ONE

BEYOND THE BASICS: THE EFFECTS OF NON-CORE CURRICULAR ENRICHMENT ON STANDARDIZED TEST SCORES AT PUBLIC SECONDARY SCHOOLS

Susan Catron and Robert W. Wassmer

Standardized test scores for school sites, like California's Academic Performance Index (API), are intended to measure student performance in what have been labelled the core learning areas of "reading, writing and arithmetic" and have become the focus of public concern. Some have argued that this focus has de-emphasized the pursuit of curricular enrichment of anything besides these "Three R's." In this paper we assess whether such a de-emphasis could hurt the pursuit of high standardized test scores. We do this through a regression analysis of factors that are expected to cause differences in recent API scores from 188 public high schools and 184 middle schools in 14 northern California counties, and find that the ratio of offerings in Health/PE, Fine Arts and Foreign Language exerts a positive effect on API scores depending on the average parent education level at a school site. The major implication is that curricular inputs must be used strategically to achieve the desired outcome. Notably, the less academic forms of enrichment are expected to enhance scores in schools serving disadvantaged student populations (as defined by a low average parent education level). These findings suggest caution in pursuing curricular changes designed to boost a school site's standardized test scores.

Susan Catron is a 2005 graduate of California State University, Sacramento with an MA in Public Policy and Administration. She is currently a fiscal and policy analyst at the California Legislative Analyst's Office. She can be reached at susancatron@msn.com. Robert W. Wassmer is a professor of Public Policy and Administration at CSU, Sacramento, where he teaches courses in applied microeconomics, regression analysis, and urban economics. He can be reached at rwassme@csus.edu. The opinions expressed here are the authors' own.

I. INTRODUCTION

The curriculum in public schools is increasingly becoming "one-size-fits-all" due in part to the standards and accountability movement. Though standards have a legitimate role to play in establishing a common framework of knowledge, there is evidence to suggest that schools are squeezing diversity out of the curriculum in an effort to boost their standardized test scores. This is particularly true of subjects that fall under the umbrella of enrichment, such as music and art. Numerous studies have documented a relationship between some forms of enrichment and academic achievement. Yet, few studies have attempted to look at enrichment as a curricular input or to establish a link between the curricular mix and aggregate student achievement. In view of the heightened focus on school-level accountability, such an analysis has the potential to inform policy and practice in a context that is relevant to the current dialogue.

With 2003-04 data from 188 public high schools and 184 middle schools in fourteen Northern California counties, our study uses multiple regression to explore the causal link between enrichment – defined here as the number of non-core courses per student – and student achievement, as measured by the Academic Performance Index. While the problem of squeezing diversity out of the curriculum appears to begin in the early grades, the emphasis here is on secondary schools for two reasons. First, ready access to data regarding instructional time spent on non-core subjects is unavailable for grades K-6. Beginning at the secondary level, however, schools report the number of courses by subject, making such an analysis possible. Second, and more importantly, secondary school reform (particularly at the high school level) is the "the next big thing" on the policy agenda, making it an opportune time to weigh in on the matter (Posnick-Goodwin 2005, 6).

It should be noted from the outset that because a California school's Academic Performance Index (API) reflects aggregate student test scores in only four core subject areas (language arts, math, science, and history), the use of the API is an attempt to test the hypothesis that participating in *enrichment* or non-core courses benefits performance in *traditional core* subjects.[1] In this context, a positive link would suggest potential instrumental effects between enrichment and academic achievement, such as enhanced motivation on the part of students or synergies between skills developed in core and non-core courses. From the policy perspective, the presence of such a link may have important implications for how the core curriculum is defined, how school performance is assessed, and whether there is a need to mitigate potential negative impacts associated with the current accountability system.

To put the present study into context, the analysis begins with some background on the standards and accountability movement followed by a brief review of the literature pertaining to the modeling of educational outcomes as well as the role of non-traditional subjects in academic achievement. The analysis continues with an overview of the regression model, including the broad causal factors, the dependent and independent variables used to operationalize these factors, the expected direction of effects, and other details about the data (e.g., sources, descriptive statistics). The fifth section offers an analysis and interpretation of the data, along with some specifics about the data analysis methods employed. Finally, the paper concludes with a discussion of the implications for educational policy and practice, and some thoughts about future research needs.

II. BACKGROUND

A catalyst for the present standards and accountability movement was the 1983 report, *A Nation at Risk*, which called attention to a decline in the educational performance among U.S. students compared to previous generations and contemporary peers in other countries.[2] In the ensuing 20 years, standards-based reform has become "the driving force behind most federal, state and local education policy" in the U.S. (EdSource Online 2005).[3] The underlying assumption is that "a rise (in) student achievement is commensurate with clearly defined student standards the attainment of which is measured by carefully chosen assessments" (NSTA 2000). As of 2004, 49 states (including California) had established content standards for at least the four core subjects, and most had implemented statewide assessment systems to gauge student progress toward grade-level mastery of each standard (Stanford University Bridge Project 2004). Standards and accountability are also the centerpiece of federal education policy as set forth in the *No Child Left Behind Act.*

California began developing its content standards in the mid-1990s, and is one of only three states to earn an "A" from the Fordham Foundation for the quality of its math, science *and* English standards (Fordham Foundation 2005). Although initially voluntary, implementation of the standards (and the associated curriculum frameworks) was in essence legislatively mandated in 1999 by the passage of the Public Schools Accountability Act, which sought to hold schools accountable for student performance relative to the new standards. A central feature of the Act was the development of a new indicator – the Academic Performance Index (API) – to measure the performance and year-to-year progress of schools on one comparable scale. While the API was initially comprised of only one test (the Stanford-9), it currently reflects the results from a variety of test-based inputs covering content in four core subjects. These include a norm-referenced test (the CAT/6), a series of standards-based tests, and the California High School Exit Exam (grade 10 results only).[4] The test scores are aggregated to produce a composite performance index for each school, with values ranging from 200 to 1,000. To give some bite to the accountability system, the Public Schools Accountability Act also provided for various rewards and sanctions linked to API scores. Not surprisingly, in the face of California's ongoing fiscal imbalance the monetary incentives quickly evaporated. The sanctions, however, did not.

It is difficult to argue that schools should not be held accountable, and standardized testing is a good way to compare achievement across schools and districts. By design, such a comparative measurement system provides an incentive for schools to improve student achievement even in the absence of monetary rewards. Yet, as is often the case, incentives can have unintended consequences. In the context of California's standards and accountability program, one such consequence appears to be a narrowing of the curriculum. At countless schools, art, music, shop and other "non-essential" courses are being cut as schools increase their emphasis on core subjects. This is not meant to imply that an enriching curriculum can no longer be found in public schools, but there appears to be a gradual shift under way as schools chip away at the fringes to make room for more basics. In the five years following the enactment of the Public Schools Accountability Act, the percentage of public school students receiving music instruction dropped nearly fifty percent statewide (Posnick-Goodwin 2005). In fact, student enrollment declined in nearly every non-core subject as evidenced in Table 1.[5]

Table 1: Subject Area Enrollments as a Percentage of Total Student Enrollment in California Public Schools

Subject Area (Includes AP/IB if applicable)	*% of Total Enrollment*		*% Change*
	1999-00	*2003-04*	
Music	*18.51%*	*9.38%*	*- 49.37%*
Humanities	*1.16%*	*0.70%*	*- 39.91%*
Computers/Tech/Voc Education	*18.74%*	*15.22%*	*- 18.77%*
Health Education	*4.43%*	*3.69%*	*- 16.85%*
Physical Education	*40.10%*	*35.90%*	*- 10.47%*
Art	*10.18%*	*9.81%*	*- 3.60%*
Foreign Language	*13.45%*	*13.65%*	*+ 1.53%*
Drama/Theater Arts	*2.30%*	*2.36%*	*+ 2.76%*
Social Science	*38.77%*	*40.93%*	*+ 5.57%*
Science	*35.08%*	*37.19%*	*+ 6.01%*
English	*52.31%*	*56.30%*	*+ 7.62%*
Math	*39.05%*	*43.18%*	*+ 10.57%*

Source: DataQuest, query on statewide enrollment by subject as a percentage of total enrollment (calculated).

Unfortunately, the finite nature of public resources necessitates difficult choices. The key is determining the basis on which educational priorities should be established. If one views public education as a means of producing "raw materials for industry," then it makes sense to focus on the essential skills that make children employable (Kozol 1991). However, even if one subscribes to such a philosophy, educational theory supports the idea that intelligence comes in many forms: linguistic, spatial, musical, kinesthetic and others (Gardner 1983). Thus, besides their aesthetic appeal, music and art may be an important pathway to learning for some children. To the extent that an enriching curriculum provides a means for all children to find a path to learning, squeezing diversity out of the curriculum may limit the pathways for some children to succeed. Moreover, teachers suggest that a lack of enrichment risks hurting all students by producing adept test-takers who are incapable of understanding "nuance, creative problem solving, working thematically across disciplines and working collaboratively" (Posnick-Goodwin 2005, 12-13).

III. LITERATURE REVIEW

There are different frameworks from which to examine questions relating to academic achievement. From the economic perspective, much of the previous research makes use of a production function meant to describe how a given set of inputs is transformed into an output. When applied to education, the question becomes how do schools (given the "quality" of students they must educate) use resources such as classrooms, teachers and the like to generate student outcomes (i.e., grades, graduation rates, test scores, etc.)? A key advantage of this economic approach – which has been widely applied in education for more than 30 years – is the ability to isolate the independent effects attributable to each variable of interest while controlling for the multitude of factors that are known to impact outcomes. Such an approach fits well with the objective of this study: to isolate the effects (if any) of non-core curricular inputs on API scores.

The broad casual factors thought to influence educational outcomes are often grouped under three categories: student inputs, family/social inputs and school inputs. The first two causal factors are so highly interrelated that they are sometimes classified together as "student endowments." This includes students' innate abilities as well as their home learning environments. Separate from school resources and quality, these factors influence students' readiness to learn and very likely set the outside parameters for students' ability to achieve. Given the available data, Wassmer (2002, 2004) defines *student inputs* as a function of the ethnic make-up of a school's student body, the percentage of students participating in free or reduced lunch programs, and the proportion designated English-language learners. Coates (2003) uses similar measures, but defines ethnicity as simply white or non-white.

With regard to *family or social inputs,* the theoretical objective is to model the learning support systems in the home environment. Specifically, do parents or guardians support and enhance classroom learning by assisting their children with homework, volunteering in the classroom, engaging their children in enrichment activities, etc.? In view of the strong relationship between parental education and the aforementioned activities, most economists and education researchers use some measure of parental education level as a proxy. From an economic perspective, this phenomenon might be labeled as a "taste" or "preference" for education; in other words, more educated parents tend to place a higher value on the importance of education. They also tend to possess greater personal and financial resources with which to demonstrate their support. Not surprisingly, numerous students show a positive link between parental education and academic achievement (see, for example, Hanushek 1989, Kostatis 1987).

The third factor that completes the production function model is *school input*s, measured in dollars or the particular resources purchased. Though student and social characteristics have been found to explain more than half the variation in API scores (Wassmer 2004, 15), the school inputs category is critically important to educational policy makers because it represents the only factor over which schools have direct influence. Unfortunately, it is also the least conclusive and most controversial. In a meta-analysis of over 100 regression studies using some version of an educational production function, Hanushek (1989) finds that most school inputs do *not* exert significant effects; and, when they do, the results are mixed. Hedges, Laine and Greenwald contradict Hanushek's findings in their own meta-analysis of these same studies, suggesting that some resources *do* show a "systematic positive pattern," including per-pupil expenditures and teacher experience (Hedges et al. 1994, 8).

As an extension of the school inputs category, many production function (or related) models go beyond purchased resources to test or control for various *structural or institutional characteristics.* Commonly used parameters include location (urban, suburban, rural), size (total enrollment), sector (public/private) and selectivity (charter or magnet status). Here again the results are mixed, although a public sector setting and urban location are generally associated with lower achievement levels (Lee & Bryk 1988, Elliott 1998). There is growing attention to school size as a factor – particularly in view of the small schools reform movement – but the findings to date are inconclusive (Lamdin 1995, Ready et al. 2004). This may be due in part to the difficulty of distinguishing schools that are small "by default" versus those that are small "by design" (Ready et al. 2004). Finally, the effect of charter status is also inconclusive, although preliminary evidence suggests that the type of charter matters; specifically, "start-up" charters (as opposed to conversions) have been associated with positive effects on academic achievement (Zimmer & Guarino 2003).

Few studies in education economics have considered the role of curricular characteristics, leading Coates to conclude that "lack of information on the amount of instruction in the subjects on which students are tested…(is a) misspecification in the literature" (Coates 2003, 274). He aims to fill this void by adding such a measure (i.e., instructional time by subject matter), though he focuses only on core subjects in one particular elementary grade level (Coates 2003). He finds a significant relationship between time spent on language arts and mathematics instruction and scores on the California Achievement Test (which is not surprising given the emphasis of these skills on the CAT test), as well as some linkages between social science and reading scores. A handful of other studies in education economics explore instruction-related effects, such as the number of math and science courses taken (Walberg et al. 1986) or the amount of homework assigned (Betts 1997). Generally speaking, these studies show that greater exposure to a subject (through coursework or homework) exerts a positive effect on student achievement. However, unlike Coates, these researchers examine the effects on achievement within the same subject (direct effects) rather than across subjects (indirect or instrumental effects); they also focus only on core academic subjects (Walberg et al. 1986, Betts 1997). To get a broader view of instructional effects, then, it is necessary to look beyond economics to the general education literature where much of the research on pedagogical issues is found. While by no means an exhaustive review, following is a summary of key findings for the four broad subject groupings relevant to the present study.

The Arts

Much of the research on arts and achievement is correlative in nature. In an effort to determine where *causal* inferences can be made, Hetland and Winner (2001) review 188 art-related studies conducted between 1950 and 1999 and find three areas in which "clear causal links" can be demonstrated. Controlling for student characteristics, Catterall also finds a positive effect from involvement in the arts (though not necessarily arts *education*) among both the general student population as well as lower socioeconomic students as a subgroup (Catterall et al. 1999). Specifically, he finds gaps in achievement on standardized math and reading tests between those with "high arts involvement" and those with "low arts involvement," where involvement is broadly defined as enrollment in arts-related courses in or out of school and/or participation in extracurricular activities such as band, chorus or dance. Finally, McNeal (1995) reinforces the link between arts-related extracurricular activities (e.g., band, drama) and positive student outcomes, though in his study the outcome of interest is student retention.

Foreign Languages

The evidence of a link between foreign language (as a curricular input) and student achievement is stronger. The College Board has consistently found a positive correlation between the study of foreign languages and SAT scores (Matts 2005). More importantly, in a statistically rigorous study controlling for differences in students' prior language arts abilities and socioeconomic status, Cooper (1987) finds that studying a foreign language positively influences SAT verbal scores. In a similar study, Olsen and Brown (1992) find higher English and math scores on the ACT exam among college applicants who studied a foreign language.

Physical Fitness

The findings linking physical education and achievement are supportive, but less compelling than other enrichment subjects. There is strong evidence of a positive correlation between physical *fitness* level (though not physical *education*) and academic achievement. On this point, a widely cited study conducted by the California Department of Education (2002) finds that higher levels of physical fitness – as indicated by scores on the Fitnessgram administered statewide in grades 5, 7 and 9 – are associated with higher levels of achievement on the SAT-9 (formerly a component of the API). The empirical evidence regarding physical *education* comes from four studies conducted since the 1950's in France, Australia, Canada and the United States. Using experimental or quasi-experimental designs and adequate statistical controls, these studies find that treatment effects associated with the particular physical education programs tested do *not* generally raise academic test scores or grades. However, in several cases a reallocation of academic class time to increased physical education had no detrimental effects on scores, suggesting that academic productivity is not hurt – and may be enhanced – by physical education (Shepherd 1997, Sallis & McKenzie 1999).

Career/Vocational Education

Finally, the evidence on vocational education is mixed. This may be due in part to the wide variety of programs that fall under the vocational umbrella (ROP, school-to-work, career and technical education, etc.), the multiple institutions responsible for such programs (schools, county offices of education) or the diversity of individuals served (secondary students as well as adults). The emphasis here is on high school-based programs, generally lumped under the "school-to-work" rubric. In a meta-analysis of more than 100 studies of school-to-work programs, Hughes and colleagues find that participating students have better attendance and graduation rates than non-participants, though the relationship to achievement (as measured by standardized test scores or grades) is inconclusive (2002).

In summary, there is ample evidence to lend legitimacy to the research question under consideration here, but there is little guidance from previous econometric or sociological research on how enrichment is best operationally defined as a school-level input or what effect we should expect curricular enrichment to have on school-wide academic achievement.

IV. Regression Model

The regression model used for this analysis is the classic educational production function, extended through the addition of certain curricular inputs. The dependent variable is the Academic Performance Index, which has been the key measure of school performance in California since the Public Schools Accountability Act was enacted. As such, it has taken center stage in curricular decisions as administrators attempt to maximize API scores. An important reminder here is that the API does not directly measure performance in non-core subjects; thus, its use in this model is an attempt to test the hypothesis that participation in non-core curricular enrichment has instrumental effects on performance in traditional core subjects.

The variable of interest (curricular enrichment) is defined in this model as the number of course offerings or "sections" per student outside the traditional core subjects.[6] The assumption made here is that more courses per student (regardless of the range of choices) means more students are able to participate.[7] In terms of specific subject areas, this definition encompasses four broad categories of courses: fine arts (i.e., studio art, production art, drama and performing arts, music), foreign languages, health/physical education and, finally, career, technical and vocational education (captured collectively under "other"). Advanced Placement and International Baccalaureate sections in these subject areas are also included. Because of the potential for these subjects to exhibit different effects on API scores, each broad subject grouping is included individually in the model, thereby providing four distinct variables of interest.

Due to the amount of data manipulation required to merge inputs from multiple Department of Education databases and compile all of the necessary curricular inputs, the samples are limited geographically to a 14-county region in Northern California including Alameda, Contra Costa, El Dorado, Marin, Napa, Placer, Sacramento, San Francisco, San Mateo, Solano, Sonoma, Sutter, Yolo and Yuba. This includes a sample size of 188 high schools and 184 middle schools, adequate to provide a heterogeneous sample (i.e., a broad range of school sizes, locations and settings) useful for predictive purposes.[8] While student composition in the sample is not representative of the state as a whole, such differences are controlled for by the inclusion of student characteristics in the model.

The broad causal model used to explore the potential link between curricular enrichment and school achievement is expressed in the following general form:

School Achievement = *f* (student inputs, family/social inputs, school inputs, and curricular inputs);

where:

School Achievement = *f* [2003 Base API Score];

Student Inputs = *f* [% African American (-), % American Indian (-), % Asian (+), % Filipino (-), % Hispanic (-), % Pacific Islander (-), % English-language learners (-)];

Family/Social Inputs = *f* [% participating in free/reduced lunch program (-), average parent education level (+)];

School Inputs = *f* [total enrollment (+/-), % of teachers with full credentials (+), students per computer (-)];

Curricular Inputs = *f* [number of sections per student in fine arts (+/-), foreign languages (+/-), health/PE (+/-), vocational/career/technical education (+/-)];

Other Control Variables = *f* [charter status (+/-), percent tested (+), percent opted out of testing (+), population density (+/-)].[9]

Explanatory Variables

For student and social characteristics, the chosen independent variables reflect those most commonly used in previous studies and are intended to serve as proxies for students' abilities and home environments. In this analysis, all non-white racial/ethnic groups (recorded by the California Department of Education) are included in the model, expressed as the proportion of a school's student body comprising each group; thus, the calculated effects represented by the regression coefficients are relative to the excluded category (% white). The model also includes the proportion of students who are designated English-language learners as well as the percentage of students participating in the free or reduced lunch program. At the other end of the socioeconomic spectrum, high average parent education is used to indicate a supportive home learning environment. It should be noted that parental education is represented on a 5-point scale where 1 equals "no high school diploma" and 5 equals "graduate school" (as reported by students), and is included as a school-wide average.

In the school inputs category, total enrollment is used to capture the effects of school size on outcomes and to control for the varying sizes of schools in the study. Teacher experience is measured as the proportion of teachers who possess full credentials. One additional measure considered was the percent of teachers with emergency credentials, which has been shown in previous studies to have an impact on test scores (Wassmer 2002, 2004); however, when included with full credentials, this parameter's effect on API was insignificant (likely due to the strong relationship between the two) and failed to contribute to the explanatory power of the equation. Finally, the number of students per computer is used as a proxy for a school's financial resources. While the ratio of students to computers is admittedly a very crude way of estimating differences in resources, it is more likely than other physical school resources to benefit from grants and private donations. In that sense, it may account for at least some of the variations in financial resources between individual schools. The final category of production inputs – curricular inputs – was operationally defined earlier.

Several additional controls are also included in the model. First is a dummy variable indicating a school's charter status. Second is the population density in the area surrounding each school, captured by a dummy variable indicating whether the school is in a highly urbanized area. From a theoretical perspective, this "large city" variable is designed to pick up effects attributable to the broader community environment that are not accounted for elsewhere in the model. The final two controls are test participation indicators shown to be significant in previous studies: percent tested and percent opted out of testing.

An important caveat at this juncture is the potential for errors in the specification of variables in a regression model. One common error is omitted variable bias, a condition in which the model fails to include one or more relevant explanatory variables. For example, this model excludes per-pupil expenditures because these data are not readily available at the school level. If an omitted variable also bears some relationship to the variable of interest (e.g., per-pupil expenditures affects both API scores and the ratio of enrichment courses offered), the potential estimation errors are more serious. In this study, we attempt to control for differences in school resources through the use of proxy variables such as teachers' credential status and selected physical assets. Moreover, these proxies show weak correlations to the variables of interest.

Predicted Effects

The expected direction of effects is indicated by the use of a "+" sign (positive effect), a "-" sign (negative effect), or a mixed "+/-" sign (a non-zero effect). As a general rule, we expect variables that serve as a proxy for the proportion of lower socioeconomic students – including minority status (except Asians) and free/reduced lunch participation – to be negatively related to achievement. This is consistent with previous research, as well as theories that such students (as a group) tend to possess weaker support systems that are less conducive to learning. One exception is Asian status, which has been positively associated with achievement in previous studies. Average parent education level, on the other hand, is a proxy for higher socioeconomic status and a more conducive home environment; therefore, it is expected to exhibit a positive effect on achievement. In other words, schools with higher average parent education levels are expected to have higher API scores. Finally, the proportion of students who are English-language learners is predicted to have a negative relationship to API scores. This expectation has empirical support, but also makes intuitive sense as API tests are administered in English and specifically test English-language skills.

With regard to school inputs, the intuitive expectation is that variables typically associated with *greater* school resources – such as the proportion of teachers with full credentials – are likely to exhibit positive effects on API. On the other hand, we expect variables typically associated with *fewer* resources – such as a higher ratio of students to computers – to exhibit negative effects on achievement. For total enrollment, we make no predication regarding the direction of the effect (denoted by "+/-"). Likewise, the effects of curricular inputs are uncertain. While the hypothesis being tested here is that enrichment opportunities are positively associated with API scores, there is little empirical evidence to support a specific prediction; thus, our hypothesis simply assumes a non-zero effect.

Finally, the predicted effects of the additional control variables are mixed. We have no specific expectation for charter status given the lack of conclusive evidence and the inability to categorize charters as start-ups or conversion schools. We expect highly urbanized status to be negatively related to API given that inner-city schools typically serve communities with lower socioeconomic status populations and higher rates of crime and other social ills that characterize a less learning-conducive community environment. Finally, we expect both indicators of test participation to have a positive association with achievement. This seems somewhat counterintuitive at first, but each variable likely reflects a different phenomenon. For example, the fact that the proportion opted out of testing is positively associated with API scores (Wassmer 2002, 2004) suggests that lower performing students are more likely to be opted out of testing. One anecdotal explanation offered by school insiders is that schools with higher opt-out rates may have larger special education populations, whose parents are more likely to exclude their children from testing. Wassmer's findings also show that the effect of percent of first-day enrollees tested on achievement is significant and positive for middle and high schools, perhaps reflecting that schools with higher test rates have lower rates of mobility and absenteeism which is characteristic of higher socioeconomic (and thus, higher performing) student populations.

The tables that follow include descriptions and sources for each variable (Table 2) as well as descriptive statistics (Table 3). Bivariate correlations are not reported here, but the analysis reveals a few highly correlated variables among some of the student/social inputs (e.g., Free/Reduced Lunches, English-language Learners and Parent Education Level).

Table 2: Variable Labels, Descriptions and Data Sources

Variable Label	*Description*	*Source*
Dependent		
API03	*Academic Performance Index 2003 Base Score*	*California Department of Education (CDE), 2003 API Base Data File at* http://api.cde.ca.gov/datafiles.asp
Independent: Student Inputs		
% African American	*Percent of students African American*	*CDE, 2003 API Base Data File*
% American Indian	*Percent of students American Indian*	*CDE, 2003 API Base Data File*
% Asian	*Percent of students Asian*	*CDE, 2003 API Base Data File*
% Filipino	*Percent of students Filipino*	*CDE, 2003 API Base Data File*
% Hispanic	*Percent of students Hispanic*	*CDE, 2003 API Base Data File*
% Pacific Islander	*Percent of students Pacific Islander*	*CDE, 2003 API Base Data File*
% English-language Learners	*Percent of students English-language Learners*	*CDE, 2003 API Base Data File*
Independent: Family/Social Inputs		
% Free/ Reduced Lunches	*Percent of students participating in free/reduced lunches*	*CDE, 2003 API Base Data File*
Parent Education Level	*Parents' average education level (5 point scale where 1=no high school and 5=grad school)*	*CDE, 2003 API Base Data File*

Independent: School Inputs		
Total Enrollment	*Total school enrollment*	*CDE, 2003 API Base Data File*
% Full Credential	*Percent of teachers with full credential status*	*CDE, 2003 API Base Data File*
Students Per Computer	*Number of students per computer*	*Ed Data (compiled via school-level query on school technology) at* http://www.ed-data.k12.ca.us
Independent: Curricular Inputs (Enrichment Opportunities)		
Fine Arts Sections Per Student	*Total number of fine arts courses per student (art, drama, music) including AP & IB classes*	*CDE, DataQuest (compiled); source available at* http://www.cde.ca.gov/ds/dq
Foreign Language Sections Per Student	*Total number of foreign language classes per student including AP & IB*	*CDE, DataQuest (compiled); source available at* http://www.cde.ca.gov/ds/dq
Health/PE Sections Per Student	*Total number of health and PE classes per student*	*CDE, DataQuest (compiled); source available at* http://www.cde.ca.gov/ds/dq
Other Sections Per Student	*Total number of other classes per student (computers, vocational, technical education)*	*CDE, DataQuest (compiled); source available at* http://www.cde.ca.gov/ds/dq
Independent: Controls		
Dummy: Charter?	*Dummy variable to capture status as a charter school (1 = yes)*	*CDE, Charter School Locator at* http://www.cde.ca.gov/ds/si/cs
Dummy: Large City?	*Population density of area in which school resides (compiled)*	*Ed Data (compiled via query on school profile) at* http://www.ed-data.k12.ca.us
Percent Tested	*Percent of total student body tested (API '03)*	*CDE, 2003 API Base Data File*
Percent Opted Out	*Percent opted out of testing by parents*	*CDE, 2003 API Base Data File*

Table 3: Descriptive Statistics

	HIGH SCHOOL SAMPLE				MIDDLE SCHOOL SAMPLE			
Variable Label	Mean	Standard Deviation	Min.	Max.	Mean	Standard Deviation	Min.	Max.
Dependent Variable								
API03	682.56	99.47	335	936	721.11	106.62	475	922
Independent: Student Inputs								
% African American	11.15	15.13	0	86	13.21	17.14	0	84
% American Indian	0.98	1.17	0	7	0.94	1.75	0	20
% Asian	12.00	13.47	0	65	11.23	13.09	0	67
% Filipino	4.42	6.86	0	45	4.22	6.44	0	49
% Hispanic	20.56	14.92	3	73	21.93	17.25	2	99
% Pacific Islander	1.12	1.73	0	13	0.96	1.39	0	9
% English – language learners	12.92	13.78	0	90	14.41	14.16	0	67
Independent: Family/Social Inputs								
% Free/Reduced Lunches	22.74	19.25	0	98	32.03	24.07	0	100
Parent Education Level	3.15	0.60	2	5	3.09	0.67	1	5
Independent: School Inputs								
% Full Credential	87.72	12.71	20.0	100.0	88.17	12.03	33.0	100.0
Students Per Computer	4.67	2.98	0.8	26.8	6.18	3.44	1.0	21.1
Total Enrollment	1092.63	562.58	21.0	3293	722.58	304.32	79	1820
Independent: Curricular Inputs								
Fine Arts Sections Per Student	.02268	.01271	.0000	.0952	.01169	.00896	.0000	.0514
Foreign Language Sections Per Student	.02324	.01069	.0000	.0926	.00487	.00606	.0000	.0412
Health/PE Sections Per Student	.02278	.01196	.0000	.0926	.02591	.00872	.0000	.0700
Other Sections Per Student	.02060	.01933	.0000	.1311	.00708	.00598	.0000	.0374
Independent: Other Controls								
Dummy: Charter?	0.07	0.254	0	1.00	0.03	0.163	0	1
Dummy: Large City?	0.16	0.371	0	1.00	0.19	0.397	0	1
Percent Tested	95.44	3.64	82.81	100.00	98.25	1.682	91.82	100.00
Percent Opted Out	1.82	2.58	0	14.22	0.71	1.082	.00	6.92

V. ANALYSIS AND INTERPRETATION OF FINDINGS

The driving question behind this study is whether there is a causal link between curricular enrichment as a school-level variable and API scores. Overall, the data provide limited support for the main hypothesis, although the findings are strengthened upon examination of the interaction between curricular inputs and selected student characteristics. In short, it appears some curricular inputs do matter under some conditions.

Before presenting the results, a few technical notes are in order. All models presented here make use of the log-log or double log form. The use of a non-linear model is consistent with previous educational research. An additional benefit of this functional form is that it facilitates interpretation by expressing the regression coefficients as elasticities (i.e., the % change in API given a 1% change in a given explanatory variable). Once properly specified, we tested the models for the two most common problems associated with cross-sectional data: multicollinearity and heteroskedasticity. On the basis of pair-wise correlation coefficients and Variance Inflation Factors (VIF),[10] we found multicollinearity between certain student and social inputs (i.e., free lunch participation, English-language Learners and Hispanic). However, the high VIFs generally did not inhibit the ability to gauge statistical significance (except in the case of Percent Hispanic), making it unnecessary to omit any variables. Another reason not to be concerned with multicollinearity is that our goal is to control for as many student and social inputs as possible in order to measure the independent effect of curricular characteristics; thus, excluding relevant variables could detract from the robustness of the model. With regard to heteroskedasticity[11] – which occurs when the variance of the error term is not constant due to a wide variation in the size of the observations (such as large and small schools) – a Park Test confirmed its presence and we corrected the results using White's heteroskedasticity-robust estimator.[12]

In the final phase of our analysis, we refined the regression results to determine whether curricular inputs exhibit significant effects when viewed in interaction with other explanatory variables. Much of the educational production function literature begins and ends with an analysis of the effect exhibited in the basic regression model (i.e., the main effects). In this case, the main effects of curricular enrichment (when viewed as individual subject groupings) provide little evidence to support the hypothesis that enrichment in all its forms matters with respect to test scores. However, as Lee and Burkham point out, basing one's conclusions solely on main effects could lead to "incomplete or incorrect conclusions" (2001, 27). Thus, for both samples, we analyzed the potential interaction effects between a school's average parent education level (the largest and most significant student/social input) and each of the curricular inputs in question. Not surprisingly, this analysis of the interaction effects adds an important dimension to the findings.

Regression Results

Table 4 provides five regression results, beginning with the high school sample (columns 2-4) followed by the middle school sample (columns 5-6). Column 2 shows the "Basic Model" for the high school sample with no interactions included as explanatory variables. The results of this regression are generally consistent with our a priori expectations. For the specific purpose of this study it is important to note that none of the measures of non-core curriculum enrichment exerts a statistically significant influence on high school API scores. Column 3 offers the basic regression model expanded to include the interaction effect between Parent Education Level and Health/PE Sections (labeled "Effect #1"), and column 4 offers the results of expanding the basic model to include the interaction between Parent Education Level and Foreign Language Sections (labeled "Effect #2"). The interaction effects reveal a different perspective on the role of curricular enrichment. When we include the interaction between the average level of parent education and the various curricular inputs, it becomes apparent that other non-core inputs do matter under some conditions.

The Effect #1 regression indicates that Parent Education Level exhibits a negative interaction with Health/PE sections per student in high school. In this case, all three component parts (i.e., the two explanatory variables and associated interaction term) are significant indicating different Y-intercepts (or starting points) and different slopes (the direction and degree of change expected). This finding suggests that the impact of increasing the number of sections per student is muted when average parent education is higher; conversely, the impact is greater for schools with student populations from less educated households.

In contrast, the Effect #2 regression offers some evidence that increasing Foreign Language sections per student in high schools is less productive when parent education levels are lower. In this case, only the interaction term is significant suggesting that as Parent Education Level and Foreign Language Sections are both increased, the regression line becomes more positive (or less negative, as the case may be). The two explanatory variables do not achieve a level of significance (likely due to multicollinearity), hence no definitive causal interpretation can be made. Nevertheless, the major reversal in the direction and size of the coefficient on Foreign Language Sections – coupled with the positive parental education interaction – suggests that expanding sections is most helpful in cases where the school composition already leans toward more educated households and may cause a depressive effect in schools with a low average parent education level.

Columns 5 and 6 present the regression results for the middle school sample, again starting with the basic model (no interactions) followed by an expanded model showing the significant interaction effect. Unlike the high school sample, the middle school data set reveals a positive main effect on API for one curricular input (Foreign Language Sections) suggesting that as Foreign Language Sections Per Student is increased in middle schools, API scores are predicted to rise. One possible explanation is the potential for spillover effects from foreign language instruction to traditional language arts. A second possible explanation is that middle schools with a higher ratio of foreign language sections move students onto a college preparatory track earlier in their academic careers, increasing the motivation to perform well in school. Finally, the effect may be developmental: earlier access may have a greater effect on test scores. It is unclear why this effect is not apparent in high schools. One plausible explanation is that, at the high school level, where Foreign Language instruction is more prevalent, participating students possess a wider range of abilities, motivation levels and support systems in the home, so the measured effect incorporates more low performers.

The middle school sample also reveals one significant interaction effect, in this case between Fine Arts Sections and Parent Education Level (see column 6). Like Health/PE Sections in high schools, the interaction is negative, indicating that the effect of Fine Arts Sections is greater for schools with student populations from less educated households.

Table 4: Regression Results with Significant Interactions

	HIGH SCHOOL SAMPLE			MIDDLE SCHOOL SAMPLE	
Variable Label **(Ln=variable is in log form)**	**Basic Model:** **No Interactions**	**Effect #1** **Parent Ed * Health/ PE**	**Effect #2** **Parent Ed * For. Lang.**	**Basic Model:** **No Interactions**	**Effect #1** **Parent Ed * Fine Arts**
Constant	5.874 (.337)	5.624 (.305)	6.008 (.326)	6.294 (.196)	6.341 (.197)
African American	-.001 (.001) VIF=3.175	-.001* (.001) VIF=3.189	-.001 (.001) VIF=3.199	-.003*** (.000) VIF=4.085	-.003*** (.000) VIF=4.190
American Indian	-.004 (.005) VIF=1.649	-.007 (.005) VIF=1.665	-.002 (.005) VIF=1.712	-.004** (.002) VIF=1.296	-.004** (.002) VIF=1.298
Asian	.002*** (.001) VIF=2.284	.002*** (.001) VIF=2.285	.003*** (.001) VIF=2.296	.002*** (.000) VIF=2.858	.002*** (.000) VIF=2.863
Filipino	-.001** (.001) VIF=1.353	-.002** (.001) VIF=1.353	-.001* (.001) VIF=1.357	-.001+ (.001) VIF=1.352	-.001 (.000) VIF=1.355
Hispanic	.001 (.001) VIF=3.781	.001 (.001) VIF=3.856	.001 (.001) VIF=3.784	-.001 (.001) VIF=6.357	-.001 (.001) VIF=6.453
Pacific Islander	-.003 (.004) VIF=1.415	.001 (.004) VIF=1.495	-.003 (.004) VIF=1.416	-.006* (.003) VIF=1.296	-.007** (.003) VIF=1.323
English-language Learners	-.004** (.002) VIF=5.211	-.004*** (.001) VIF=5.223	-.004** (.002) VIF=5.217	-.003*** (.001) VIF=6.532	-.003*** (.001) VIF=6.657
Free/Reduced Lunches	-.002** (.001) VIF=4.882	-.002*** (.001) VIF=4.939	-.002** (.001) VIF=5.033	-.002*** (.001) VIF=6.918	-.002*** (.001) VIF=6.967
Parent Education Level (Ln)	.225*** (.062) VIF=4.584	.416*** (.098) VIF=7.318	.077 (.095) VIF=11.539	.122** (.062) VIF=4.414	.177*** (.056) VIF=7.022
Full Credential (Ln)	.072 (.064) VIF=2.264	.078 (.057) VIF=2.266	.078 (.064) VIF=2.274	.059** (.029) VIF=2.585	.040 (.030) VIF=2.738
Students Per Computer (Ln)	-.002 (.017) VIF=1.344	.013 (.014) VIF=1.434	-.004 (.018) VIF=1.349	-.001 (.007) VIF=1.378	-.002 (.007) VIF=1.388

Total Enrollment (Ln)	.017 (.012) VIF=2.154	.023** (.010) VIF=2.202	.017 (.012) VIF=2.155	-.000 (.011) VIF=2.044	-.003 (.010) VIF=2.085
Fine Arts Sections	.385 (.611) VIF=1.163	.164 (.515) VIF=1.628	.330 (.591) VIF=1.617	.216 (.477) VIF=1.855	4.551** (2.040) VIF=26.563
Foreign Language Sections	.847 (.979) VIF=1.696	.784 (1.025) VIF=1.697	-4.119 (2.858) VIF=24.383	2.362*** (.882) VIF=1.686	2.881*** (.875) VIF=1.848
Health/PE Sections	1.330 (1.106) VIF=1.551	9.784*** (3.473) VIF=20.712	1.550 (1.102) VIF=1.606	-.146 (.507) VIF=1.664	-.091 (.500) VIF=1.668
Other Sections	-.415 (.570) VIF=2.209	-.055 (.453) VIF=2.300	-.318 (.507) VIF=2.237	-.387 (.528) VIF=1.279	-.561 (.505) VIF=1.296
Dummy: Charter?	.018 (.022) VIF=2.180	.020 (.024) VIF=2.181	.023 (.023) VIF=2.193	.035 (.034) VIF=2.320	.026 (.034) VIF=2.358
Dummy: Large City?	-.024 (.023) VIF=2.207	-.011 (.021) VIF=2.250	-.028 (.024) VIF=2.225	.002 (.014) VIF=3.001	.010 (.014) VIF=3.154
Percent Tested (Ln)	.946*** (.240) VIF=3.330	1.024*** (.232) VIF=3.347	.944*** (.242) VIF=3.330	.604* (.310) VIF=2.910	.642** (.312) VIF=2.917
Percent Opted Out	1.11*** (.354) VIF=3.500	1.200*** () VIF=3.509	1.030*** (.344) VIF=3.357	1.07* (.620) VIF=2.558	1.048* (.605) VIF=2.559
SIGNIFICANT INTERACTION EFFECTS					
High School Effect #1: Parent Ed * Health/PE	--	-3.270*** (1.142) VIF=22.133		--	--
High School Effect #2: Parent Ed * Foreign Language	--	--	1.671** (.805) VIF=34.199	--	--
Middle School Effect #1: Parent Ed * Fine Arts	--	--	--	--	-1.422** (.644) VIF=31.710
R-Squared	.831	.853	.836	.921	.923
Adj. R-Squared	.811	.834	.815	.911	.913
F-Statistic (Crit-F = 2.35 @ 90% level)		F = 16.03***	F = 3.57**		F = 2.81*

* Significant at 90% level. ** 95% level *** 99% level. (All two-tailed tests.) -- not significant/relevant for this data set.

[1] VIFs are from pre-heteroskedasticity corrected regressions.

To facilitate direct comparisons (since some variables cannot be converted to logarithms), we transform the linear variables to elasticities. In this form, each variable has the same interpretation (i.e., the percentage change in the dependent variable given a one-percent change in the explanatory variable holding the other variables constant. This translation makes it easier to discern the relative magnitude of the effects for each variable. A second consideration for policy making is the confidence interval surrounding each variable, which reveals the range of possible outcomes. Table 5 provides the elasticities and confidence intervals for all significant main effects along with the necessary formulas to make these calculations. Elasticities for the interaction effects are presented separately in Table 6, as they require a different interpretation.

Table 5: Elasticities and Confidence Intervals – No Interactions

HIGH SCHOOL SAMPLE					
Variable Name (Ln = variable in log form)	**OLS Log-Log Coefficient**	**Elasticity[1]**	**Coefficient Confidence Interval (90%)[2]**	**Elasticity Confidence Interval (90%)[3]**	**Effect of 10% Increase**
Asian	.002*** (.001)	.029	.0016 to .0032	.0197 to .0389	0.29%
Filipino	-.001** (.001)	-.006	-.0026 to -.0003	-.0114 to -.0013	-0.06%
English-language Learners	-.004** (.002)	-.052	-.0068 to -.0012	-.0882 to -.0149	-0.05
Free/Reduced Lunch	-.002** (.001)	-.042	-.0031 to -.0005	-.0715 to -.0118	-0.42%
Parent Education Level (Ln)	.224*** (.062)	.224	.1222 to .3268	.1222 to .3268	2.24%
Percent Tested (Ln)	.946*** (.240)	.946	.5523 to 1.3403	.5523 to 1.3403	9.46%
Percent Opted Out	1.114*** (.354)	2.029	.5321 to .1.6956	.9692 to 3.0888	20.29%
MIDDLE SCHOOL SAMPLE					
Variable Name (Ln = variable in log form)	**OLS Log-Log Coefficient**	**Elasticity[1]**	**Coefficient Confidence Interval (90%)[2]**	**Elasticity Confidence Interval (90%)[3]**	**Effect of 10% Increase**

African American	-.003*** (.000)	-.035	-.0033 to -.0021	-.0433 to -.0273	-0.35%
American Indian	-.004** (.002)	-.004	-.0069 to -.0009	-.0065 to -.0009	-0.04%
Asian	.002*** (.000)	.023	.0014 to .0027	.0158 to .0301	0.23%
Pacific Islander	-.006* (.003)	-.024	-.0110 to -.0002	-.0465 to -.0009	-0.24%
English-language Learners	-.003*** (.001)	-.040	-.0041 to -.0015	-.0584 to -.0212	-0.40%
Free/Reduced Lunch	-.002*** (.001)	-.051	-.0024 to -.0008	-.0785 to -.0241	-0.51%
Parent Education Level (Ln)	.122** (.062)	.122	.0205 to .2231	.0205 to .2231	1.22%
Full Credential (Ln)	.059** (.029)	.059	.0114 to .1074	.0114 to .1074	0.59%
Foreign Language Sections	2.362*** (.882)	.012	.9120 to 3.8122	.0044 to .0186	0.12%
Percent Tested (Ln)	.604* (.310)	.604	.0945 to 1.1140	.0945 to 1.1140	6.04%
Percent Opted Out	1.071* (.620)	.759	.0515 to 2.0897	.0365 to 1.4824	7.59%

* Significant at the 90% level. ** Significant at the 95% level *** Significant at the 99% level. (All two-tailed tests.).

Formulas:

[1] Coefficient x Mean (if linear); otherwise coefficient = elasticity (if logged);

[2] Coefficient +/- (Standard Error * Critical-t); critical-t f= 1.684;

[3] Point Confidence Intervals * Mean (if linear); otherwise, interval for elasticity = interval for coefficient (if logged).

Once parental education interactions are added to the model, the magnitudes of impact at the mean parent education level are still relatively small (see Table 6). The implication is that increasing sections in the average school is likely to have a negligible effect. However, when a range of predictions are made at different parental education levels, it becomes clear that the direction and magnitude of the effects vary significantly along the parent education continuum. For example, increasing Health/PE Sections Per Student by 20% (about 5 sections) in a high school with an average parent education level of 2.0 (equivalent to "high school graduate") is predicted to raise API scores by 1.48% or about 10 points.[13] On the other hand, at an average parent education level of "some college" or higher, the increase is completely extinguished and becomes increasingly negative as the education level rises. A similar pattern is seen with respect to Fine Arts Sections in middle schools, though the magnitude is lower.

Table 6: Elasticities for Significant Interaction Effects Only (range of predictions for assumed values of a school's average parent education level)

HIGH SCHOOL SAMPLE								
	OLS Log-Log Coefficient	**Elasticity[1]**	**Effect of 10% Increase**	**Effect of 20% Increase**	**OLS Log-Log Coefficient**	**Elasticity[1]**	**Effect of 10% Increase**	**Effect of 20% Increase**
Predicted Value by Ave. Parent Education Level	**Effect #1: Health/PE Sections Per Student**				**Effect #2: Foreign Lang. Sections Per Student**			
Mean Level (3.15)	-0.516	-.012	-0.12%	-0.24%	1.065	.025	0.25%	0.50%
Range of Predictions:								
2.0 (High School Grad)	3.244	.074	0.74%	1.48%	-0.857	-.020	-0.20%	-0.40%
3.0 (Some College)	-0.026	-.001	-0.01%	-0.20%	0.814	-.019	-0.19%	-0.38%
4.0 (College Grad)	-3.296	-.075	-0.75%	-1.50%	2.485	.058	0.58%	1.16%
5.0 (Graduate School)	-6.566	-.150	-1.5%	-3.00%	4.156	.096	0.95%	1.92%
MIDDLE SCHOOL SAMPLE								
Predicted Value by Ave. Parent Education Level	**Effect #1: Fine Arts Sections Per Student**				**No Other Significant Effect**			
Mean Level (3.09)	0.157	.002%	0.02%	0.04%	N/A			
Range of Predictions:								
2.0 (High School Grad)	1.707	.020%	0.20%	0.40%				
3.0 (Some College)	0.285	.003%	0.03%	0.06%				
4.0 (College Grad)	-1.137	-.013%	-0.13%	-0.26%				
5.0 (Graduate School)	-2.559	-.030%	-0.30%	-0.60%				

Formula: [1] Coefficient x Mean

The opposite effect occurs with Foreign Language Sections in high schools: an increase in the number of sections per student exerts a negative effect until parent education level reaches an average of "college graduate" or higher. Since the main effects are not significant, further testing is needed to confirm this result. If the effect is confirmed, it would suggest that implementing a mandatory foreign language requirement (as implied by the default "A-G" curriculum advocated by California's Superintendent of Schools (O'Connell 2004) is likely to have a depressive effect on API scores in low-performing schools which tend to serve students from less educated households. This is not meant to imply such a proposal has no merit, only that the API as a litmus test does not bear out this strategy.

It should be noted there are several other variables that show a large, significant relationship to API, including Parent Education Level (in isolation of curricular interactions), Percent Tested and Percent Opted Out of testing. From a policy perspective, changes to these factors seem to offer a more certain payoff. Yet, none offers a viable strategy for increasing scores. Raising parental education level holds promise as a policy option – and schools have a legitimate role to play here – but any effort to impact this variable is clearly a long-term proposition. Interestingly, the two test participation controls exhibit the strongest impacts of all. However, any attempt to be strategic about influencing these factors to manipulate test scores at the school-level crosses an ethical line. (Although, human behavior would suggest that this sometimes occurs.) This is not to say there is no case to be made for systematically instituting different assessments or thresholds for certain segments of the student population (such as special education students) as some have passionately argued, only that such changes should be consistent to allow for meaningful comparisons across schools.

As the results demonstrate, some non-core inputs do matter with respect to API scores. Specifically, parental education moderates the effect of three non-core subjects: Health/PE, Foreign Language and Fine Arts. In terms of impact, the effect of curricular enrichment on test scores is relatively small, though the presence and magnitude of particular subject-matter effects are highly sensitive to school composition. The major implication of this finding is that curricular inputs must be used strategically to achieve the desired outcome. Notably, we predict that the less academic forms of enrichment will enhance scores only in schools serving disadvantaged student populations (as defined by a low parent education level).

VI. Conclusion and Policy Implications

The purpose of this study is to assess the effects of curricular enrichment on a school's Academic Performance Index using multivariate regression techniques. The results show a significant positive relationship between one subject category tested (Foreign Language Sections) and school-wide performance at the middle school level. One interpretation of the singular main effect is that the case for a causal link is strongest for only the most academic of subjects considered. However, an examination of the interactions between a school's average parent education level and the number of non-core sections per student across a variety of subjects adds an important dimension to the story: enrichment matters more under some conditions. Specifically, a "parental education effect" is evident for Health/PE and Foreign Language Sections at the high school level and Fine Arts at the middle school level.

The intention here is not to test or prescribe a specific reform initiative for secondary schools, but to determine if there is any empirical basis to give pause to the incremental changes in curricular enrichment occurring partly as a consequence of the emphasis on accountability.

With that framework in mind, the results support the hypothesis that some forms of curricular enrichment do matter with respect to school-wide standardized test scores under certain conditions. From an intuitive perspective, these findings seem all too obvious. Yet, there is little evidence in the academic literature to support a causal link between curricular enrichment and API scores, particularly in relation to school-level variables under the influence of school professionals. Provision of curricular enrichment opportunities does not necessarily come at the expense of higher test scores and, under some conditions, may actually improve scores. Still, in the absence of a comparative study to gauge the relative benefits of increasing core or non-core subjects, applying these findings to educational policy and practice requires some degree of restraint.

Ultimately, this study should not be interpreted as supporting the efficacy of a particular policy proposal; rather, the findings are more suggestive of some general themes regarding curricular design and school accountability.

✓ **Some cause for concern about a one-size-fits-all approach to curricular design.** There is little evidence here to advocate a position regarding the role of enrichment in the core curriculum. With the exception of Foreign Language Sections in middle schools, no other non-core subject exhibits a significant overall effect on test scores. Yet, there is a valuable lesson to draw from the interaction between student characteristics and enrichment course offerings. If different populations have different needs or benefit disproportionately from exposure to a given subject, a strategic approach to the curricular mix is likely to be more effective.

✓ **A need to explore different intervention strategies.** In contrast to the role of enrichment in the core curriculum, this study has more direct implications for academic intervention. The findings lend support to the notion that broader access to certain non-core subjects is most beneficial in schools serving students from less supportive or less learning-conducive households – precisely the schools where enrichment is being cut in an effort to combat low performance. The fact that manipulation of the curricular mix in these presumably "non-essential" subjects can positively impact school-wide test scores is an important outcome, particularly in view of other studies showing higher rates of obesity and lower rates of extracurricular enrichment among disadvantaged students. Together, these findings offer empirical evidence to give pause to a tendency to over-prescribe participation in core subjects for low performing or at-risk students if such an intervention comes at the expense of efforts to improve physical health, enrich the mind and (potentially) create a stronger attachment to the school experience. Perhaps the provision of some forms of enrichment – coupled with remedial education – constitutes a more effective strategy for ensuring student success in the long run? Rather than pursue one limited approach and await the outcome (though, such untested pendulum swings are typical of public education), a more concerted effort to evaluate different intervention strategies has more potential to enlighten policy and practice.

✓ **Limited gains in test scores are likely, though a targeted approach to provision of curricular enrichment improves the odds.** Despite the positive causal links between some forms of curricular enrichment and test scores (under some conditions), it must be noted that the potential gains are small and a strategic approach is required to achieve the desired effect. Thus, enhancing enrichment opportunities in all schools holds little promise as a strategy for boosting test scores. This is not entirely unexpected. Perhaps the bigger message here is that there is little evidence that non-academic forms of enrichment are a significant drag on scores. If it can be demonstrated that access to curricular enrichment has other benefits – such as enhanced motivation, retention or propensity to continue one's education – the case for enrichment is that much stronger.

Still, there may be other policy options that hold more promise for improving student achievement, such as increasing student readiness through targeted or universal preschool, expanding after-school intervention programs or providing incentives to draw more experienced teachers to low performing schools. If money is a limiting factor – and it usually is – initiatives with a greater potential payoff deserve a higher priority for public funding. Nevertheless, short-term fiscal constraints need not limit the long-term vision for K-12 education. If the goal is a vibrant, well-rounded experience that develops critical competencies and stimulates a love of lifelong learning that is essential to society's long-term productivity, there are ways to expand curricular options while minimizing costs. (This hearkens back to Hanushek, who suggests it is how money is spent that matters). One way is through greater use of institutional mechanisms such as magnet, charter and small schools. Another is through targeted institutional incentives, such as earmarking some portion of existing federal vocational funds to support more coordinated vocational sequences.[14] Beyond fiscal considerations, advancing the long-term vision also requires attention to accountability.

✓ **A need to align accountability mechanisms with the broader goals of public education**. A vital lesson here is that incentives matter, thus attention must be given to the incentives inherent in the accountability system and how well the resulting behaviors are aligned with the broader goals of public education. Absent such alignment, the drive to boost test scores in core subjects may induce curricular choices that are counterproductive under some conditions as well as unintended by policy makers and parents.

Future Research Needs

Recognizing that one limited study does not provide sufficient grounds for policymaking, the following are some thoughts about future research avenues in regards to this issue. One area for further exploration – which may be even more apropos in the context of high schools – is the effect of curricular enrichment on retention. Even if some forms of enrichment have little impact on test scores, students who are more engaged and interested in what they are learning may be more likely to graduate. In fact, the motivational gains may be particularly beneficial to those who struggle with core subjects and are at risk of failure. To that end, one obvious follow-up is a cross-sectional analysis using dropout rates as the dependent variable to isolate the role of curricular factors in the retention function.

A second area for further study is the interaction between non-core sections and other inputs. While the regression results presented here suggest that the primary benefit of certain inputs – such as physical education and fine arts – may be to schools with less educated parents, exploring these and other interactions in a more narrowly defined study with a larger sample could reveal other patterns undetected here. The analysis could also be strengthened by the use of more sophisticated techniques for modeling interactions between variables, including centering variables, testing for variations in the effects using polynomial forms, and exploring the presence of multi-way interactions.

Finally, there is a need to better understand the relative merits of different intervention strategies for low performers. Given the evidence that students from less learning-conducive households may benefit actually most from exposure to certain non-core subjects, more experimentation and comparative analyses are needed to support an evidence-based approach to intervention that is sensitive to student needs.

Clearly, the book is not yet closed in terms of understanding how curricular enrichment interacts with student success. While there is evidence here to suggest a positive causal link between API scores and some forms of curricular enrichment under some conditions, more research is needed to resolve the issue. In the meantime, now is the time to bring more clarity with respect to the goals of public education and the system by which progress toward those goals is measured. As part of that visioning process, the interaction between curricular enrichment and student success – in the broader sense of the term – is an issue that deserves further consideration.

REFERENCES

Betts, J. (1997). "*The role of homework in improving school quality*" cited in Betts, J. and Costrell, R. (2000, August). "*Incentives and Equity Under Standards-Based Reform*," University of California, San Diego, Discussion Paper 2000-20.

California Department of Education. 2002, December 10. *"State Study Proves Physically Fit Kids Perform Better Academically."* CDE, 2002. Retrieved from http://health.ocde.us/physed.asp

California Department of Education. 2004, March. "Summary of the 2003 API Base."

Catterall, J., Chapleau, R. and Iwanaga, J. 1999. *"Involvement in the Arts and Human Development: General Involvement and Intensive Involved in Music and Theatre Arts."* The Imagination Project, UCLA. Retrieved from http://www.gseis.ucla.edu/faculty/publications/100535351554386100.pdf

Coates, D. 2003. "Educational Production Functions Using Instructional Time as an Input." *Education Economics,* vol. 11, no. 3.

Cooper, T. 1987. "Foreign Language Study and SAT Verbal Scores." *Modern Language Journal*, vol. 71, no. 4, pp. 381-387.

Coulton, C. and Chow, J. 1993. "Interaction Effects in Multiple Regression." *Journal of Social Service Research,* vol. 16, no. 1/2, pp. 179-199.

EdSource Online. 2005. Standards and Curriculum Overview. Retrieved from http://www.edsource.org/edu_sta.cfm on August 1, 2005.

Elliott, M. 1998, July. "School Finance and Opportunities to Learn: Does Money Well Spent Enhance Students' Achievement?" *Sociology of Education,* vol. 71, no. 3, p. 223-245.

Fordham Foundation. 2005. *The State of State Math Standards 2005* and *The State of State English Standards 2005*, *The State of Science Standards 2005.* Retrieved on February 15, 2005 from http://www.edexcellence.net/foundation/publication/index.cfm?topic=Testing%20%26%20Accountability

Gardner, H. 1983. *Frames of Mind: The Theory of Multiple Intelligences.* USA: Basic Books.

Hanushek, E. 1989, May. "The Impact of Differential Expenditures on School Performance." *Educational Researcher,* v. 18.

____. 1996. "Measuring Investment in Education." *Journal of Economic Perspectives,* vol. 10, no. 4, Autumn, pp. 9-30.

Hedges, L., Laine, R. and Greenwald, R. 1994, April. "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher,* v. 23.

Hetland, L. and Winner, E. 2001, May. "The Arts and Academic Achievement: What the Evidence Shows," *Arts Education Policy Review,* vol. 102, no. 5, pp. 3-6.

Hughes, K., Bailey and T., Karp, M. 2002, December. "School-to-Work: Making a Difference in Education," *Phi Delta Kappan,* vol. 84, issue 4.

Kostakis, Anastasia. 1987, October. "Differences Among School Outputs and Educational Production Functions." *Sociology of Education,* vol. 60, pp. 232-241.

Kozol, J. 1991. *Savage Inequalities*. New York, NY: Crown Publishers, Inc.

Lamdin, D. 1995. "Testing for the effect of school size on student achievement within a school district." *Education Economics,* vol. 3, issue 1, p. 33.

Lee, V. and Bryk, A. 1988. "Curriculum Tracking as Mediating the Social Distribution of High School Achievement." *Sociology of Education,* vol. 61, no.2., pp. 78-94.

Lee, V. and Burkham, D. 2001. *"Dropping Out of High School: The Role of School Organization and Structure,"* Cambridge, MA: Harvard Civil Rights Project.

Matts, T. 2005. Memo to Chief State School Officers, May 5, 2005.

McNeal, Jr., R. 1995, January. "Extracurricular Activities and High School Dropouts," *Sociology of Education,* vol. 68, January, pp. 62-81.

National Science Teachers Association (NSTA). 2000. NSTA Position Statement. Accountability. http://www.nsta.org/positionstatement&psid=1

O'Connell, Jack. 2004. Give all high school students course loads of college-bound. *Sacramento Bee*. March 14.

Olsen, S. and Brown, L. 1992, Spring. "The Relationship between High School Study of Foreign Languages and ACT English and Mathematics Performance," *ADFL Bulletin,* v. 23, no. 3, pp. 47-50. Retrieved from http://www.adfl.org/bulletin/v23n3/233047.htm

Posnick-Goodwin, S. 2005, June. "Is it curtains for the arts in California's public schools?" *California Educator,* vol. 9, issue 9, pp. 6-13.

_____. 2005, April. "High schools on the hot seat." *California Educator,* vol. 9, issue 7. Retrieved from http://www.cta.org/CaliforniaEducator/v9i7/Feature_1.htm

RAND Corporation. 2003. *"How California Charter Schools Operate and Perform."* RAND Research Brief. Retrieved from http://www.rand.org/publications/RB/RB8022/RB8022.pdf

Ready, D., Lee, V. and Welner, K. 2004, October. "Educational Equity and School Structure: School Size, Overcrowding and Schools-Within-Schools." *Teachers College Record,* vol. 106, issue 10, p 1989.

Sallis, J. and McKenzie, T. 1999, June. "Effects of health-related physical education on academic achievement: Project SPARK," *Research Quarterly for Exercise & Sport,* vol. 70, no. 2, pp. 127, 8 p.

Shephard, R. 1997. "Curricular Physical Activity and Academic Performance," *Pediatric Exercise Science,* 9, pp. 113-126.

Stanford University Bridge Project. 2004, March 4. *"Analyzing State and Regional K-16 Policies: A Toolkit for Researchers and Policy Analysts."* Retrieved from http://www.stanford.edu/group/bridgeproject/policytoolkit/index.html

Stock, J. and Watson, M. 2003. *Introduction to Econometrics.* Boston, MA: Addison Wesley.

Studenmund, A. 2001. *Using Econometrics: A Practical Guide.* U.S.: Addison Wesley Longman, Inc.

Walberg, H., Fraser, B. and Welch, W., 1986, January-February. "A Test of a Model of Educational Productivity Among Senior High School Students," *Journal of Educational Research,* vol. 79, no. 3, pp. 133-39.

Warren, Paul. 2005. *"Improving High School: A Strategic Approach."* Legislative Analyst's Office, Sacramento, CA. Retrieved from http://www.lao.ca.gov/PubDetails.aspx?id-1322

Wassmer, R. 2002, 2004. *Lessons from California's Public Elementary Schools Where Performance Exceeds Expectations*. California: Senate Office of Research. (Full report: 2002, Journal version: 2004). Retrieved from http:/www.csus.edu/indiv/w/wassmerr/wpaperou.htm

ENDNOTES

[1] First, note that "core" is used in a specific way here to denote only the four academic subjects measured in the API. Thus, the use of the term enrichment to describe non-core courses is simply a means of differentiating non-core subjects from the core subjects reflected in the API. It is *not* meant to imply that core courses cannot be enriching. Finally, it should also be noted that foreign languages and fine arts are technically considered "core" subjects in California; however, neither is covered in state testing nor is specifically required for graduation. (Currently, students may choose one course in *either* subject.)

[2] The report, *A Nation at Risk,* was published in 1983 by the National Commission on Excellence in Education, created at the behest of the U.S. Secretary of Education. For a complete copy, see http://www.ed.gov/pubs/NatAtRisk/index.html.

[3] A more complete overview of standards-based reform (from which this quote is taken) is available from EdSource Online, 2005.

[4] The Academic Performance Index (API) is comprised of a norm-referenced test (formerly the SAT-9, currently the CAT-6), a series of California Standards Tests or CSTs (one for each core subject), and the California High School Exit Exam. (Source: California Department of Education, Summary of the API Base, 2004). Norm-referenced tests are referenced to a national average (for predictive validity), while standards-based tests are criterion-referenced (for content validity).

[5] The decline in some subjects is more pronounced at the elementary level, but secondary schools are not immune. For example, during the period between 1999-00 and 2003-04, the number of physical education courses offered per student dropped 6% in middle schools and 7% in high schools statewide. Middle schools have also seen a steep decline in the number of fine arts and foreign language courses offered per student (down 16-17%, respectively).

[6] Note that "core" is used in a very specific way in this analysis, defined here as the four subjects on which students are tested for purposes of calculating the API (i.e., language arts, math, science, social science).

[7] Consideration was also given to using a measure of "depth" (i.e., the number of unique choices per subject area). While this is an interesting angle on the topic, the present study is more concerned with the level of participation/access. For an interesting discussion on depth and the arts, see Catterall, et.al., 1999.

[8] Since there is some variability in middle school configurations, middle schools in this sample were limited to those with the most common grade configuration (grades 6-8).

[9] Some consideration was given to including magnet status as an institutional control, but the available data set was outdated (6 years old) and appeared unreliable. Year-round status was also excluded as none of the high schools studied operate on such a schedule. The population

density variable ultimately reflects only "large city" schools since no other population density category was significant during preliminary testing. One other geographic control was tested (county). Only San Francisco County was significant, but this effect was extinguished after adding the large city dummy and correcting for heteroskedasticity.

[10] As Studenmund explains, "The variance inflation factor (VIF) is a method of detecting the severity of multicollinearty by looking at the extent to which a given variable can be explained by all the other explanatory variables in the equation" (Studenmund, 2001, p. 256-7). A VIF > 5 indicates severe multicollinearity.

[11] When heteroskedasticity is present, OLS misrepresents the variances and standard errors making the t-scores unreliable for hypothesis testing (Studenmund, 2001).

[12] The t-statistic is -4.230, which is significant at the 99% level; for details, see Appendix B.

[13] In an average-size high school with a parent education level of 2.0, a 20% increase in sections per student (from .0228 to .0273) results in 5 new sections with a commensurate increase of 10 points on the API.

[14] Credit for this latter recommendation goes to Paul Warren at the LAO (Warren, 2005, p. 47). Redirection of state funds is another viable source. To that end, Governor Schwarzenegger signed SB70 allocating $20 million from existing Prop. 98 funds toward expanded vocational programs. (For a copy of the bill, see www.leginfo.ca.gov)