2006

Meeting the Standards: An Analysis of Eight Grade Educational Assesment Test Scores in Maine

Michael Donihue
Colby College

Joseph Mattos
Colby College

Caroline Theoharides
Colby College

Charlotte Tiffany
Colby College

Goldfarb Center for Public Affairs and Civic Engagement

WORKING PAPER SERIES

Meeting the Standards: An Analysis of Eighth Grade Educational Assessment Test Scores in Maine

Michael Donihue
Joseph Mattos
Caroline Theoharides
Charlotte Tiffany

Working Paper *2006-003*
August 2006

Colby College
5310 Mayflower Hill
Waterville, ME 04901–8848

Meeting the Standards: An Analysis of Educational Assessment Test Scores in Maine
Michael Donihue, Joseph Mattos, Caroline Theoharides, Charlotte Tiffany
Goldfarb Center Working Paper No. *2006-003*
August 2006

Abstract

This paper examines the impact of socioeconomic factors on eighth grade achievement test scores in the face of federal and state initiatives for educational reform in Maine. We use student-level data over a five year period to provide a framework for understanding the policy implications of these initiatives. We model performance on standardized tests using a seemingly unrelated regressions approach and then determine the likelihood of meeting the standards defined by the adequate yearly progress requirements of the No Child Left Behind Act and Maine Learning Results initiatives. Our results indicate that the key factors influencing a student's test scores include the education of a student's parents, special services received for learning disabilities, and alternative measures of academic achievement.

Michael Donihue[*] '79
Associate Professor of Economics, Colby College, Waterville, ME
Michael.Donihue@colby.edu

Joseph Mattos '73
Assistant Superintendent, Waterville Public School District, Waterville, ME
jmattos@fc.wtvl.k12.me.us

Caroline Theoharides '06
Colby College, Waterville, ME
cbtheoha@alum.colby.edu

Charlotte Tiffany '00
Colby College, Waterville, ME
charlottetiffany@hotmail.com

[*]Corresponding author. We gratefully acknowledge the data collection efforts of Colby College students Katie Gagne '05, Andriy Avramenko '04, and Ben Schreiner '01 on this project. Any errors remain the responsibility of the authors.

Meeting the Standards:

An Analysis of Educational Assessment Test Scores in Maine

Introduction

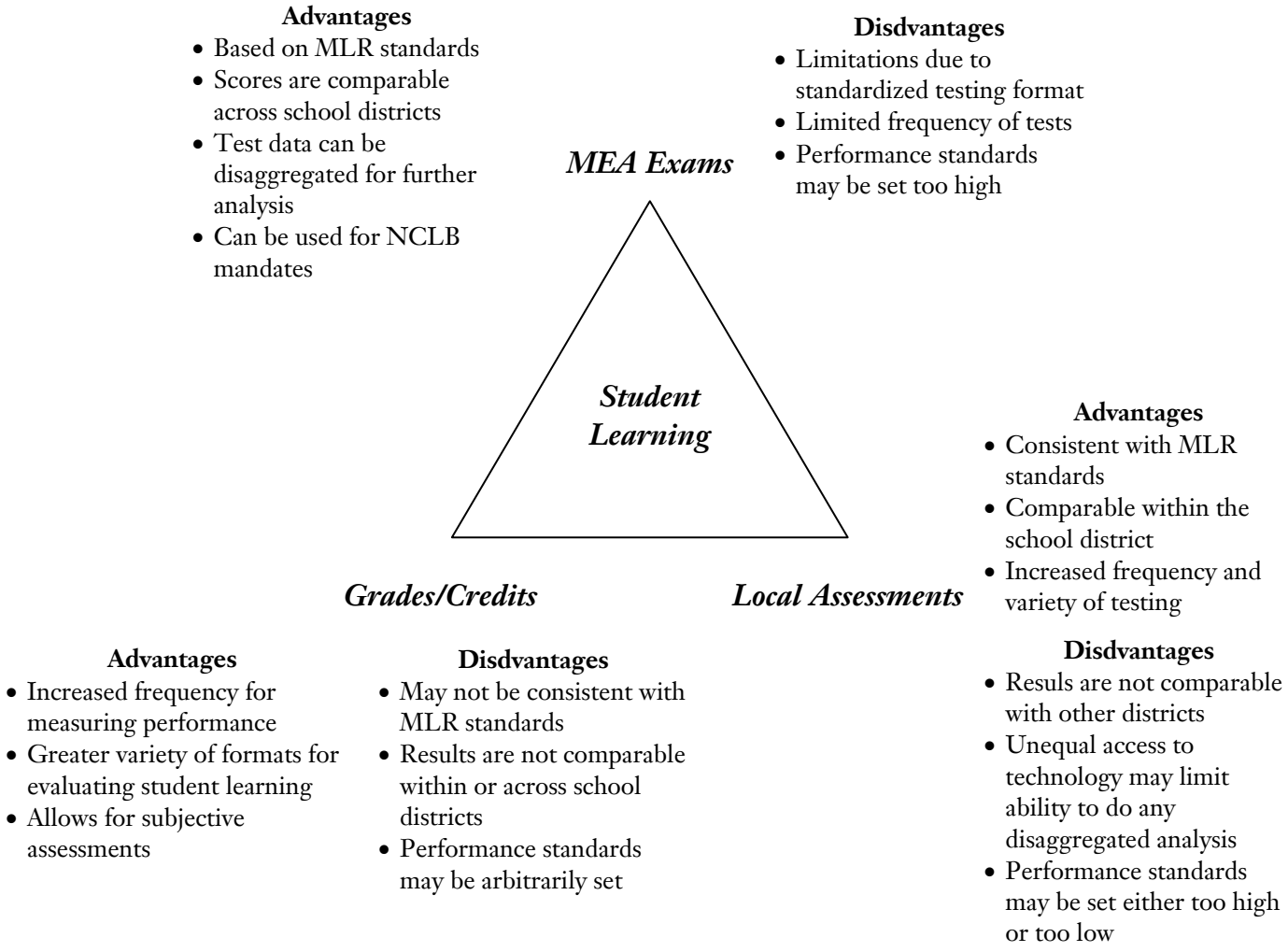
The use of standardized tests as a measure of student learning in the primary and secondary grade levels in the United States has long been a topic for debate among parents, school administrators, and state departments of education. In 2002 the No Child Left Behind Act (NCLB) gave national prominence to this debate as states began working to develop methods for complying with the mandates of this legislation. Many states have a significant history of using standardized tests for measuring student learning and are adapting their methods of evaluating the results of these tests to bring them in line with the requirements of the NCLB mandates for adequate yearly progress toward the educational reform goals set forth in the act.

Maine provides a good example of a state that has been pursuing significant educational reforms over the past decade for which standardized testing has been a predominant feature under an initiative entitled *Maine Learning Results* (MLR). From the perspective of NCLB, Maine's Educational Assessment (MEA) tests represent the primary means for evaluating student performance across school districts. However, these standardized tests are only one aspect of evaluating student learning under the reforms embodied by MLR.

Figure 1 provides an analytical framework for thinking about assessments of student learning and illustrates the context in which school administrators in Maine think about the role of standardized testing. The process of evaluating student learning encompasses three interrelated components: state-wide standardized tests represented by the MEA exams; grades and/or credits awarded in each school district; and local assessments that may or may not be formally linked to educational goals across school districts. Each method of evaluation has its advantages and disadvantages. For Maine, the MEA exams provide important information for determining what types of student work and performance levels represent specific achievement standards. The MEA exams also provide useful sub-group information that helps determine whether or not there is educational equity in the student population. The primary concern with using standardized test results in this fashion,

however, has been in the definition of appropriate levels of achievement for purposes of determining excellence in education and accountability for the school district.

Figure 1
Assessments of Student Learning in Maine



Local assessments confront many of the same issues as state-wide exams, however they allow for more flexibility in terms of both testing frequency and format in determining the MLR goals of educational excellence and equity. Comparability across school districts can be a problem as incentives exist to tailor the local assessment tools and the goals to be evaluated to reflect local conditions.

Grades and credits provide for the greatest variety of assessments and the most flexibility in evaluating student learning. However, there is much less comparability, even within a particular school, as standards for excellence are often subjectively determined by teachers and driven by the curriculum.

We found it difficult to apply our framework for evaluating student learning using the existing literature on socio-economic factors contributing to student achievement. Some of the work on educational attainment looks at aggregate scores by school districts using a production function approach introduced by Hanushek (1986, 1989) and applied more recently by Chakraborty, Biswas, and Lewis (2001) and by Hanushek and Raymond (2004). In this context, the problem becomes one of maximizing output, represented by some measure of educational achievement, given a set of behavioral inputs and technology, which may be represented by physical capital or features unique to a particular curriculum.

Primont and Domazlicky (2006) provide an interesting linear programming alternative to the production function approach in the evaluation of student achievement and efficiency in Missouri schools in light of the NCLB mandates. Using broad demographic aggregates and pass rates in 355 school districts on three standardized tests, Primont and Domazlicky estimate behavioral equations and simulate alternative scenarios to draw conclusions about the impact of the NCLB mandates on managerial efficiency among schools.

Alexander (2000), in a study that predates NCLB, uses a broad data set for school districts in New York in focusing on the state Regents exams to evaluate the impact of curriculum standards on student performance. Due to the nature of the data set, Alexander has few social indicators to work with, but does develop some insightful results that indicate that curriculum standards can improve student performance, but do not necessarily improve educational equity.

More broadly, there exist a large number of studies that employ data from national longitudinal surveys of students in a variety of contexts. Most of these studies predate NCLB and deal somewhat differently with educational reform initiatives at the state level. A representative example of this literature is the article by Bishop and Mane (2001) in which the authors examine a representative sample of eighth graders from 1988 to 1994 found in the National Educational Longitudinal Study to look. The goal of their study was to

quantify the impact of state graduation requirements on subsequent academic performance in high school. In this article, they conclude that, in general, such requirements do not seem to have positive effects.

Among high school students, there is another longitudinal database maintained by the Department of Education entitled High School and Beyond (HS&B) that is favored by many researchers. There have been several studies of educational attainment published in specific disciplines that figure prominently in the literature using the HS&B data. One example is the article by Jones (1987) who finds a significant marginal gain from additional math courses in overall math competency, controlling for various demographic sub-groups.

In the analysis that follows we have chosen to focus on student-level data over a five-year period to evaluate standardized test scores and take a somewhat different analytical approach to the issue of performance standards and student learning. In our analysis we model performance on the MEA exams in five subject areas for five class years of students in a rural school district in Maine, controlling for social and demographic characteristics of the students. We evaluate student performance in terms of both exam scores and in terms of the factors that determine the likelihood that a student will achieve the standards set forth by MLR which are consistent the mandates for adequate yearly progress required by NCLB. We believe that by targeting our study in this way we can provide a more useful framework for school administrators in understanding the policy implications that arise from the use of standardized tests for evaluating student performance. More importantly, the statistical methods and results in this paper can provide teachers and administrators with a means for comparing, and perhaps validating, competing measures of student learning.

The analysis that follows begins with an overview of the educational reform initiatives represented by MLR and the role of the MEA exams in defining standards for student learning. We then describe the school district from which we gathered the data used in our analysis. Next we estimate student achievement on five of the MEA exams for a panel of 479 students from five class years using a seemingly unrelated regressions framework. We then estimate the likelihood that a student from this sample will meet the MLR standards using a variety of demographic and social control variables and present our conclusions.

Maine's Learning Results

The historical underpinnings for Maine's Learning Results lie in the Maine Education Reform Act (MERA) of 1984 which was the state's response to the *Nation at Risk Report* published in 1983 by the National Commission on Excellence in Education. The MERA contained initiatives for educational reform that focused on improvements in the quality of teaching and student learning and the equity of educational opportunities across the state. One of the initiatives included in the MERA was the creation of the MEA exams. In 1989 a state-wide task force was commissioned by the governor to build upon the educational reforms in MERA and identify targets for educational outcomes for Maine's students. In 1990 Maine's *Common Core of Learning* emerged from this task force. This report provided a non-disciplinary organization of knowledge, skills, and attitudes for high school graduates based on an integrated approach to teaching and learning.

Significant progress on educational reform continued in 1993 when the State Board of Education, at the direction of the Maine Legislature, established the Task Force on Learning Results. This task force was charged with developing long-range educational goals and standards for both student achievement and school performance. Three years later the task force presented its recommendations in the form of proposed legislation that included a set of specific performance standards and a plan for implementing them. With this task force report, the issue of direct accountability for student learning became an integral part of the educational reform process in Maine. From the principles and guidelines contained in the task force report, the Legislature directed the Department of Education and the State Board of Education to develop "Learning Results" in eight areas of learning. Subsequent legislation in 1997 provided for accountability of student achievement of the Learning Results through a combination of state and local assessments. The MEA exams were redesigned in 1998 to align the exams with the educational achievement goals and standards set forth in Maine's Learning Results.

Beginning with the 1998-1999 academic year, the redesigned MEA exams have been administered to students in the fourth, eighth, and eleventh grades. In our study we examine student performance on the MEA exams in Mathematics, Science and Technology, Social Studies, Reading, and Writing. Scores on the exams are tabulated by the

Department of Education and scaled to align the results with tests in previous years and the standards for performance defined by MLR. Scores are reported to parents and school administrators along with summary information and classifications according to the MLR standards. These classifications include identification for students who do not meet the standards, who partially meet the standards, who meet the standards, and who exceed the standards.

For the purposes of the NCLB mandates, adequate yearly progress in Maine is determined by the percentage of students who achieve the "meet the standards" level on the Math and Reading exams. Each year, the percentage of students who achieve this level is compared to Maine's target benchmarks established by NCLB. These targets increase each year with a goal of reaching 100 percent achievement by the year 2014. As in other states, the NCLB adequate yearly progress mandates in Maine apply to sub-categories of students as well, including different ethnic groups, students with disabilities, and students coming from economically disadvantaged families.

In the analysis that follows we chose to examine only eighth grade MEA scores for two reasons. First, by the eleventh grade the majority of students who will drop out of high school will have already left, thus biasing the sample of students remaining to take the exams. Anecdotally, it was also determined that many students in rural school districts in Maine pay less attention to the MEA exams in the eleventh grade than they do to other exams occurring at the time, like the SAT exams for example.

Second, the eighth grade MEA exam scores, as well as grades and a variety of local assessments, are important tools in the advising process. During a student's high school years, the MEA exam scores provide supplemental information to help determine whether intervention programs, like study halls, literacy classes, tutorials, or summer programs, might be warranted for students at risk of dropping out of high school or failing to meet the MLR standards necessary for certification for a high school diploma.

A Rural Maine School District

The school district chosen for this analysis is located in rural Maine serving a community of about 15,600 people. In many ways this community is typical of small towns

throughout New England. The past decade has seen the closure of major manufacturing employers in the area including a paper mill, shoe manufacturers, and a shirt factory. However, the school district has some advantages over other small communities in the state. There are two four-year colleges in the community, one offering graduate degrees in business and computer technology. There is a state community college nearby and a branch campus of the University of Maine in the area. The major employers in the school district include two hospitals and the two colleges. There are two parochial schools in the area offering classes through grade six. In terms of the economic climate in the area, local employment fell by 3.6% from 1999 to 2004 and the town had an unemployment rate nearly one percentage point above the state average in 2004. According to the 2000 census, median household income in the school district was $26,816, about thirty percent less than that for all of Maine.

Average scores on all five eighth-grade MEA exams in the public school district were generally at or above those for the state as a whole for the five class years we studied. The school district we studied does not 'track' students in the sense of separating certain boys or girls into a more rigorous set of classes. There are, however, different levels of mathematics taught in the sixth through eighth grades, including an "advanced math" curriculum. The mathematical abilities of students are evaluated by teachers beginning in the fourth grade and qualified students are placed into an advanced curriculum with opportunities for switching in or out of the classes up to and including the eighth grade year of study. Students completing the advanced math class in the eighth grade are then eligible to begin a more advanced curriculum in math and sciences when they enter the high school.

Student Data

We collected information from a total of 840 student records across five class years using student records made available by the administrators of the school district with the approval of the school board. For each student, data were also collected directly from MEA score reports, from school yearbooks, from report cards, and from participation lists provided by school administrators. Background information on a student's parents and

number of siblings in the household was obtained from registration cards completed when the student entered the school system.

Because we relied on school records, rather than direct surveys of students or parents, and the school district we studied was a rural district without computerized record keeping, there were limitations in our ability to use the data available to us. For many of the students the background information we collected on their family was recorded when the student first entered the school system and never updated. Some students who transferred into the school system did not have registration cards with this background information. For other students the registration cards were incomplete. And in some cases a student took only a subset of the five MEA subject tests we chose to focus on.

Definitions for the twenty-five variables for which we collected data are presented in Table 1 and summary statistics for our entire database are provided in Tables 2 and 3. Students in the high school graduating class of 2003 took their eighth grade MEA exams during the 1998–1999 academic year. Test score results were subsequently made available to parents and school administrators during the fall of 1999. Similarly, the eighth grade MEA exams for the final class in our sample, the class of 2007, were administered to students during the 2002–2003 academic year.

Table 1
Variable Definitions

advm	Equals 1 if the student is enrolled in the 8th grade advanced math class; 0 otherwise
divorce	Equals 1 if the student's parents are divorced; 0 otherwise
dnm_math	Equals 1 if the student scored 520 or below on the MEA Math exam; 0 otherwise
dnm_read	Equals 1 if the student scored 520 or below on the MEA Reading exam; 0 otherwise
dnm_sci	Equals 1 if the student scored 520 or below on the MEA Science and Technology exam; 0 otherwise.
dnm_socs	Equals 1 if the student scored 520 or below on the MEA Social Studies exam; 0 otherwise
dnm_writ	Equals 1 if the student scored 520 or below on the MEA Writing exam; 0 otherwise
eduf	Years of education of the student's father
edum	Years of education of the student's mother
female	Equals 1 if the student is female; 0 if male
hhr	Equals 1 if the student achieved high honors in the 8th grade; 0 otherwise
hldb	Equals 1 if the student was ever held back to repeat a grade level; 0 otherwise
lunch	Equals 1 if the student participated in the federally subsidized school lunch program; 0 otherwise.
math	MEA Mathematics exam score
migr	Equals 1 if the student is a child of a migrant worker family; 0 otherwise
occf	Equals 1 if the father is employed in a professional occupation;* 0 otherwise
occm	Equals 1 if the mother is employed in a professional occupation;* 0 otherwise
paro	Equals 1 if the student previously attended a parochial school; 0 otherwise
read	MEA Reading exam score
sci	MEA Science and Technology exam score
sibl	Number of siblings in the student's family
socs	MEA Social Studies exam score
specs	Equals 1 if the student received special services; 0 otherwise
sport	Equals 1 if the student played a team sport during the 8th grade; 0 otherwise
writ	MEA Writing exam score
y03	Equals 1 if the student is a member of the high school graduating class of 2003; 0 otherwise
y04	Equals 1 if the student is a member of the high school graduating class of 2004; 0 otherwise
y05	Equals 1 if the student is a member of the high school graduating class of 2005; 0 otherwise
y06	Equals 1 if the student is a member of the high school graduating class of 2006; 0 otherwise
y07	Equals 1 if the student is a member of the high school graduating class of 2007; 0 otherwise

*Professional occupation was determined using US Bureau of Labor Statistics definitions.

Table 2
Summary Statistics

Variable	Valid Observations	Mean	Standard Deviation	Minimum	Maximum	Number of Students
advm	840	0.21	0.41	0	1	180
divorce	762	0.40	0.49	0	1	301
dnm_math	747	0.38	0.49	0	1	287
dnm_read	766	0.10	0.30	0	1	76
dnm_sci	751	0.23	0.42	0	1	171
dnm_socs	747	0.21	0.40	0	1	154
dnm_writ	758	0.09	0.29	0	1	68
eduf	601	13.92	2.74	8	22	
edum	678	13.52	2.38	1	20	
female	840	0.48	0.50	0	1	407
hhr	840	0.21	0.41	0	1	180
hldb	840	0.13	0.34	0	1	112
lunch	840	0.30	0.46	0	1	249
math	747	528.04	15.44	502	566	
migr	840	0.02	0.13	0	1	15
occf	611	0.32	0.47	0	1	193
occm	670	0.29	0.45	0	1	191
paro	840	0.14	0.34	0	1	115
read	766	538.41	12.63	502	572	
sci	751	531.31	12.88	502	566	
sibl	717	2.04	1.37	0	11	
socs	747	533.01	13.45	502	570	
specs	840	0.16	0.37	0	1	134
sport	840	0.52	0.50	0	1	436
writ	758	535.76	10.70	504	562	
y03	840	0.26	0.44	0	1	216
y04	840	0.20	0.40	0	1	165
y05	840	0.20	0.40	0	1	170
y06	840	0.17	0.37	0	1	140
y07	840	0.18	0.38	0	1	149

As indicated by the summary statistics in Table 2, of the 840 student records we examined 48% (407) were for females and 52% (433) were for eighth-grade boys. In our analysis we considered *advm* as a 'tracking' indicator variable, distinguishing students in the advanced math classes deemed by their teachers to be capable of higher academic

achievement. Twenty-one percent (180) of the students in our study were enrolled in the advanced math class in eighth grade. Of these, 48% (87) were female.

We were able to obtain the parents' marital status for 91% (762) of the students from their registration cards and of these, 40% (301) had parents who were divorced. Slightly fewer records had information on the years of schooling for the student's parents at the time of registration. On average, however, both mothers and fathers completed at least one year of post-high school education. Twenty-one percent of the students for whom we collected data achieved high academic honors during their eighth grade year; fifty-eight percent of these were in the advanced math class. Thirteen percent (112) of the students repeated an academic year prior to completing the eighth grade. Only 2% (15) of the students came from migrant worker families while 14% (115) transferred into the public school system from parochial schools. Over half of the students participated in a team sport in the eighth grade.[1]

We had no income data for student households. We were, however, able to identify students who participated in the federally subsidized school lunch program. Eligibility for the program is based on an income threshold. Thirty percent (249) of the eighth grade students in our dataset came from households that chose to participate in the federally subsidized school lunch program administered by the school district.[2]

We tried to classify the employment status of a student's parents according to the type of job they had, but here the limited nature of the data available on the registration card presented difficulties. We ultimately decided to subjectively assign parents into a professional occupation using the US Bureau of Labor Statistics definitions of professional jobs as a guideline. Unemployed mothers or fathers were assigned a value of zero for *occm* and *occf*, respectively. Slightly less than one-third of the parents in our study were employed as professionals.

Sixteen percent (134) of the 840 students received 'special services' at some point during their academic years prior to completing the eighth grade. Most of these involved students

[1] Our definition of team sports included cheerleading.

[2] During the time period of our study, the total number of students participating in the federally subsidized school lunch program in the school district increased each year. By 2004 approximately fifty percent of the K-12 students in the district qualified for the program.

with significant learning disabilities. However, some students may have received short-term help for speech impediments, temporary health conditions, or reading difficulties in earlier grade levels. Slightly more than 13% (95) of the students who took all five MEA subject exams had received special services.

Eighty-four percent (708) of the students had scores for all five MEA exams among their records. In terms of meeting the MLR standards, 287 students scored at or below the required threshold of 520. All but four of these students were not in the advanced math curriculum. Thirty-eight percent of the students in our sample did not meet the MLR standards in math. Ten percent of the students did not meet the MLR standards in reading while only 9 percent did not meet the MLR standards in writing. On the MEA Science and Technology exam, 23% did not meet the MLR standards. Twenty-one percent failed to meet the MLR standards in social studies. The class of 2007 had the highest average scores on the MEA Mathematics and Writing exams. The class of 2005 had the highest average MEA Science and Technology exam score. The class of 2004 scored the highest on both the Social Studies and Reading exams.

Table 3 provides a correlation matrix for the variables in our database. Most of the MEA exam scores are highly correlated with each other (Math and Writing being the exception), indicating that a student who does well on one exam is likely to do well on the other subject tests. Among the family background characteristics we observed a fairly high correlation between parents' education levels as well as the education and occupation levels of both fathers and mothers, as might be expected. There also exists a consistent negative correlation between MEA performance on all subject tests and those students who participated in the federally subsidized school lunch program, and those who received special services.

Table 3: Correlation Matrix

	advm	div	eduf	edum	female	hhr	hldb	lunch	math	migr
divorce	-0.175									
eduf	0.379	-0.214								
edum	0.363	-0.174	0.595							
female	-0.001	0.016	0.018	-0.033						
hhr	0.477	-0.216	0.390	0.322	0.179					
hldb	-0.162	0.143	-0.102	-0.146	-0.009	-0.119				
lunch	-0.218	0.258	-0.353	-0.350	-0.024	-0.206	0.183			
math	0.648	-0.198	0.435	0.411	0.056	0.536	-0.156	-0.302		
migr	-0.049	0.049	-0.064	-0.053	0.049	-0.027	0.026	0.129	-0.101	
occf	0.235	-0.084	0.608	0.400	-0.007	0.237	-0.018	-0.186	0.323	-0.068
occm	0.215	-0.055	0.397	0.566	0.012	0.161	-0.091	-0.260	0.197	-0.039
paro	0.138	-0.144	0.186	0.152	-0.012	0.121	-0.044	-0.221	0.139	-0.054
read	0.473	-0.246	0.420	0.419	0.215	0.506	-0.184	-0.323	0.693	-0.004
sci	0.534	-0.234	0.435	0.450	-0.003	0.464	-0.163	-0.308	0.739	-0.116
sibl	-0.113	-0.024	-0.073	-0.117	0.024	-0.029	0.101	0.206	-0.089	0.275
socs	0.481	-0.216	0.453	0.442	0.098	0.495	-0.188	-0.327	0.719	-0.068
specs	-0.220	0.104	-0.076	-0.103	-0.123	-0.164	0.240	0.159	-0.346	0.064
sport	0.259	-0.188	0.201	0.253	-0.025	0.166	-0.078	-0.267	0.203	-0.068
writ	0.408	-0.169	0.345	0.330	0.340	0.443	-0.171	-0.288	0.490	-0.050
y03	-0.082	0.022	-0.022	-0.030	-0.014	-0.135	0.066	-0.006	-0.073	-0.038
y04	0.005	0.012	0.065	0.047	-0.036	0.187	-0.150	-0.013	0.012	0.001
y05	0.004	-0.063	-0.033	-0.028	0.016	-0.032	0.064	0.101	0.063	-0.001
y06	0.047	-0.070	0.030	0.031	-0.012	0.000	-0.006	-0.059	-0.015	-0.036
y07	0.039	0.088	-0.043	-0.017	0.049	-0.007	0.020	-0.028	0.020	0.079

	occf	occm	paro	read	sci	sibl	socs	specs	sport	writ
occm	0.319									
paro	0.186	0.148								
read	0.268	0.239	0.146							
sci	0.310	0.241	0.168	0.691						
sibl	-0.047	-0.037	-0.013	-0.027	-0.146					
socs	0.333	0.261	0.158	0.753	0.781	-0.091				
specs	-0.040	-0.035	-0.060	-0.354	-0.329	0.035	-0.356			
sport	0.144	0.158	0.182	0.295	0.225	-0.074	0.222	-0.160		
writ	0.159	0.189	0.148	0.631	0.484	-0.129	0.565	-0.334	0.267	
y03	0.028	0.055	0.019	-0.020	-0.095	0.309	-0.070	0.056	0.136	-0.108
y04	-0.044	0.022	-0.040	0.064	0.027	0.106	0.110	-0.052	-0.046	-0.063
y05	0.020	-0.014	-0.045	-0.031	0.050	-0.123	0.071	-0.106	-0.079	0.035
y06	0.039	-0.036	0.082	-0.040	0.034	-0.206	-0.131	0.049	-0.004	-0.002
y07	-0.043	-0.039	-0.013	0.025	-0.008	-0.161	0.017	0.053	-0.021	0.149

Pairwise sample sizes.

Explaining MEA Scores

We chose to model student performance on each assessment test using a linear regression framework. Across the five class years in our data set we identified a balanced panel of 479 students who completed all five exams and for whom we have a complete set of information for a relevant subset of the explanatory variables described above. In each of the five regression equations we focused on a common core of explanatory variables including: our tracking variable, *advm*; the education level of the student's parents, *edum* and *eduf*; the student's sex; a summary measure of academic achievement, *hhr*; our income proxy, *lunch*; and whether or not the student received special services, *specs*. In addition, we considered other variables that might prove relevant to performance on individual exams, but might not apply to all tests. We also controlled for fixed effects using class year dummy variables.

In terms of anticipated effects among our core explanatory variables, we expect that students enrolled in the advanced math curriculum and those achieving high honors during the eighth grade will, all else equal, perform better on all five exams. We also expect that the more years of education a student's parents have the higher their test scores. Students receiving special services are expected to have lower exam scores on average. We also anticipate that students participating in the federally subsidized school lunch program, because this variable serves as a proxy for income disadvantaged households, will have lower average exam scores. We make no a priori assumptions about the effect of gender on exam scores.

For the five equations in our model each observation corresponds to the same student. The data are organized such that the dependent variable in each equation represents a different test score for the same student, and the independent variables in each equation explain the variation in student performance on the corresponding assessment exam. Using this framework, it is likely that the prediction errors for each equation will be correlated. In other words, the unexplained variation in a particular student's writing test score, for example, is likely to be correlated with the prediction error for her reading test score, and to varying degrees with her science, math, and social studies scores as well. To account for this correlation, and thus improve the efficiency of the parameter estimates in our model, we

employ a seemingly unrelated regressions (SUR) estimation approach. Our results are presented in Table 4.

Our results are broadly consistent with the anticipated effects of our core explanatory variables. Students enrolled in the advanced math curriculum in the eighth grade perform significantly better on all exams; with the greatest effect appearing on the MEA Math exam where the scores for these students are 15 points higher on average than students not enrolled in advanced math. Not surprisingly, the difference appears to be smallest, although still significant, on the MEA Writing exam. We also find that high honor roll students perform significantly better on all exams, all else equal.

The coefficients on our parent's years of education variables, *eduf* and *edum*, provide a measure of what is referred to as the intergenerational transfer of human capital effect.[3] As expected, we find that the more years of schooling a student's parents have the higher the student's test scores, on average. For all five exams, the effect of an additional year of schooling for a student's mother appears to be larger than that for the father.[4]

According to our results, while an additional year of schooling for the mother provides a statistically significant, positive effect on student performance on each exam, an additional year of education for the father appears to have no statistically significant effect on a child's performance on the MEA Writing exam. A father's education level does contribute positively to performance on all of the other exams, however.

[3] See the working paper by de Walque (2005) for a good exposition of this issue in the context of economic development in Africa.

[4] However, we conducted appropriate chi-squared hypothesis tests for each pair of coefficient estimates and found that the difference in the effects between a mother and a father's education levels is not statistically significant in any of the equations

Table 4
Explaining MEA test scores

Explanatory Variables	Mathematics		Science and Technology		Social Studies		Reading		Writing	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
advm	15.00***	1.22	7.17***	1.13	5.67***	1.15	4.96***	1.04	2.94***	0.93
eduf	0.55**	0.22	0.58***	0.20	0.73***	0.21	0.44**	0.19	0.24	0.17
edum	0.77***	0.27	0.68***	0.23	0.79***	0.23	0.65***	0.21	0.50***	0.19
female	-0.29	0.94	-1.27	0.87	1.50*	0.89	4.24***	0.80	5.25***	0.72
hhr	7.68***	1.27	5.69***	1.18	7.08***	1.20	6.62***	1.09	5.43***	0.97
lunch	-2.07*	1.17	-1.91*	1.07	-2.33**	1.09	-2.54**	1.00	-1.36	0.89
specs	-8.13***	1.46	-6.80***	1.36	-6.77***	1.38	-7.30***	1.25	-5.25***	1.12
sibl	0.79**	0.35								
occm	-2.00**	1.00								
sport							1.40**	0.68	2.56***	0.72
y03	-1.19	1.46	-1.87	1.31	-1.68	1.33	-1.29	1.21	-6.44***	1.09
y04	-1.54	1.48	-1.15	1.35	-0.67	1.38	-1.37	1.25	-6.56***	1.11
y05	2.17	1.60	1.55	1.48	1.81	1.51	-0.93	1.37	-3.00**	1.22
y06	0.97	1.52	0.81	1.41	-4.99***	1.43	-1.61	1.30	-4.18***	1.16
intercept	505.8***	3.74	514.1***	3.22	512.1***	3.28	522.0***	2.98	525.7***	2.66
R^2	0.578		0.432		0.476		0.482		0.456	

Seemingly Unrelated Regression estimates for a balanced panel of 479 students.
Two-tailed levels of statistical significance: *** 1 percent level; ** 5 percent level; * 10 percent level.

We found that students who participate in the federally subsidized school lunch program have lower test scores, on average. However, the effect is not as statistically significant as some of our other core variables. On the MEA Writing exam there appears to be no statistically significant difference in performance between income disadvantaged students and those who do not participate in the subsidized school lunch program, all else equal.

In terms of other factors that contribute to explaining variations in performance on the MEA exams, the number of siblings and the mother's profession were both statistically significant factors in explaining variations in student performance on the MEA Mathematics exam. According to our results, the more brothers and sisters a student has the better their performance on the MEA Mathematics test, on average. This may reflect the availability of greater resources at home as math is one subject area where siblings can more easily learn from each other or are more likely to have invested in learning tools such as a calculator or flash cards.

Somewhat surprisingly, if a mother is employed in a professional occupation, her child's score on the MEA Mathematics exam was two points lower, on average, than the test scores for children of mothers not employed in a professional occupation according to our results. We also found that students who participated in at least one team sport performed better than students who did not participate in sports on both the Reading an Writing MEA exams.

Finally, in terms of the fixed effects in our model the only clear difference across class years was that the class of 2007 outperformed the others in the MEA Writing exam by a statistically significant margin. In addition, scores on the MEA Social Studies exam for the class year of 2006 were significantly lower than for the class of 2007.

Meeting the MLR Standards

An important question for public school administrators is how best to devote scarce resources to meet the educational needs of the community and comply with the goals of MLR and the mandates of NCLB. To provide some guidance in answering this question we first identified those students who did not meet the standards in terms of their

performance on the MEA exams. Then, to determine which factors explained the likelihood that a student would not meet the MLR standard on a particular exam, we estimated logistic regressions in which the value of the dependent variable was set equal to 1 for those students who failed to meet the standard and 0 for students whose score was greater than the MLR threshold.

The results in Table 5 examine the likelihood that a student will fail to meet the MLR standards in the five subject areas. Each year the majority of the students in the school district met or exceeded the MLR standards in every subject area. Thus, for this part of our analysis we confined the list of potential explanatory variables to the core variables defined above in order to maximize the number of student observations. The results in Table 5 are largely consistent with our anticipated results for the predictive test score equations in Table 4.[5] For the MEA Mathematics exam, 35% of the 547 students with valid observations for all variables in this regression failed to meet the MLR standard. As expected, students enrolled in the advanced math curriculum were less likely to fail to meet the standard. Students with highly educated parents were also more likely to meet the MLR standard, as were students who achieved high honors in the eighth grade. There was a statistically significant likelihood that students receiving special services would fail to meet the MLR standard, all else equal. We also found a slight statistical advantage for eighth grade girls in meeting the MLR math standard. However, there appeared to be no statistically significant difference related to the subsidized school lunch program using the core set of control variables for this sample of students.

[5] Among the explanatory variables in Tables 5 and 6, *advm* could not be included in the MEA Science or Writing regressions because none of the students who failed to meet the MLR standards were enrolled in the advanced math curriculum. Similarly, *hbr* was excluded from the MEA Reading regression because none of the students who failed to meet the MLR standard achieved high honors.

Table 5
Likelihood a student failed to meet the MLR standards[5]

Explanatory Variables	Mathematics		Science and Technology		Social Studies		Reading		Writing	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
advm	-2.27***	0.54			-0.43	0.59	-1.75*	1.04		
eduf	-0.14***	0.06	-0.14*	0.07	-0.23***	0.08	-0.26**	0.11	-0.18	0.12
edum	-0.14**	0.06	-0.31***	0.08	-0.29***	0.08	-0.06	0.10	-0.16	0.11
female	-0.40*	0.22	0.18	0.26	-0.70**	0.29	-0.71*	0.38	-1.18**	0.50
hhr	-1.56***	0.43	-3.41***	1.04	-2.76***	1.06			-1.46	1.07
lunch	0.24	0.25	0.33	0.27	0.53*	0.29	0.71*	0.37	0.51	0.45
specs	1.18***	0.32	1.78***	0.33	1.49***	0.34	1.11***	0.39	2.18***	0.45
y03	0.22	0.33	0.40	0.38	0.32	0.41	-0.23	0.52	3.22***	1.11
y04	0.34	0.35	0.65	0.41	0.20	0.45	-0.21	0.54	3.20***	1.13
y05	-0.60	0.39	0.24	0.44	-0.79	0.56	-0.23	0.62	2.78**	1.18
y06	0.02	0.35	-0.40	0.45	1.36***	0.43	0.15	0.53	0.43	1.47
intercept	3.64***	0.92	3.85***	1.19	4.95***	1.27	1.69	1.62	-1.11	2.11
Students	547		549		547		548		544	
Number who did not meet MLR standard	193 (35%)		101 (18%)		94 (17%)		41 (7%)		33 (6%)	

Logistic regression estimates.
Two-tailed levels of significance for normally distributed z-scores of each coefficient estimate: *** 1 percent level; ** 5 percent level; * 10 percent level.

Looking at the pattern of results in Table 5 across the other subject areas it is clear that students who have received special services from the public school system are more likely to be unsuccessful in meeting the MLR standards. Unlike the results in Table 4, however, a statistically significant gender effect emerged in our results for students who did not meet the MLR standards in all but the Science and Technology subject areas. In all of the other areas of learning girls were more likely than boys to achieve the MLR standards. Parent education levels do not appear to explain in a statistically significant fashion the outcome for the 6% of the students who failed to meet the MLR standards in writing. And, according to our results, a father's level of education is statistically more important than the mother's years of schooling in explaining success in meeting the MLR standards in reading although both contribute negatively to failure.

As a guide for policy makers, Table 6 presents the marginal effects for the regression results in Table 5 and thus provides an indicator of where the most important contributions lie in meeting the MLR standards. From the magnitude of the calculations in Table 6 we see that students who have received special services at some point during their time in the public school system appear to be at the greatest risk of not meeting the MLR standards on all exams. For the MEA Mathematics exam, thirteen percent of the 527 students in this sample received special services. All else equal, on a scale of 0 to 1, the likelihood that a student who has received special services will fail to meet the MLR standards in mathematics increases by 0.248 relative to other students according to our results. Not surprisingly, students in the advanced math curriculum appear to be at the lowest risk of failing to meet the MLR standards in mathematics as the likelihood falls by 0.293. Twenty-six percent of the 524 students in this sample were enrolled in the advanced math curriculum in the eighth grade.

Table 6
Increase or decrease in the likelihood a student failed to meet the MLR standards[5]

Explanatory Variables	Mathematics		Science and Technology		Social Studies		Reading		Writing	
	Marginal Effect	Sample Mean	Marginal Effect	Sample Mean	Marginal Effect	Sample Mean	Marginal Effect	Sample Mean	Marginal Effect	Sample Mean
advm	-0.293***	0.26			-0.020	0.26	-0.043*	0.26		
eduf	-0.025***	14.01	-0.009*	14.01	-0.012***	14.01	-0.008**	14.05	-0.002	14.04
edum	-0.025**	13.67	-0.020***	13.68	-0.015***	13.67	-0.002	13.72	-0.002	13.74
female	-0.070*	0.47	0.012	0.46	-0.036**	0.46	-0.023*	0.47	-0.016**	0.48
hhr	-0.220***	0.26	-0.152***	0.26	-0.098***	0.26			-0.015	0.27
lunch	0.043	0.24	0.024	0.25	0.030*	0.25	0.028*	0.25	0.008	0.24
specs	0.248***	0.13	0.206***	0.13	0.128***	0.13	0.055***	0.13	0.072***	0.13
y03	0.040	0.25	0.029	0.25	0.018	0.25	-0.007	0.25	0.127***	0.24
y04	0.062	0.23	0.051	0.23	0.011	0.23	-0.006	0.23	0.129***	0.23
y05	-0.094	0.16	0.017	0.16	-0.032	0.16	-0.007	0.16	0.113**	0.16
y06	0.003	0.18	-0.024	0.18	0.105***	0.18	0.005	0.18	0.006	0.18
Students	547		549		547		548		544	
Number who did not meet MLR standard	193 (35%)		101 (18%)		94 (17%)		41 (7%)		33 (6%)	

Marginal effects equal the derivatives computed using the logistic regression estimates in Table 5 and evaluated at the sample mean of each explanatory variable. The indicated levels of statistical significance are reproduced from Table 5.

For both the MEA Science and Technology and Social Studies exams, students who achieved high honors in the eighth grade were most likely to meet the MLR standards on average. Eighth grade girls were most likely to meet the MLR Writing standards, all else equal. In the case of the MEA Reading exam, boys who received special services and were not in an advanced math class appear to be at greatest risk for not meeting the MLR standards. And for the MEA Science and Technology exam, it would appear that a student who received special services, did not achieve high honors, and whose parents have less than the average number of years of schooling is most at risk for failing to meet the MLR standards.

Conclusions

Administrators in rural school districts across the country have come under increasing pressure during the past decade as they attempt to implement educational reform mandates imposed by state legislatures and the federal government in the face of declining enrollments, reductions in state and federal funding for education, and diminished economic opportunities in the communities they serve. In this paper we provide a framework for addressing the resource allocation dilemma faced by public school administrators attempting to meet the educational needs of the community and comply with these mandates by examining the experience of a rural public school district in Maine.

Across the country assessments of student learning generally include grades, other forms of local assessments, and some form of state-wide standardized tests. In Maine, the state legislature has set goals for educational attainment based on standardized assessments of student performance under the Maine Learning Results educational reform initiative. The MLR guidelines were implemented during the 1998-1999 school year using standardized test scores in five subject areas as the primary measure of how well these goals were being met. Our analysis focuses on the determinants of performance on these exams for five classes of eighth graders during the period of 1998 through 2003.

Unlike previous studies, our model uses student-level data to provide a framework for school administrators to better understanding the policy implications that arise from the use

of standardized tests for evaluating student performance. In some cases, grades, standardized tests, and local assessments result in conflicting conclusions regarding the quality of student learning. The statistical methods and results in this paper provide teachers and administrators with a means for validating competing measures of student learning.

According to our results, the most important factors contributing to a student's success at meeting the MLR standards include the educational attainment of the parents, whether a student received help for a learning disability, and their academic achievement in other areas of the school curriculum. We find some evidence that boys are at greater risk than girls in failing to meet the MLR standards. In addition, students who are eligible for the federally subsidized school lunch program generally perform worse on the standardized exams according to our results.

Public school teachers and administrators care deeply about the quality of their students' learning. Low levels of student achievement have generally been regarded as a sign of failure in a school system. Yet finding the root causes of these failures is a difficult process. Our research suggests that a careful analysis of the important determinants of performance on standardized tests can provide insight into the results of other assessments of student learning and provide direction for improving student achievement.

References

Alexander, Nicola A. (2000). "The missing link: an econometric analysis on the impact of curriculum standards on student achievement," *Economics of Education Review*, 19, 351–361.

Bishop, John H. and Ferran Mane (2001). "The impacts of minimum competency exam graduation requirements on high school graduation, college attendance and early labor market success," *Labour Economics*, 8(2), 203–222.

Chakraborty, Kalyan, Basudeb Biswas, W. Cris Lewis (2001). "Measurement of technical efficiency in public education: A stochastic and nonstochastic production function approach." *Southern Economic Journal*, 67(4), 889–905.

Coladarci, Theodore (1997). *State of Maine Learning Results*, Maine Department of Education, Augusta, ME.

de Walque, Damien (2005). "Parental Education and Children's Schooling Outcomes: Is the Effect Nature, Nurture, or Both?" *World Bank Policy Research Working Paper 3483*.

Hanushek, Eric A. (1986) "The economics of schooling: Production and efficiency in public schools," *Journal of Economic Literature*, 24, 1141–1177.

Hanushek, Eric A, Margaret E. Raymond (2004). "Does School Accountability Lead to Improved Student performance?" NBER *Working Paper W10591*.

Hanushek, Eric A., Lori L. Taylor (1990). "Alternative assessments of the performance of schools," *The Journal of Human Resources*, 25(2), 1990, 179–201.

Jones, Lyle V. (1987). "The Influence on Mathematics Test Scores, by Ethnicity and Sex, of Prior Achievement and High School Math Courses, 18(3), 180–186.

National Commission on Excellence in Education (1983). *A Nation At Risk: The Imperative Of Educational Reform*. Washington, DC: US Government Printing Office.

Primont, Diane F., Bruce Domazlicky (2006). "Student achievement and efficiency in Missouri schools and the No Child Left Behind Act," *Economics of Education Review*, 25(1), 77–90.