《人文及社會科學集刊》
第二十一卷第一期（98/3），pp. 1-24


# Do Public Exams Raise Student Performance? A Cross-National Difference-in-Differences Analysis*

Min-hsiung Huang
Associate Research Fellow
Institute of European and American Studies, Academia Sinica

## ABSTRACT

Is a system of public exams conducive to student learning? To address this question, I employ a difference-in-differences methodology which isolates the effects of public exams from the effects of unmeasured country-level features, using data from the Trends in International Mathematics and Science Study (TIMSS) for students in Grade 4 and Grade 8 in 1995 and 2003. Findings of this study suggest that students in countries with a system of public exams perform higher in math, but not in science. The positive effect of public exams on math performance found in this study, however, is only half as large as what was previously reported. Furthermore, when countries with public exams are extended to include Cyprus and Portugal, public exams no longer have a positive effect on student math performance.

**Key Words:** public exams, student assessment, difference-in-differences, student performance, TIMSS

* This research was supported by the National Science Council [grant number 93-2412-H-001-023]. Correspondence may be directed to Min-Hsiung Huang, Academia Sinica, Institute of European and American Studies, Nankang, Taipei, 11529, Taiwan (email: mhhuang@sinica.edu.tw; tel: 886-2-3789-7246).
Received: March 13, 2008; Accepted: June 23, 2008

# 1. Introduction

Student academic performance is often regarded as an important measure of a nation's human capital, which is a strong force driving economic growth (Bishop, 1989; Hanushek and Kimbo, 2000; Barro, 2001). It is a common goal for policy makers around the world to improve the quality of schools and to promote the level of student academic performance. To achieve these goals, one traditional approach is to increase educational spending and enrich the resource endowments of schools. The sufficiency of school resources is usually indicated by total expenditure per student, class size, student-teacher ratio, and the adequacy of instructional materials. Whether or not more resource inputs can consistently and cost-effectively lead to higher academic performance outputs has been widely researched. This line of research, while controversial, generally suggests that increasing school resources does not significantly lead to higher student performance, especially in developed countries (Hanushek, 1996, 2003; Gundlach et al., 2001; Woessmann and West, 2006).[1]

While resource-based educational policies seem ineffective and costly, some researchers suggest that educational institutions may have a large impact on student performance (Woessmann, 2004). To improve the quality of schools, therefore, is to change the institutional structure of the schooling system. While many educational institutions may affect student performance, such as school autonomy (Chubb and Moe, 1990), competition between public and private schools (Epple and Romano, 1998), school choice (Rouse, 1998; Greene et al., 1999), and the degree of teacher unionization (Hoxby, 1996); this study focuses on one aspect of the institutional structure: the effects of large-scale curriculum-based external examinations (henceforth, public exams) on student performance. As practiced in many countries, public exams are administered to students in primary or secondary

1 Some studies suggest that school facilities have a stronger effect on student performance in developing countries in which basic educational resources are inadequate and very unequally distributed (Heyneman and Loxley, 1983; Fuller and Kapakasa, 1991). Other studies, however, do not find a stronger resource effect in developing countries (Baker et al., 2002; Hanushek, 1995). In developed countries, many studies do not find a resource effect, but some find just such an effect. For example, Greenwald et al. (1996) found resources were positively related to student outcomes.

schools and used to make important decisions about (a) whether a student will be assigned to a particular curricular program, school, or class; (b) whether a student will be promoted to the next grade level; and (c) whether a student will receive a high school diploma (Heubert and Hauser, 1999).

Relative to some resource-based reforms, such as class-size reduction, the implementation of a public exam system is relatively inexpensive (Hoxby, 2002; Linn, 2000). A system of public exams also seems more readily mandated and implemented by external authorities than other institutional reforms (Linn, 2000). If a system of public exams has a large positive impact on student performance, the implementation of such a system would present a highly attractive policy alternative, given its cost-effectiveness and applicability. Therefore, whether or not a system of public exams is conducive to student learning is an important question to be pursued, and the question the present study aims to address.

This paper begins with a review of (a) the definition of public exams, (b) a theory about why public exams are conducive to student learning, and (c) the findings and limitations of previous studies in assessing the effects of public exams on student performance. To evaluate the effects of public exams on student performance, I employ the difference-in-differences (henceforth, DD) methodology to isolate the effects of public exams from the effects of unmeasured country-level features, using data from the Trends in International Mathematics and Science Study (TIMSS) for students in Grade 4 and Grade 8 in 1995 and 2003.

## 1.1. What is a public exam?

Students are constantly assessed by their teachers in class. Public exams, however, are different from teacher-grading. Public exams are conducted by an external organization, usually supported by provincial or national governments, taken by a large number of students in a given grade, used to screen students for promotion and graduation, and are regarded as highly influential on students' future educational and occupational opportunities (Mullis et al., 2000). In addition, public exams are curriculum-based, relative to an external standard, typically competitive, and often administered at the end of an educational program (Bishop, 1997). Unlike 'national assessments' (such as the National Assessment of Educational Progress in the U.S.), which provide information on how well a group of students perform or how efficiently an educational system works, public exams certify the level of achievement for each student and provide educational and occupational rewards to students who perform well.[2]

An example of a public exam is the 'Basic Competence Test for Junior High School Students' in Taiwan. The test is administered to most Taiwanese ninth-graders for admission to senior high schools. Taiwanese students compete ferociously for places in the best public academic high schools based on the results of the test. Taiwanese students who score low on the test usually end up going to private vocational high schools not geared toward 4-year academic universities. Indeed, many Taiwanese students who end up going to private vocational high schools end their formal educational experience there and do not pursue higher education. Not all public exams are as competitive as those in Taiwan. In the Netherlands, for example, most primary school students take tests developed by the National Institute for Educational Measurement (CITO) to guide secondary school choice; these CITO tests assess individual student achievement at the end of primary schooling, and students who perform poorly on these tests and have a poor school record may be required to repeat a grade (Silva, 2002). In Iceland, all tenth graders are required to sit nationally coordinated tests, a local form of public exams. Whether students are qualified to be placed in the academic branch of upper secondary school depends on the results of nationally coordinated tests taken in Grade 10 (The National Academic Recognition Information Center, 1996; Stefansson and Karlsdottir, 2007).

### I.2. Why do public exams have a positive impact on student performance?

A theory was proposed by Bishop (1997, 1999, 2006) and extended by Woessmann (2003, 2005) to explain why a system of public exams may be conducive to learning. When a system of public exams is absent, students are usually assessed by teachers within their classrooms. Therefore, student performance is not comparable across classrooms or schools. How well one class performs relative to another is usually impossible to determine in such a situation. In such a system, there is no fair way to assess the performance of individual teachers and school administrators. Teachers and administrators who do a poor job are not punished, while those who are successful go unrewarded. Without effectively monitoring, teachers and school adminis-

2 For a thorough discussion of the nature of public exams, please visit http://www1.worldbank.org/education/exams/nature.asp (October 13, 2008)

trators are likely to use school resources in pursuit of their own interests rather than to improve student performance. When there are no public exams, college admissions are based on class rank and students are usually graded within their classes on a curve. Students tend to persuade each other not to study hard because it is easier for all if no one in the classroom makes an extraordinary effort. Students who work hard face the risk of losing friends because their hard work makes it more difficult for other students to get good grades. The collective behavior of students in the classroom, in turn, pressures teachers to lower standards and give higher grades than their students deserve.

Unlike within-school or within-classroom tests, a public exam provides information on how individual students perform relative to all other students in the nation (or region). Student performance becomes transparent and well-signaled when a system of public exams is in place. For students, the rewards of learning become more apparent, especially when exams involve consequences for individuals, such as entering a better school. The better means of signaling student achievement provided by public exams makes it possible for educational institutions and employers to emphasize academic achievement and to reward those who perform well on public exams. A public exam provides an absolute national measure of achievement and makes relative ranking within classrooms less important. Consequently, the problem of peer pressure against study becomes less serious. A public exam significantly improves the signaling of student achievement, making incompetent teachers and principals more visible and likely to be punished. Competent teachers and school administrators, on the other hand, are more likely to be rewarded. Therefore, public exams hold teachers and schools accountable for student performance and provide incentives for students, teachers, and school administrators to focus on learning.

## 2. A Survey of Research

One way to examine the effects of public exams on student achievement is to conduct a cross-national study. In such cross-national studies, countries can be divided into those that have a system of public exams and those that do not. Given that all of the relevant factors are controlled, differences in achievement between these two groups of countries may be taken as indicative of the effects of public exams on student achievement. In the same way, in countries in which some provinces or states have a system of public exams, a provincial-level analysis can also be conducted to examine the effects of public exams on learning.

Using country-level data from the 1995 TIMSS for seventh and eighth graders in 39 countries/regions, Bishop (1996, 1997) found countries with a 'curriculum-based external exit examination system' (CBEEES) tend to have higher national mean scores in math and science, after controlling for country differences in per capita gross domestic product (GDP) and a dummy variable for students from East-Asian countries. A CBEEES is highly consequential for students' career opportunities, relative to an external (beyond school level) standard, curriculum-based, signaling multiple levels of achievement, and covering most secondary-school students. Bishop (1997) suggests that the impact of CBEEES is about one U.S. grade-level equivalent in mathematics and 1.2 grade-level equivalents in science (one U.S. grade-level equivalent represents the amount of math or science knowledge learnt in an entire school year for the average American student).

Bishop and his colleagues also conducted cross-state analyses within the U.S. to support the positive impact of a public exam on students' learning outcomes. For example, Bishop, Moriarty, and Mane (2000) found that students in New York State, the only state with a CBEEE system in the nation, perform one grade level ahead of socio-economically comparable students from other states in the mean eighth-grade NAEP (National Assessment of Educational Progress) math scores and SAT (Student Achievement Test) scores. Using data from the National Educational Longitudinal Study (NELS), Bishop and Mane (2001) also found that students from 22 states requiring a minimum competency high school graduation exam tend to earn more after high school graduation and are more likely to enter college. Overall, Bishop and his colleagues found a consistent positive effect on student performance on public exams between cross-national analyses and cross-state analyses.

Limited by a small sample size, Bishop's country-level analysis—using data from TIMSS—suffered from omitted-variable bias. This shortcoming was partially overcome by Woessmann's student-level analyses, which allow for an adequate set of independent variables, including background characteristics of the student and his/her family, school resources, teacher characteristics and some institutional features of the school system. By combining data on eighth-grade students from TIMSS 1995 and TIMSS 1999, Woessmann (2003) analyzed student-level data from 54 countries and found that public exams had a positive effect on student math and science performance. Using data from the Program for International Student Assessment (PISA) additionally, Woessmann (2005) suggests that the performance gap between countries with public exams and those without was the equivalent of approximately one grade, which is consistent with the findings reported by Bishop (1997).

The cross-national analyses of Bishop and Woessmann raise some methodological concerns. The first is whether their analyses suffer from a bias due to unmeasured omitted variables. For example, the performance gap between countries with a system of public exams and those without may be due to some unobserved or uncontrolled country features. Unless these unobserved factors are taken into account, the effects of public exams are likely to be confounded with the effects of other factors. These unobserved factors could include cultural values concerning achievement in school, levels of economic development, degree of centralization in the educational system and the policy environment.

The second concern is whether some countries in the analyses of Bishop and Woessmann should really be treated as countries without public exams. According to Bishop and Woessmann, the university entrance examinations in Greece, Portugal, Spain, and Cyprus should not be regarded as public exams because students on a vocational track do not take them and 'thus do not represent an integral part of the education system' (Woessmann, 2005). In fact, these four countries should probably be treated as countries with public exams. In addition to a university entrance examination, Cyprus has final compulsory external examinations at the end of secondary schooling on which students must perform well in order to receive a graduation certificate and enter college (The National Academic Recognition Information Center, 1996; Eurydice, 2004; Papanastasiou, 1995: 256; Mullis et al., 2000: 153; Martin et al., 2000: 165; Martin et al., 2004: 183; Mullis et al., 2004: 166). In Portugal, 'National achievement tests are administered across all schools at the end of basic education and at the end of secondary education. These are used for deciding upon the award of a certificate' (Rau, 1995: 807). The final examination at the end of secondary school in Portugal not only decides whether or not students may receive a graduation certificate, but also affects students' chances of pursuing higher education (Rau, 1995: 807). In Greece, as of 1991, '77.5 percent of the students who graduated from upper secondary education, representing approximately 48.5 percent of the 17-18 age group, qualified for participation in tertiary education entry examinations' (Kontogiannopoulou-Polydorides et al., 1995: 364). The percentage of

Greek students participating in university entrance examinations was even higher in 1995 and 2003 when TIMSS assessments were administered. Finally, in Spain, admission to higher education is not entirely dominated by university entrance examinations because GPA (Grade Point Average) in high school is considered an equally important factor. Nevertheless, the percentage of high school graduates participating in the university examination is likely to be very high, given that 66 percent of the population of tertiary age were in tertiary education as of 2005, according to the Institute for Statistics at the UNESCO. In sum, a review of the examination systems in these four countries suggests that Cyprus and Portugal, which are regarded by Bishop and Woessmann as countries without high-school exit exams, do have such exams. Greece and Spain may not have end-of-school exams, but they have college entrance exams which are taken by the majority of students. Because these four countries are low-performing countries, treating them as countries without public exams would yield a spurious finding that public exams have a positive effect on student performance.

To deal with the potential problem of omitted-variable bias, this study employs a difference-in-differences method which controls for unobserved country-level features. To examine the consequences of treating Cyprus and Portugal as countries with public exams, this study provides two sets of results for comparison: one categorizes all countries as Bishop and Woessmann have categorized them, while the other differs from the categorization of Bishop and Woessmann in that it treats Cyprus and Portugal as countries with public exams.

## 3. Data

The application of DD methodology requires two observational points within a single country: one where public exams have not significantly influenced student learning, and the other where public exams have already affected student performance. Therefore, this study uses data from fourth- and eighth-grade students on the TIMSS 1995 and 2003 mathematics and science assessments for countries participating at both grade levels in each survey year. This results in a total of 26 countries from TIMSS 1995 and a total of 25 countries from TIMSS 2003 for analysis. To increase the number of countries for analysis, this study analyzes two sets of data: one consists of 26 countries from TIMSS 1995 and 10 non-overlapping countries from TIMSS 2003, with a total of 36 countries; the other consists of 25 countries from TIMSS 2003 and 11 non-overlapping countries from TIMSS 1995, with a total of 36 countries. Canada and Latvia have public exams only in some provinces/regions, and so, as they cannot be categorized into either group of countries with or without public exams, these two countries are excluded from the analysis. All Japanese students in TIMSS 1995 did not respond to the question of the number of books at home, this yields a total of 33 countries for one set of data and a total of 34 countries for the other set of data for the present analysis.

For a fair international comparison, the TIMSS achievement test was designed to reflect as many curricula as possible in all participating countries. To ensure that TIMSS test items were appropriate and reflected different curricula among countries, the construction of math and science tests were based on the TIMSS curriculum frameworks, which were set by consensus and endorsed by all participating countries. In addition, a test-curriculum matching analysis was conducted, and the results showed that the general pattern of achievement results across all countries is unchanged when test items inappropriate to specific countries were omitted (Beaton and Gonzalez, 1997).

The TIMSS test included a series of multiple-choice test items and open-ended response questions requiring short or more elaborate explanations. Not all students answered the same test questions. To minimize the response burden on individual students, matrix-sampling techniques were used to divide the test item pool so that each sampled student responded to only a portion of the test questions, but a portion which still covered all subject areas (Garden, 2000). Therefore, students answered different test items depending upon which test booklets they received. Based on item response theory, student responses were scaled to provide accurate estimates of achievement which could be compared across countries (Yamamoto, 2000). In addition, a 'plausible values' method was used to produce proficiency scores in Math and Science. The achievement scores, therefore, are available as a set of five plausible values for every individual student. All regression analyses appearing in this study are undertaken five times, each using a different plausible value. Each regression estimate reported in the study is the mean of five estimates.

In the TIMSS base year 1995, the international proficiency scores (Math or Science) were scaled to have a mean of 500 and a standard deviation of 100, and the metric of the 1995 scale has been preserved, so test scores in later survey years can be compared with those in 1995. To make achievement test scores comparable between Grade 4 and Grade 8 for this analysis, test scores of each grade level from all countries are standardized at the individual level to a mean score of 500 and a standard deviation of 100. For 24 countries from TIMSS 1995 and 10 non-overlapping countries from TIMSS 2003, Table A1 in the appendix presents the number of students, average test scores and the presence of a system of public exams in each country. Table A2 reports the same statistics for 24 countries from TIMSS 2003 and 10 non-overlapping countries from TIMSS 1995.

## 4. Method

To estimate the treatment effect when treatment is randomly assigned to some units, a simple method would be to compare the treated units before and after treatment. This way, however, tends to give biased results because the treatment effect is confounded with the effects of other factors that take place around the time of treatment. The DD methodology deals with this problem by using a control group to difference out these confounding factors and isolate the treatment effect (Meyer, 1995). Specifically, the DD methodology models the treatment effect by estimating the difference between outcome measures at two time points for both the experimental and the control groups and then calculating the difference between these two groups.

The DD methodology requires an experimental group and a control group, with repeated measures of the same units from two time points: before and after treatment. A panel survey with data collected on the same individuals at different time points meets this requirement, but repeated cross-sections, such as two national random survey samples separately collected, can also be used in a DD analysis.[3] For example, Hanushek and Woessmann (2006) took advantage of the facts that no country tracked students before the fourth grade, and that some countries did track their students at a higher grade level, they adopted a DD methodology by comparing differences in achievement between younger (the fourth graders) and older students (eighth, ninth, or tenth graders) among tracking and non-tracking countries. Hanushek and Woessmann estimated the effect of tracking by comparing the average achievement gain from grade 4 to grade 8 (or higher)

3 See Meyer (1995) for a discussion of DD approach to cross-sectional and other types of data.

in tracked countries to that in untracked countries. The younger and older student samples used by Hanushek and Woessmann were not from a panel survey, but were collected separately from two cross-sectional national surveys.

The present study adopts a methodological strategy similar to that used by Hanushek and Woessmann (2006), and therefore faces some common methodological limitations. For example, the DD models used by Hanushek and Woessmann would present unbiased results only when there were no grade-specific educational policies that were correlated with whether or not a country adopted a policy of tracking. Similarly, the DD methodology used in this study is based on the assumption that there were no grade-specific policies related to whether or not a country adopts a system of public exams.

The use of DD methodology in this paper is based on an additional assumption that the performance of fourth graders is not yet affected by public exams, but the effects of public exams may increase with grade levels as a public exam approaches. This assumption is partially supported by the findings of Woessmann (2005), who used data from TIMSS 1995 and reported that the impact of public exams on math performance is 17.5 percent of a standard deviation larger in the eighth grade than in the seventh grade. A similar finding was reported for student performance in science. Because most countries have public exams at the end of lower or upper secondary schooling, public exams may not have influenced the performance of fourth graders. For eighth graders, however, who are much closer to the timing of a public exam, the impact of a public exam on student learning should have already taken place.

To identify the effects of public exams, Bishop (1997) and Woessmann (2002, 2005) assessed the performance gap between countries with public exams and countries without public exams for students at Grade 8. However, the presence of public exams may be correlated with other cross-country features also affecting student performance. For this reason, this study employs the DD methodology to difference out these confounding factors and to isolate the effects of public exams. Calculating average achievement for countries with public exams and countries without public exams separately for fourth and eighth grade students gives four outcomes which, as shown in the following equation, can be used to obtain a DD estimate of the effect of public exams:

$$\mathrm{DD} = [\mathrm{E}(Y_{g8}) - \mathrm{E}(Y_{g4}) | \textit{exam countries}] - [\mathrm{E}(Y_{g8}) - \mathrm{E}(Y_{g4}) | \textit{non-exam countries}] \quad (1)$$

where $Y$ is the average math or science achievement in a country. As Equation 1 demonstrates, the DD model compares changes within countries, which limits the bias caused by unobserved differences between countries. Simply comparing achievement in countries with and without public exams, as Bishop and Woessmann have done, is not the best approach to evaluate the effects of public exams. This is because country differences in student achievement are strongly driven by a number of social and educational factors other than the presence of a system of public exams. Unless a model is capable of controlling for these unobserved country factors, the effects of public exams cannot be correctly estimated. By using the DD approach to estimate the effects of public exams on student performance, country differences in early achievement and other stable features are taken into account.

The DD estimator appearing on the left-hand-side of Equation 1 can be obtained in a regression framework as well (Wooldridge, 2006: 456). A regression framework makes possible for additional controls. Specifically, this study uses the following regression model to obtain the DD estimate of the effect of public exams:

$$T_{iscg} = \beta_1 + \beta_2 FB_{iscg} + \beta_3 AG_{iscg} + \beta_4 FM_{iscg} + \beta_5 AD_c + \beta_6 GR_g + \beta_7 PE_c + \beta_8 (GR_g * PE_c) + \varepsilon_{iscg} \quad (2)$$

where $T_{iscg}$ is the TIMSS math or science test score $T$ of student $i$ in school $s$ in country $c$ in grade level $g$. $FB_{iscg}$ is student family background measured by the number of books at home, $AG_{iscg}$ is a set of student age dummy variables, and $FM_{iscg}$ represents female student.[4] $AD_c$ indicates non-overlapping countries added from the other TIMSS survey; for example, when analyzing data mainly from TIMSS 1995, $AD_c$ represents non-overlapping countries added from TIMSS 2003. $AD_c$ controls for any systematic differences between TIMSS 1995 and TIMSS 2003. $GR_g$ denotes eighth-graders, $PE_c$ indicates countries with a system of public exams, and $GR_g * PE_c$ is the interaction term of $GR_g$ and $PE_c$. The coefficient on the interaction term, $\beta_8$, gives the DD estimate of the effect of public exams,

4 See Schuetz, Ursprung, and Woessmann (2005) for a discussion on the use of the number of books at home in measuring student family background.

which is the major point of interest of this research. Because the interaction includes the variable of public exams, which is measured at the country level, the regression analysis clusters the standard errors at the country level and reports clustering-robust standard errors. This reflects the fact that the number of independent observations on the variable of public exams is not the number of students, but the number of countries. The regression analyses conducted in this study use students' sampling probabilities as weights. Between countries, the weights give equal weight to each country.

A very small number of students did not report their gender and age, and these students are excluded from the analysis. About three percent of the respondents did not report the number of books at home; therefore, every missing value on the number of books at home is replaced by the median of the valid cases in the respective classroom. However, all Japanese students in TIMSS 1995 did not report the number of books at home and are excluded from the analysis entirely.

## 5. Result

In Table 1, the determination of whether or not a country has a system of public exams is based on the reports of Bishop (1999) and Woessmann (2002), who regarded Cyprus and Portugal as countries without public exams. Whether or not Cyprus and Portugal can be regarded as countries without public exams, however, is open to question. Therefore, Table 2 reports the results when Cyprus and Portugal are treated as countries with public exams.

The results presented in Table 1 suggest that students perform significantly higher in math when a system of public exams is present. Two different sets of data, one with most countries from TIMSS 1995 and the other with most countries from TIMSS 2003, generate the same finding with similar effect size. When a system of public exams is in place, students score 19 points higher in math. The size of the effect of public exams on math performance, 19 points, is only half as large as that reported by Woessmann (2005) and Bishop (1997) who found that the effect of public exams on math performance was about 40 points, roughly as large as one grade-level equivalent in the TIMSS studies. The positive effect of public exams on science achievement is smaller than that on math achievement and it does not reach statistical significance, as demonstrated in Table 1.[5]

When Cyprus and Portugal are treated as countries with public exams, as Table 2 shows, the presence of public exams no longer has a significant effect on math achievement and the effect of public exams on science achievement remains insignificant. This suggests that the findings of cross-national studies conducted by Bishop (1997, 1999) and Woessmann (2003, 2005) should be interpreted with great caution. Because very few countries do not have a system of public exams, the findings reported by Bishop and Woessmann can be unreliable. By treating low-performing countries like Cyprus and Portugal as countries without public exams, Bishop and Woessmann may have overestimated the positive effect of public exams.

## 6. Conclusion

An institutional feature like the presence of public exams usually does not vary within a country. For this reason, some researchers analyzed international data of student achievement and used cross-national variation in the policy of implementing a public exam. This study also adopts the international approach, but it differs from previous international studies by employing a DD methodology to control for unobserved country features, which also affect student performance and are correlated with whether or not a country has a system of public exams. With the DD methodology, the effect of public exams can be isolated from the effect of unobserved or uncontrolled country features. This study finds that the positive effect of public exams on student math performance is only half as large as what previously reported, and it also finds that public exams do not have a significant impact on student performance in science. Furthermore, when countries with public exams are extended to include Cyprus and Portugal, public exams no longer significantly affect math performance.

The results of this analysis suggest that the use of public exams in educational selection may not be an important factor in explaining why students in some countries perform better in math and science. An analysis of the effects of the presence versus absence of public exams on academic performance is not the best strategy because very few countries will be found in

5 A simple DD model, as presented in Equation 1, generates similar results with respect to effect size. Countries with public exams have a mean math score 24 points higher and a mean science score 14 points higher than do countries without public exams.

the category of not having public exams. Such a dichotomous approach does not help us understand why countries with public exams differ so greatly in academic achievement. The vast achievement differences among countries which hold public exams may have to do with the fact that countries differ considerably in the particular way they administer and use such exams. Therefore, future research should take note of why countries that commonly have public exams differ so much in academic performance. Those who propose a system of public examinations to improve students' academic achievement should consider the fact that most countries have such a system, but do not all realize high levels of academic performance.

International variation in the presence of public exams has not been helpful in identifying the effects of public exams; future research may be more fruitful if consideration is given to different forms of public examination across countries and their differing effects on student performance are examined. A public exam which has important consequences for students of all ability levels may have a stronger positive effect on student performance than a public exam of consequence to only certain students. Country differences in the total number of public exams taken by students during primary and secondary schooling may explain why some countries perform better than others. The extent to which admission to the next level of schooling is dependent on exam results may influence student motivation and, ultimately, achievement. When the effects of different forms of public exams on learning are closely examined, policy makers will be able to make an informed decision about how to implement—not just whether to implement—a system of public exams.

## References

Baker, David P., Brian Goesling, and Gerald K. LeTendre

2002 "Socioeconomic Status, School Quality, and National Economic Development: A Cross-national Analysis of the 'Heyneman-Loxley Effect' on Mathematics and Science Achievement," *Comparative Education Review* 46(3): 291-312.

Barro, Robert J.

2001 "Human Capital and Growth," *American Economic Review Papers and Proceeding* 91(2): 12-17.

Beaton, Albert E. and Eugenio J. Gonzalez
1997 "TIMSS Test-curriculum Matching Analysis," in Michael O. Martin and Dana L. Kelly (eds.), *Third International Mathematics and Science Study: Technical Report, Volume 2*. Chestnut Hill, MA: Boston College.

Bishop, John H.
1989 "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review* 79: 178-194.
1996 "The Impact of Curriculum-based External Examinations on School Priorities and Student Learning," *International Journal of Education Research* 23 (8): 653-752.
1997 "The Effect of National Standards and Curriculum-based Exams on Achievement," *The American Economic Review* 87(2): 260-264.
1999 "Are National Exit Examinations Important for Educational Efficiency?" *Swedish Economic Policy Review* 6(2): 349-401.
2006 "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn-Externalities, Information Problems and Peer Pressure," in Eric Alan Hanushek and Finis Welch (eds.), *Handbook of the Economics of Education*. Amsterdam: North-Holland.

Bishop, John H. and Ferran Mane
2001 "The Impacts of Minimum Competency Exam Graduation Requirements on High School Graduation, College Attendance and Early Labor Market Success," *Labor Economics* 8: 203-222.

Bishop, John H., Joan Y. Moriarty, and Ferran Mane
2000 "Diplomas for Learning, Not Seat Time: The Impacts of New York Regents Examinations," *Economics of Education Review* 19(4): 333-349.

Chubb, John E. and Terry M. Moe
1990 *Politics, Markets, and America's Schools*. Washington, DC: Brookings Institution Press.

Epple, Dennis and Richard E. Romano
1998 "Competition between Private and Public Schools, Vouchers, and Peer-group Effects," *American Economic Review* 88(1): 33-62.

Eurydice
2004 "The Education System in Cyprus," *Information Database on Education Systems in Europe*. http://194.78.211.243/Eurybase/Application/frameset.asp?country=CY&language=EN (July 16, 2007)

Fuller, Bruce and A. Kapakasa
1991 "What Factors Shape Teacher Quality? Evidence from Malawi," *International Journal of Educational Development* 11(2): 119-127.

Garden, Robert A.
2000 "TIMSS Test Development," in Michael O. Martin, Kelvin D. Gregory, and Steven E. Stemler (eds.), *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College.

Greene, Jay P., Paul E. Peterson, and Jiangtau Du
1999 "The Effectiveness of School Choice: The Milwaukee Experiment," *Education and Urban Society* 31(2): 190-213.

Greenwald, Rob, Larry V. Hedges, and Richard D. Laine
1996 "The Effect of School Resources on Student Achievement," *Review of Educational Research* 66(3): 361-396.
Gundlach, Eric, Ludger Woessmann, and Jens Gmelin
2001 "The Decline of Schooling Productivity in OECD Countries," *The Economic Journal* 111(471): C135-C147.
Hanushek, Eric A.
1995 "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer* 10(2): 227-246.
1996 "School Resources and Student Performance," in Gary T. Burtless (ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, DC: The Brookings Institution.
2003 "The Failure of Input-based Schooling Policies," *The Economic Journal* 113(485): F64-F98.
Hanushek, Eric A. and Dennis D. Kimbo
2000 "Schooling, Labor-Force Quality, and the Growth of Nations," *American Economic Review* 90: 1184-1208.
Hanushek, Eric A. and Ludger Woessmann
2006 "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries," *The Economic Journal* 116(510): C63-C76.
Heubert, Jay Philip and Robert Mason Hauser
1999 *High Stakes: Testing for Tracking, Promotion, and Graduation*. Washington, DC: National Academy Press.
Heyneman, Stephen P. and William A. Loxley
1983 "The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High- and Low-Income Countries," *The American Journal of Sociology* 88(6): 1162-1194.
Hoxby, Caroline Minter
1996 "How Teachers' Unions Affect Education Production," *Quarterly Journal of Economics* 111(3): 671-718.
2002 "The Cost of Accountability," Working Paper 8855. Cambridge, MA: NBER.
Kontogiannopoulou-Polydorides, Georgia, T. Mylonas, J. Solomon, and D. Vergidis
1995 "Greece," in T. Neville Postlethwaite (ed.), *International Encyclopedia of National Systems of Education* (2nd ed.). Oxford, UK: Pergamon.
Linn, Robert L.
2000 "Assessment and Accountability," *Educational Researcher* 29(2): 4-16.
Martin, Michael O. et al.
2000 *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.
Martin, Michael O., Ina V. S. Mullis, Eugenio J. Gonzalez, and Steven J. Chrostowski
2004 *TIMSS 2003 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: Boston College.

Meyer, Bruce D.
1995 "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics* 13(2): 151-161.
Mullis, Ina V. S. et al.
2000 *TIMSS 1999 International Mathematics Report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: Boston College.
Mullis, Ina V. S., Michael O. Martin, Eugenio J. Gonzalez, and Steven J. Chrostowski
2004 *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: Boston College.
Papanastasiou, Constantinos
1995 "Cyprus," in T. Neville Postlethwaite (ed.), *International Encyclopedia of National Systems of Education* (2nd ed.). Oxford, UK: Pergamon.
Rau, J. M.
1995 "Portugal," in T. Neville Postlethwaite (ed.), *International Encyclopedia of National Systems of Education* (2nd ed.). Oxford, UK: Pergamon.
Rouse, Cecilia Elena
1998 "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics* 113 (2): 553-602.
Schuetz, Gabriela, Heinrich W. Ursprung, and Ludger Woessmann
2005 "Education Policy and Equality of Opportunity," IZA Discussion Paper 1906. Bonn: Institute for the Study of Labor.
Silva, M. M.
2002 "Netherlands," in Rebecca Marlow-Ferguson (ed.), *World Education Encyclopedia: A Survey of Educational Systems Worldwide* (2nd ed.). Detroit, MI: Gale Group.
Stefansson, T. and R. Karlsdottir
2007 "Iceland," in Wolfgang Hoerner (ed.), *The Education Systems of Europe*. Dordrecht, Netherlands: Springer.
The National Academic Recognition Information Center
1996 *International Guide to Qualification in Education* (4th ed.). London, UK: Mansell.
The World Bank
2008 "The Nature of Public Examinations," http://www1.worldbank.org/education/exams/nature.asp (October 13, 2008)
Woessmann, Ludger
2002 *Schooling and the Quality of Human Capital*. Berlin, Germany: Springer.
2003 "Central Exit Exams and Student Achievement: International Evidence," in Paul E. Peterson and Martin R. West (eds.), *No Child Left Behind? The Politics and Practice of School Accountability*. Washington, DC: Brookings Institution Press.
2004 "Institutional Comparisons in Educational Production," *Journal for Institutional Comparison* 2(4): 3-20.
2005 "The Effect Heterogeneity of Central Examinations: Evidence from TIMSS,

TIMSS-Repeat and PISA," *Education Economics* 13(2): 143-169.

Woessmann, Ludger and Martin R. West

2006 "Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS," *European Economic Review* 50(3): 695-736.

Wooldridge, Jeffrey M.

2006 *Introductory Econometrics: A Modern Approach* (3rd ed.). Mason, OH: South-Western.

Yamamoto, Kentaro

2000 "Scaling Methodology and Procedures for the TIMSS Mathematics and Science Scales," in Michael O. Martin, Kelvin D. Gregory, and Steven E. Stemler (eds.), *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College.

**Table 1. The Effects of Public Exams on Math and Science Performance, Difference-in-Differences Estimates, TIMSS 1995 and 2003**

| | Mostly TIMSS 1995 | | | Mostly TIMSS 2003 | | |
|---|---|---|---|---|---|---|
| | *Coeff.* | *S.E.* | *p value* | *Coeff.* | *S.E.* | *p value* |
| | | | Math | | | |
| Constant | 323.13 | 19.47 | 0.000 | 329.35 | 24.64 | 0.00 |
| Girl | −5.59 | 1.56 | 0.001 | −3.87 | 1.84 | 0.04 |
| Number of books at home | 21.40 | 2.68 | 0.000 | 22.26 | 2.77 | 0.00 |
| Grade 8 | 56.45 | 12.43 | 0.000 | 64.65 | 11.80 | 0.00 |
| TIMSS 2003 | −2.28 | 21.33 | 0.916 | −14.04 | 17.25 | 0.42 |
| Public exam | 23.15 | 22.09 | 0.303 | 18.59 | 22.31 | 0.41 |
| Grade 8 * Public exam | 19.44 | 7.88 | 0.019 | 19.12 | 8.32 | 0.03 |
| R-squared | | .15 | | | .16 | |
| | | | Science | | | |
| Constant | 304.68 | 17.90 | 0.000 | 306.28 | 25.70 | 0.00 |
| Girl | −12.90 | 2.24 | 0.000 | −9.25 | 2.29 | 0.00 |
| Number of books at home | 25.74 | 1.96 | 0.000 | 24.58 | 2.58 | 0.00 |
| Grade 8 | 68.01 | 13.63 | 0.000 | 78.05 | 14.17 | 0.00 |
| TIMSS 2003 | −13.44 | 18.74 | 0.478 | −7.32 | 13.33 | 0.59 |
| Public exam | 26.84 | 18.55 | 0.158 | 29.96 | 19.22 | 0.13 |
| Grade 8 * Public exam | 13.85 | 9.36 | 0.149 | 12.18 | 8.76 | 0.17 |
| R-squared | | .19 | | | .18 | |
| Number of countries | | 33 | | | 34 | |
| Number of students | | 247,468 | | | 276,752 | |

Note. This table reports clustering-robust standard errors, using countries as level of clustering. The regression analyses use students' sampling probabilities as weights. Between countries, the weights give equal weight to each country. All models control for a set of student age dummy variables. Data set named 'Mostly TIMSS 1995' consists of 23 countries from TIMSS 1995 and 10 non-overlapping countries from TIMSS 2003. Data set named 'Mostly TIMSS 2003' consists of 24 countries from TIMSS 2003 and 10 non-overlapping countries from TIMSS 1995.

**Table 2. The Effects of Public Exams on Math and Science Performance When Cyprus and Portugal Are Regarded as Countries with Public Exams, Difference-in-Differences Estimates, TIMSS 1995 and 2003**

| | Mostly TIMSS 1995 | | | Mostly TIMSS 2003 | | |
|---|---|---|---|---|---|---|
| | *Coeff.* | *S.E.* | *p value* | *Coeff.* | *S.E.* | *p value* |
| | Math, Cyprus & Portugal have public exams | | | | | |
| Constant | 319.48 | 24.17 | 0.000 | 324.87 | 28.20 | 0.00 |
| Girl | −5.58 | 1.60 | 0.002 | −3.84 | 1.86 | 0.05 |
| Number of books at home | 21.94 | 2.58 | 0.000 | 22.60 | 2.72 | 0.00 |
| Grade 8 | 57.86 | 12.82 | 0.000 | 68.49 | 11.21 | 0.00 |
| TIMSS 2003 | 0.67 | 20.94 | 0.974 | −16.45 | 17.58 | 0.36 |
| Public exam | 19.11 | 26.29 | 0.472 | 20.37 | 26.73 | 0.45 |
| Grade 8 * Public exam | 13.22 | 8.36 | 0.124 | 8.92 | 7.28 | 0.23 |
| R-squared | | .14 | | | .16 | |
| | Science, Cyprus & Portugal have public exams | | | | | |
| Constant | 303.34 | 20.96 | 0.000 | 303.30 | 27.87 | 0.00 |
| Girl | −12.88 | 2.26 | 0.000 | −9.18 | 2.30 | 0.00 |
| Number of books at home | 26.14 | 1.88 | 0.000 | 24.94 | 2.55 | 0.00 |
| Grade 8 | 68.68 | 14.34 | 0.000 | 81.74 | 13.17 | 0.00 |
| TIMSS 2003 | −9.81 | 18.77 | 0.605 | −11.16 | 13.86 | 0.43 |
| Public exam | 18.86 | 20.71 | 0.369 | 28.08 | 21.51 | 0.20 |
| Grade 8 * Public exam | 9.54 | 9.36 | 0.316 | 2.95 | 7.98 | 0.71 |
| R-squared | | .17 | | | .17 | |
| Number of countries | | 33 | | | 34 | |
| Number of students | | 247,468 | | | 276,752 | |

Note. This table reports clustering-robust standard errors, using countries as level of clustering. The regression analyses use students' sampling probabilities as weights. Between countries, the weights give equal weight to each country. All models control for a set of student age dummy variables. Data set named 'Mostly TIMSS 1995' consists of 23 countries from TIMSS 1995 and 10 non-overlapping countries from TIMSS 2003. Data set named 'Mostly TIMSS 2003' consists of 24 countries from TIMSS 2003 and 10 non-overlapping countries from TIMSS 1995.

# Appendix

**Table A1. Number of Students, National Average Test Scores, and Public Exams: 24 countries from TIMSS 1995 and 10 non-overlapping countries from TIMSS 2003**

| | Number of Students | | Math Scores | | Science Scores | | Public Exams | |
|---|---|---|---|---|---|---|---|---|
| | Grade 4 | Grade 8 | Grade 4 | Grade 8 | Grade 4 | Grade 8 | Math | Science |
| Armenia | 5,672 | 5,695 | 459 | 480 | 451 | 457 | 1 | 1 |
| Australia | 6,486 | 7,214 | 512 | 506 | 538 | 514 | 1 | 1 |
| Austria | 2,578 | 2,694 | 525 | 517 | 538 | 528 | 0 | 0 |
| Belgium-Flemish | 4,712 | 4,970 | 549 | 539 | 524 | 515 | 0 | 0 |
| Cyprus | 3,350 | 2,913 | 467 | 449 | 445 | 430 | 0 or 1 | 0 or 1 |
| Czech Rep. | 3,267 | 3,326 | 533 | 535 | 530 | 545 | 1 | 1 |
| England | 3,046 | 1,776 | 477 | 482 | 524 | 521 | 1 | 1 |
| Greece | 3,044 | 3,984 | 457 | 461 | 468 | 468 | 0 | 0 |
| Hong Kong | 4,248 | 3,316 | 555 | 561 | 506 | 495 | 1 | 1 |
| Hungary | 2,885 | 2,755 | 516 | 515 | 507 | 526 | 1 | 1 |
| Iceland | 1,791 | 1,766 | 439 | 467 | 474 | 465 | 1 | 0 |
| Iran | 3,360 | 2,544 | 393 | 401 | 382 | 447 | 1 | 1 |
| Ireland | 2,871 | 3,075 | 515 | 505 | 512 | 504 | 1 | 1 |
| Israel | 1,820 | 1,320 | 497 | 502 | 479 | 498 | 1 | 1 |
| Italy | 4,282 | 4,278 | 504 | 486 | 521 | 488 | 1 | 1 |
| Japan | 4,266 | 5,102 | 564 | 575 | 549 | 545 | 1 | 1 |
| Korea | 2,812 | 2,920 | 577 | 574 | 572 | 535 | 1 | 1 |
| Kuwait | 3,700 | 1,568 | 369 | 324 | 373 | 389 | 0 | 0 |
| Lithuania | 4,216 | 4,569 | 534 | 503 | 519 | 519 | 1 | 1 |
| Moldova | 3,979 | 4,033 | 505 | 462 | 504 | 469 | 1 | 1 |
| Morocco | 4,234 | 2,865 | 356 | 388 | 332 | 388 | 1 | 1 |
| Netherlands | 2,216 | 1,950 | 546 | 516 | 532 | 530 | 1 | 1 |
| New Zealand | 2,419 | 3,682 | 463 | 485 | 502 | 496 | 1 | 1 |
| Norway | 2,191 | 3,265 | 465 | 483 | 502 | 500 | 1 | 0 |
| Philippines | 4,542 | 6,870 | 367 | 379 | 357 | 367 | 1 | 1 |
| Portugal | 2,850 | 3,389 | 440 | 430 | 450 | 453 | 0 or 1 | 0 or 1 |
| Russian Fed. | 3,963 | 4,667 | 531 | 510 | 531 | 513 | 1 | 1 |
| Scotland | 3,286 | 2,815 | 486 | 477 | 509 | 485 | 1 | 1 |
| Singapore | 7,125 | 4,636 | 591 | 605 | 520 | 574 | 1 | 1 |
| Slovenia | 2,532 | 2,704 | 519 | 519 | 519 | 530 | 1 | 1 |
| Taiwan | 4,661 | 5,379 | 562 | 589 | 553 | 574 | 1 | 1 |
| Thailand | 2,985 | 5,767 | 456 | 503 | 442 | 495 | 1 | 1 |
| Tunisia | 4,331 | 4,931 | 349 | 411 | 341 | 395 | 1 | 1 |
| United States | 7,296 | 7,082 | 510 | 476 | 540 | 498 | 0 | 0 |

**Table A2. Number of Students, National Average Test Scores, and Public Exams: 24 countries from TIMSS 2003 and 10 non-overlapping countries from TIMSS 1995**

| | Number of Students | | Math Scores | | Science Scores | | Public Exams | |
|---|---|---|---|---|---|---|---|---|
| | Grade 4 | Grade 8 | Grade 4 | Grade 8 | Grade 4 | Grade 8 | Math | Science |
| Armenia | 5,672 | 5,695 | 459 | 480 | 451 | 457 | 1 | 1 |
| Australia | 4,300 | 4,789 | 500 | 507 | 526 | 527 | 1 | 1 |
| Austria | 2,578 | 2,694 | 525 | 517 | 538 | 528 | 0 | 0 |
| Belgium-Flemish | 4,712 | 4,970 | 549 | 539 | 524 | 515 | 0 | 0 |
| Cyprus | 4,323 | 4,001 | 510 | 461 | 490 | 435 | 0 or 1 | 0 or 1 |
| Czech Rep. | 3,267 | 3,326 | 533 | 535 | 530 | 545 | 1 | 1 |
| England | 3,582 | 2,827 | 531 | 501 | 544 | 545 | 1 | 1 |
| Greece | 3,044 | 3,984 | 457 | 461 | 468 | 468 | 0 | 0 |
| Hong Kong | 4,607 | 4,966 | 572 | 590 | 545 | 558 | 1 | 1 |
| Hungary | 3,319 | 3,302 | 528 | 532 | 534 | 544 | 1 | 1 |
| Iceland | 1,791 | 1,766 | 439 | 467 | 474 | 465 | 1 | 0 |
| Iran | 4,342 | 4,917 | 396 | 412 | 430 | 448 | 1 | 1 |
| Ireland | 2,871 | 3,075 | 515 | 505 | 512 | 504 | 1 | 1 |
| Israel | 1,820 | 1,320 | 497 | 502 | 479 | 498 | 1 | 1 |
| Italy | 4,282 | 4,278 | 504 | 486 | 521 | 488 | 1 | 1 |
| Japan | 4,202 | 4,822 | 564 | 573 | 548 | 554 | 1 | 1 |
| Korea | 2,812 | 2,920 | 577 | 574 | 572 | 535 | 1 | 1 |
| Kuwait | 3,700 | 1,568 | 369 | 324 | 373 | 389 | 0 | 0 |
| Lithuania | 4,216 | 4,569 | 534 | 503 | 519 | 519 | 1 | 1 |
| Moldova | 3,979 | 4,033 | 505 | 462 | 504 | 469 | 1 | 1 |
| Morocco | 4,234 | 2,865 | 356 | 388 | 332 | 388 | 1 | 1 |
| Netherlands | 2,936 | 3,035 | 539 | 539 | 530 | 537 | 1 | 1 |
| New Zealand | 4,307 | 3,799 | 495 | 496 | 525 | 519 | 1 | 1 |
| Norway | 4,342 | 4,133 | 455 | 463 | 477 | 492 | 1 | 0 |
| Philippines | 4,542 | 6,870 | 367 | 379 | 357 | 367 | 1 | 1 |
| Portugal | 2,850 | 3,389 | 440 | 430 | 450 | 453 | 0 or 1 | 0 or 1 |
| Russian Fed. | 3,963 | 4,667 | 531 | 510 | 531 | 513 | 1 | 1 |
| Scotland | 3,935 | 3,516 | 492 | 500 | 509 | 510 | 1 | 1 |
| Singapore | 6,668 | 6,018 | 590 | 609 | 566 | 581 | 1 | 1 |
| Slovenia | 3,069 | 3,578 | 481 | 495 | 499 | 520 | 1 | 1 |
| Taiwan | 4,661 | 5,379 | 562 | 589 | 553 | 574 | 1 | 1 |
| Thailand | 2,985 | 5,767 | 456 | 503 | 442 | 495 | 1 | 1 |
| Tunisia | 4,331 | 4,931 | 349 | 411 | 341 | 395 | 1 | 1 |
| United States | 9,829 | 8,912 | 518 | 507 | 539 | 527 | 0 | 0 |

# 校外公開考試是否提高學生學習表現? 運用雙重差分模型的跨國分析

黃敏雄

中央研究院歐美研究所副研究員

## 摘 要

實施校外公開考試制度是否會提高學生的學習表現？為了回答這個問題，筆者使用 1995 年與 2003 年的「國際數學與科學教育成就趨勢調查（TIMSS）」資料庫，以同時有四、八年級學生參與的國家為分析樣本，運用雙重差分迴歸模型來控制未觀測到的國家間差異，以有效評估校外公開考試對學生學習的影響程度。結果發現，有校外公開考試制度的國家，學生有較好的數學表現，但沒有較好的科學表現。雖然本研究發現校外公開考試對學生的數學表現有正面影響，但影響程度只有過去研究發現的一半。此外，當賽普勒斯與葡萄牙被歸類為有校外公開考試的國家時，校外公開考試對學生數學表現的正面影響就不再顯著。

關鍵字：校外公開考試、學生評量、雙重差分、學生學習表現、TIMSS