



Cornell University  
ILR School

Cornell University ILR School  
**DigitalCommons@ILR**

---

Articles and Chapters

ILR Collection


---

3-1-1998

## The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement

John H. Bishop  
*Cornell University*, [jhb5@cornell.edu](mailto:jhb5@cornell.edu)

Follow this and additional works at: <https://digitalcommons.ilr.cornell.edu/articles>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Human Resources Management Commons](#), and the [Labor Relations Commons](#)

**Thank you for downloading an article from DigitalCommons@ILR.**

**Support this valuable resource today!**

---

This Article is brought to you for free and open access by the ILR Collection at DigitalCommons@ILR. It has been accepted for inclusion in Articles and Chapters by an authorized administrator of DigitalCommons@ILR. For more information, please contact [catherwood-dig@cornell.edu](mailto:catherwood-dig@cornell.edu).

If you have a disability and are having trouble accessing information on this website or need materials in an alternate format, contact [web-accessibility@cornell.edu](mailto:web-accessibility@cornell.edu) for assistance.

---

# The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement

## Abstract

[Excerpt] Two presidents, the National Governors Association, and numerous blue-ribbon panels have called for the development of state or national content standards for core subjects and examinations that assess student achievement of these standards. The Competitiveness Policy Council, for example, advocated that "external assessments be given to individual students at the secondary level and that the results should be a major but not exclusive factor qualifying for college and better jobs at better wages." It is claimed that curriculum-based external exit exam systems (CBEEESs) based on explicit content standards will improve the teaching and learning of core subjects. What evidence is there for this claim? Outside the United States, such systems are the rule, not the exception. What impacts have such systems had on school policies, teaching, and student learning?

## Keywords

ILR, center, human resource, job, worker, advanced, labor market, satisfaction, employee, work, manage, management, training, HRM, employ, model, industrial relations, examination, teacher, student, parent, school, achievement, wage, United States

## Disciplines

Educational Assessment, Evaluation, and Research | Human Resources Management | Labor Relations

## Comments

### Suggested Citation

Bishop, J. H. (1998). The effect of curriculum-based external exit exam systems on student achievement [Electronic version]. *Journal of Economic Education*, 29, 171-182.  
<http://digitalcommons.ilr.cornell.edu/articles/24>

### Required Publisher Statement:

*Journal of Economic Education*, "The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement," John H. Bishop, Vol. 29, #2, pp. 171-182, Spring 1998. Reprinted with permission of the Helen Dwight Reid Educational Foundation. Published by [Heldref Publications](#), 1319 18th Street, NW, Washington, DC 20036-1802. Copyright © (1998).

# The Effect of Curriculum-Based External Exit Exam Systems on Student Achievement

John H. Bishop

Two presidents, the National Governors Association, and numerous blue-ribbon panels have called for the development of state or national content standards for core subjects and examinations that assess student achievement of these standards. The Competitiveness Policy Council (1993, 30), for example, advocated that "external assessments be given to individual students at the secondary level and that the results should be a major but not exclusive factor qualifying for college and better jobs at better wages." It is claimed that curriculum-based external exit exam systems (CBEEESs) based on explicit content standards will improve the teaching and learning of core subjects. What evidence is there for this claim? Outside the United States, such systems are the rule, not the exception. What impacts have such systems had on school policies, teaching, and student learning?

## WHAT IS A CURRICULUM-BASED EXTERNAL EXIT EXAMINATION SYSTEM?

A curriculum-based external exit examination system has several characteristics:

1. It produces signals of student accomplishment that have real consequences for the student.
2. It defines achievement relative to an *external* standard, not relative to other students in the classroom or the school. Thus, fair comparisons of achievement across schools and across students at different schools are possible. Costrell's (1994) analysis of the optimal setting of education standards concluded that a more centralized setting of standards (state or national achievement exams) results in higher standards, achievement, and social welfare than a decentralized setting of standards (i.e., teacher grading or schools' graduation requirements).
3. It is organized by discipline and keyed to the content of specific course sequences, which focuses responsibility for preparing the student for particular exams on one or a small group of teachers.
4. It signals multiple levels of achievement in the subject. If only a pass-fail signal is generated by an exam, the standard will have to be set low enough to allow almost everyone to pass and this will not stimulate the great bulk of students to greater effort (Kang 1985; Costrell 1994).

---

*John H. Bishop* is an associate professor of economics at Cornell University (e-mail: [jhb5@cornell.edu](mailto:jhb5@cornell.edu)).

5. It covers almost all secondary school students. Exams for elite schools, advanced courses, or college applicants will influence standards at the top of the vertical curriculum but will probably have limited effects on the rest of the students. The school system as a whole must accept responsibility for how students do on the exams. A single exam taken by all is not essential. Many nations allow students to choose which subjects to be examined in and offer high- and intermediate-level exams in the same subject.

Commercially prepared achievement tests such as the CAT, CTBS, ITBS, and ITED are not CBEEES, because they fail the first characteristic. Students have no stake in doing well on these tests. Where stakes are attached to student performance, teachers and school administrators experience the consequences, not individual students.

The minimum competency exams that many American states require students to pass to graduate from secondary school are not a CBEEES because they fail characteristics three and four. These tests focus on basic skills taught in primary and the lower grades of secondary schools. Although minimum competency exams have apparently reduced the numbers of students with very low basic skills levels (Lerner 1991), the passing standard is quite low. The tests are typically taken in 9th or 10th grade, and most students pass on the first sitting. High school transcripts indicate only whether the student eventually passes the test, not achievement levels above the minimum. Thus, for the great majority of students who pass them on the first try, the tests do not stimulate further study. Incentive effects are focused on the small number who fail on the first try and must repeat the test. Minimum competency exams can be a useful part of a CBEEES, but other more demanding curriculum-based exams requiring higher levels of performance are essential.

Characteristic 4 is essential because of its impact on the incentive effects of exams. By age 13, students differ dramatically in their levels of achievement. On the National Assessment of Educational Progress test, 7–9 percent of 13-year-olds were four or more grade-level equivalents (GLE) behind students of their age, and 15–17 percent were four or more grade-level equivalents ahead of students of their age. With achievement differentials this large, incentives for effort are stronger for most students if the full range of achievement is signaled rather than just whether the individual has passed some absolute standard. When only a pass-fail grade is generated by a test, many students pass without exertion and are, thus, not stimulated to greater effort by the reward for passing. Some of the least well prepared students will judge the effort required to achieve the standard to be too great and the benefits too small to warrant the effort, and they give up on meeting the standard. Few students will find the reward for exceeding a single absolute cutoff an incentive for greater effort (Kang 1985). Costrell agrees: “The case for perfect information [making scores on external examinations available rather than just whether the individual passed or failed] would appear to be strong, if not airtight: for most plausible degrees of heterogeneity, egalitarianism, and pooling under decentralization, perfect information not only raises GDP, but also social welfare” (1994, 970).

The SAT-I reasoning tests are not CBEEESs because they do not meet characteristics three and five. They are not organized around school subjects and fail to assess most of the material—history, science, economics, civics, literature, foreign languages, and the ability to write an essay—that high school students are expected to learn. The SAT was designed to minimize any effects on high school teaching and student study habits. Richard Gummere, Harvard College's admissions director at the time the machine-scored multiple-choice SAT replaced the curriculum-based essay-style College Board Examinations, was candid about why the SAT had been adopted: "Learning in itself has ceased to be the main factor [in college admissions]. The aptitude of the pupil is now the leading consideration" (Gummere 1943, 5). The subject-specific SAT-II achievement tests meet some of the requirements of a CBEEES but, because colleges admit on the basis of the SAT-I, not SAT-II, the stakes are low and few students take them.

The Advanced Placement (AP) examinations are the one exception to the generalization that the United States lacks a national CBEEES. Although growing rapidly, AP is still a very small program. In 1995, only 3.2 percent of juniors and seniors took AP English or AP history exams, and only 2 percent took AP calculus or science exams (National Education Goals Panel 1995). Low participation means that AP exams fail characteristic five and are, consequently, not a CBEEES. They can, however, serve as a component of a larger system.

### **WHY ARE CBEEESs HYPOTHESIZED TO INCREASE ACHIEVEMENT?**

A CBEEES improves the signaling of academic achievement. As a result, colleges and employers are likely to give greater weight to academic achievement when they make admission and hiring decisions, so the rewards for learning should grow and become more visible. A CBEEES also shifts attention toward measures of absolute achievement and away from measures of relative achievement such as class rank and teacher grades.

Grading on a curve or basing college admissions on class rank gives students a personal interest in persuading each other not to study. The serious student makes it more difficult for others to get top grades. When exams are graded on a curve, joint welfare is maximized when no one studies. Side payments of friendship and respect and punishments of ridicule and harassment enforce the cooperative "don't study" solution. When learning is assessed relative to an outside standard, students no longer have a personal interest in getting teachers off track or persuading each other to refrain from studying.

In the absence of a CBEEES, school reputations are largely outside the control of school staff; they are determined instead by the social class of the student body, mean SAT scores, and numbers attending prestigious colleges. When a CBEEES is in place, exam results displace social class as the primary determinant of school reputations and, in turn, school staff are induced to give enhanced learning higher priority. Teachers will upgrade curricula and assign more homework and parents will demand better science labs and more rigorous teaching. School administrators will be pressured to increase the time devoted to examination subjects and hire more qualified teachers. The competition for teachers is likely to bid up salaries.

A further benefit of a CBEEES is the professional development that teachers receive when they are brought to centralized locations to grade the extended-answer portions of the exams. In May 1996, I interviewed a number of Alberta teachers about their experiences as members of grading committees. They all said that having to discuss and agree with their colleagues about what constituted an excellent, good, adequate, poor, and failing response to essay questions had been their best professional development experience since entering the profession.

## **DO CBEEESs INCREASE ACHIEVEMENT? A LOOK AT THE EVIDENCE**

I tested the hypothesis that CBEEESs improve achievement by comparing nations, states, and provinces that do and do not have such systems. Three different data sets were examined: science and mathematics achievement in the 40-nation Third International Math and Science Study (TIMSS), science and math scores on the International Assessment of Educational Progress (IAEP) for 16 nations, and SAT test scores for New York State versus other states.

### **Third International Mathematics and Science Study**

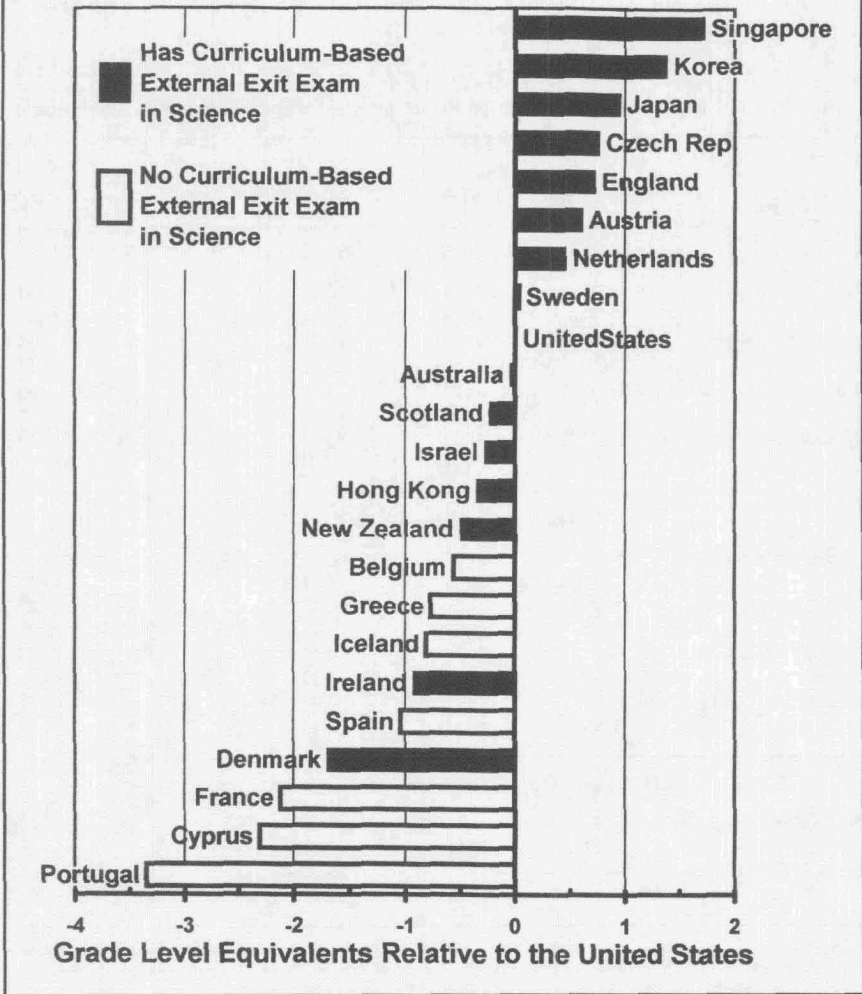
The just-released TIMSS provides 1994–95 data for seventh and eighth graders in 40 countries. I reviewed comparative education studies, government documents, and education encyclopedias and interviewed embassy personnel and Cornell graduate students from the countries to determine which of the TIMSS nations have CBEEEs in secondary school.<sup>1</sup> Twenty-two national school systems were classified as having CBEEEs for math and science in all parts of the country: Austria, Bulgaria, Columbia, Czech Republic, Denmark, England, Hong Kong, Hungary, Ireland, Iran, Israel, Japan, Korea, Lithuania, the Netherlands, New Zealand, Russia, Scotland, Singapore, Slovak Republic, Slovenia, and Thailand. Three countries—France, Iceland, and Romania—had CBEEEs in mathematics but not in science. Five countries—Australia, Canada, Germany, Switzerland, and the United States—had CBEEEs in some states or provinces but not in others. Norway had regular exit examinations in mathematics but examined science only every few years. Latvia was assigned a .5 on the CBEEES variable because it was in transition away from a CBEEES. The countries having a CBEEE in either subject were Belgium (both Flemish- and French-speaking systems), Cyprus, Greece, the Philippines, Portugal, Spain, and Sweden. The university entrance examinations in Greece, Portugal, Spain, and Cyprus, and the ACT and SAT in the United States were not considered to be CBEEESs (Madeus and Kelleghan 1991). University entrance exams should have much smaller incentive effects because students headed into the job market do not take them and teachers can avoid responsibility for their students' exam results by arguing that not everyone is college material or that examiners set unreasonably high standards to limit enrollment in higher education.

Twenty-three of the TIMSS countries are ranked by the science achievement of their 13-year-olds in Figure 1.<sup>2</sup> The United States ranks number 9 in science.

The gaps between the vertical grid lines represent one United States grade-level equivalent—the difference between seventh and eighth grade TIMSS test score means for the United States. Achievement differentials across nations are very large. In science, Singapore and Korea are more than one GLE ahead of the United States, and Portugal, Cyprus, and France are more than two GLEs behind.

The mean seventh and eighth grade science and mathematics test scores for all 40 TIMSS countries were regressed on per capita gross domestic product (GDP) for 1987 and 1990, deflated by a purchasing power parity price index, a dummy for East Asian nations, and a dummy for CBEEES. The results indicate that test

**FIGURE 1**  
TIMSS Science Achievement at Age 13: Children of High School Graduates



scores were significantly higher in nations with higher per capita product, East Asian nations, and nations with a CBEEES (Table 1).

The analysis of achievement at a particular grade level may have been biased, however, by differing policies regarding grade retention, age of school entry, and which grade was chosen for assessment. CBEEESs, for example, might have been associated with high rates of grade retention. Therefore, a preferable dependent variable is a measure of student achievement at some fixed age. The third and fourth rows of the first two panels in Table 1 present estimates of the median test score for each nation's 13-year-olds (Beaton et al. 1996a; Beaton et al. 1996b, Table 1.5). For countries not included in this table, the 13-year-old median was estimated by age adjusting the seventh and eighth grade means.<sup>3</sup> Switching to the age-constant achievement reduced the estimated impact of a CBEEES. The CBEEES coefficient had a *p* value of .08 (two-tailed test) in the mathematics model and a *p* value of .01 (two-tailed test) in the science model. The estimated impacts are substantively important: 1.3 U.S. GLEs in science and 1.0 U.S. GLEs in mathematics.

**TABLE 1**  
The Effect of Curriculum-Based External Exams on Science and Mathematics Achievement

Variable	External exit exam	LnGDP/Pop 1987 & 90	East Asia	K-12 spending /GDP	Adj R <sup>2</sup> RMSE
<i>TIMSS Science-1994 (U.S. GLE = 26)</i>					
Mean, 7th graders	38.0*** (2.93)	33.8*** (3.44)	20.1 (1.24)		.317 35.4
Mean, 8th graders	42.4*** (3.40)	36.2*** (3.80)	14.4 (.92)		.364 34.4
Median, 13-yr.-olds	34.9*** (.277)	45.0*** (4.68)	21.5 (1.35)		.402 34.7
Median, 13-yr.-olds	32.0*** (2.57)	38.0*** (3.71)	33.7* (2.01)	13.6* (1.86)	.442 33.9
<i>TIMSS Math-1994 (U.S. GLE = 24)</i>					
Mean, 7th graders	29.6** (2.09)	46.6*** (4.60)	66.0*** (4.01)		.469 36.2
Mean, 8th graders	36.0** (2.54)	48.7*** (4.81)	62.0*** (3.75)		.476 36.6
Median, 13-yr.-olds	24.7* (1.82)	56.0*** (5.77)	69.4*** (4.37)		.537 35.1
Median, 13-yr.-olds	21.5 (1.55)	53.9*** (5.07)	75.9*** (4.41)	5.7 (.75)	.545 35.1
<i>IAEP-1991 (U.S. GLE = 6 for Science and 8 for Math)</i>					
Science % correct	4.3 (1.72)	1.7 (.61)	9.6** (2.81)		.436 4.0
Math % correct	15.7*** (3.85)	3.7 (.25)	16.1** (2.81)		.641 6.0

Note: Numbers in parentheses are *t* values. TIMSS = Third International Math and Science Study. GLE = grade-level equivalent. IAEP = International Assessment of Educational Progress.

\**p* < .01, two-tailed test; \*\**p* < .05, two-tailed test; \*\*\**p* < .001, two-tailed test.



CBEEESs may improve achievement by inducing greater social investments in education. Row 4 (Table 1) presents results of regressions that add the share of GDP spent on education to the standard model. Coefficients on this variable were positive for both outcomes and significantly so for science. The estimated impact of spending was modest, however. A one percentage point increase in the share of GDP devoted to education increased the science achievement of 13-year-olds by one-half a GLE.

### **Analysis of the 1991 International Assessment of Educational Progress**

The 1991 International Assessment of Educational Progress (IAEP) data set can also be tested for CBEEES effects. Fifteen nations were available for the analysis: England, France, Hungary, Ireland, Israel, Emilia Romagna/Northern Italy, Korea, Portugal, Scotland, Slovenia, Soviet Union, Spain, Switzerland, Taiwan, and the United States.

I regressed the mean percentage correct (adjusted for guessing) for 13-year-old students on the same set of variables as in the analysis of the TIMSS data (Table 1). For mathematics, the effect of CBEEES was highly significant and quite large. Because the U.S. standard deviation was 26.8 percentage points in mathematics, the CBEEES effect on math was more than one-half of a U.S. standard deviation or about two U.S. GLEs. CBEEES had smaller nonsignificant effects on science achievement. East Asian students scored significantly higher than students in Europe and North America. Coefficients on per capita GDP were positive but not statistically significant. These results are consistent with the causal hypotheses presented above. Causation is not proved, however, because other explanations can no doubt be proposed.

Other sources of variation in curriculum-based exams need to be analyzed. Best of all would be a study that held national culture constant. For example, students in Canadian provinces with a CBEEES outperform students in the rest of Canada (Bishop 1996).

### **THE IMPACT ON NEW YORK STATE REGENTS EXAMINATIONS**

In the early 1990s, New York State was the only state with a CBEEES. It has been administering curriculum-based Regents examinations to high school students since June 1878. Sherman Tinkelman, Assistant Commissioner for Examinations and Scholarships, described these exams in a 1966 report.

The Regents examinations are closely related to the curriculum in New York State. They are, as you can see, inseparably intertwined. One supports and reinforces the other. . . . These instruments presuppose and define standards. . . . They are a strong supervisory and instructional tool—and deliberately so. They are effective in stimulating good teaching and good learning practices. (1966, 12)

Students took these examinations throughout their high school years. A student taking a full schedule of college preparatory Regents courses would typically take Regents exams in mathematics and earth science at the end of 9th grade; mathematics, biology, and global studies exams at the end of 10th grade;

mathematics, chemistry, English, and foreign language at the end of 11th grade; and physics at the end of 12th grade.

In 1993, about 56 percent of 9th graders took the Mathematics Course 1 exam and, of these, 24 percent scored below the 65 percent passing grade. Similar proportions of 10th and 11th graders took the global studies, biology, and English exams. Failure rates were 20 percent in global studies, 18 percent in biology, and 13 percent in English. Students not taking Regents exams were typically in courses that were considerably less challenging. A system of minimum competency tests in specific subjects set a minimum standard for these students but, as in other states, the passing standard was low.

### **Impacts on Achievement**

New York's students are more disadvantaged, more heavily minority, and more likely to be foreign born than students in most other states. Among northern states, only Maryland, Delaware, and Illinois have a larger share of African American pupils. Nationwide, only California has a greater share of its population foreign born, and only California, Texas, Arizona, New Mexico, and Colorado have larger Hispanic population shares. Literacy levels among New York adults are substantially below the national average (National Education Goals Panel 1995, vol. 2).

Consequently, when one compares student achievement levels, family background must be taken into account. Considering the high incidence of at-risk children, New York students do remarkably well. The proportions of students taking algebra, calculus, chemistry, and physics are generally above national averages. A larger proportion of New York's 11th and 12th graders take and pass (9.4 percent) AP exams in English, science, math, or history than any other state except Utah (NGEP 1993, vol. 2).

Graham and Husted's (1993) analysis of SAT test scores in the 37 states with reasonably large test-taking populations found that New York students did better than comparable students in other states. They did not, however, test the statistical significance of the New York State effect and used an unusual log-log specification.

The results of a linear regression predicting 1991 mean SAT-M + SAT-V test scores for the 37 states for which data are available are presented in Table 2. With the exception of the dummy variable for New York State (NYS), all right-hand-side variables of the equation are proportions—generally the share of the test-taking population with the characteristic described. Clearly, New Yorkers did significantly better on the SAT than students of the same race and social background living in other states. When this model was estimated without the NYS dummy variable, New York had the largest positive residual in the sample. The next largest (Wisconsin's) positive residual was 87 percent of New York's residual. Illinois and Nevada had positive residuals that were about 58 percent of New York's value. Arizona, California, Colorado, Florida, New Mexico, Ohio, Rhode Island, Texas, and Washington had negative residuals greater than 10 points. Many of these states had large populations of Hispanics and recent immigrants, a trait that was not controlled for in the analysis. New York's achievement is all

**TABLE 2**  
**Determinants of Mean Total SAT-I Scores for States**

Dependent variable	Coefficients of explanatory variables										<i>R</i> <sup>2</sup> RMSE
	NYS	Partic. rate	Parents AA-BA+	Private school	Prop. black	Large school	3+Math courses	3+Eng. courses	lnTeach /stud	lnExpend /stud	
SAT-I <sup>a</sup>	46** (2.7)	-68** (2.6)	370*** (6.4)	60 (1.6)	-135 (3.2)	-44* (1.8)	85 (1.3)	-36 (.3)			.926 14.8
SAT-I <sup>a</sup>	.35* (2.0)	-88*** (3.3)	367*** (6.6)	69* (1.9)	-113 (2.6)	-36 (1.5)	45 (.7)	-45 (.4)	48* (1.7)	13 (.8)	.933 14.2
Mean	.027	.414	.581	.207	.078	.129	.617	.797	-2.822	1.648	
SD	.164	.240	.097	.082	.064	.113	.067	.038	.113	.215	

*Note:* Numbers in parentheses are *t* values.

<sup>a</sup>The dependent variable has a mean of 925 and a standard deviation of 55.

\**p* < .01, two-tailed test; \*\**p* < .05, two-tailed test; \*\*\**p* < .001, two-tailed test.

the more remarkable when one considers that Hispanics and immigrants are a large share of its school children.

For individuals, the summed SAT-V + SAT-M has a standard deviation of approximately 200 points. Consequently, the differential between New York's SAT mean and the prediction for New York based on outcomes in the other 36 states is about 20 percent of a standard deviation or about three-quarters of a GLE.

Adding the teacher-pupil ratio and spending per pupil to the model reduces the NYS coefficient by 25 percent. It remains significantly greater than zero, however. The significant coefficient on teacher-pupil ratio suggests that heavy investment in K-12 schooling in New York State (possibly stimulated in part by the Regents exam system) may be one of the reasons why New York state students perform better than comparable students in other states.

### **Does New York State Invest More in K-12 Education?**

The theory predicts that the existence of its CBEEES will induce New York State to spend more on K-12 education and focus that spending on instruction. Indeed, New York's ratio of K-12 teacher salaries to college faculty salaries is significantly above average. New York teachers are also more likely to have master's degrees than the teachers of any state except Connecticut and Indiana. New York ranks number 7 in both the teacher-pupil ratios and the ratio of per pupil spending to gross state product per capita (Bishop 1996).

Clearly, New York invests a great deal in its K-12 education system. If the cause of the high spending was a strong general commitment to education or legislative profligacy, one would expect spending to be high on both K-12 and higher education. This is not the case. Relative to other states, New York invests much more in K-12 education than in higher education. New York ranks number one in the ratio of K-12 spending per pupil to higher education spending per college student.

The Regents exams are currently medium-stakes tests, not high stakes tests. Exam grades count for less than a third of the final grade in the course and influence only the type of diploma received. Employers ignore these exam results when they make hiring decisions. During the 1980s, scholarships sponsored by the Regents were based on aptitude test scores, not Regents exam results. Passing scores on Regents exams are not necessary for admission to community colleges or out-of-state colleges. Students are aware that they can avoid Regents courses and still go to college. Indeed some perceive an advantage to avoiding them.

My counselor wanted me to take Regents history and I did for a while. But it was pretty hard and the teacher moved fast. I switched to the other history and I'm getting better grades. So my average will be better for college. Unless you are going to a college in the state, it doesn't really matter whether you get a Regent's diploma. (Ward 1994, 2)

Indeed, the small payoff to taking Regents exams may be one of the reasons why so many students have not been taking Regents courses.

This is about to change. The Board of Regents has announced that students graduating in the year 2000 must take a new six-hour Regents English examination and pass it at the 55 percent level. The class of 2001 has the additional require-

ment of passing a six-hour examination in algebra and geometry. The class of 2002 must also pass two six-hour Regents examinations in global studies and American history. When laboratory science exams are required, the phase-in of all five new required Regents exams will be completed with the graduating class of 2003. Once the system has adjusted to the new exams, the Regents intend to raise the scores necessary to pass from the 55 percent to 60 percent and then to 65 percent.

## CONCLUSIONS

A review of the evidence suggests that advocates of standards and examination-based reform of American secondary education may be right. The countries and Canadian provinces with such systems outperform other countries at a comparable level of development. New York State, the only state with a CBEEES, does remarkably well on the SAT test when student demography is held constant. In addition, Bishop and Moriarty (1977) show that on NAEP mathematics tests, New York's eighth graders outperform students with comparable socio-economic backgrounds in other states by roughly a grade-level equivalent. CBEEESs are, however, probably not the most important determinant of achievement levels. CBEEESs are common in developing nations where achievement levels are often quite low (e.g., Columbia and Iran). Belgium, by contrast, has a top-quality education system without having a CBEEES. More research on the effects of CBEEESs is clearly in order.

What are the implications of these findings for the National Council on Economic Education's efforts to upgrade economics education in America's secondary schools? The primary implication is that high school students should be offered externally examined economics courses. Externally examined high school courses—Advanced Placement, International Baccalaureate, Regents courses—are growing in popularity. Both the International Baccalaureate and Advanced Placement programs have externally examined economics courses. The council has developed curriculum content standards for economics to serve as a guide for the planned NAEP assessment of economics and for states seeking to promulgate standards for economics. The next logical step is to create a secure external examination reflecting these content standards that can be used for certifying the economics competence of high school students who are not seeking college credit for the course. This step should attract additional high school students to the subject, upgrade the quality of economics instruction, and improve student study incentives.

## NOTES

1. A list of the interviews conducted and the books and documents consulted is available from the author.
2. The sample of countries in Figure 1 are those that either have or do not have a CBEEE in science throughout the country and that have per capita GDP of at least one-third of the U.S. level.
3. The Philippines, for example, had a math score mean of 399 in eighth grade and a mean of 386 in seventh grade. The mean age of eighth graders was 14, and the mean age of seventh graders was 12.9. The math score for 13.5-year-olds was estimated by interpolation between seventh and eighth grade means.  $\text{Math}_{13.5} = 386 + (399 - 386) * [(13.5 - 12.9) / (14 - 12.9)]$ .

## REFERENCES

- Beaton, A. E., I. V. S. Mullis, M. O. Marten, E. J. Gonzalez, D. L. Kelly, and T. A. Smith. 1996a. *Mathematics achievement in the middle school years: IEA's third international mathematics and science study*. Boston, Mass.: CSTEPP, Boston College. //:<http://wwwwsteep.bc.edu/TIMSS>.
- . 1996b. *Science achievement in the middle school years: IEA's third international mathematics and science study*. Boston, Mass.: CSTEPP, Boston College. //:<http://wwwwsteep.bc.edu/TIMSS>.
- Bishop, J. H. 1996. The impact of curriculum-based external examinations on school priorities and student learning. *International Journal of Education Research* 23 (8): 653–752
- Bishop, J. H., and J. Moriarty. 1997. *Diplomas for learning; not seat time: The effects of New York's Regents examinations*. Paper presented at the Regents Forum in Albany, N.Y., Oct. 28.
- Costrell, R. 1994. A simple model of educational standards. *American Economic Review* 84 (Sept.): 956–71.
- Graham, A., and T. Husted. 1995. Understanding state variation in SAT scores. *Economics of Education* 12 (3): 197–202.
- Gummere, R. 1943. The independent school and the post war world. *Independent School Bulletin* 4(April) (quoted in Arthur Powell, *Standards*, 1995, chap 6).
- International Assessment of Educational Progress. 1992. *IAEP technical report*, vol. 1. Princeton, N.J.: Educational Testing Service.
- Kang, S. 1985. A formal model of school reward systems. In *Incentives, learning and employability*, ed. John Bishop. Columbus, Ohio: National Center for Research in Vocational Education.
- Lerner, B. 1990. Good news about American education. *Commentary* 91 (March): 19–25.
- Madeus, G. 1991. The effects of important tests on students: Implications for a national examination or system of examinations. American Educational Research Association Invitational Conference on Accountability as a State Reform Instrument, Washington, D.C., June: 1–19.
- Madeus, G., and T. Kellaghan. 1991. *Examination systems in the European Community: Implications for a national examination system in the United States*. Contractor report for the Office of Technology Assessment, U.S. Congress, Washington, D.C., March 27. 1–100.
- National Center for Educational Statistics. 1992. *The digest of education statistics: 1992*. Washington, D.C.: U.S. Department of Education.
- National Education Goals Panel. 1995. *Data for the National Education Goals Report: 1995*. 1, Washington, D.C.: U.S. Government Printing Office.
- Steinberg, L. 1996. *Beyond the classroom*. New York: Simon and Schuster.
- Tinkelman, S. N. 1966. *Regents examinations in New York state after 100 years*. Albany, N.Y.: The University of the State of New York, the State Education Department: 1–15.
- Ward. 1994. A day in the life. *N.Y. Teacher* (Jan.). Albany, N.Y.: 1–4.