

Will No Child Left Behind Lead to Barbarianism?
The Effect of Accountability on Arts Education

An Honors Thesis for the Department of Economics

By Joshua Hyman

Tufts University, 2005

Will No Child Left Behind Lead to Barbarianism?
The Effect of Accountability on Arts Education
Joshua Hyman
Tufts University
2005

Abstract

With the recent implementation of No Child Left Behind (NCLB), school accountability systems based on student testing have become standard throughout United States public schools. Using administrative data from Wyoming and Kentucky, this paper explores the effects of school accountability on arts and music education. The hypothesis is that the focus of NCLB on reading, math and science, will draw resources away from untested subjects such as music, visual and performing arts. I find that in Wyoming, where they only test in reading, writing and math, NCLB causes provision of arts education to decline. In Kentucky, where there is testing in the arts and incentives are attached to performance, I find that NCLB has no effect on the provision of arts education.

Joshua Hyman
joshua.hyman@tufts.edu

Acknowledgments

Many thanks to Shannon Cranmore of the Wyoming Department of Education for taking the time to first convert to the proper format and then send me the numerous spreadsheets that I used for my Wyoming analysis. Thanks also to Kevin Hill of the Kentucky Department of Education for both supplying me with and helping me to retrieve the data I used for my Kentucky analysis.

I greatly appreciate helpful comments from Larry Getzler of the Virginia Department of Planning and Budget as well as from audience members at the 2005 conference of the American Education Finance Association. I appreciate the support of Jane Hannaway of the Urban Institute and Professor David Figlio of the University of Florida. I am grateful for the financial support of the Tufts Summer Scholars Program.

Thank you to Professors Linda Loury and Dorothy Pilla, the second and third members of my thesis committee, for meeting with me and offering me your assistance. Your comments and constructive criticisms were very important in helping me to revise and complete my thesis.

Finally, I would like to express my deepest gratitude to my thesis advisor, Professor Thomas Downes, for his tireless efforts and unwavering support over the past year. His level of commitment far surpassed my expectations of what a thesis advisor could possibly offer me. The amount of time and attention he was able to give made me feel like I was his only student. I am thankful for his good-humored nature and I truly appreciate all that I have learned under his mentorship.

Table of Contents

Abstract	2
Acknowledgments	3
Introduction	5
Literature Review	7
Data	
Wyoming.....	24
Kentucky.....	26
Methodology	
Wyoming.....	32
Kentucky.....	34
Results	
Wyoming.....	37
Kentucky.....	42
Conclusions	45
Tables	
1. Wyoming Summary Statistics.....	49
2. Kentucky Summary Statistics.....	51
3. Kentucky Questionnaire Summary Statistics.....	53
4. Wyoming Regression Results.....	55
5. Kentucky Fourth Grade Regression Results.....	57
References	58

From New York to Washington state, administrators have proposed cutbacks to other subjects as they struggle to meet the demands of state accountability programs and the federal No Child Left Behind Act.... Such decisions, exacerbated by budget cuts and a shortage of qualified teachers, have left less time and fewer resources for the other subjects in the core curriculum that are excluded from accountability measures.

- Education Week, 11/5/03

With the implementation of No Child Left Behind (NCLB) in 2002, statewide school accountability systems based on student testing have become ubiquitous throughout United States public schools. While there is considerable agreement amongst policy makers that the quality of American public schooling must improve, controversy arises over whether school accountability is an effective means of improvement. Accountability has been shown to lead to improved performance, but also there is evidence of a wide array of unintended effects produced by school accountability. Increased class time being spent on learning test-taking skills, suspending poor test takers during testing periods, even reshaping the testing pool by classifying poor test takers as disabled, all have been proven effects of school accountability. With the focus of NLCB on math, science and reading, attention and school resources will be shifted away from untested subjects. Yet we know very little about the impact of this change in focus.

Both Zastrow (2004) and The National Association of State Boards of Education (2003) show education in America is moving towards a more narrowed curriculum, one

that does not include an amount of instruction in the arts and music deemed sufficient by experts in the field. This research on trends in arts education, however, has used various surveys taken of students, teachers and principals as evidence. No research has been published using comprehensive non-survey data, and no papers whatsoever in the field of economics have yet been written exploring the effects of school accountability on arts education.

In this paper, a wealth of school and district level administrative data from two states, Wyoming and Kentucky, are used to quantify the effects of No Child Left Behind on arts education. Using data from Wyoming on the salary and number of arts and music teachers, I examine changes in staffing in the arts post-NCLB. Since Kentucky tests in the arts, I use these data to see how the trends in test scores have changed after the implementation of NCLB.¹ For each state I supplement these measures of arts provision with information on finances, staffing and demographics. By taking into account these other factors, I make sure that the changes in test scores and staffing are not being caused by something other than school accountability.

My results indicate that in Wyoming, where the arts are not included in the accountability system, provision of arts education has declined significantly in the aftermath of No Child Left Behind. Resources are reallocated away from the arts and toward subjects to which incentives have been attached, such as math and science.

In Kentucky, however, there is testing in the arts, upon which incentives are attached just as with any other subject. In this situation, my results illustrate that there is

¹ For this paper I assume that student test scores in the arts give an accurate measure of the quality of arts education. This is an arguable assumption. For example, as noted later in the paper, much literature has been written detailing unintended ways in which schools or teachers artificially inflate scores.

no measurable decline in provision of arts education relative to other subjects.² These results lead to the conclusion that a possible way to ensure that the growth of test based school accountability in America does not cause arts education to disappear is to include the arts as a tested subject and to attach to it the same incentives as with math, reading and science.

The following section of this paper provides a review of the literatures discussing both the effects of school accountability and the importance of quality arts education. This section ends with a brief discussion about the current state of arts education in America. The next section then supplies a more in depth look at the data from Wyoming and Kentucky and the recent educational reforms in these states. After this, we develop a model appropriate for analyzing the data and then provide the results generated from estimation of the parameters of this model. The paper ends with an evaluation of the results produced, final conclusions and areas for future research.

Literature Review

School accountability has recently been an incredibly hot topic in the media and amongst parents and educators. The centerpiece of the accountability movement is the No Child Left Behind Act (NCLB), federal legislation signed into law in January of 2002 that mandates that every state, to receive federal aid, must put into place a set of standards in math, English/language arts and science along with a testing plan to make

²As discussed in the Results section of this paper, given the cumulative nature of arts education, I would not be able to observe fully a shifting of resources away from arts education because it would take time to see the effects of NCLB show up on test scores.

sure that the standards are being met. Schools that continually fail to make adequate yearly progress toward these standards are subject to corrective actions. NCLB extends federal authority over state and local schools, essentially making school accountability a reality for every state in the U.S. (Rudalevige, 2003).

In NCLB there is no mention of testing in arts or music education. NCLB, however, does define the fine arts as a core subject and nearly every state currently has academic standards for visual and performing arts, as well as for music. Some states even have statewide arts assessments, though most are not included in the results for which schools are held accountable.

After outlining the basics of accountability and further discussing No Child Left Behind, this section will explore some issues relating to the design and implementation of school accountability systems. It will then go on to give some examples of systems that states have developed and assess their effectiveness. Next, some unintended effects of school accountability will be examined. This section will close with a discussion of arts education in America.

In “Sorting Out Accountability Systems” Eric Hanushek and Margaret Raymond (2002) explain the basics of school accountability. The authors describe the basic skeleton of accountability systems as involving four main parts: goals, standards, measurement and consequences. The start of any accountability system is to lay out goals for what the system is intended to accomplish. Usually the goals are rather general and ambiguous, for example “to ensure that all students are provided with high quality schooling that prepares them for entry into society.” The vague wording usually helps

secure legislative approval but can sometimes lead to problems involving ambiguity over who should do what.³

Standards are the next section in accountability systems and provide clear and precise details about what is expected (Hanushek and Raymond, 2002). The standards indicate the level of mastery expected of the students in specific subjects. One major innovation of the recent school reform movement is the shift to focus on setting of standards and the examination of “outcomes” or student learning, as opposed to inputs such as decreasing class size or increasing teacher salaries.

Standards-based reform is based on the notion that by setting high academic standards for student achievement and measuring performance, students and schools can be held accountable, and rewards and penalties can be utilized to change behavior and improve performance. Setting standards, though, is a complicated and difficult task. Standards should be high, of course, to encourage increases in learning, but setting them too high can be unrealistic and lead to failure in attaining the standards and poor student performance. Setting the standards too low, on the other hand, accomplishes nothing and is seen as not challenging.

The next part of an accountability system is measurement (Hanushek and Raymond, 2002). Measurement is the means that a system uses to gauge compliance with the standards. Much of the debate that takes place when a statewide accountability system is created concerns the tools to use for measurement and the positioning of the cut-off point for meeting standards. While there is disagreement over how to test performance, be it with multiple-choice questions, essays or evaluations of portfolios, most observers

³ Some common goals in the world of arts education are ensuring that students can adequately create and critique art and that they have an understanding of both art history and aesthetics (Dobbs, 1998).

agree that a successful measurement tool must directly measure and be closely aligned with the standards put forth by the system.

The final aspect of school accountability systems as described by Hanushek and Raymond (2002) is the consequences to schools and students for their ability to meet or not meet the expected level of competency. The very core of the standards and accountability movement is to align standards, teaching and the performance of students. The lever that induces this alignment is the issuing of consequences – both positive and negative. If schools or students expect no actions as a result of their performance there is little motivation to pay attention to outcomes. Consequences for students are usually negative; for example, failure to pass certain tests would result in not passing to the next grade or even graduating high school. Consequences to schools and teachers can be positive, in the form of bonus money and pay if the schools do well, or negative, in the form of taking away funds or taking over the school if it performs poorly.

The various accountability systems implemented in each state in response to No Child Left Behind follow the description of the prototypical system given by Hanushek and Raymond (2002).⁴ The goals of NCLB are intended to close the achievement gap between students of low and high-income families and to provide the opportunity for all children in America to learn and succeed. NCLB instructs states to develop their own standards and assessments. There must be academic standards and assessments for math and English/language arts for the third through eighth grades. There must also be academic standards and assessments for science in elementary, middle and high school. Assessments of English language proficiency are mandated as well as participation in the

⁴ Though testing in the arts falls outside the scope of NCLB, those states that choose to do so should follow this prototypical system closely. This will hopefully ensure an accurate and effective accountability system.

NAEP (National Assessment of Educational Progress) assessments for reading and math (The NCLB Act of 2001, Public Law 107-110).

The timeline imposed on the assessments is as follows: up to and including the 2004-2005 school year, states must administer annual assessments in math and English/language arts at least once during grades three through five, six through nine, and ten through twelve. Beginning no later than the 2005-2006 school year, states must administer those same tests in each of grades three through eight and at least once in grades ten through twelve. Beginning no later than the 2007-2008 school year, states must administer annual assessments in science at least once in grades three through five, six through nine, and ten through twelve (The NCLB Act of 2001, Public Law 107-110). Students in schools that fail these assessments or fail to make adequate yearly progress may transfer to another school in the same district. If a school fails persistently, it is subject to corrective action such as staff replacement or school restructuring. (Peterson and West, 2003).

Though No Child Left Behind is presently the major force behind current school accountability systems, its positioning in the history of school accountability makes it a natural step in the educational reform process. Since the beginning of American public schooling in the early nineteenth century, there have been limited types of accountability in public school systems in such forms as the elections of school boards. Until the 1970's, however, the focus was on "inputs", or resources, and schools were held accountable only to provide such inputs as qualified teachers, modern facilities and textbooks. Policymakers operated under the assumption that by allocating more money to schools, schooling would naturally improve (Cuban, 2004).

In the 1970's however, with the growing availability of student test scores, policymakers realized that despite consistent increases in spending per-pupil since the beginning of the century, student performance was, at best, stagnant (Betts, 1998).⁵ As more information became available confirming the poor performance of American schools, officials came under pressure to somehow improve student achievement. They argued that what mattered most was results. Thus, the attention of education reform was shifted to "outputs" or student achievement (Ravitch, 2002).

In the early 1980's expectations developed that schools should become accountable for their results (Ravitch, 2002). In 1983 a national commission appointed by the Secretary of Education released a report titled "A Nation at Risk" claiming the poor quality of American schools was making the country vulnerable to foreign competition (Peterson and West, 2003). This report helped to speed along the process of school accountability, but states were slow to implement systems. South Carolina implemented the first state system as part of a broader piece of legislature in 1984. By 1993, Georgia, Indiana, Kentucky, North Carolina and Texas all had accountability systems in place, while many others had systems developing. (Ladd and Clotfelter, 1996).

In 1994, President Clinton passed Goals 2000, federal legislation urging states to develop standards and assessment systems for students on a voluntary basis (Rudalevige, 2003). By the time No Child Left Behind was put into effect in 2002, nearly every state had academic standards and accountability systems already in place. In this regard, NCLB was just a natural step toward centralizing the efforts of the individual states. In addition, because states have been implementing school accountability systems for some

⁵ While the preponderance of evidence reveals that higher per-pupil expenditures in the form of smaller class size and more highly educated and trained teachers have had an insignificant effect on student learning, countervailing evidence does exist (Krueger and Whitmore, 2001).

time, there has been much literature written documenting various effects of school accountability.

Reliable accountability systems need to satisfy a number of requirements. For example, systems should look at different types of data, such as dropout rates, attendance, and student portfolios, and not only student achievement, when assessing a school. In order to ensure that each area gets proper attention, it may be necessary to test in all areas.⁶ All students should be tested and included in the results, even those more difficult to test, such as disabled or English deficient students. Of course, such students should receive the proper accommodations. If high school exit exams are utilized, students should be given multiple tries to pass the test. An appeals process should be available to those students who fail the test. Results of school testing should be made public in a timely fashion so that teachers can use it to help improve instruction (Baker and Linn, 2004).

Given a statewide school accountability system is well designed and efficient, there are still other things that must be done to ensure the effectiveness of the system in increasing student learning. School systems need to build capacity in terms of financial, material and human resources. Accountability operates on the principle that resources are available and that students, teachers and schools just need the incentive to work more efficiently. This is not always the case, especially in urban schools that often serve minority and low-income students. Funds need to be made available to assist these schools if the accountability system is to work effectively (Hamilton and Stecher, 2004).

⁶ There are tradeoffs to testing in all areas that policy makers may need to consider. For example, time spent on testing takes away from instructional time. My results indicate that the benefits of testing in all areas may outweigh the costs.

Funding also needs to be made available for professional development, which is crucial for the effectiveness of accountability. If teachers are to be expected to learn new curriculum and new ways of teaching, they must be properly trained. A great deal of research has shown that at present the conditions under which teachers work do not help to increase their performance or moral. Teachers work in isolation, get little meaningful feedback for their work, and are unsupported in their induction into the profession. Elementary teachers teach too many subjects and need legitimate expertise in each of them. High school teachers have too many students to get to know them all well enough. If accountability systems are to increase student learning, increased attention must be paid to teachers (Wasley, 2004).

Given these extensive requirements for effective school accountability systems, a common critique is that they are not worth the cost to the public. One of the most surprising and encouraging aspects of school accountability, however, is its cost. In “The Cost of Accountability”, Caroline Hoxby (2002) calculates the per-pupil cost of statewide accountability systems using data on both the revenue from the standardized testing industry and the expenditures of individual states. Using the revenue from the standardized testing industry she calculates a per-pupil cost of \$5.81. For the twenty-five states for which she calculated per-pupil costs, the range was from \$1.79 in South Carolina to \$34.02 in Delaware. The large variation in state costs is due to the varying depth of the systems. To put these costs into perspective, in the 2000-01 school year per-pupil spending was \$8,157 on average. In other words, the cost of accountability is less than one tenth of one percent of school spending.⁷

⁷ The per-pupil cost calculated from the revenue of the standardized testing industry is likely an underestimate because it does not include other costs of accountability besides paying for tests. Hoxby’s

Of course, a small price tag would be nothing to brag about if school accountability was ineffective as a means of improving student performance. There are a few states and districts that have had an accountability system in place for enough time now that its impact has been evaluated. Dallas, Texas implemented its accountability system in the 1991-92 school year. Ladd (1996) shows that, on average, in the years 1991 through 1995, students in Dallas significantly improved their reading and writing test scores on the Texas Assessment of Academic Skills (TAAS) compared to the state average. The author also finds significant decreases in the dropout rate compared to other large cities in Texas.

A more recent study (Deere and Strayer, 2001) of Texas schools shows that from 1994 through 1999, student performance increased sharply. The fraction of all students who passed the math test rose from 58% to 85% and the fraction of those passing the reading test increased from 74% to 86%. Over the same period the fraction of schools that received the “exemplary” rating, the highest available, rose from 1.3% to 11.7%. Similarly, the percentage of schools given the “recognized” rating, the second highest, rose from 13.1% to 36.8%.

The study proceeds to explore what caused the increases in scores from 1994 through 1999. More time and resources were allocated to those subjects included in the accountability ratings, such as math and reading, and less were allocated to those that were not, such as science and social studies. Also, more resources were focused on those

state by state calculations include other expenses such as running an accountability office or the cost of publishing report cards. The real cost of accountability, however, is not just its price but also the forgone activities that would receive more time and attention if student testing and accountability were not implemented.

students at or below the passing level, improving their performance the most, while neglecting the more advanced students (Deere and Strayer, 2001).

The results seen in Texas, where schools do whatever necessary to improve scores, often causing unintended negative effects, are rather common. Teaching to the test, narrowing curriculum, and increasing time allocated to test taking skills are all common side effects of accountability. One way that schools improve aggregate test scores and school ratings is by over-classifying low performing students into special education categories that are exempt from taking the test. Students of low socioeconomic status are more likely than other students to be reclassified. Placement into special education, when not necessary, negatively impacts the schooling experience for those students affected. It also increases costs for schools and reduces efficiency of allocation of funds (Figlio and Getzler, 2002).

Another common unintended effect of accountability is teachers cheating to improve their rating or that of their school's. Cheating can take many forms and has recently been documented in New York, Massachusetts, Texas and California. Filling in questions left blank, changing student's responses, allowing students extra time to complete tests, providing students with correct answers, and even illegally attaining copies of the test before test time and teaching specific problems are all examples of cheating. In a survey given to elementary school teachers, nearly ten percent reported they believed that teachers in their school "often" or "frequently" gave students answers. Another study, done in North Carolina, showed that 35 percent of teachers had witnessed

teacher cheating. Empirical work has shown that at least four to five percent of classrooms experience teacher cheating (Jacob and Levitt, 2002).⁸

One of the goals of current education reform and of No Child Left Behind is to decrease the learning gap between children of high and low socioeconomic status (SES). Yet school accountability systems often weaken already poorly performing schools. Many urban schools in communities serving low-income and minority students are desperately short on materials, facilities and qualified teachers. Under NCLB, when these students perform poorly and are then labeled as failing, the schools serving these students are subject to negative consequences such as replacement of staff, and eventually reconstitution or state takeover. These consequences often lead to higher qualified teachers looking to teach elsewhere, which only make matters worse. Many states are now also implementing high school exit exams that students must pass in order to receive a diploma. A higher proportion of low-income and minority students fail these exams and are thus at a disadvantage when leaving high school (Oakes, Blasi, and Rogers, 2004).

There are even more ways that schools unethically improve their test scores. Schools use discipline as a means of improving test scores by increasing suspension duration of low achieving students during test taking periods and decreasing duration for high achieving students. This ensures more low performing students will be absent on test taking days. This behavior has been found only to occur in grades with high stakes testing (Figlio, 2003).

Another rather peculiar way in which schools have been shown to artificially improve results is by altering school lunch menus to fit test taking schedules. The link

⁸ If the arts are to be brought into accountability systems, given the unavoidable ways in which schools and teachers attempt to inflate scores, it is crucial that standards be well developed and the test be properly aligned in order to minimize inaccuracies.

between nutrition and cognitive ability has been well established, and it has been found that schools give students significantly higher calorie lunches on test taking days to improve performance (Figlio and Winicki, 2002).

The Chicago Public Schools (ChiPS) demonstrated sharp increases in student performance due to its accountability system but also experienced many unintended and negative consequences. The ChiPS system utilized the Iowa Test of Basic Skills (ITBS), which it had been giving to students for several years but to which high stakes were attached only in 1996-97. Jacob (2002) shows that the performance in reading and math of Chicago students declined slightly in the early nineties and was then stagnant until 1996 when it rose dramatically. This increase was extremely significant when compared to other midwestern cities without accountability systems.

The Chicago Public School System demonstrated many of the unintended effects described above. Teachers were found cheating, and the number of students receiving special education in grades with high stakes testing grew dramatically beginning in 1997. Students entering grades with high stakes testing were retained at an increasing rate, thus excluding them from taking the test. Also, scores in tested subjects not included in the school's rating, such as science and social studies, did not improve nearly as fast as math and reading (Jacob, 2002). One of the most dramatic effects of the school accountability system in Chicago, however, was the drain of resources away from non-tested areas, *especially* art and music. From 1995 until 2000, the average spending on arts and music declined by twelve percent. In schools in the bottom quartile of performance, in which the incentives for improvement are the greatest, arts and music spending decreased by 21

percent (Jacob, 2003). These potential negative affects on arts and music education brought about by accountability and No Child Left Behind is the focus of this paper.

There are many reasons why reductions in the quality of arts education could be problematic. Music, visual and performing arts engage multiple skills and abilities, developing cognitive, personal and social growth. Creating art is a personal experience and students must draw upon their own creativity. This enables students to feel an emotional investment in their work. When students engage in the arts, they are motivated not by grades or test results, but by the learning experience itself. Art changes the dynamic of a learning environment and increases the enjoyment students receive from school. Students who are bored and misbehave in class often find comfort and learn better in arts classes (Fiske, 1999).

Numerous studies have shown that involvement in the arts drastically increases academic achievement. Raw SAT scores during 2000, 2001, and 2002 were approximately fifty points higher in verbal and thirty points higher in math for students taking classes in the arts (The College Board, 2000, 2001, and 2002). Caution must be taken when interpreting these results, however, because it is probable that those schools that offer extensive arts classes serve students of higher socioeconomic status who would tend to score higher on standardized tests naturally.

One study in which socioeconomic status was “reasonable-accounted for” followed 25,000 students for nearly ten years and found that amongst eighth graders, 67.3% of students highly involved in the arts scored in the top two quartiles of the standardized test distribution, while only 49.6% of those not involved in the arts achieved similar success. The results for tenth graders were 65.7% versus 47.5% and for twelfth

graders, 57.4% versus 39.3%. Despite the nominal differences staying roughly the same (about 18 percentage points), these numbers show that the percentage advantage for students involved in the arts increases over time. Eighth graders involved in the arts have a 36% advantage, tenth graders, 38% and twelfth graders, 46% (Catterall, Chapleau, and Iwanaga, 1999).

Not only does learning in the arts during elementary and secondary education improve academic performance, it also reduces the gap between students of high and low SES. The arts can reach students who are not otherwise being reached and can provide a reason for poor and minority students, often disengaged from and disenchanted with school, to become excited again. The above-mentioned study also shows that among low SES students, 37.7% of eighth graders highly involved in the arts scored in the top two quartiles of the standardized test distribution, while only 29.8% of those not involved in the arts did the same. The results for tenth graders were 35.2% versus 28.1% and for twelfth graders, 30.9% versus 23.4%.⁹ Far fewer low SES students have access to the arts outside school than do high SES students, which is why arts curriculum is essential in public schools. (Catterall, Chapleau, and Iwanaga, 1999).

The same study also finds a significant correlation between the study of music and high achievement in math. It shows that amongst high SES twelfth graders, 48% of those students heavily involved in instrumental music scored highly (level 4 or 5) on the math standardized tests used in the study, while only 38.6% of those with no instrumental music involvement scored the same. Amongst low SES students, the results were 33.1%

⁹ Though a comparison of the above numbers seems to suggest that high SES students would gain more from arts instruction than low SES students, this study shows that it is almost twice as likely for a high SES student to be highly involved in the arts as it is a low SES student. Thus, an increase in arts provision would increase performance for more low SES students than high, and the socioeconomic gap would decrease.

for those highly involved in instrumental music versus 15.5% for those with no instrumental music involvement (Catterall, Chapleau, and Iwanaga, 1999).¹⁰

In the Chicago Public Schools, where there is a very high percentage of low income and minority students, the Chicago Arts Partnership in Education (CAPE) was founded in 1992 as an effort to help increase the level of arts education. By 1998 more than 60% of sixth graders in schools participating in CAPE performed at grade level on the ITBS, while only 40% the remainder of sixth graders in Chicago performed at grade level.¹¹ Similar results were found amongst Chicago eighth graders. Surveys given to Chicago teachers showed that on scales for school climate, quality of relationships with parents and the community, professional development and instructional practices, CAPE schools outperformed non-CAPE schools in every category (Catterall and Waldorf, 1999).

Despite the evidence that arts education closes the socioeconomic gap, improves students' personal growth, attitude toward school and academic performance, time and resources are being diverted away from arts education. No Child Left Behind has focused the nation's attention toward math, reading and science, while fiscal stress has forced schools throughout the nation to make budget cuts. Speculation has existed for some time now that arts education in American public schools has been suffering, but it was not until recently that research was done to support these claims (Zastrow, 2004; Jacob, 2002).

¹⁰ There is an abundance of literature showing that improved performance in math leads to greater future earnings. Thus, an argument can be made that the results of this study indicate that involvement in instrumental music leads to greater future earnings.

¹¹ Caution is needed in interpreting these results because, while the CAPE schools are distributed throughout the city, it is possible that they were not selected at random.

A survey was given to 956 elementary and secondary school principals in the fall of 2003 asking them to report on the state of nine academic subjects such as reading, writing, math, science, and the arts. While the overwhelming majority of principals reported an increase in instructional time and professional development given to reading, writing and math, the outcome was much more bleak for the arts. Of all principals, 25% reported decreased instructional time given to the arts, while 33% anticipated future decreases.¹² Results were even worse for high minority schools, where 36% of principals reported decreases in instructional time, while 42% anticipated future decreases. Similar results were given indicating a decreasing number of arts teachers and reduction in professional development in the arts (Zastrow, 2004).

This is not the only survey to have been done examining the current state of arts education in America. The National Center for Education Statistics has implemented a nationwide series of “Fast Response Survey System” (FRSS) surveys, the first given in 1995 and the most recent in 2000. These surveys “provide a national profile of the status of arts education in the nation's regular public schools...” (<http://nces.ed.gov/surveys/frss/publications/2002131/>).¹³ They present information on the availability of instruction in the arts, funding and staffing.

In 1995 and 2000, music and visual arts were available at the vast majority of elementary schools (97% and 85% in 1995, and 94% and 87% in 2000, respectively). Of these schools, far fewer employed full time specialists in the field (70% and 43% in 1995, and 72% and 55% in 2000, respectively). Secondary schools produced similar results. For

¹² In comparison, only 8% of all principals reported increased instructional time in the arts, and only 7% anticipated future increases.

¹³ Regular school is defined as a public elementary/secondary school that does not focus primarily on vocational, special, or alternative education.

both elementary and secondary schools, dance and theater was offered substantially less frequently than music and visual arts. The 2000 survey also provides percentages for schools that serve high minority populations. In 2000, 91% of elementary schools that had a minority enrolment greater than or equal to 50% offered music and 81% offered visual arts. Thus, prior to NCLB, a student attending an elementary school with a significant minority enrolment was only slightly less likely to have the opportunity to take music or visual arts than was a student in the average elementary school (<http://nces.ed.gov/surveys/frss/publications/2002131/>).

As illustrated by the comparison between the 1995 and 2000 FRSS surveys, there was not much change in provision of arts education in the late nineties. Unlike the Zastrow survey given in the fall of 2003, neither of the FRSS surveys seems particularly discouraging. This seems to support the argument that it is No Child Left Behind, enacted in 2002, that is causing arts education to suffer in all schools and suffer particularly in schools serving significant minority populations. Although logical, this can not be assumed simply from the Zastrow survey or others like it. They do illustrate decreases in arts provision, but they do not differentiate the effects of NCLB from the fiscal effects of a recession and shrinking state budgets. In order to get an accurate picture of the effects of NCLB on arts and music education, the fiscal effects of budgetary strain must be isolated and separated from that of No Child Left Behind.

DATA

Wyoming

For Wyoming, the data analyzed in this paper begin with the 1998-99 school year and continuing through the 2003-04 school year. Major school reform was being put into effect in Wyoming at the start of this period. The impetus for these changes was a court decision resulting from a suit brought in October of 1993 by four of Wyoming's larger school districts and the Wyoming Education Association, alleging that the state's public school finance system was unconstitutional. An initial ruling by the district court partially upheld the constitutionality of the finance system, but in November of 1995, the Wyoming Supreme Court unanimously ruled that the entire state system for funding public schools was unconstitutional.

Major school reform legislation was enacted in June of 1997, but because of insufficient time before the upcoming school year, a large share of the reform was not implemented until the 1998-99 school year, the first year that this paper examines. Along with a new and equitable finance system, the reform called for sufficient building facilities, improved access to technology, and a statewide student assessment system. The Wyoming Comprehensive Assessment System (WyCAS), which tests reading, writing and math in grades four, eight and eleven was administered for the first time in April of 1999 (Christenson, 2001).

Legislative action caused further reforms to be slowly put into practice over the next couple years, smoothing out the implementation of the new finance model and addressing such issues as teachers' salaries and maintenance costs. Beginning in 2002, the State Board of Education commenced work on compliance with the new federal

requirements created by No Child Left Behind. Though the WyCAS was a good start, NCLB requires mandatory testing in grades 3-8 and school improvement programs through rewards and sanctions, neither of which Wyoming had at the time (2002 Wyoming Joint Education Interim Committee).

In order to examine the state of arts education before and after the implementation of No Child Left Behind, this paper utilizes panel data from 1998 until 2003. All of the Wyoming data analyzed in this paper were obtained from the Wyoming Department of Education website, specifically the Reports and Data section.¹⁴ Within the Reports and Data section, the data on WyCAS test results were found in the Historical WyCAS section. Much of the financial, demographic and nonacademic information was retrieved from the statistical reports series, while the remaining data were found in the school district profiles.

Whereas the majority of the Wyoming data reaches to the 2003-2004 school year, some of the financial data unfortunately were only available up to the 2002-2003 school year, limiting the regressions involving those data to 1998-2002. Similarly, our data on special education counts were only available back until the 1999-2000 school year, not 1998-1999 like the majority of our data, limiting its use as well.

The WyCAS is given in grades four, eight and eleven and so these are the grades for which I am using data. Information on full time equivalent staffing levels are available at the secondary school level only. Thus, any estimated effects on staffing in the arts hold only for high school staff and do not necessarily extend to elementary or middle

¹⁴ <http://www.k12.wy.us>

school.¹⁵ The other secondary staff variables represent base and total salary. Base salary is the number of dollars per year that a teacher or staff member earns. Total salary is an estimate of the true value of the salary including all benefits such as health insurance or money that goes toward retirement.

The Wyoming summary statistics can give a partial glimpse at how provision of the arts might be changing due to the implementation of No Child Left Behind. As seen in Table 1, the mean full time equivalent number of teachers in the fine arts fluctuates over the period of analysis, ending slightly higher. Language arts seems to increase more rapidly than the fine arts, and both math and science see a huge jumps, dwarfing the gains of the other two. Clear trends upward in math FTE and total expenditure per-pupil, and downward in students-per-teacher, all point toward more resources coming in throughout the time period.

Kentucky

Kentucky is the second source of state level data. A Kentucky Supreme Court decision in 1989 ruled that the entire state's education system was unconstitutional. This led to the passage of the Kentucky Education Reform Act of 1990 (KERA), completely overhauling the curriculum, organization and financing of K-12 education in Kentucky. This major school reform was enacted seven years earlier than the reform in Wyoming,

¹⁵ If in Wyoming, in a post NCLB period, resources in the arts were shifted from away from grades 9-12 toward K-8, then I would be overstating the impact of NCLB on arts provision because I would be mistakenly seeing a reallocation of resources as a reduction. Given the information in the surveys cited earlier, I think this situation is very unlikely. In fact, it's also possible I could be understating the effects of NCLB because I am not taking parallel changes in K-8 arts provision into account. As is true in any subject, education in the arts is cumulative and if resources are taken away in grades K-8, learning might be less effective at the secondary level.

so whereas in Wyoming the major changes in public education policy were implemented just before our period of analysis begins, the changes in Kentucky were implemented well in advance.

Reforms to the education finance system were phased in over five years beginning in 1990 and have not been significantly altered since. By 1996, a set of detailed curriculum had been published and school governance had been reformed such that school-based decision making councils were in place to set school policies. The Kentucky Instruction Results Information System (KIRIS), Kentucky's first accountability system since the passage of KERA, was put into place for the 1991-92 school year as a temporary system. KIRIS broke new ground in holding schools accountable by imparting monetary rewards and penalties based on changes in a single accountability index, much like the accountability system used by Kentucky today (Flanagan and Murray, 2004).

To accompany the KIRIS, Kentucky introduced the Comprehensive Test of Basic Skills (CTBS-5) in 1995, which tests students in grades three, six and nine in reading, language arts and math. In 1998, the first year of our period of analysis, the Commonwealth Accountability Testing System (CATS), Kentucky's current accountability system, was put into place. The CATS is very similar to previous accountability systems and utilizes both the CTBS-5 and the newly designed Kentucky Core Content Test (KCCT) to measure students' performance. The KCCT tests students in reading, writing, mathematics, science, social studies, arts and humanities, practical living and vocational studies.

The KCCT is given once in elementary school, once in middle school and once in high school. Because there are so many subjects being tested, though, they are split up so that all the subjects are tested in grades 4 or 5, 7 or 8, and 10 or 11 (except for writing which gets tested in grade 12).

For each school, percentages of students scoring at determined proficiency levels in all of these subject areas is combined with scores on the CTBS-5 as well as with nonacademic indicators, such as a school's attendance, dropout and retention rates, and the post-graduation experience of its students to create the Kentucky Accountability Index.¹⁶ Scores on this index are reported annually, but schools are evaluated over a two-year period. Each school is given a performance target and schools that exceed this target receive reward money. Schools falling short of their target receive a scholastic audit, assistance from a distinguished educator and eligibility for improvement funds (KY Dept. of Education Website, 2005).

As part of the KCCT testing, a questionnaire is given for each subject area asking questions about students' experiences both learning in that area and taking the respective KCCT subject test. There is a broad range of questions addressing time spent on specific activities or using specific equipment, and attitudes towards how well or how hard the student tried on the test. Students' answers to these questions do not affect their score on the KCCT or the index of the school.

¹⁶ Each subject area on the KCCT is given a specific weight on the academic index. The weighting formula varies slightly depending on what grade is being tested, but in general reading and math are weighted the heaviest, science, social studies and writing are in the middle, and arts/humanities and practical living/vocational studies are weighted the lightest.

All of the Kentucky data analyzed in this paper were drawn from the Kentucky Department of Education Website.¹⁷ A large portion of the data was found in the Reports section under Testing and Reporting. The majority of the nonacademic data were found in the nonacademic data section of Reports. Most of the test results from the KCCT and the CTBS-5 were found in their respective sections under Reports. The remainder of the data was obtained from Max, the Kentucky Department of Education's data warehouse.

Most of the Kentucky data were available from 1998-2003. I have CTBS-5 test results back through the 1996-1997 school year. Having the CTBS-5 results for 1996 and 1997 allows me to use them as a pretest, creating a picture of where the students are before they enter this paper's period of analysis. Other data, unfortunately, such as demographics and financial information were only available from 2000 until 2002 or from 2000 until 2003. The limited span of these data limits their usefulness in my analysis.

I only have questionnaire data for one year, 2002, so I regretfully cannot do any regression analysis on this data. I do, however, provide the results of a few questions in Table 3. The first question I look at is asked in all the subject areas except for writing. It asks the students how many of the questions on the test tested things they learned in school. This question seems to address the degree to which the test is properly aligned, an important requirement for a successful accountability system. If the test is testing strictly what is being taught in school, then it is well aligned. If this is not the case, then either the test is poorly aligned, or the test is aligned with the proper standards and curriculum, but teachers are teaching the wrong material. Either way, the validity of the test results would have to be questioned.

¹⁷ <http://www.education.ky.gov>

By looking at Table 3, it becomes very clear that the arts and humanities test is least aligned of all the tests. In the high school and middle school KCCT exam, the arts questionnaire has the most preferential answers, C and D, filled in the least of any other subject. In the elementary KCCT exam, only the reading test fares worse than the arts. This puts into question not only the validity of the test results of scores on the arts section of the KCCT, but more specific to this paper, the accuracy of using the arts score on the KCCT as a indicator of arts provision. If the test is not properly aligned, then as arts scores trend up or down, learning in the arts might now necessarily follow. Obviously, these data are supplied by students and very subjective. These observations based on the results seen in Table 3 should be considered accordingly.

The next question in Table 3 is only on the arts questionnaire and addresses time spent on arts education. The question asks the student, during a typical workweek, how much time do you spend on arts and humanities? Based on the numbers seen in the table, it seems that in elementary school, though it is not likely that a student will receive an inordinately high amount of arts instruction, it is even more certain that students will at least receive some level of arts education. Only 3.42 percent of elementary student claim to spend no time at all, and this number climbs through middle school and to high school, where 22.8 percent of students spend no time on the arts. At the same time, there is a greater number of students spending over 4 hours on arts. This shows that whereas in elementary school and, to an extent, middle school, students are assured of getting a foundation in the arts, in high school arts education is left up more to the interest of the students.¹⁸

¹⁸ The average amount of time spent on arts education seen in this table is relatively consistent with previous findings from the national surveys cited earlier in this paper.

The last question I look at and report in Table 3 asks students, if, during the school year, they had the opportunity to perform/create in each of the areas of dance, drama/theatre, music, visual arts, and literature. Students respond yes the majority of the time, but decreasingly so as they go from primary up to secondary schooling. The final two parts of Table 3 yield the general impression that provision of arts education might be stronger in earlier years and thought of as less crucial during high school.

As with Wyoming, many of the variables in the Kentucky data are self-explanatory. A few variables, especially the test score indices, are worthy of explanation in some detail. The eight KCCT index variables give school and district scores on the seven subject areas of the KCCT and also a total academic index, which is the average of the all seven subject scores. Each of these scores is out of one hundred, and every school is expected to reach a perfect score on every index by the year 2014.

The other test score variables are for the CTBS-5, the national norm referenced test given in Kentucky since 1995. There are eight CTBS-5 test score variables, four of which report the mean score of the school or district in reading, language, math and total score. These scores range from 1.0 to 99.0. The other four CTBS-5 test score variables report the national percentile that the school or district's mean score places them into, also ranging from 1.0 to 99.0. There is also a variable indicating reasons why a school did not administer tests during a specific testing period.

The wealth variable is simply local taxes collected divided by the levied equivalent tax rate. Because the local taxes data were available for only years 2000-2002, the wealth variable is also only available for that period.

Methodology

Wyoming

Using the data from the Wyoming Department of Education, I am seeking to explore if the provision of arts education changes after the implementation of No Child Left Behind, and if so, in what manner. I do this by running regressions on the arts provision variables using the various academic, demographic and financial variables, as well as year dummy variables, as regressors. I create an NCLB variable that is effectively a year dummy for both 2002 and 2003, the years included in the period of my analysis in which NCLB began to take effect in schools. By examining the coefficient on the NCLB regressor, I can see what effect NCLB is having on the quantity of arts and music teachers in Wyoming public schools, as well as on their salaries.

My review of the literature indicates that the concentration of high stakes accountability systems on certain subjects, such as math, science and reading, will draw resources away from untested subjects or tested subjects upon which high stakes are not attached. The question I want to ask in Wyoming is, holding overall staffing levels constant, what is happening to staffing in the arts. I must develop a model that finds these effects but controls for staffing levels and staffing mix. To keep staffing levels constant I include a pupil per-teacher variable. I use several variables to control for staffing mix including students' ethnicity percentages and assessed value per-pupil, a measure of community wealth.

It is not enough to find the effects of NCLB on the absolute level of resources allocated to arts provision. I must also find its effects on the provision of the other areas of study and compare those effects to those on the arts. For example, if the regressions

showed NCLB to have had no effect on arts provision but increased that of every other subject, then provision of the arts would have been reduced relative to other subjects. I ran regressions using as the dependant variable full time equivalent (FTE) staffing, and subject salary per pupil for each of the eleven academic areas. The *subject salary per pupil* variable was created by taking the product of the FTE and total salary per FTE divided by average high school enrolment. This variable takes into account both quantity and salary of staff. In all of my regressions, I use the logs of those variables representing financial data, such as the subject salary per pupil variables, assessed value, and expenditures.

I use a standard fixed effects regression model of the form,

$$y_{it} = \alpha_i + \tau_t + \beta x_{it} + \varepsilon_{it}$$

a model often used when doing regressions involving panel or longitudinal data. Here α_i is a district-specific effect and τ_t is a year-effect common across all districts. Panel data sets, such as the ones I use, consist of large numbers of cross sectional units observed at a few points in time. Panel data provide a well-suited environment for controlling for endogeneity problems that might be involved in a cross-section setting.

Issues of endogeneity arise when there is something unobservable correlated with the regressors and affecting the dependent variable in a regression. If not dealt with properly, this endogenous error term skews the results of a regression. If the unobservable factor causing endogeneity is stable over time (or at least throughout the period of the analysis), then the fixed effects regression model corrects the endogeneity

and produces an accurate regression. In the above model, the unobservable characteristic causing endogeneity is represented by α_i .

One example of a characteristic that could lead to endogeneity problems in my analysis is the tastes of the people living in each school district. Assuming a Tiebout world where people with the same preferences live in the same communities, their preferences in regards to education will influence provision of staffing, tests scores and other regressors in my model. Tastes are effectively unobservable and although certain measurable factors such as income could conceivably be worthy proxies for tastes, they may not fully account for unobserved tastes. Also, it seems reasonable to assume that the tastes of a school district will not change in the relatively short period of six years that this paper examines. Since tastes are unobservable and time stable, the fixed effects regression model is effective at correcting the endogeneity error.

Kentucky

I use a similar but somewhat more complicated methodology in Kentucky. As in Wyoming, I use panel data ranging across a number of years for my analysis. For the reasons described above, I continue to use a fixed effects model as the basis of my regressions. This resolves the potential endogeneity problems for the Kentucky data, as it did for Wyoming. Since I use the same model and type of data, the regressions for Kentucky will at first glance look very similar to those in Wyoming. Because I am looking at test scores and not staffing, however, the methodology I use for Kentucky is quite different.

For my analysis on Wyoming, I attempt to discover what is happening to staffing in the arts, holding overall staffing levels constant. My methodology in doing this, described above, involves including variables that isolate the effects of NCLB on staffing in the arts. In Kentucky, my interest lies not in staffing, but in test scores. The question I want to answer in Kentucky is how has student performance in the arts changed in comparison to other subjects as a result of NCLB. I answer this using both results from the KCCT, the test that includes arts, and the CTBS-5. The KCCT is taken in grades 4, 8 and 11, and the CTBS-5 is taken in grades 3, 6 and 9. Since I have the CTBS-5 data ranging back two years before the KCCT data, I can use the CTBS-5 scores in conjunction with the KCCT scores as a means of showing where the students' level of achievement was before the KCCT exams and also the *value added* by the years of schooling between KCCT exams.

Value-added models of determining school achievement give a much more accurate picture than just observing how well a school scores on a particular exam. They capture the influence of student and family characteristics that may effect academic achievement and control for variation in schooling quality in years prior to the pre-test. Methodologies that use value-added techniques look at the performance of a specific cohort of students before and after the grade in which they are tested. The inclusion of the performance before entering the grade is the key to the value added method. By including a pre-test score as a control variable, an accurate picture of how much achievement the school actually contributed during that grade can be portrayed.

An example of why value-added techniques are essential to include in school achievement and school accountability indices is that a school with high test scores may

not be a better school than one with low test scores. A mediocre school with ineffective teachers in an affluent suburban neighborhood may consistently score higher on accountability exams than a better school with experienced teachers in an impoverished inner city neighborhood simply because of the demographics of the student body. The more advantaged suburban youth will have had certain benefits, such as their parents reading to them before bed or having more resources outside of school, to which the poorer children in the inner city never had exposure. Value-added measurements would show that the school in the inner city is in fact the better school.

I develop my value-added model by merging the CTBS-5 data with the KCCT data, so that every cohort of students has a CTBS-5 score and a KCCT score that applies to them. For example, fourth graders taking the KCCT in 1998 took the CTBS-5 as third graders in 1997. I took the 1997 third grade scores and assigned it to that cohort. Eighth graders taking the KCCT in 2000 took the CTBS-5 as sixth graders in 1998, so I assigned the 1998 sixth grade CTBS-5 scores to that cohort of students. In 2003, eleventh graders taking the KCCT took the CTBS-5 as ninth graders in 2001, and I used the same method with this cohort.

I run regressions for all of the seven areas tested for each grade - four, eight and eleven. Whereas the KCCT tests in seven academic areas, the CTBS-5 tests in only three so I include the CTBS-5 test score variable that most accurately represents the KCCT subject being tested. The math and science indices are matched with the CTBS-5 math score. The reading and writing indices are matched with the reading and writing scores from the CTBS-5. The arts and humanities, social studies and practical living indices are

all matched with the total score variable of the CTBS-5, which is an average of the three subject scores.

Among the regressors I include are student ethnicities, certified staff per-pupil, and revenue and expenditure information. I also run a parallel set of regressions that include the wealth variable but only produce results for the period 2000-2002 because of the limited range of the wealth variable. I do not report the results of these regressions because of the limited length of the period of analysis.

Results

Wyoming

Table 4 represents the results from the Wyoming FTE staffing and subject salary regressions. There are three fixed effects regressions for each group of dependent variables, each adding more independent variables.

In the staffing regressions, the dependent variables for the four regressions reported from each set are the number of full time equivalent staff per-pupil in fine and performing arts, language arts, math and science.¹⁹ In the first and most basic set of regressions, the coefficient on the No Child Left Behind variable in the fine and performing arts regression is statistically insignificant, indicating that NCLB has had no effect on the provision of staffing in the arts in Wyoming. The only variable that has a

¹⁹ Results for other subject areas are available on request. In other subject areas well outside the focus of No Child Left Behind, the coefficient of the NCLB variable is generally negative.

statistically significant effect on staffing in the arts for this regression is the percentage of students in the district who are Native American.

The coefficient for the percent of students who are Native American in this regression is 0.0009799. This means that, all else being equal, districts with a higher Native American population will have a significantly higher provision of staffing in the arts. If, for example, a district happens to have the mean percentage of Native American students for 2003-04 of nine percent, then the number of full time equivalent staff per-pupil in the arts would be 0.0088191 higher for this district than a district with no Native American student population. This translates into a 69 percent higher provision of staffing in the arts in the district with the larger Native American population.

The likely reason that districts with higher Native American populations see a greater provision of staffing in the arts is that districts that serve large populations of Native American students receive impact aid. Impact aid is aid from the federal government awarded to school districts that are affected by federal activities and often serve low-income and minority student populations. Areas with sizeable Native American populations are considered to be affected by federal activities and thus qualify for impact aid. It is often the case that state funding formulas for school districts do not take impact aid into account. Therefore, these districts, though usually rather poor, have extra funds allocated to them from which it appears they spend a fair share on staffing in the arts.

The No Child Left Behind variable, though insignificant in the fine and performing arts and language arts regressions, is statistically positive for the math and science regressions. This is saying that the passage of NCLB has increased provision of

staffing in math and science. The likely reason for this result is that these two subjects are two of the three subjects included as part of the NCLB legislation. Schools that do not reach proficiency in these subjects are subject to negative consequences such as staff replacement or school restructuring. As confirmed by the results of this regression, the existence of NCLB provides incentive for school districts to allocate more resources toward the subjects upon which high stakes are attached, such as math and science. Districts accomplish this by increasing staffing levels for those subjects.

The second set of FTE staffing regressions reported in Table 4 includes the average teacher experience and pupils per-teacher independent variables. The pupils per-teacher variable is a measure of overall staffing levels and including this variable controls for fluctuations in fiscal status and accounts for fiscal constraint.²⁰ Districts with higher pupil-teacher ratios have fewer resources. The coefficient on the pupils per-teacher variable is negative across all subject areas and is significant for nearly all, including fine and performing arts, language arts, math and science. This implies that, as expected, districts with more resources, and thus a higher teacher/pupil ratio, will have higher staffing in all areas. Put another way, increases in overall staffing are applied across all subject areas.

In the second set of regressions, which includes the pupils per-teacher variable, the coefficient on the No Child Left Behind variable has become statistically significant for the fine and performing arts regression. The coefficient on the NCLB variable for this regression is -0.0016789, revealing that once overall staffing levels are held constant, we see that NCLB has had a significantly negative effect on the provision of arts education

²⁰ I have run regressions including other variables such as total expenditure per-pupil and total revenue per-pupil that, in addition to the pupils per-teacher variable, account for fluctuations in fiscal status. The differences were minimal between those regressions and the regressions including only pupils-per-teacher.

in Wyoming. The NCLB coefficient shows that NCLB has caused the number of full time equivalent staff per-pupil in the arts to go down by 0.0016789. When comparing this decrease to the mean number of full time equivalent fine and performing arts staff per-pupil in 2001 of 0.0129, we see that holding all else equal, No Child Left Behind has caused a 13.01 percent decrease in provision of staffing in the arts.

By looking at the same numbers for math, we see contrasting results. The coefficient on the NCLB variable in the math regression is 0.0017873, implying that No Child Left Behind has caused an increase in the number of full time equivalent staff per-pupil in math by 0.0017873. Using the mean full time equivalent math staff per-pupil in 2001 of 0.0135, we see a 13.24 percent increase in staffing in math caused by NCLB. These estimated effects are consistent with the conclusions drawn from earlier papers cited in the literature review (Jacob, 2003 and Deere and Strayer, 2001) that the focus of incentives on certain academic subject areas draws resources away from those subjects that are untested or upon which incentives are not attached.

The average teacher experience variable has a significantly negative coefficient for both the math and science regressions. These results are not surprising as there has always been more turnover in math and science than in other subjects. This occurs because math and science teachers have more employment options outside of teaching (Hanushek, Kain and Rivkin, 2004). This higher turnover rate means more rapid hiring of new teachers, which brings down the mean level of teacher experience for the district.

The coefficients on the variable representing the percentage of students who qualify for free or reduced lunch are also significant for the math and science regressions, but in this case are both positive. An interpretation of these results is that poorer areas

that are likely to be at a higher risk of failing to reach proficiency levels in tested subject areas will reallocate more toward tested areas as an attempt to avoid sanctions or censure. This interpretation would be consistent with conclusions drawn from Zastrow (2004) that the schools with less resources and higher percentages of minority students are the schools where the arts suffer the most.

The third and final set of FTE staffing regression results included in Table 4 incorporates the WyCAS proficiency variables for reading, writing and math. The results are extremely similar to that of the second set of regressions.

The last set of regression results reported in Table 4 is for the subject salary per pupil variables. For this regression, the coefficient on the NCLB variable for the fine arts is now higher than language arts (0.2439369 and 0.1953941 respectively) but still lower than both math and science. The results are less stark because adding the salary information to the FTE staffing dampens the effects of NCLB. Even though staffing in the fine arts in Wyoming is being reduced, the teachers that stay are the ones with the higher salaries. As the current art and music teachers rise up the salary schedule, and without new ones with low salaries to bring down the mean, the average salaries climb higher. Given the nature of salary schedules and the existence of tenure, it is much harder to reduce salaries than it is to reduce quantity of staffing levels.

Kentucky

Table 5 represents the results from the fixed effects regressions run on the Kentucky data for grade four.²¹ The dependent variables reported from each regression are the district level KCCT score for each of the tested subjects. Besides year dummies and the NCLB variable, the only independent variables in the first set of regressions are the district level CTBS-5 score for reading, language and math. The coefficients on the NCLB variable are positive and significant for all subject areas, including arts and humanities, implying that the implementation of No Child Left Behind has increased the performance of students in Kentucky across all tested areas.

In all likelihood, these increases in test scores were being driven not by No Child Left Behind but by the incentives put into place by the Commonwealth Accountability Testing System (CATS) of Kentucky in 1998. Because test results are based on cumulative experiences, it takes time for adjustments in incentives to show up in test scores. Thus, the most likely explanation of this result is that the coefficient on the NCLB variable is not showing the effects of NCLB, but rather the effects of the CATS implemented four years prior. Similar caution is not needed in the interpretation of the Wyoming results because staffing levels, my indicator of arts provision in Wyoming, would respond much quicker to a shifting of incentives than would test results.

The probable explanation for the seemingly contradictory estimates of the impact of accountability on the arts between Wyoming and Kentucky lies in the design of the CATS. The CATS is structured so that the scores from all subject areas, including the

²¹ The results were similar for the regression results for grades 8 and 11. There were small differences in magnitude of the coefficient on the NCLB dummy for the different grades. The coefficients of the NCLB variable in the grade 8 regressions were slightly smaller and in grade 11, were moderately larger. There were also some differences between grades in magnitude of the coefficients on the CTBS-5 variables.

arts, are averaged into a single accountability index that determines whether schools either receive rewards or get audited. Thus in Kentucky, unlike in Wyoming, incentives are attached to many subjects including arts and humanities. Because the incentives apply to all areas and do not just focus on a few specific academic subjects such as math and reading, resources are not reallocated away from the arts. Referring back to Table 2, a possible explanation of why the coefficient on the NCLB variable in the arts and humanities regression is actually higher than other subjects is that the mean arts score of 40.7116 in 1998 is by far the lowest of the KCCT scores. This leaves more room for improvement than in other subjects, and schools might see the arts as an easy opportunity to raise their accountability index.

The coefficient on the NCLB variable for the reading regression is another surprising result in that it is significant only at the ten percent level, as opposed to the one percent level seen for the regressions on every other subject. This could possibly be explained using the same logic as above. Notice that the mean reading score in 1998 of 78.8808 for the KCCT is clearly the highest of any subject. This leaves less room for improvement than in other subjects. As of 2003, the scores of every subject, having all improved but at different rates, are clustered more tightly together, showing that the rate of improvement was less for reading and more for arts than for other subjects.

The coefficients on the CTBS-5 variables are also all positive but only significant for the reading, language and math variables. For those subjects where I use the total combined CTBS-5 score as a control since reading, language arts or math did not seem appropriate as a pretest, the coefficients on the CTBS-5 scores are insignificant. It would make sense that if a district scored well on the pre-test then it would also score well on

the KCCT, leading to a positive and significant coefficient for the CTBS-5 variable. It therefore seems possible that the combined CTBS-5 score is not an appropriate pre-test for the arts and humanities, social studies and practical living sections of the KCCT.

The second and final set of regressions reported includes a teacher per-pupil variable, a special education percentage variable, and a classroom staff and superintendent salary variable.²² The inclusion of these has little effect on the coefficients on the independent variables from the first set of regressions. One difference is that the coefficient on the NCLB dummy in the regression for KCCT reading is now completely insignificant, exacerbating the irregularity between the reading and other subject scores. As with the Wyoming regressions, the teacher per-pupil variable is a measure of overall staffing levels, and its inclusion controls for fluctuations in fiscal constraint.

The coefficient on the teacher per-pupil variable for the arts and humanities regression is 184.8544 and significant. This means that all else being equal, districts with a higher teacher-pupil ratio will perform better in the arts. For example, compared to a district that has the mean level of teachers per-pupil in 2001 of 0.0773, a district with 0.0847 teachers per-pupil, one standard deviation higher, would score 1.37 points higher on the arts section of the KCCT. Because districts with higher teacher/pupil ratios have more resources, it would be expected that performance in all subjects would improve as the number of teachers per-pupil increases. The only regression that produces a positive significant coefficient for the teacher per-pupil variable, however, is the arts and humanities regression. This may imply that, as opposed to the core subjects that qualify

²² I have run a third set of regressions including some variables on revenue and expenses. These regressions did not significantly alter the results estimated in the first two sets of regressions.

as necessities, the arts are somewhat of a luxury that, similar to with an income effect, respond dramatically to increases in resources.

Conclusions

I have estimated the effects of accountability on the provision of art and music education in Wyoming and Kentucky using a wealth of scholastic, demographic and financial data. My results indicate that in Wyoming, where the arts are not included in student testing, resources are reallocated away from the arts and toward subjects to which incentives are attached. This leads to a decline in the provision of arts education. These estimated results are consistent with the findings of previous survey based papers written in the field of education. As the full effects of No Child Left Behind continue to take effect over the next decade and states scramble to reach proficiency levels, this negative trend in the provision of arts education threatens only to get worse.

While the results estimated in Wyoming support my hypothesis that accountability is having a harmful effect on arts and music education, in Kentucky I find different, more optimistic results. In Kentucky, where they test in the arts and attach incentives for schools to increase student performance in the arts, my results show no change in the provision of arts education relative to other academic areas in the years after the implementation of No Child Left Behind. This is likely a result not of NCLB but of the accountability system implemented in Kentucky previously. Because the federal legislation of NCLB does not mandate that states test in the arts, the vast majority of

states do not do so. Kentucky goes beyond the norm, including arts education in their accountability system, thus ensuring it will not be neglected.

If linking arts with testing provides an incentive for improved performance in the arts, then it makes the arts less susceptible to being hurt disproportionately compared to other subjects in periods of economic downturn or fiscal constraint. Advocates of strong arts education would greatly prefer the arts to move away from resembling a luxury good, and more toward being a necessity. As long as there is a large income effect on arts education, arts provision will always be at the mercy of the cyclical rise and fall of K-12 resources.

The results seen in Kentucky provide a model for one possible solution to the national decline in arts education caused by No Child Left Behind. If art and music in America's public school system are to receive the time and resources they deserve, states must include them in their accountability systems. They must attach incentives as they do currently with math, reading and science. This gives incentive to the school districts to maintain quality arts education because they will be held accountable for the performance of their students.

It might be worth asking, however, if the cause of the system in Kentucky working out well and upholding arts education was because there were easy gains in arts scores available. As scores in the arts catch up with those in other subjects and these easy gains are exhausted, it seems probable that incentives would shift towards math, science and reading, subjects to which incentives are attached not just through the state accountability system, but the federal one as well. Thus, individual statewide accountability systems that include arts and music might not be enough to maintain

adequate provision of arts education. A more permanent solution would be to create a federal accountability system, similar to that in Kentucky, which would include arts and music. This would eliminate contradictory incentives between state and federal school accountability systems.

Given my results, further research should be done to verify that testing in the arts and attaching incentives to student performance would help to stop the decline of arts provision. In this paper I simply observe the results estimated in Kentucky and reason that they are caused by its unique accountability model. Work should be done examining the effects of including arts and music or other typically untested subjects in accountability systems. Also, it is important that research be carried out analyzing the efficiency of standardized testing and accountability in the arts, examining if current tests are effective in assessing a student's knowledge of the arts, or if critiques of portfolio work, or other methods of assessment would be more effective. If arts are to be included in accountability systems, we must first make sure that the ways in which we currently assess and report student's knowledge of the arts are accurate and meaningful.

In this paper I have presented my findings that when the arts are not included in an accountability system, resources are redirected elsewhere. I have not attempted to make the normative argument that this is a bad thing. A clear distinction needs to be made between the two claims. Further research would have to be done exploring the normative argument, and would have to rest on whether the loss that students experience due to the decline in arts education is greater than the gain they experience from a concentration in resources elsewhere.

For example, suppose hypothetically that earnings, employment and post-high school educational prospects depend solely on math and reading scores. In this case, a decline in arts education may not make students worse off if the resources taken away from arts are directed toward math and reading. Earlier in this paper I cited research showing that arts education increases academic performance. Further work would have to be done to see if resources spent on arts education would increase math and reading scores to the same extent as an equal amount spent directly on math and reading. It seems unlikely that this would be the case. By these standards students would be better off with reduced arts education.

The normative argument rests on the costs and benefits of a reduction in arts education as a whole, which future earnings and employment may not fully address. Questions must be addressed not just about the tradeoff of resource allocation between the arts and other subjects, but about the tradeoff between a society that values future earnings and one that also sees value in a basic understanding of and appreciation for arts, music and culture. This paper provides some insight into which direction our nation might be heading, and it is up to future debate as to whether this is the best path, not just for our GDP, but also for our quality of life as a society and as a civilization.

Table 1: Wyoming Summary Statistics

	1998		1999		2000		2001		2002		2003	
	Mean	Std. dev.	Mean	Std dev.	Mean	Std. dev.	Mean	Std. dev.	Mean.	Std. dev.	Mean	Std. dev.
Number of full time equivalent staff in the fine arts per-pupil	0.0124	0.0039	0.0140	0.0076	0.0129	0.0047	0.0129	0.0036	0.0139	0.0068	0.0128	0.0064
Number of full time equivalent staff in math per-pupil	0.0126	0.0035	0.0131	0.0040	0.0145	0.0068	0.0135	0.0051	0.0154	0.0066	0.0152	0.0058
Number of full time equivalent staff in reading per-pupil	0.0159	0.0075	0.0167	0.0070	0.0168	0.0063	0.0157	0.0060	0.0165	0.0064	0.0160	0.0064
Number of students per-teacher	13.8163	2.0263	13.0190	2.0025	12.8675	2.0695	12.7051	2.0012	12.5712	2.1175	12.3180	2.0745
Percentage of students in district who are Hispanic	5.1080	4.5193	5.1628	4.3754	5.2526	4.5815	5.3934	4.6241	5.8623	4.8410	6.1905	5.1763
Percentage of students in district who are Native American	8.3856	24.3940	8.2061	24.0254	8.6449	24.6260	8.2973	24.0093	8.7551	24.9197	8.9979	24.9108

Percentage of students in district/school who are eligible for subsidized school lunches	29.9591	13.5328	30.3387	12.4928	29.9728	12.0864	30.3475	13.0289	32.1828	12.9167	33.1497	12.9411
Per-pupil assessed value	94420.9	88116.5	88212.4	86436.4	105370.6	109588.8	156475.1	187091.0	172081.5	236106.8	143722.9	185335.5
Total expenditure per-pupil	7650.29	1611.65	8512.04	2124.83	9028.73	2332.09	10189.97	2901.57	10924.20	3486.03		
State revenue per-pupil	4562.16	2115.14	4996.49	2415.67	4998.52	2827.88	5075.55	3172.53	5536.19	3136.50		
Percentage of students scoring at or above proficiency level in reading portion of WyCAS test	49.6364	11.7989	44.3696	10.8758	50.5217	11.8991	46.7907	12.2874	46.6304	9.1223	50.7500	12.5700
Percentage of students scoring at or above proficiency level in math portion of WyCAS test	34.7273	11.5222	38.8696	11.0385	41.6087	13.6634	40.3023	13.7694	44.8044	11.5232	44.5909	11.8603

Table 2: Kentucky Summary Statistics

	1998		1999		2000		2001		2002		2003	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Score of school/district on arts and humanities portion of KCCT	40.7116	8.7564	43.2193	9.2585	43.4831	8.5178	48.0917	9.9300	54.4158	10.6501	56.5106	10.4325
Score of school/district on reading portion of KCCT	78.8808	8.7167	80.2104	8.3542	80.7729	8.4952	82.0584	8.2058	83.3111	9.0214	86.4616	8.0319
Score of school/district on math portion of KCCT	55.9149	10.4839	58.6202	10.6707	61.8924	10.6253	63.7286	11.4177	66.1792	10.4578	75.3554	11.6477
Score of school/district on science portion of KCCT	70.4939	8.9992	73.3760	8.8799	77.0232	9.1179	77.4331	9.5912	81.3368	9.7558	87.8608	9.9001
Score of school/district on social studies portion of KCCT	65.6116	11.3207	66.2063	11.4717	67.1450	10.9453	69.7799	11.9451	73.6495	10.8862	80.7913	12.7560

Score of school/district practical living portion of KCCT	68.3917	10.6829	69.2842	10.0010	71.2099	9.8592	72.6407	10.4065	76.5986	9.3298	81.5807	11.1514
Score of school/district on writing portion of KCCT	52.9356	10.4935	54.7369	10.3505	58.7480	10.0604	62.3288	9.9919	68.3105	9.0570	72.4711	8.7258
Percentage of students in school/district who are black	5.1230	8.4445	5.3354	9.1746	5.7876	9.9690	5.9081	10.0968	5.6127	9.9384		
Percentage of students in school/district who are Hispanic	0.5585	1.2013	0.6180	0.9633	0.6664	1.3026	0.7685	1.6828	1.0421	1.9691		
Number of full time equivalent certified staff per-pupil	0.0728	0.0069	0.0751	0.0071	0.0761	0.0075	0.0773	0.0074	0.0778	0.0083		
Total expenses per-pupil	6398.44	717.99	6698.32	735.63	7162.34	883.02	7179.89	887.87	7409.22	961.27		
State revenue per-pupil	4097.77	646.02	4175.70	686.97	4386.75	752.75	4319.40	735.29	4522.56	779.76		

Table 3: Kentucky Questionnaire Summary Statistics

a) How many of the questions tested things you learned in school?

Percent of students responding...	Elementary Schools				Middle Schools				High Schools			
	A	B	C	D	A	B	C	D	A	B	C	D
Arts questionnaire	1.85	22.11	47.03	28.30	4.75	38.13	42.06	14.08	8.96	47.95	31.95	7.76
Reading questionnaire	4.90	23.35	44.31	26.26	3.13	25.39	51.60	18.74	6.46	40.42	43.43	8.34
Math questionnaire	0.90	7.75	46.75	43.83	1.21	12.25	52.43	33.16	2.60	32.51	48.26	15.41
Science questionnaire	1.40	13.81	49.20	34.45	1.69	24.30	55.16	17.76	3.63	43.79	42.91	8.39
Social studies questionnaire	0.91	11.75	52.70	33.99	1.54	20.62	56.12	20.67	2.34	27.06	53.63	15.36
Practical living/vocational studies questionnaire	2.54	19.26	41.77	34.73	5.34	31.75	43.58	16.14	8.50	47.43	34.35	7.06

Answers: A = none of the questions

B = some of the questions

C = most of the questions

D = all of the questions

b) During a typical workweek, how much class time do you spend on arts and humanities?

Percent of students responding...	Elementary Schools					Middle Schools					High Schools				
	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
Arts questionnaire	3.42	43.23	36.47	9.86	6.12	8.85	31.93	29.03	13.73	15.48	22.80	22.85	23.71	9.20	18.12

Answers: A = no time

B = less than 1 hour

C = 1 – 2 hours

D = 3 – 4 hours

E = more than 4 hours

c) During the school year have you had the opportunity to perform/create in each of the areas of dance, drama/theatre, music, visual arts, and literature?

Percent of students responding...	Elementary Schools		Middle Schools		High Schools	
	A	B	A	B	A	B
Arts questionnaire	78.07	19.87	70.19	27.49	59.91	34.36

Answers: A = yes

B = no

Table 4: Wyoming Regression Results²³

Regression 1:

Independent Variable	fine_perf_arts_tfte_pp	language_arts_tfte_pp	math_tfte_pp	science_tfte_pp
total_amer_ind_alsk_dist_tot_pct	0.0009799 (5.25)***	0.0001696 (0.94)	0.0001469 (0.72)	0.0003310 (2.18)**
total_hispanic_dist_total_pct	0.0000366 (0.18)	0.0004970 (2.48)**	-0.0002089 (-0.92)	-0.0001792 (-1.06)
nclb	0.0006550 (0.99)	0.0002154 (0.34)	0.0020474 (2.84)***	0.0011786 (2.19)**
pct_comb_free_redu	0.0000494 (0.71)	-0.0000413 (-0.62)	0.0001579 (2.08)**	0.0003106 (5.49)***
assessed_value_pp	0.0000000 (-1.53)	0.0000000 (-1.17)	0.0000000 (-0.02)	0.0000000 (2.02)**

Regression 2:

Independent Variable	fine_perf_arts_tfte_pp	language_arts_tfte_pp	math_tfte_pp	science_tfte_pp
total_amer_ind_alsk_dist_tot_pct	0.0009638 (5.51)***	0.0001652 (0.93)	0.0001015 (0.51)	0.0002977 (2.01)**
total_hispanic_dist_total_pct	0.0001474 (0.74)	0.0005777 (2.86)***	-0.0003360 (1.49)	-0.0002637 (-1.57)
nclb	-0.0016789 (-2.22)**	-0.0011509 (-1.5)	0.0017873 (2.09)**	0.0008617 (1.35)
pct_comb_free_redu	0.0000729 (1.12)	-0.0000266 (-0.4)	0.0001522 (2.06)**	0.0003080 (5.58)***
assessed_value_pp	0.0000000 (-1.58)	0.0000000 (-1.13)	0.0000000 (-0.22)	0.0000000 (2.27)**
num_pupls_per_teach	-0.0016850 (-5.56)***	-0.0009292 (-3.02)***	-0.0006915 (-2.02)**	-0.0005885 (-2.30)**
avg_teacher_exp	0.0002255 (1.04)	0.0002210 (1.00)	-0.0007584 (-3.08)***	-0.0005287 (-2.88)***

²³ Year dummies are included in the regressions but left out of the results for the sake of brevity.

Regression 3:

Independent Variable	fine_perf_arts_tfte_pp	language_arts_tfte_pp	math_tfte_pp	science_tfte_pp
total_amer_ind_alsk_dist_tot_pct	0.0009646 (5.38)***	0.0002249 (1.25)	0.0001297 (0.65)	0.0002847 (1.88)*
total_hispanic_dist_total_pct	0.0001502 (0.75)	0.0005824 (2.88)***	-0.0003559 (-1.6)	-0.0002603 (-1.54)
nclb	-0.0016927 (-2.03)**	-0.0008646 (-1.03)	0.0026026 (2.81)***	0.0006903 (0.98)
pct_comb_free_redu	0.0000745 (1.13)	-0.0000183 (-0.28)	0.0001484 (2.03)**	0.0003079 (5.52)***
assessed_value_pp	0.0000000 (-1.54)	0.0000000 (-1.06)	0.0000000 (-0.27)	0.0000000 (2.24)**
num_pupls_per_teach	-0.0016912 (-5.51)***	-0.0008778 (-2.85)***	-0.0006345 (-1.87)*	-0.0006073 (-2.35)**
avg_teacher_exp	0.0002237 (-1.02)	0.0002292 (1.04)	-0.0007772 (-3.20)***	-0.0005279 (-2.85)***
read_pct_profcent	-0.0000020 (-0.07)	-0.0000128 (-0.46)	0.0000754 (2.45)**	-0.0000102 (-0.43)
writing_pct_profcent	0.0000086 (0.35)	0.0000086 (0.35)	-0.0000501 (-1.84)*	0.0000096 (0.46)
math_pct_profcent	0.0000028 (0.09)	-0.0000474 (-1.53)	-0.0000611 (-1.78)*	0.0000175 (0.67)

Regression 4:

Independent Variable	log_fine_perf_arts_salary_pp	log_language_arts_salary_pp	log_math_salary_pp	log_science_salary_pp
total_amer_ind_alsk_dist_tot_pct	0.0722625 (6.09)***	0.0171234 (1.33)	0.0304492 (2.48)**	0.0204111 (1.75)*
total_hispanic_dist_total_pct	-0.0111783 (-0.89)	0.0352939 (2.59)**	-0.0290081 (-2.24)**	-0.0001871 (-0.02)
nclb	0.2439369 (5.07)***	0.1953941 (3.73)***	0.3094531 (6.22)***	0.2784326 (5.88)***
pct_comb_free_redu	0.0008251 (0.2)	-0.0007855 (-0.17)	0.0076235 (1.76)*	0.0143927 (3.48)***
log_assessed_value_pp	-0.0349289 (-0.47)	-0.0676008 (-0.83)	-0.0941411 (-1.22)	-0.1385317 (-1.88)*

Table 5: Kentucky Fourth Grade Regression Results²⁴

Regression 1:

Independent Variable	idx_ah	idx_rd	idx_ma	idx_sc	idx_ss	idx_pl	idx_wr
nclb	13.6753 (20.98)***	0.9405 (1.77)*	9.5028 (14.05)***	7.5388 (12.03)***	7.4963 (9.98)***	8.1587 (11.67)***	13.9062 (21.00)***
mean_tot	0.0054 (0.09)				0.1007 (1.39)	0.0090 (0.13)	
mean_rd		0.7178 (13.18)***					
mean_lg							0.3670 (5.22)***
mean_ma			0.1316 (2.13)**	0.5708 (9.99)***			

Regression 2:

Independent Variable	idx_ah	idx_rd	idx_ma	idx_sc	idx_ss	idx_pl	idx_wr
total_amer_indn_alsk_pct	0.3122 (0.72)	0.0572 (0.15)	0.2828 (0.65)	0.0590 (0.14)	0.5560 (1.16)	0.9926 (2.16)**	0.2047 (0.44)
total_asian_pac_isl_pct	0.2069 (0.54)	0.6754 (2.03)**	-0.2549 (-0.66)	0.8727 (2.32)**	-0.1624 (-0.38)	-0.8701 (-2.13)**	0.8098 (1.95)*
total_black_pct	-0.1575 (-1.81)*	-0.0057 (-0.08)	-0.0126 (-0.14)	-0.1214 (-1.43)	-0.1626 (-1.70)*	-0.1411 (-1.53)	-0.1721 (-1.83)*
total_hispanic_pct	0.0258 (0.1)	-0.0547 (-0.24)	0.6488 (2.49)**	-0.1588 (-0.62)	0.2972 (1.04)	0.1765 (0.64)	0.0393 (0.14)
nclb	9.4177 (4.78)***	1.8353 (1.09)	6.2433 (3.17)***	7.8808 (4.09)***	5.0751 (2.34)**	10.0800 (4.82)***	15.9881 (7.56)***
total_certfd_staff_fte_pp	184.8544 (2.30)**	87.0653 (1.26)	59.3644 (0.74)	73.0101 (0.93)	124.3163 (1.41)	138.9011 (1.63)	-28.1561 (-0.32)
exceptnl_chldrn_pct	-66.9418 (-1.06)	-82.4754 (-1.51)	-66.7604 (-1.05)	-94.7577 (-1.53)	-67.5199 (-0.97)	21.1121 (0.31)	-72.0238 (-1.05)
avg_classrm_salary	0.0007506 (2.21)**	0.0001409 (0.48)	0.0007636 (2.25)**	-0.0003458 (-1.04)	0.0003964 (1.06)	-0.000275 (-0.76)	0.0003979 (1.09)
superntndnt_salaries	0.0000040 (0.15)	-0.000017 (-0.72)	0.0000018 (0.07)	0.0000285 (1.07)	-0.0000007 (-0.02)	-0.000025 (-0.87)	-0.000045 (-1.54)
mean_tot	0.0354 (0.44)				0.0435 (0.49)	-0.0426 (-0.5)	
mean_rd		0.5457 (7.69)***					
mean_lg							0.1935 (2.13)**
mean_ma			0.0629 (0.84)	0.3990 (5.44)***			

²⁴ Year dummies are included in the regressions but left out of the results for the sake of brevity.

References

- 1) Baker, Eva L. and Linn, Robert L. "Validity Issues for Accountability Systems" in Fuhrman, Susan and Elmore, Richard (editors) Redesigning Accountability Systems for Education. Teachers College, Columbia University, 2004.
- 2) Betts, Julian R. "The Two-Legged Stool: The Neglected Role of Educational Standards in Improving America's Public Schools" Economic Policy Review – Federal Reserve Bank of New York: 97-116 March 1998.
- 3) Catterall, James S., Chapleau, Richard and Iwanaga, John. "Involvement in the Arts and Human Development: General Involvement and Intensive Involvement in Music and Theater Arts" in Fiske, Edward B. (editor) Champions of Change: The Impact of Arts on Learning. The Arts Education Partnership, 1999.
- 4) Catterall, James S. and Waldorf, Lynn. "Chicago Arts Partnership Evaluation: Summary Evaluation" in Fiske, Edward B. (editor) Champions of Change: The Impact of Arts on Learning. The Arts Education Partnership, 1999.
- 5) Cristenson, Chris. (2001) Wyoming. In C.C. Sielke, J. Dayton, C.T. Holmes, and A.L. Jefferson (Comps.), Public School Finance Programs of the United States and Canada: 1998-99 (NCES 2001-309). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- 6) Cuban, Larry. "Looking Through the Rearview Mirror at School Accountability" in Sirotnik, Kenneth A. (editor) Holding Accountability Accountable: What Ought to Matter in Public Education. Teachers College, Columbia University, 2004.
- 7) Deere, Donald and Strayer, Wayne. "Putting Schools to the Test: School Accountability, Incentives and Behavior" Texas A&M University, 2001.
- 8) Dobbs, Stephen M. "Learning In and Through Art: A Guide to Discipline-Based Art Education" The Getty Education Institute for the Arts, L.A. California, 1998.
- 9) Figlio, David N. "Testing, Crime and Punishment." University of Florida, November 2003.
- 10) Figlio, David N. and Getzler, Lawrence S. "Accountability, Ability and Disability: Gaming the System." National Bureau of Economic Research Inc., Cambridge, MA, NBER Working Paper 9307, October 2002.

- 11) Figlio, David N. and Winicki, Joshua. "Food for Thought: The Effects of School Accountability Plans on School Nutrition." National Bureau of Economic Research Inc., Cambridge, MA, NBER Working Paper 9319, November 2002.
- 12) Fiske, E. "Champions of Change: The Impact of Arts on Learning." The Arts Education Partnership, 1999.
- 13) Flanagan, A. and Murray, S. "A Decade of Reform: The Impact of School Reform in Kentucky" in Yinger, John (editor) Helping Children Left Behind: State Aid and the Pursuit of Educational Equity. The MIT Press, MIT, 2004.
- 14) Hamilton, Laura. and Stecher, Brian. "Responding Effectively to Test-Based Accountability." Phi Delta Kappan v85: 578-583 April 2004.
- 15) Hanushek, Eric A., with others. "Making Schools Work: Improving Performance and Controlling Costs". Washington, D.C.: Brookings Institution, 1994.
- 16) Hanushek, Eric A., Kain, John F. and Rivkin, Steven G. "Why Public Schools Lose Teachers." Journal of Human Resources 39(2): 326-354, 2004.
- 17) Hanushek, Eric A. and Raymond, Margaret E. "Sorting Out Accountability Systems" in Evers, Williamson M. and Walberg, Herbert J. (editors) School Accountability. Hoover Institution Press, 2002.
- 18) Hoxby, Caroline. "The Cost of Accountability" in Evers, Williamson M. and Walberg, Herbert J. (editor) School Accountability. Hoover Institution Press, 2002.
- 19) Jacob, Brian A. "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools" National Bureau of Economic Research Inc., Cambridge, MA, NBER Working Paper 8968, June 2002.
- 20) Jacob, Brian A. "Getting Inside Accountability: Lessons From Chicago." Brookings-Wharton Papers on Urban Affairs: 41-81 2003.
- 21) Jacob, Brian A. and Levitt, Steven D. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." Quarterly Journal of Economics, v18 no. 3: 843-77 August 2003.
- 22) Krueger, Alan B. and Whitmore, Diane M. "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR." The Economic Journal, 111: 1-28 Jan 2001.
- 23) Ladd, Helen F. "The Dallas School Accountability and Incentive Program: An Evaluation of Its Impacts on Student Outcomes" Economics of Education Review, 18: 1-16 1999.

- 24) Ladd, Helen F. and Clotfelter, Charles T. "Recognizing and Rewarding Success in Public Schools" in Ladd, Helen F. (editor) Holding Schools Accountable. The Brookings Institution, Washington D.C., 1996.
- 25) Oakes, Jeannie, Blasi, Gary and Rogers, John. "Accountability for Adequate and Equitable Opportunities to Learn." in Sirotnik, Kenneth A. (editor) Holding Accountability Accountable: What Ought to Matter in Public Education. Teachers College, Columbia University, 2004.
- 26) Peterson, Paul E. and West, Martin R. "The Politics and Practice of School Accountability" in Peterson, Paul E. and West, Martin R. (editors) No Child Left Behind? The Politics and Practice of School Accountability. The Brookings Institution, Washington D.C., 2003.
- 27) Ravitch, Diane. "Testing and Accountability, Historically Considered" in Evers, Williamson M. and Walberg, Herbert J. (editors) School Accountability. Hoover Institution Press, 2002.
- 28) Rudalevige, Andrew. "No Child Left Behind: Forging a Congressional Compromise." in Peterson, Paul E. and West, Martin R. (editors) No Child Left Behind? The Politics and Practice of School Accountability. The Brookings Institution, Washington D.C., 2003.
- 29) The College Board. Profile of College Bound Seniors National Report for 2000, 2001, 2002.
- 30) *The No Child Left Behind Act of 2001*, Public Law 107-110, 107th Congress, Jan. 8 2002.
- 31) Wasley, Patricia A. "Responsible Accountability and Teacher Learning." in Sirotnik, Kenneth A. (editor) Holding Accountability Accountable: What Ought to Matter in Public Education. Teachers College, Columbia University, 2004.
- 32) Zastrow, Claus von. "Academic Atrophy: The Condition of the Liberal Arts in America's Public Schools." Council for Basic Education, March 2004.