped by: Per Capita Assessed Value		% Unionized Manufacturing Workers		% Poverty	
Community Factor										
College completion rate		−6.26		−2.96		0.07		0.68		0.82
		(−2.7)		(1.4)		(0.1)		(0.3)		(0.3)
School Factors										
Teacher–pupil ratio	−0.98	2.32	−13.8	−9.64	6.62	6.69	−1.41	−2.00	15.6	14.7
	(−0.3)	(0.7)	(−1.1)	(−0.8)	(0.9)	(0.8)	(−0.1)	(−0.1)	(1.8)	(1.7)
Teacher salary	0.30	−0.73	−1.65	−2.13	−1.95	−2.00	−1.41	−1.24	−0.93	−0.81
	(0.2)	(−0.5)	(−2.1)	(−2.5)	(−1.3)	(−1.2)	(−1.0)	(−0.8)	(−0.6)	(−0.5)

Note: n = 46 pseudo-states. t-statistics are in parentheses.

grouping variable, and the reported parameter estimates and standard errors are averages over the thirty replications.[33]

The creation of divisions provides a method of preserving variance in the underlying state resource characteristics—something that would not be done with purely random assignment of schools to pseudo-states. Consequently, the aggregation of schools by pseudo-state should reduce any attenuation due to measurement error in exactly the same way that aggregation within actual states reduces such attenuation. The key difference is, however, that random assignment within artificial division substantially weakens the link between schools and unmeasured state-specific factors.

Each of these five grouping variables substantially weakens the link between schools and their specific state. Therefore, if aggregation by state inflated the school resource coefficients by reducing measurement error, we would expect that aggregation by pseudo-state should produce estimates quite similar to the aggregate coefficients reported in tables 6 and 8. On the other hand, if aggregation by state increased the school resource coefficients by exacerbating omitted variables bias, we would expect that the new pseudo-state aggregate estimates would more closely resemble the school-level coefficients in tables 5 and 7.

Table 9 presents the aggregate test score results; table 10 presents the aggregate school attainment results. The results are clearest in the case of school attainment, the outcome measure most strongly related to the school resources in the aggregate state-level specifications. The estimated effects of a higher teacher salary and a higher teacher–pupil ratio in table 10 are quantitatively very close to those in the school-level estimates (table 7) but very much smaller than those in the state aggregates (table 8). The only teacher salary coefficients that exceed the school-level estimates are produced by the teacher–pupil ratio and the state unionization rate groupings, while the only teacher–pupil ratio coefficients that exceed the school-level estimates are produced by the state poverty rate and the state unionization rate groupings. Even in these cases, the estimated effects lie much closer to the school-level estimates in table 7 than to the aggregate state-level estimates in table 8. State groupings by the percentage unionized and the poverty rate produced teacher–pupil coefficients slightly above those produced by the school-level specifications, but again these estimates lie

[33] The t-statistics reported in the tables are the ratio of average coefficients to average standard errors from the thirty replications.

TABLE 10.—ALTERNATIVE GROUPED ESTIMATES (IV) OF SCHOOL ATTAINMENT

	Pseudo-states Grouped by Teacher–Pupil Ratio		Pseudo-states Grouped by Teacher Salary		Random Pseudo-States within "Divisions" Grouped by: Per Capita Assessed Value		% Unionized Manufacturing Workers		% Poverty	
Community Factors										
In-state public tuition		−0.0003		0.0006		0.0003		0.00008		−0.00001
		(−0.9)		(2.0)		(1.1)		(0.2)		(−0.1)
Wage rate		−0.0064		−0.0096		−0.0043		−0.0029		−0.0041
		(−1.7)		(3.7)		(−1.3)		(−1.0)		(−1.2)
Unemployment rate		8.51		10.16		4.33		6.12		4.6
		(1.9)		(2.4)		(1.0)		(1.7)		(1.2)
College completion		5.48		3.25		2.35		2.02		1.64
		(2.9)		(1.9)		(1.2)		(1.1)		(0.8)
School Factors										
Teacher–pupil ratio	4.04	1.67	−2.12	−2.72	1.62	1.28	5.45	5.04	2.19	2.64
	(1.5)	(0.6)	(−0.3)	(−0.4)	(0.2)	(0.1)	(0.9)	(0.8)	(0.4)	(0.3)
Teacher salary	0.48	0.54	0.35	−0.20	0.34	0.02	0.70	0.50	0.44	0.03
	(0.4)	(0.5)	(0.7)	(0.3)	(0.3)	(0.04)	(0.7)	(0.4)	(0.4)	(0.01)

Note: n = 38 pseudo-states. t-statistics are in parentheses.

much closer to the school estimates in table 7 than to the aggregate state estimates in table 8.

The test score results in table 9 also offer very little evidence in favor of the measurement error explanation. Similar to both the school and state-level results, all teacher salary coefficients are negative. The teacher–pupil ratio estimates fluctuate noisily depending upon the grouping variable: three grouping variables produce positive coefficients and two produce negative coefficients. There is no clear pattern in support of the measurement error explanation, particularly in light of the fact that the actual state aggregate coefficients in table 6 are small and statistically insignificant at any conventional level once the observed community factors are included.

There is no confirmation that the inflated coefficients generated by the actual state-level aggregate specifications were simply the result of correcting measurement errors at the individual school level. Specific measurement error corrections that break the correlation with omitted state-level variables yield resource results that are consistent with the school-level estimates. Thus the overall evidence strongly supports the position that the difference between school-level and aggregate-level results in the estimated effects of school resources arises from omitted variables bias that is aggravated by aggregation as opposed to a simple measurement error problem.

VI. Conclusions

The results in this paper provide evidence that problems of omitted variables bias tend to increase along with the level of aggregation, causing analyses that use more aggregated data to overestimate systematically the influence of school expenditure related characteristics on student attainment. Aggregate analyses of student performance, particularly at the state level, typically have very crude measures of school and family factors. They never employ value-added models. Moreover, aggregate analyses drawing data from different states generally neglect potentially important financing, organizational, and regulatory features of states. In short, they are subject to extensive specification problems. This situation is exactly where aggregation bias is most important, and the review of past analyses shows a pattern of results that is entirely consistent with the presence of substantial specification error in aggregate studies of school resource effects. Investigation of the competing hypothesis that aggregation is beneficial because it reduces biases from measurement error provides no support for the alternative. In contrast, studies which contain more information about community characteristics and which use less aggregated data are likely to produce more reliable estimates of the true impact of school expenditure on attainment.

Our findings are consistent with the view that increases in school expenditure used to reduce the teacher–pupil ratio and raise teacher salaries have had little systematic impact on student attainment. Therefore, further reductions in the teacher–pupil ratio or further increases in teacher salary by themselves are unlikely to generate improvements in the performance of students who attend United States public elementary and secondary schools. Additionally, nothing suggests that resources are more important in assuring the students attain more schooling than they are in determining the pattern of cognitive achievement.

Significant differences in school quality exist. These differences, however, are not systematically related to school resources—a finding that introduces added complexity into the development of educational policies. Policies aimed at altering the incentive structure and, thus, the ways in which resources are used appear much more likely to succeed than policies aimed simply at adding more resources to schools (Hanushek with others (1994)).

REFERENCES

Aitkin, M., and N. Longford, "Statistical Modelling Issues in School Effectiveness Studies," *Journal of the Royal Statistical Society* A 149, pt. 1 (1986), 1–26.

Betts, Julian R., "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth," this REVIEW 77 (May 1995), 231–247.

———, "Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature," in Gary Burtless (ed.), *Does Money Matter: The Effect of School Resources on Student Achievement and Adult Success* (Washington, D.C.: Brookings, 1996), 141–191.

Bishop, John, "The Impact of Academic Competencies of Wages, Unemployment, and Job Performance," *Carnegie-Rochester Conference Series on Public Policy* 37 (Dec. 1992), 127–194.

Borjas, George J., "Self-selection and the Earnings of Immigrants," *American Economic Review* 77 (Sept. 1987), 531–553.

Bryk, Anthony S., and Stephen W. Raudenbush, *Hierarchical Linear Models: Applications and Data Analysis Methods* (Newbury Park, CA: Sage Publications, 1992).

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

———, "School Quality and Black–White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107 (Feb. 1992b), 151–200.

———, "The Economic Return to School Quality: A Partial Survey," Working Paper No. 334, Industrial Relations Section, Princeton University (Oct. 1994).

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Equality of Educational Opportunity* (Washington, D.C.: U.S. Government Printing Office, 1966).

Ferguson, Ronald, "Paying for Public Education: New Evidence on How and Why Money Matters," *Harvard Journal on Legislation* 28 (Summer 1991), 465–498.

Gold, Steven D., David M. Smith, Stephen B. Lawton, and Andrea C. Hyary, *Public School Finance Programs of the United States and Canada: 1900–91* (New York: The Nelson A. Rockefeller Institute of Government, State University of New York, 1991).

Grunfeld, Yehuda, and Zvi Griliches, "Is Aggregation Necessarily Bad?" this REVIEW 42 (Feb. 1960), 1–13.

Hanushek, Eric A., "Efficient Estimators for Regressing Regression Coefficients," *The American Statistician* 28 (May 1974), 66–67.

———, "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *Journal of Human Resources* 14 (Summer 1979), 351–388.

———, "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24 (Sept. 1986), 1141–1177.

———, "The Impact of Differential Expenditures on School Performance," *Educational Researcher* 18 (May 1989), 45–51.

———, "Money Might Matter Somewhere: A Response to Hedges, Laine, and Greenwald," *Educational Researcher* 23 (May 1994), 5–8.

——, "Assessing the Effects of School Resources on Student Performance: An Update," Working Paper no. 424, Rochester Center for Economic Research, University of Rochester (1996a).

——, "A More Complete Picture of School Resource Policies," *Review of Educational Research* (Fall 1996b).

Hanushek, Eric A., and Steven G. Rivkin, "Understanding the 20th Century Growth in U.S. School Spending," *Journal of Human Resources* 32 (Winter 1997).

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor, "The Identification of School Resource Effects," *Education Economics* 4 (Aug. 1996).

Hanushek, Eric A., and Lori L. Taylor, "Alternative Assessments of the Performance of Schools: Measurement of State Variations in Achievement," *Journal of Human Resources* 25 (Spring 1990), 179–201.

Hanushek, Eric A., with others, *Making Schools Work: Improving Performance and Controlling Costs* (Washington, D.C.: Brookings Institution, 1994).

Heckman, James J., Anne S. Layne-Farrar, and Petra E. Todd, "Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings," this REVIEW 78 (Nov. 1996).

Hedges, Larry V., Richard D. Laine, and Rob Greenwald, "Does Money Matter? A Meta-analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher* 23 (Apr. 1994), 5–14.

Hoxby, Carolyn, "Teachers Unions and the Effectiveness of Policies Designed to Improve School Quality," mimeo, Department of Economics, Harvard University (1995).

Johnson, George E., and Frank P. Stafford, "Social Returns to Quantity and Quality of Schooling," *Journal of Human Resources* 8 (Spring 1973), 139–155.

Maddala, G. S. *Econometrics* (New York: McGraw-Hill, 1977).

Moulton, Brent R., "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics* 32 (1986), 385–397.

Pakes, Ariel, "On the Asymptotic Bias of Wald-type Estimators of a Straight Line When Both Variables Are Subject to Error," *International Economic Review* 23 (June 1982), 491–497.

Peterson, *Peterson's Guide to Four Year Colleges* (Princeton, NJ: Peterson's Guide, 1982).

Rivkin, Steven G., "Schooling and Employment in the 1980s: Who Succeeds?" Ph.D. Dissertation, University of California, Los Angeles (1991).

——, "School Desegregation, Academic Attainment and Employment," mimeo, Department of Economics, Amherst College (1995).

Speakman, Robert, and Finis Welch, "Does School Quality Matter?—A Reassessment," mimeo, Texas A&M University (Jan. 1995).

Summers, Anita, and Barbara Wolfe, "Do Schools Make a Difference?" *American Economic Review* 67 (Sept. 1977), 639–652.

Swamy, P. A. V. B., "Efficient Inference in a Random Coefficient Regression Model," *Econometrica* 38 (Mar. 1970), 311–323.

Theil, Henri, *Principles of Econometrics* (New York: John Wiley & Sons, 1971).

Tiebout, Charles M., "A Pure Theory of Local Expenditures," *Journal of Political Economy* 64 (Oct. 1956), 416–424.

Welch Finis, "Measurement of the Quality of Schooling," *American Economic Review* 56 (May 1966), 379–392.

White, Halbert, "Instrumental Variables Regression with Independent Observations," *Econometrica* 50 (Mar. 1982), 483–499.

APPENDIX

TABLE A1.—DESCRIPTIVE STATISTICS

Variable	Mean	Standard Deviation
Outcome Measures		
Years of schooling completed	13.7	1.7
12th grade test score	15.1	5.4
Explanatory Factors		
10th grade test score	13.9	5.0
Father's education	13.1	2.6
Mother's education	12.8	2.1
% father's education unknown	6.8	
% mother's education unknown	3.8	
Family income <$12,000	20.8	
Family income $12–20,000	39.1	
Family income >$20,000	40.1	
% female	52.2	
% Black	13.4	
Unemployment rate (%)	6.0	3.3
Weekly wage ($)	232.90	77.90
Local college completion rate (%)	17.7	8.0
In-state public tuition ($)	946.70	303.40
Student–teacher ratio	19.0	4.4
Starting teacher salary/college graduate salary (%)	71.4	9.4

TABLE A2.—WITHIN-GROUP PARAMETER ESTIMATES FOR TEST SCORE PRODUCTION AND SCHOOL ATTAINMENT

	12th Grade Test Score		Years of Schooling Completed	
Variable	School Fixed Effects	State Fixed Effects	School Fixed Effects	State Fixed Effects
Prior Test Score	0.86	0.86	0.17	0.16
	(134.9)	(139.7)	(19.8)	(20.8)
Father's Education	0.10	0.11	0.10	0.11
	(7.10)	(8.58)	(6.16)	(7.06)
Father's Education Unknown	0.91	1.08	1.14	1.18
	(4.34)	(5.30)	(4.50)	(4.94)
Mother's Education	0.08	0.09	0.11	0.11
	(5.31)	(5.83)	(5.40)	(6.12)
Mother's Education Unknown	0.56	0.72	1.16	1.08
	(2.17)	(2.85)	(2.87)	(2.84)
Family Income between $12,000 and $20,000			−0.08	−0.12
			(−0.88)	(−1.40)
Family Income over $20,000			0.28	0.27
			(2.82)	(2.88)
Female	−0.33	−0.33	0.16	0.16
	(−5.89)	(−6.06)	(2.42)	(2.69)
Black	−0.58	−0.59	0.39	0.47
	(−5.11)	(−6.65)	(3.46)	(5.57)
Observations	11,386	11,386	2,309	2,309

Note: t-statistics are in parentheses.